Metaphor and Relevance Theory: A New Hybrid Model

by

Hanna Stöver

A thesis submitted for the degree of Doctor of Philosophy
of the University of Bedfordshire

October 2010
Metaphor and Relevance Theory: A New Hybrid Model

by

Hanna Stöver

Abstract

This thesis proposes a comprehensive cognitive account of metaphor understanding that combines aspects of Relevance Theory (e.g. Sperber & Wilson 1986/95; Carston 2002) and Cognitive Linguistics, in particular ideas from Conceptual Metaphor Theory (e.g. Lakoff & Johnson 1980; Lakoff 1987; Johnson 1991) and Situated Conceptualization (e.g. Barsalou 1999; 2005). While Relevance Theory accounts for propositional aspects of metaphor understanding, the model proposed here additionally accounts for nonpropositional effects which intuitively make metaphor feel ‘special’ compared to literal expressions. This is achieved by (a) assuming a further, more basic processing level of imagistic-experiential representations involving mental simulation patterns (Barsalou 1999; 2005) alongside relevance-theoretic inferential processing and (b) assuming processing of the literal meaning of a metaphorical expression at a metarepresentational level, as proposed by Carston (2010).

The approach takes Tendahl’s ‘Hybrid Theory of Metaphor’ (2006), which also combines cognitive-linguistic with relevance-theoretic ideas, as a starting point. Like Tendahl, it incorporates the notion of conceptual metaphors (Lakoff & Johnson 1980), albeit in a modified form, thus accounting for metaphor in thought. Wilson (2009) suggests that some metaphors originate in language (as previously assumed by Relevance Theory) and others originate in thought (as previously assumed within Cognitive Linguistics). The model proposed here can account for both. Unlike Tendahl, it assumes a modular mental architecture (Sperber 1994), which ensures that the different levels of processing are kept apart. This is because each module handles only its own domain-specific input, here consisting of either propositional or imagistic-experiential representations. The propositional level, which remains the dominant processing route in utterance
understanding, as in Relevance Theory, receives some input from the imagistic-experiential level. This is mediated at a metarepresentational level, which turns the imagistic-experiential representations into propositional material to be processed at the inferential level in the understanding of literal expressions. In metaphor understanding, however, the literal meaning is not processed as meaning-constitutive content. As a result, the imagistic-experiential aspects of the literal meaning in question are not processed as propositional input. Rather, they are held at the metarepresentational level and experienced as strong impressions of the kind that only metaphors can communicate.
Contents

Chapter 1: Introduction 8
1.1. Relevance Theory and Cognitive Linguistics: A comparison 12
1.2. The proposal 16
1.3. Thesis summary 18

Chapter 2: Metaphor in Relevance Theory 31
2.1. The main principles of Relevance Theory 32
2.2. Ad hoc concept construction 40
2.3. Adoptions from other approaches: Fodor and Barsalou 48
2.4. Modularity 52
   2.4.1. Fodor: The Modularity of Mind 53
   2.4.2. Sperber: The Massive Modularity Hypothesis 57
2.5. Issues with Relevance Theory 64
   2.5.1. Ad hoc concept construction in poetry:
       Issues with extended metaphors 65
   2.5.2. The nature of propositions:
       The ‘inferential fitness’ issue 73

Chapter 3: Alternative approaches 82
3.1. Recanati: Primary and secondary pragmatic processes 83
3.2. Corazza & Dokić: Incomplete propositions 87
3.3. Conceptual Metaphor Theory 89
   3.3.1. Lakoff & Johnson: ‘Metaphors We Live By’ 90
   3.3.2. Grady: Primary conceptual metaphors
       and metaphorical compositionality 94
   3.3.3. Barsalou’s mental simulations 97
   3.3.4. Refining the definition:
       Conceptual metaphors as schematic patterns 103
   3.3.5. The multilingual perspective:
       Applications in second language acquisition 108
### Chapter 4: Language and thought

4.1. Linguistic Relativity
- 4.1.1. The history of linguistic relativism
- 4.1.2. Spatial reasoning
- 4.1.3. ‘Thinking for Speaking’
- 4.1.4. Metaphor and gesture

4.2. Awareness of metaphoricity
- 4.2.1. The role of the literal meaning
- 4.2.2. The literal meaning in metaphor processing:
  - A psycholinguistic study
- 4.2.3. The multilingual perspective:
  - Metaphorical translation asymmetries

### Chapter 5: Combining approaches

5.1. Steen: Symbolic and behavioural approaches
- 5.1.1. The three-dimensional model
- 5.1.2. Combining the three-dimensional model with Relevance Theory?
- 5.1.3. The five-step method
- 5.1.4. Combining the five-step method with Relevance Theory?

5.2. Tendahl’s Hybrid Theory of Metaphor

### Chapter 6: The new hybrid model:

- **A relevance-theoretic perspective**
- 6.1. A hybrid model within a modular framework
- 6.2. Carston: The metarepresentational level
- 6.3. Resolving issues with Tendahl’s model
- 6.4. Representational formats
- 6.5. Conceptual metaphors in the new hybrid model
- 6.6. Processing routes
Acknowledgements

I am grateful to the Arts and Humanities Research Council UK for awarding me the scholarship that made this thesis possible.

Many thanks to Adrian Huge whose support enabled me to come to the UK and go to university in the first place.

Most of all, I would like to thank my supervisors Billy Clark and Vladimir Žegarac for all their generous and wonderful support, advice, help and encouragement.

I would also like to thank Robyn Carston for providing her latest ideas on metaphor, which served as substantial input for the thesis.

Many thanks to Alan Cienki for inspiring discussions.

And a special thanks to Deirdre Wilson for believing in my abilities.
Chapter 1
Introduction

The main aim of this thesis is to propose an account of metaphor comprehension which explains metaphor in language, metaphor in thought, and how the two are related. The model combines ideas from Relevance Theory (e.g. Sperber & Wilson 1986/95; 1987; Carston 2002) about language processing and the (modular) nature of cognition, and Conceptual Metaphor Theory (e.g. Lakoff & Johnson 1980; Lakoff 1987; Johnson 1991), about metaphor in thought. While the model is based on the notion that comprehension is inference-driven, it integrates ideas about nonmodular associative simulation (e.g. Barsalou 2005) by assuming the involvement in comprehension of propositional as well as imagistic-experiential representations. This is made possible by the integrative properties afforded by a model of massive modularity (Sperber 1994).

An account of metaphor comprehension should be able to explain a range of aspects. It should have the potential to explain how comprehension of different types of linguistic metaphor works, such as novel and creative metaphors as well as conventionalised, or even lexicalised metaphors. It should be a cognitive account that is able not only to explain how linguistic metaphorical meaning is worked out in context, but also to include potential ways of metaphorical thinking in the explanation. It should thus cater for metaphor in language as well as metaphor in thought, and how the two relate to each other in comprehension. Further, it should account for the nonpropositional effects of a metaphorical expression as well as its propositional meaning in context. That is, it should explain, at least to a certain extent, how metaphors are experienced at a sensory level, as compared to literal expressions which, intuitively, are less intensely felt. Consequently, it should also explain how such nonpropositional effects may impact to a certain degree on the proposition expressed.

Pilkington (2000) suggests that the focus within Relevance Theory on propositional representations limits the extent to which it can explain how we qualitatively experience poetic metaphors. Thus, intensely felt sensory impressions which accompany the comprehension of poetic, or more generally
creative, metaphorical expressions (often termed ‘qualia’ in the philosophical literature), and which often seem to impact on the proposition expressed, do not seem to receive the attention they deserve. In my MA dissertation (Stöver 2006) I investigated the issue, and made some observations about the strongly felt impressions which accompany the comprehension of extended creative metaphors in particular. I suggested that semantic priming processes play a significant role in the process, since referring to related concepts in succession intensifies how they are experienced. While this is not a particularly novel finding, it does offer a limited and partial answer to the question of how to account for nonpropositional aspects, such as impressionistic feelings, within a propositional account of metaphor comprehension. This thesis proposes a way of including nonpropositional representations in a relevance-theoretic account. This may lead to a solution to the problem, in that nonpropositional effects are allocated a role that is more than that of a mere side effect.

Developing an account of metaphor understanding that can fulfil all of these requirements is an ambitious endeavour. I have approached it by looking at a range of approaches from different schools of thought and focusing on two models which each fulfil part of the requirements of a comprehensive account, and which, taken together in an appropriate way, might lead to a comprehensive account. These are Relevance Theory and, from the field of Cognitive Linguistics, Conceptual Metaphor Theory. Combining ideas from these different schools of thought was originally suggested by Markus Tendahl (2006). This thesis takes Tendahl’s proposed ‘Hybrid Theory of Metaphor’ (2006) as a starting point and proposes a new model which also seeks to get the best out of both theoretical points of view. The model proposed here develops Tendahl’s proposal further by observing that the different representational formats thought to be involved in comprehension are each allocated their appropriate processing mechanisms. This is made possible by a modular framework. Furthermore, while assuming relevance-theoretic comprehension procedures in inferential terms, the model accounts for the intuitive impression that metaphor is ‘special’, by way of envisaging a fine-tuned interplay between different levels of processing. This idea of a ‘specialness’ of metaphor is explained in the following.
There are various definitions of ‘metaphor’ around, ranging from a label for specific rhetorical devices to a term for a way of thinking, or expressing oneself in art, etc. Thus it is important to define what I mean when I claim that the subject of this thesis is ‘metaphor’. Traditionally, literal meaning in language has been seen as the norm, with metaphor playing the role of an unusual and deviant rhetorical device. In recent decades, it has been acknowledged that metaphor is ubiquitous in everyday language use (cf. e.g. Gibbs 1994), and claiming that metaphorical expressions fall out of the norm would simply be inaccurate. Hence the more recent descriptions of metaphorical use of language, in particular in Cognitive Linguistics and in Relevance Theory, have taken care to emphasise that views rooted in traditional rhetoric, which claim that metaphor is deviant, do not reflect the actual state of affairs. Thus, the message has been that metaphor is not special, no more than any other use of language. With the backdrop of this historical development, I feel it is time to take a less polarised view, taking for granted that metaphor is not a deviant form of language use and focusing on what it might be that makes metaphor different from other forms of figurative language, such as hyperbole, proverbs, alliteration, assonance, onomatopoeia, etc. All these forms of expression have evocative power and create mental imagery and other sensory mental experiences. And while the traditional view of metaphor as a deviant rhetorical device is, arguably, no longer valid, it seems that it must have developed for some reason, and that this must have been to do with how people felt about metaphor. Thus, this thesis is not intended to re-establish the traditional view, but rather to investigate what it is that creates the widespread intuition that metaphor is ‘special’, and I shall argue that it is the relative degree of ‘metaphoricity’ that metaphorical use of linguistic expressions creates. This notion is central to this thesis, and here defined in terms of a relationship between an encoded concept and interpretation in a specific context.

The notion of metaphoricity is relevant, for example, in cases where we assume that a word with a literal encoded meaning, that is, conventional in a given speech community, which represents a real-world referent (e.g. the word pig refers to a pink/beige farm animal with a short trunk and curly tail), and that this word is used to refer to something that is not an actual instance of that encoded
referent (e.g. a human rather than a pig). Then we have a difference between the encoded concept and the communicated concept. So when pig does not refer to the animal but to a person with certain stereotypical pig-like attributes, we may represent both concepts, that associated with the animal as well as that associated with a type of person. The metaphoricity comes about through an awareness of this difference between the two concepts at a metarepresentational level, and we can talk of a certain tension that is felt between these two representations. Cognitive linguists label the encoded meaning involved in a metaphorical expression ‘source domain’ and the communicated meaning ‘target domain’. In this thesis, it is seen to be of significance whether the meaning encoded in the literal expression comes to bear on the proposition expressed, i.e. the propositional meaning which the addressee takes as the communicated outcome. With expressions that are used in such a way that the latter is conceivably different from the encoded literal meaning, such that we can talk of a tension between the word meaning and the communicated meaning, we have a degree of metaphoricity, according to the definition used here. Coming from the philosophy of language, Elizabeth Camp (2008), for example, describes this as an ‘intuitively felt gap between literal and intended meaning, where the first provides the perspective for constructing the second.’ (2008: 14). Of course, it could be argued that other tropes, such as metonymy, may also create such a gap. For example, people regularly refer to certain politicians when talking of The White House. Here, there is a gap between the encoded meaning (a building that is white on the outside) and the communicated meaning (the politicians working there). But this is not a case of metaphoricity as defined here. Metaphoricity involves not only a gap between two meanings, both of which we are consciously aware of, but also another characteristic that makes metaphoricity, which I suggest has to do with the kind of relationship cognized between the two representations involved. Thus, an important aspect in creating metaphoricity is to take (often stereotypical) features from the source domain and to attribute them to the target domain, as when a person is called a pig, because they happened to drop their food, for example. Here, features relating to stereotypical pig behaviour, e.g. rolling around in the mud and thus making a mess, are attributed to that person. This is not the
case with metonymy, where, in our example, we do not attribute features of a white building to the politicians working there. Thus, it is assumed that the specific cognitive tension involved the processing of different representations and their metaphorical relationship to one another, i.e. their metaphoricity, play a role in making metaphor feel ‘special’, since we are aware of this tension and it accordingly has an effect on us. And while it remains to be empirically tested whether people in general, or just some people, perceive metaphoricity as it is outlined here, this thesis begins by investigating the theoretical implications of postulating such ‘specialness’.

An important aspect pertaining to the perceived specialness of metaphoricity is that we need to be aware of its use (while not necessarily having to be able to label it) and to metarepresent it accordingly. So it makes a difference whether a metaphorical expression is used ‘deliberately’ or not, where ‘nondeliberate’ use usually leads to no metaphoricity at all, and so the same linguistic expression can sometimes be a metaphor, as defined here, and sometimes not (Steen 2008). This aspect of awareness is closely related to a cognitive process of metarepresenting the relevant literal meanings of a metaphorical expression, which has been discussed by Carston (2009; 2010). These definitions will be elaborated on further below and constitute important elements of the proposal outlined in this thesis.

1.1 Relevance Theory and Cognitive Linguistics: A comparison

It has been established for some decades now that metaphors fulfil more functions than just being ornamental extras in literature or special rhetorical devices: they are used frequently in everyday life in ordinary acts of communication. This observation is shared by scholars in Relevance Theory and Cognitive Linguistics alike. Both schools of thought have a cognitive outlook and investigate how language works within a framework that looks at cognition as a whole. However, each is based on different principles and, although they might be looking at the same phenomena, they look at them from different angles. For this reason, neither cognitive linguists nor relevance theorists seem to have shown much interest in
each other’s work, simply because the assumptions underlying the two approaches seem too different. Only recently have some theorists within Cognitive Linguistics begun to compare the relevance-theoretic treatment of metaphor with that of Cognitive Linguistics, in particular with Conceptual Metaphor Theory, and Blending Theory (e.g. Fauconnier & Turner 1995), and they point out that the apparently incompatible approaches might be complementary instead (Tendahl 2006, Gibbs & Tendahl 2006, Tendahl & Gibbs 2008). This in turn has been discussed by relevance theorist Deirdre Wilson (2009). A major difference between the two schools of thought is that approaches in Cognitive Linguistics focus on the role metaphor may play in cognition, and they understand metaphor to be a way of thinking which finds expression in language. By contrast, relevance theorists see metaphors as arising in communication, where they are employed to convey thoughts, which are not necessarily metaphorical but may be most conveniently expressed with a metaphor (e.g. Wilson 2009: 2).

In comparing Relevance Theory (e.g. Sperber & Wilson 1986/95; Carston 2002) and Conceptual Metaphor Theory (e.g. Lakoff and Johnson 1980; Johnson 1987; Lakoff 2008), Tendahl & Gibbs (2008) observe that the respective approaches look at different types of linguistic metaphor: Relevance Theory traditionally focuses on nominal metaphors of the form ‘X is a Y’ (‘Robert is a bulldozer’, ‘You are the cream in my coffee’). By comparison, in Conceptual Metaphor Theory, and in Cognitive Linguistics in general, metaphor is seen to underlie almost all language, even expressions which for most people seem to be literal. Such expressions, e.g. ‘I’m feeling down’, are in Cognitive Linguistics understood to be metaphorical because etymologically they are assumed to be metaphorically motivated. That is, their metaphorical meaning has lexicalised over time, and it seems to be a matter of viewpoint whether we label them as metaphorical (because of their origin) or polysemous (because they are not perceived, by ordinary speakers, as metaphorical). However, in Cognitive Linguistics, the story goes that even an expression like ‘I’m up for it’ is cognitively processed as a metaphor, even though, to my knowledge, this remains to be empirically proven. Nevertheless, the basic idea in Conceptual Metaphor Theory, namely that cognition is structured by patterns of underlying mappings
between domains (conceptual metaphors, explained in section 3.3), is an insightful assumption which, if taken moderately, deserves investigation. The focus in Relevance Theory, on the other hand, is on how the hearer works out metaphorical meaning in communication, and this usually involves looking at novel metaphors, or at least metaphors that are conventional up to only a certain degree, that is, non-lexicalised metaphors (e.g. ‘Sally is a block of ice’). It seems that not much attention has been devoted to the comprehension of metaphors at the very conventional, or even lexicalised, end of the spectrum, which is, by contrast, the main focus in Conceptual Metaphor Theory, with view to the assumed underlying metaphorical thought patterns. In Relevance Theory, the idea is that pragmatic mechanisms are at work, constrained by the presumption of optimal relevance (explained in section 2.1), to get the communicated meaning across. What the underlying thought patterns may be, as concerns the content of the thought that gets communicated, does not lie in the scope of explaining how it is communicated.

However, Tendahl & Gibbs (2008) point out that Relevance Theory presents itself as a theory of cognition as such, and suggest as a criticism that in spite of this, there seems to be a tendency in Relevance Theory to view cognition as composed of unrelated instances of utterance understanding. The emphasis, they claim, is on metaphor as a purely linguistic phenomenon, without paying adequate attention to the more holistic aspects of the human mind. Therefore Tendahl & Gibbs argue that it would make sense for Relevance Theory to put stronger emphasis on ways of thinking, such as analogical thinking, when it comes to explaining linguistic metaphorical phenomena. This would include acknowledging the wider networks of metaphorical expressions as discussed in the cognitive-linguistic literature, rather than viewing metaphorical expressions more or less in isolation from one another. Conceptual Metaphor Theory, by comparison, can cater for this as it puts a strong emphasis on the perceived semantic connectedness of various metaphors.

At first glance, it might appear that Relevance Theory is better suited to explaining the comprehension of novel metaphors, while Conceptual Metaphor Theory could be seen as adequately accounting for conventional metaphors. It is
possible to think that each approach will be applicable to a relatively greater or lesser extent, corresponding to the degree of metaphoricity and/or lexicalisation of the linguistic expression at hand. Arguably, there is a continuum of types of metaphorical expression running from literal at one end, via conventional (‘frozen’ or ‘dead’) metaphorical (e.g. ‘I’m feeling low’, where ‘low’ does not denote a physical position in space but is a metaphor for a negative psychological state), idiomatic expressions (e.g. ‘That was the straw that broke the camel’s back’), to novel/creative metaphorical expressions (e.g. ‘My heart is made of marzipan’). Etymologically, it might be assumed that if an expression is often used metaphorically, the metaphorical meaning can lexicalise and become a further encoded meaning of the expression, usually alongside the original literal meaning, resulting in polysemy. For example, the English word cool now has a psychological meaning alongside its physical meaning, and the word crane now has a meaning referring to a machine used in construction work alongside its BIRD meaning (Crystal 1997: 332). We can find expressions whose metaphorical meaning is lexicalised to different degrees, i.e. such a meaning can be more or less accessible in memory as a ‘fixed expression’.

However, it seems that the main difference between the two approaches does not pertain to whether one approach can explain conventional metaphors and the other approach novel metaphors. While it goes without saying that each approach claims to be able to explain all types of metaphors, the main difference between them, as observed e.g. by Wilson (2009), lies in their assumptions concerning the origin of metaphors: Conceptual Metaphor Theory assumes the origin to be in thought, while Relevance Theory assumes it to be in language use. Wilson tentatively suggests that it could be that we actually have both types of metaphor, some arising in language and others in thought (2009: 3), since some metaphorical expressions become conventional and may lexicalise through frequent use (origin in language), and others might surface from cognitive patterns that result from metaphorical thinking, as proposed in Conceptual Metaphor Theory (origin in thought). This thesis proposes an account that aims to explain both types.

---

1 I follow the convention of using small capital letters for conceptual content.
1.2 The Proposal

The account presented here is partly based on Tendahl’s (2006) proposal, which combines cognitive-linguistic approaches with relevance-theoretic elements. Inspired by this move of taking the complementary potential of the approaches seriously and consequently incorporating both in an inclusive approach, I also develop a hybrid model of metaphor comprehension that combines elements from Cognitive Linguistics, in particular Conceptual Metaphor Theory, with Relevance Theory. However, my approach has its starting point in Relevance Theory, rather than Cognitive Linguistics. I take a different route from Tendahl, who adopts the relevance-theoretic comprehension heuristic based on the presumption of optimal relevance, plus a version of the relevance-theoretic notion of ad hoc concept construction as additions to a Cognitive Linguistics model in order to enhance the latter. The problem I see with this is that approaches rooted in Cognitive Linguistics, including Tendahl’s, are based on the assumption that comprehension takes place employing general reasoning abilities, while relevance-theoretic comprehension mechanisms employ inferential procedures that best work within a mental architecture based on modularity (e.g. Sperber 1994). There are several other aspects of Tendahl’s model which, from a relevance-theoretic perspective, require some scrutiny. These will become apparent in due course.

Nevertheless, this thesis proposes combining Relevance Theory with ideas from Cognitive Linguistics approaches, because the latter offer advantages which Relevance Theory can benefit from. As Wilson mentions, it might just be that some metaphors arise in language, while others arise in thought. Relevance Theory can account for the former, and Conceptual Metaphor Theory for the latter, so it should be fruitful to have both. Furthermore, the model proposed here introduces elements from approaches that favour the simulation view of processing, based on the notion of mental re-enactment (Barsalou 1999; 2005; 2009). These can incorporate all aspects of human experience, which has the advantage that nonpropositional aspects, such as sensory impressions, feelings, etc. can be accounted for in an inclusive way. Approaches within Relevance Theory could benefit from explaining such effects more holistically.
Therefore, I propose a hybrid model which adheres to relevance-theoretic principles, such as modularity and the assumption that in inferential processing propositional representations are the dominant cognitive structures. Nonpropositional representations are here understood to be subservient to the former. These principles ensure a model rooted in rationality, such that propositional (or factual) representations are clearly kept apart from the more impressionistic and associationist elements in the mind. Thus, one thing which sets my approach apart from that of Relevance Theory is that I propose a further route of cognitive processing based on simulation processes as outlined by the psychologist Lawrence Barsalou (2005; 2009). This involves imagistic-experiential representations, including conceptual metaphors, as proposed in Conceptual Metaphor Theory. My account is inclusive of metaphorical thought patterns, and is able to account for metaphors of both linguistic and cognitive origin. The notion of conceptual metaphors, however, undergoes a redefinition (or a refinement) in my account to be moderately understood as very schematic patterns based on embodiment, which have a low-level guidance function in cognition. I consider this redefinition of imagistic-experiential representations to be a way of introducing nonpropositional effects, which, although they do not influence logical thinking directly, still deserve to be assigned a role in cognition. After all, they appear to be involved in our intuition that metaphor is ‘special’ due to a felt intensity that metaphors often evoke. In addition, my account conceives of conceptual metaphors as experiential representations which belong to this category of nonpropositional representations. They may be accessed in comprehension, without interfering with the propositional representations, i.e. those representations that can be evaluated for truth, and may thus be a cause of some metaphors arising in thought.

While highlighting the role of additional, more intuitive knowledge structures within a broadly relevance-theoretic model is influenced by Cognitive Linguistics in general and Tendahl’s model in particular, the idea of dual processing has been inspired by Robyn Carston (2009; 2010). She proposes processing of metaphorical expressions on two levels, whereby the literal meaning is held mentally on a metarepresentational level during the course of
comprehension, with the function of informing the outcome of figurative interpretation. This seems likely, and it also supports the intuition that metaphors are experienced cognitively in an ‘intense’ way, as this partly explains the imagistic quality of metaphors. However, Carston’s proposal still remains in the propositional realm, as is typical of Relevance Theory. Thus, I propose that, while in agreement with Carston’s employment of a second processing route at the metarepresentational level, we must also have a nonpropositional level in addition, to further cater for the ‘special effects’ that experiencing metaphor brings with it. Therefore I assume that, on the one hand, the two different formats (propositional versus imagistic-experiential) are processed in different ways (inference for the former and simulation for the latter), and a metarepresentational module that mediates between the processes, on the other hand. This in turn can be explained with the notion of massive modularity (e.g. Sperber 1994; see also Sperber 2001; Carruthers 2006; Hirschfeld & Gelman 1994; Samuels 1998), where the metarepresentational module specialises in representations of representations.

1.3 Thesis summary

A large part of the present proposal is based on relevance-theoretic principles and ideas. Therefore, chapter 2 is devoted to explaining how Relevance Theory explains communication and cognition, and in particular how it deals with metaphor understanding. Section 2.1 explains the cognitive principles according to which communication is here understood: a balancing out of cognitive effort and effects, which is driven by the presumption of optimal relevance. This is linked to the notion of implicatures in communication, which are derived by a hearer as more or less strongly implicated meanings of an utterance. Thus, rather than conceiving of communication as encoding and decoding meaning, in Relevance Theory the communicated meaning is worked out by the hearer, whereby the speaker provides the hearer with the required indications as to the meaning intended. Metaphorical expressions are here understood as a variety of loose use, and the meaning is thought to be weakly implicated. This leads to a
wide array of weak implicatures, as opposed to one or very few strong implicatures that are communicated with literal expressions. These many weak implicatures lead to poetic effects, which are thus derived at a propositional level: all the potential meaning outcomes, vaguely communicated by a metaphor, are arrived at employing inferential mechanisms that are logically plausible. This is an important point in comparison with cognitive-linguistic approaches, which explain metaphor on the basis of holistic association, and where the condition that the input to inference needs to be propositional is not given. Thus, cognitive-linguistic approaches do not appear to explain how comprehension output can be accurate regarding true facts in the world. However, their account of nonpropositional effects seems to be foregrounded due to an inclusion of all modalities, whereas in Relevance Theory this aspect is not as highlighted, due to being rooted in propositionality.

Section 2.2 outlines the more recent relevance-theoretic treatment of metaphor understanding, involving the notion of ad hoc concept construction. It aims to explain the local cognitive steps involved in loose use, employing an interplay of implicatures and explicatures in mutual parallel adjustment. Here, the encoded concept of an expression is adjusted such that it fits the context, and the more figuratively an expression is used, the more the encoded meaning gets modified in the process. This will be illustrated with step-by-step examples.

Section 2.3 discusses how Relevance Theory has drawn on elements from other theories and how these elements have been developed to fit the overall purpose of a theory of cognition and communication. The motivation to follow up the theory’s historical influences in this way is to show how it has been possible to integrate elements of others that seem only to fulfil local requirements but might not fit with the respective overall principles. The hybrid model developed in this thesis is based on the idea that it is possible to integrate apparently disparate ideas for the purpose of an overall more comprehensive outcome, in favour of an approach that can explain more aspects of the phenomena investigated. To support this view, productive ways of mutual fertilisation are outlined here, in particular with Relevance Theory, since it is the approach out of which the hybrid model has been developed to a large part.
Relevance Theory has been influenced by some of Jerry Fodor’s (1975; 1981; 1983) and Lawrence Barsalou’s ideas (1983; 1987). Thus, section 2.3.1 briefly explains Fodor’s notion of ‘atomic concepts’ and outlines his model of cognition in the medium of a ‘Language of Thought’ (sometimes also called ‘mentalese’), the latter playing a substantial role in Relevance Theory in general, and in the discussion to unfold in this thesis, in particular. Further, I explain how Barsalou’s work on ad hoc categories influenced Relevance Theory, despite the fact that the main parts of his theory, developed further in later years (e.g. Barsalou 1999; 2005), do not seem to be compatible with relevance-theoretic principles, at least at first glance. His overall approach does not have much in common with the relevance-theoretic framework, but it will turn out that in my own framework, outlined in chapter 6, it is possible to conceive of the two as compatible after all.

Section 2.4 outlines the model of a modular mental architecture, which the relevance-theoretic account is based on. It assumes that communicative mechanisms are carried out by a mental module, which assures automatic and speedy processing. This is in sharp contrast to the mental architecture assumed in Cognitive Linguistics, which is nonmodular and holistic in that communication is thought to function by employing general reasoning capacities. Tendahl (2006) rightly points out that, if we want to merge these two approaches, we will have to decide on either a modular or a nonmodular mental architecture within which to combine elements of the approaches. A merger of the two mental architectures themselves is not possible, since they are mutually exclusive in principle. Tendahl chooses the nonmodular picture, while by contrast the account proposed here will be developed within a modular framework. Tendahl’s main argument against a modular architecture is that it is, apparently, not flexible enough to accommodate metaphorical thinking. He seems to assume that mental modules are only conceivable in the way Fodor (1983) describes them. Fodor’s model is described first, in section 2.4.1. However, as the later description of massive modularity (Sperber 1994) in section 2.4.2 makes clear, it offers a highly flexible framework, within which information flow in all directions is possible. At the same time, it also endows the organism with the constraints it needs in order to keep different types of representation apart. This will be the key argument supporting the idea of
nonpropositional as well as propositional representations to be, directly or indirectly, involved in comprehension. With a general reasoning device it is questionable how different types of processing (such as deductive and non-demonstrative inference versus simulation processes) are kept separate, as well as how the mind would deal with the enormous stream of information that is dealt with in communication.

Section 2.5 takes a step back from relevance-theoretic assumptions in order to give room to a more critical view. Section 2.5.1 addresses weaknesses of the ad hoc concept construction account, in particular that it appears inherently difficult to apply it to more complex and extended metaphors (e.g. Carston 2009, 2010; Tendahl & Gibbs 2008). The issues that surface in analysis are demonstrated by applying the approach to a poem by John Donne, which contains complex metaphorical conceits, and further by showing the expressiveness which can be achieved with extended metaphors, exemplified by a Shakespearean monologue (Lady Macbeth). These analyses show that ad hoc concept construction can neither account for the relationships in meaning between the constituents of complex metaphorical conceits, nor for the extra layer of meaning that seems to be carried by metaphors that extend over longer stretches of text, involving an underlying conceptual domain. The latter seem to convey more meaning than the sum of the single metaphors involved, which seems to require a different kind of analysis altogether. It is then suggested that a version of conceptual metaphor might be the right candidate to cater for these phenomena. Alternatively, Carston’s (2009; 2010) recent contemplations on the nature of metaphor processing could be the way forward here. As will become clear in later chapters, both suggestions appear adequate as well as compatible. Thus, it will be shown that a combination of approaches can account for metaphorical complexity and the nonpropositional effects that come about in poetic metaphor.

Section 2.5.2 aims to shed some light on issues relating to the cognitive processes involved in communication, in particular regarding how inference and associative processes are thought to work. Convictions differ between scholars, depending on underlying assumptions about what kinds of content get worked out in communication, and how. In Relevance Theory, inferential processes require
propositional input, and this section considers what requirements a representation has to fulfil to count as propositional (sometimes referred to as ‘inferential fitness’). This relates to the question of how much content is encoded in a word, if any, before it is processed, and how to determine which contents are simply accessed in comprehension, and which are pragmatically worked out. Notions like linguistically encoded meaning (LEM) (Carston 2008) and the varying definitions of ‘what is said’ (e.g. Recanati 2004) are explained and discussed. The discussion serves the purpose of providing an understanding of how the finer mechanics in Relevance Theory are thought to work. This is followed by critical contemplations about these mechanisms from a relevance-theoretic perspective (Carston 2008) and a camp that seems not as remote to Relevance Theory as Cognitive Linguistics, sharing a similar pragmatics perspective (Recanati 2002; 2004). The purpose of this excursion into the debate about ‘what is said’ and ‘inferential fitness’ is to indicate that it may well be possible to conceive of communicative and cognitive processes in various ways, and thus it may also possible to change established notions in order to create a new account that is flexible enough to include (but not to conflate) various notions.

Chapter 3 discusses approaches other than Relevance Theory, with a focus on Conceptual Metaphor Theory in section 3.3. First, in section 3.1, the discussion from the previous chapter is resumed and given more depth by outlining Recanati’s (2002; 2004) account. This account has many features in common with Relevance Theory, such as the assumption that comprehension is highly context-dependent and that much of the communicated meaning is worked out pragmatically. However, in contrast to Relevance Theory, Recanati’s account includes an association level. Thus, he assumes that at a first stage in comprehension, sentential constituents are arrived at associatively, preceding an inferential stage where the propositional meaning of the constituents is fleshed out situation-specifically. This is an example of how it is possible to envisage a pragmatic account that is not entirely based on deductive and non-demonstrative inference, which is of interest here in terms of developing the idea of an approach that incorporates different types of cognitive processing. In a similar vein, the suggestions of Corazza & Dokic (2007) described in section 3.2 pertain to the
‘inferential fitness’ issue in that they propose a different way of conceiving of propositionality, which ultimately has the potential to alter the ways in which we think that propositional representations function. They propose incomplete propositions, i.e. representations that function in cognition without being fully defined. This points to the idea that in general, what can count as input for cognitive processes may not need to be as rigidly defined as has often been assumed, paving the way for conceiving of cognition from new perspectives.

In outlining Conceptual Metaphor Theory, the main ideas introduced in Lakoff and Johnson’s canonical work ‘Metaphors we live by’ (1980) are described and critically discussed in section 3.3.1. A major issue with conceptual metaphors as originally proposed was that, while they were thought to reflect basic patterns of thinking, the ‘basicness’ of these representations as such was not at all clearly defined. This problem was remedied, to a certain extent, by Joseph Grady (1997), who proposed that the most basic conceptual metaphors are combined in a compositional fashion, to make up less basic, and more culture-specific, conceptual metaphors (section 3.3.2). While this may partly explain the issue, I am not entirely convinced that this is the whole story, but rather that many of these more complex conceptual metaphors might actually not refer back to more basic conceptual metaphors. They might not be rooted in this way but are likely to have come about through frequent use of the same linguistic expressions, i.e. they may have their origin in language rather than in thought. Metaphors that have their origin in thought, on the other hand, presumably go back to very basic experiential schematic patterns, originating in embodiment. Thus, section 3.3.4 proposes a refinement of the notion of conceptual metaphor, which is based on Barsalou’s (1999; 2005; 2009) model of mental simulation and conceptual re-enactment, outlined in section 3.3.3. Interestingly, the latter is the part of Barsalou’s work that seems not have met an interest with relevance theorists, which is not surprising, since it is nonmodular. However, in the model proposed in this thesis, Barsalou’s processes of simulation can be implemented in a modular model by conceiving of them as a subpart of the overall cognitive structure. This move is the key to introducing nonpropositional elements as an active part in comprehension, as will be explained in chapter 6. In order to supply the
theoretical discussion with an empirical element, and to support the notion of conceptual metaphors as potentially psychologically real, section 3.3.5 reports on ways in which it has been fruitfully applied as a tool in second language acquisition (Boers 2000), and a proposal of how awareness of underlying conceptual metaphors could systematically facilitate acquisition of grammatical case in second language learning (deKnop 2008). Further, a study on prepositional meaning incongruities in second language learning (Ijaz 1986) demonstrates how the teaching and learning of languages is often accompanied by the tacit assumption that learning another language involves thinking in different ways, along the lines that the grammar of the new language dictates. Ijaz calls this ‘conceptual restructuring’. This conception of the relationship between language and thought is known as linguistic relativism, and it is often criticised due to a lack of empirical evidence that there is such an influence of language on thought. Conceptual Metaphor Theory, in particular, is notorious among its critics for its ‘circularity’: in order to prove that language influences thought, it is not sufficient to give only linguistic examples of these apparent manifestations.

For the purpose of investigating this and related issues pertaining to the relationships between language and thought, chapter 4 presents research into linguistic relativity, and issues relating to conscious awareness of metaphor and metaphoricity (as defined above). The discussion is vital here, because an important aspect of the cognitive mechanisms that constitute parts of the model proposed in chapter 6 pertains to issues of conscious awareness of metaphoricity. This idea is based on questions of how language, thought and an awareness thereof interrelate. Thus, chapter 4 provides for relevant research context, and it plays an integral role in preparing the ground for the development of the proposed model, since the latter hinges on the relationship between language and thought.

Section 4.1 focuses on linguistic relativity and the inherent difficulty in finding conclusive evidence for the strong intuition many people have, that language ought to manifest itself in cognition. The main focus here is on spatial reasoning, which seems to lend itself well as an experimental field in the area (section 4.1.2). However, comparison between studies investigating the same issues reveal that results seem to depend largely on the kind of experimental setup
employed, so that the resulting evidence does not seem to support any claims with certainty. This is followed by a description of a more moderate version of linguistic relativism, Slobin’s (1996; 2003) ‘Thinking for Speaking’ (section 4.1.3). Here, the relationship between language and thought is understood in a rather direct way along the following lines: given that we prepare utterances in our head before we utter them, it is likely that we do this cognitive work so that the content to be communicated is readily ‘packaged’ for the specific language we speak.

While section 4.1.2 described some of the difficulties pertaining to finding conclusive non-linguistic evidence of an influence of language on thinking, section 4.1.4 is devoted to a field that seems more promising in that respect: gesture studies. It has been shown that people’s abstract spatial reasoning is reflected in the spontaneous gesturing accompanying their speech, including cultural differences in spatial reasoning (Cienki & Müller 2008). This type of gesturing has also been observed to occur independently of language, and is thus understood as non-linguistic evidence of cognitive structuring. Furthermore, these observations are seen as evidence of underlying conceptual metaphors, in the form of ‘metaphoric gestures’, such as those expressing time-space mappings, e.g. THE FUTURE IS AHEAD. Thus, while different combinations of co-occurrences of metaphor and speech have been observed, metaphoric gestures may also occur without any co-occurring metaphorical language. Cienki (2000) takes this line of thinking a step further by analysing data from spontaneous gesturing and speech, making the observation that in some cases, metaphoric gestures precede related speech, and the metaphor (presumably) underlying the gesture is taken up in speech, initiating a mutual reinforcement of a metaphorical pattern that supports the development of the idea that is being expressed. The analysis is based on Slobin’s (1996) ‘Thinking for Speaking’ (outlined in section 4.1.3), with the addition that, in preparing an idea for expression in speech, metaphoric gestures are incorporated on the assumption that they have an influence on the process. Closely related to what will be developed as the core part of my proposal in chapter 6, the important aspect here is the apparent interplay of imagistic-experiential reasoning, as expressed in metaphoric gestures, and ‘thinking proper’,
which involves inference, i.e. the derivation of warranted conclusions, as expressed in language.

Section 4.2 discusses whether, or how, an awareness of the role of the literal meaning of a metaphor (the vehicle), plays a role in comprehension. Section 4.2.1 discusses what Carston (2009; 2010) calls ‘the lingering of the literal’, which could be described as a gap we perceive, often consciously, between the metaphor source and its target. It seems to be a case of continuous activation of the literal meaning during processing, past the point of its contribution to the propositional meaning outcome. This has been shown in a psycholinguistic study (Rubio 2008), which is described in section 4.2.2. To illustrate these processes, section 4.2.3 describes how comprehension might go with examples of translation asymmetries, i.e. expressions that are used metaphorically in one language but have lexicalised in another. For the second language learner, this state of affairs often results in a conscious awareness of the metaphoricity of a lexicalised expression, which the native speaker might not perceive. Hence these examples demonstrate how an awareness of the metaphorical origin of an expression in a second language, compared with a lack thereof in a first language, can be seen to reflect how conscious awareness of metaphoricity may impact on comprehension. This awareness can be assumed to be mentally represented in an imagistic way, i.e. a nonpropositional ‘experiencing’ of the literal meaning. This supports the idea that an account of metaphor needs to be inclusive of different types of representation; that it ought to explain how comprehension proceeds at different levels, in addition to the propositional level.

Chapter 5 discusses attempts at combining cognitive-linguistic approaches to metaphor with Relevance Theory, or elements of the latter. The discussion is intended to lead up to my own proposal of a combination of accounts in chapter 6. Section 5.1 is devoted to describing some of the work of the cognitive linguist Gerard Steen (2007; 2008; 2009; 2010), and a discussion of how it might be possible to integrate elements of Relevance Theory into his framework. It shows how this endeavour comes up against some difficulties with integrating relevance-theoretic context-sensitive online-processing principles and Steen’s aim
predominantly to define what specific metaphors mean, as opposed to how an expression may become metaphorical in communication and how its meaning is worked out. That is, conflicting research questions seem to have led to contradictory elements within the attempted combinatorial approaches. The approaches discussed and, in particular, analysis of the theoretical conflicts involved, demonstrate the importance of observing an important aspect in merging different theoretical strands: there are theoretical constraints on combining different models. With such a merger we can only go as far as compatibility allows, which might not be given at all if the perspectives involved are based on research questions that are too disparate. Thus, in retrospect, the theoretical endeavours described in sections 5.1.2 and 5.1.4 provide some guidance with respect to the theoretical constraints observed in the proposal developed in chapter 6.

Steen (2008) contributes valuable insights with regard to metaphor in communication. He draws attention to the fact that metaphors are processed in different ways, depending on whether they are used deliberately or nondeliberately, which is an important point with regard to metaphor awareness as discussed in section 4.2. This is outlined in his three-dimensional model, described in section 5.1.1, which assumes an interplay between the levels of thought, language, and communication. Furthermore, Steen (2007) provides a comprehensive overview of different approaches to metaphor, where he draws a clear line between symbolic accounts of ‘language as such’ and behavioural accounts, which are concerned with language use. While it may seem that these different types of account are mutually exclusive, due to their different perspectives, Steen’s three-dimensional model aims to combine these perspectives, by way of redefining them as different ‘dimensions’ of the same phenomena. However, this seems to be more of a terminological change than a change in approach, and it appears that Steen’s approach is to be categorised as ‘symbolic’ after all, since it seems to assume relatively stable meanings based on predetermined metaphorical mappings. The only way in which the discourse context plays a defining role here is in terms of rhetoric, i.e. the distinction between whether a metaphor is used deliberately for rhetoric effect, or
nondeliberately to go relatively unnoticed in discourse. Section 5.1.2 describes an attempt to introduce the relevance-theoretic element of implicature derivation into the communicative dimension of Steen’s three-dimensional model, illustrated with reference to an example of a metaphorical expression. It demonstrates that, ultimately, an integration of symbolic and behavioural approaches is not possible due to their incompatibility and thus clearly demarcates a line where the combining of approaches comes up against theoretical constraints.

Section 5.1.3 describes Steen’s five-step method of identifying metaphorical content in language. It is based on the approach employed by the Pragglejaz Group (2007) to analyse large corpora with regard to metaphorical content. The five-step method has been developed in particular for longer stretches of metaphorically related text, such as poetry. Thus, it might be assumed that this could be a suitable alternative to the relevance-theoretic ad hoc concept construction account, which does not seem to lend itself very well to extended metaphors, as demonstrated in section 2.5.1. However, the five-step method, like the three-dimensional model, is built on symbolic principles with assumptions of ‘language outside use’, as it were, so that a merging of approaches is inherently difficult. An attempt at this, i.e. at enhancing the five-step method with an online-processing element in relevance-theoretic terms, is outlined in section 5.1.4, where it becomes very clear that the approaches are again incompatible. It is, however, worthwhile to embark on projects like this, even if they do not result in success. It enhances our knowledge of the range of perspectives on language and metaphor in general. In particular, the attempt suggests what the preconditions for merging approaches with different outlooks should be. That is, they need to have in common that they are either symbolic, as in Steen’s account, or behavioural, as in Relevance Theory, since our exercise has shown that a combinatorial model cannot be both symbolic and behavioural at the same time.

Section 5.2 outlines Tendahl’s (2006) combination of approaches, which appears to be a much more successful project. It therefore functions as a model case for the approach developed in chapter 6. It combines cognitive-linguistic representational formats, such as source and target domains, image schemas and conceptual spaces, while employing the relevance-theoretic principles of online
processing. Thus, it can account for context-sensitive ad hoc concept construction as well as domain mappings and the resulting conceptual metaphors. Since Tendahl seems to maintain that the relevance-theoretic comprehension procedure can be unproblematically implemented in a nonmodular mental architecture, and that inferential processes can take all kinds of representational format as input, there is nothing that seems to stand in the way of this merger. It can account for various types of metaphor, novel and conventional, short and extended, and also for nonpropositional effects, due to the employment of image schemas and other experiential representations. However, the model does not incorporate the difference between propositional and nonpropositional representations; they are all treated in the same way theoretically, and by extension are thought to be processed in the same way psychologically. Thus, the model ultimately does not distinguish between the proposition expressed and nonpropositional effects. This distinction is important, however, if we want to explain how an organism can tell the difference between true facts in the world and subjective impressions. Furthermore, it cannot account for the difference between metaphors that originate in language and those that originate in thought, since the distinction is not made in the first place. The model is, however, an inspiring piece of work that achieves a comprehensive understanding of how metaphor works from a cognitive-linguistic point of view. However, from a relevance-theoretic point of view, which is the initial position taken in this thesis, the model better adhered to the relevance-theoretic principles of modularity, and of inference as requiring propositional input. The new hybrid model, developed in chapter 6, does adopt a modular approach, which, unlike Tendahl’s model, provides the preconditions that make the relevance-theoretic comprehension heuristic psychologically plausible. Furthermore, it makes possible the incorporation of Barsalou’s simulation processes in addition to inferential processes. The former take representational input formats different from the latter, and, according to the modular approach, are kept separate, while they still provide each other with informational input.

Section 6.1 explains how a modular mental architecture is seen to function as the framework of the hybrid model developed here. Section 6.2 explains Carston’s (2010) recently developed approach to metaphor within Relevance
Theory, which involves the metarepresentational level as a second processing route. Thus, this approach aims to account for the ‘lingering of the literal’, whereby the encoded meaning of a metaphorical expression is ‘mentally held’ in front of the ‘inner eye’, and can thus accompany the proposition expressed without interfering at the explicit level. The new hybrid model incorporates this view as well as the nonpropositional representational level it introduces. The representational formats involved and the processing routes they take in comprehension are explained in sections 6.4 to 6.6. This discussion is preceded by an explanation of how the development of this new hybrid model resolves some of the issues which seem arise with Tendahl’s model if viewed from a relevance-theoretic perspective, such as the conflation of different representational formats (section 6.3).

In addition to the phenomena Tendahl’s model can account for, the new hybrid model of metaphor proposed here can account for conceptual metaphor as arising in thought in contrast to metaphorical patterns that arise in language. This is because a distinction is drawn between two representational formats, each having their own distinctive processing mechanism. On the one hand, there are imagistic and experiential representations which can be assumed to motivate basic metaphorical thought patterns, and which arise from mental simulation processes; these can be thought to motivate conceptual metaphors originating in thought. On the other hand, there are propositional representations which are processed by means of inference, i.e. the derivation of warranted conclusions, dominating the communicative functions. Here, frequently used linguistic metaphors might become entrenched and hence accessed more readily, resulting in metaphors that have their origin in language. While these two levels do not interfere with each other such that they could be confused, they do inform each other through the informational chains afforded by modular systems. Further, metarepresentational processing affords an elevated position, as it were, of the imagistic impressions evoked by the literal meaning of a metaphorical expression, which gives the experience of metaphor understanding this extra quality that makes it feel special. Chapter 7 concludes by addressing some anticipated criticisms of the new hybrid model, and a general discussion.
Chapter 2

Metaphor in Relevance Theory

This chapter offers an overview of Relevance Theory (e.g. Sperber & Wilson 1986/95; 1987; Carston 2002). The hybrid model developed in chapter 6 takes elements from Cognitive Linguistics as well as Relevance Theory, while the underlying theoretical principles of Relevance Theory play a dominant role. It is therefore important to explain them initially with some care. Section 2.1 outlines the main relevance-theoretic principles, such as the communicative principle of relevance, the role of implicatures in comprehension, and how metaphorical meanings are thought to be derived by wide arrays of weak implicatures. Section 2.2. explains and exemplifies the more recent account of metaphor understanding, involving ad hoc concept construction, which, in conjunction with the account that preceded it, involving the derivation of weak implicatures, accounts for interpretation of metaphorical expressions in context. Section 2.3 explains how Relevance Theory has drawn on other approaches (e.g. Fodor 1975; 1981; Barsalou 1983; 1987). Here, the focus is on the notion of the Language of Thought, which inference in the relevance-theoretic sense is based on. This is followed by a brief description of how Barsalou’s work on categories has inspired the idea of ad hoc concept construction. Relevance Theory is an approach that assumes cognition within a modular mental architecture, so section 2.4 is dedicated to a description and discussion of the modularity account. Section 2.5 is concerned with criticisms and discussions of relevance-theoretic ideas. A common criticism is that ad hoc concept construction cannot account for complex and extended metaphors. Therefore, examples of the latter are analysed according to an ad hoc concept construction interpretation, in order to reveal the difficulties with this. Section 2.5.2 is based on a discussion of various views on what constitutes ‘what is said’ (Carston 2008), i.e. the question of which meaning elements are encoded in a linguistic expression and which are pragmatically derived in context. This helps to clarify the relevance-theoretic position. The related discussion of propositionality and how inference is thought to work (the ‘inferential fitness’ issue) gives a first idea of what a proposition-based approach,
like Relevance Theory, can and cannot explain. Thus it fulfils the role of preparing the ground for the discussion in later chapters, on whether Relevance Theory should be complemented by approaches that are not proposition-based.

2.1 The main principles of Relevance Theory

Engaging in the act of communication involves two main intentions: the informative intention, which is to inform the hearer of something, and the communicative intention, which is to inform the hearer of the speaker’s informative intention. The latter is important because it involves the tacit assumption that the speaker has said what they have said, and in the way they have said it, because it is relevant to the hearer. In other words, an act of ostensive communication (i.e. involving the communicative intention) automatically communicates a presumption of relevance, i.e. it is mutually manifest to speaker and hearer that the speaker intends the utterance to be relevant to the hearer. This prompts the hearer to work out what the speaker may have meant by the utterance. Since communication is a collaborative act, the speaker will have anticipated which conclusions the hearer is most likely to derive, and the hearer knows that the speaker has that in mind. This is the basis of the relevance-theoretic comprehension procedure, the idea being that communication is expected to be efficient and economical.

The Definition of Relevance:

(a) An assumption is relevant to an individual to the extent that its contextual effects in that context are large.

(b) An assumption is relevant to an individual to the extent that the effort required to process it in that context is small.

Sperber & Wilson 1987: 703

On this picture, the cognitive processes involved in communication are based on the principle of least effort for a maximum of effects, whereby effects are achieved by way of new information enriching, in some way, the addressee’s view
of the world. Working out whether an utterance interpretation is consistent with the communicative principle of relevance, i.e. whether the addressee achieves a maximum of effects for investing the least possible effort, should not be understood as literally calculating effort and effect in order to balance them out for optimal relevance. That would cost the brain too much effort in itself. Rather, as Sperber (2005) explains, the processes involved in arriving at an approximate value of optimal relevance are to do with brain chemistry, in the form of genetically specified non-cognitive procedures that regulate energy allocation for brain activity. He explains that this is comparable to muscle chemistry that regulates energy allocation for muscular work. The notion of cognitive effects can be described as follows: successful communication is understood as resulting in a change of one’s cognitive environment: contextual assumptions are either strengthened by being confirmed by the utterance, or contradicted, or a combination of old and new assumptions takes place to yield a new conclusion. These changes of one’s cognitive environment are called ‘contextual (cognitive) effects’ (Carston 2002: 377).

Sperber & Wilson (1987), in referring to the literature that addressed theories of discourse (at the time), see it as psychologically unrealistic that relevance should be represented as a variable to be assessed in a predetermined context. On the contrary, they assert that context is chosen in the process by each individual in order to maximise relevance (i.e. to achieve the most cognitive effects for least processing effort). Therefore, in this framework, relevance is regarded as a given to be maintained, and context is a variable subordinate to the goal of maximising relevance.

Presumption of optimal relevance:

(a) The set of assumptions I which the communicator intends to make manifest to the addressee is relevant enough to make it worth the addressee’s while to process the ostensive stimulus;

(b) The ostensive stimulus is the most relevant one the communicator could have used to communicate I.

Sperber & Wilson 1986/95: 267
The communicator can be expected to use the stimulus that leads the addressee to accessing the hypothesis that was intended first, rather than have them test several others before getting to the right one. Otherwise it would not make it worth the addressee’s while, and the participants tacitly know that.

In Relevance Theory, metaphor interpretation in communication was initially explained with the notion of ‘loose use’, where meaning and effect of a metaphorical expression are thought to come about through a wide array of weak implicatures (as opposed to fewer and stronger implicatures with non-figurative expressions) (Sperber & Wilson 1987/95; 1987; Carston 2002). Later, the account was developed further to explicate in detail the cognitive processes employed to recover utterance meaning. One could say it ‘zoomed in’ on the cognitive mechanisms recruited, such as mutual adjustment processes of implicatures and explicatures in ad hoc concept construction (Sperber & Wilson 2006; Wilson & Carston 2006; 2007; Vega Moreno 2005), which will be elaborated on further below. In general, on this approach, metaphorical expressions and other tropes, such as approximation and hyperbole, are thought to be interpreted by employing essentially the same processes as with literal expressions, but to differing degrees, depending on the extent of departure from the encoded content. This is known as the ‘continuity account’, since the various expression types are thought to lie on a continuum between literal and figurative. On this approach, metaphor is not seen as special compared to other expression types, since all expressions undergo the same processes in comprehension, and the only differences are the degrees to which these processes are applied in each case (ibid; Wilson 2003; 2009).

The role of implicatures, and how they are thought to come about in communication, is important in this context. Implicatures are based on non-demonstrative inferences, i.e. the truth of the premises makes the truth of the conclusions probable, rather than being a proof of the communicator’s informative intentions. This is best explained with an example.

(1) A: Have you seen the new Leonardo DiCaprio film?
   B: I don’t like formulaic blockbuster films.
The implicatures most likely to be recovered here are that B thinks that Leonardo DiCaprio films are formulaic blockbuster films, and that he therefore has not seen the film and does not intend to go and see it, since he doesn’t like this kind of film. These are relatively strong implicatures, and A can be fairly sure that B intended to convey them. However, if put on the spot, B would still be able to deny to have claimed that Leonardo DiCaprio films were formulaic blockbuster films, since he did not say so explicitly.

Sperber & Wilson consider it inadequate to treat all assumptions that are recoverable in communication as explicitly intended by the speaker. They explain that this is an idealization that cannot reflect how communication works. This is because in many cases the speaker is being vague, which means that the hearer is prompted to do a lot of inferential work in order to arrive at the intended meaning. This can be illustrated with the following example.

(2)  

A: Are you coming out with us tonight?  
B: I am completely broke.

Apart from the explicit content of B’s utterance, that B is broke, and the strong implicature that B is not intending to go out, there are many more assumptions that are vaguely communicated, such as that B might want to consider to go out if someone paid for her, that it is unpleasant to have no money etc. While these are potential candidates for implicatures, it is also possible that B did not intend to communicate any of them. This kind of vague communication is rather common, and Relevance Theory can account for it. How?

A major point of departure in Relevance Theory is the task of working out how it comes about that the addressee arrives at the intended message. Relevance theorists reject the idea of a properly shared context and mutual knowledge, which is often presented as the norm, as an unrealistic idealization. Hence Sperber & Wilson (e.g. 1987) explain that the relevance-theoretic approach does not assume a symmetry of operations at the emitting and receiving ends, that is, it allows for the possibility (which naturally often occurs), that the speaker and the hearer do not have identical sets of premises in a given discourse situation, but their
cognitive environments are merely mutually manifest, which is weaker than ‘shared’. Sperber & Wilson point out that the problem with assuming that speaker and hearer always share the same context is that, for the participants, there is no way of knowing which contextual assumptions they share and which ones they do not share. Relevance Theory works with the notion of a ‘cognitive environment’, which consists of facts that are manifest to an individual, i.e. which the individual is capable of mentally representing at the time in question and accepting them as true or probably true. So it also includes facts that the individual is capable of becoming aware of, in addition to the ones that they in fact are aware of. Manifestness is a relative notion, and ‘mutual manifestness’ is understood here to be psychologically more plausible than ‘mutual knowledge’, because it is not at a level of certainty that both or more participants represent the same or similar things, but rather of probability. Technically speaking, ‘the greater the mutual manifestness of the informative intention to make manifest some particular assumption, the more strongly this assumption is communicated.’ (1987: 706).

The approach thus allows for various strengths of communicating implicatures, which at the most extreme includes a wide array of weakly communicated weak implicatures. This may sometimes lead to an indeterminacy of meaning, as is the case with poetic metaphors, for example, which do not seem to lend themselves to any explicit interpretation. In such cases, it is indeterminate which of the weak implicatures the speaker specifically intended. None of them is explicit, and their number is indefinite. But the overall picture thus communicated creates a strong impression that is not paraphrasable, which is a typical characteristic of metaphorical utterances and the poetic effects they create (cf. Sperber & Wilson 2006). On this view, the strength and probable quantity of implicatures depends on the degree of departure from the encoded meaning. Thus there is a continuum, with expressions whose meaning is derived almost entirely inferentially (i.e. the most figurative), at one end, and expressions whose meanings are more conventionalised, or lexicalised (i.e. the most literal), at the other. Hence strength of implicature is a matter of degree. In comparison, any approach that sees implicatures to be either strong and determinate, or not derivable at all, falls short of (a) including any vaguely communicated
assumptions and (b) accounting for differences in interpretation between individuals, which cannot be done if context is envisaged as a given set, as compared to the relevance-theoretic picture of individuals constructing and reconstructing context to fit optimal relevance.

A strong implicature is arrived at when an utterance prompts the hearer to supply a very specific premise or conclusion. A range of weak implicatures can be derived in cases where it is possible to give several different interpretations, each based on different sets of implicated premises and conclusions, all of which are related and all of which are consistent with the communicative principle of relevance, so that there is no predetermined hierarchical order between them. The same cognitive environment may lead different people to different interpretations, depending on numerous personal factors etc., many of which cannot really be measured or tested (illustrated below). The weaker an implicature, the less sure the hearer can be as to how far his interpretation matches the speaker’s intention. For example, if A utters (3), preparing B for a blind date with C, whom A knows but B doesn’t, B needs to do a lot of inferential work in order to gather what A might mean, resulting in a wide array of weak implicatures.

(3) A: C is made of candy floss.

Is C very sweet? B might wonder, or too sweet? Is he sticky so you cannot get rid of him easily (as with the sugar that candy floss leaves on your hands), is he without any substance of character, like the airiness of candy floss, or is he pleasant on the eye, but maybe too fluffy? There are many implicatures for B to derive, but none of them can be thought to be intended by A with any certainty. Thus, A cannot be made fully accountable for any of the interpretations that B arrives at.

It is important to recognise that there is a potential relative difference between the thought that the speaker intends to convey and the thought the hearer recovers in interpretation. Sperber & Wilson explain that the aim of communication is ‘to increase the mutuality of cognitive environments and thereby the similarity of thoughts, rather than to guarantee a (generally
unreachable) strict duplication of thoughts’ (1987: 706). On this view, then, any approach that equates (a) what a communicator intends the addressee to recover with (b) the content that the addressee actually does recover, distorts the state of affairs by idealising away from psychological reality (if the aim is to give a cognitive account, that is). An utterance is here understood as an interpretation of a thought of the speaker. This is called ‘interpretive resemblance’ (Sperber & Wilson 1986/95; Carston 2002), which means that an utterance can only ever resemble a thought but never convey it a hundred percent. The thought and its interpretation share a number of implications, and the speaker must produce an utterance whose explicit content logically or contextually implies these implications. Thus, interpretive representation is one kind of representation by resemblance, in this case resemblance by virtue of sharing logical and contextual implications. If it were possible to convey a thought a hundred percent in communication, then the expression were ‘fully literal’, which would mean that all of the logical and contextual implications were shared. On this view, however, ‘fully literal’ is an idealization that is not attainable in reality. Often a speaker chooses an utterance that implies many other assumptions she does not intend to communicate, in which case the hearer needs a way to recognise the right ones. In Relevance Theory, this goes as follows:

Relevance-theoretic comprehension procedure:
(a) Follow a path of least effort in computing cognitive effects. In particular, test interpretive hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.
(b) Stop when your expectations of relevance are satisfied.

Sperber & Wilson 2002: 24

Note that the potential of individual differences in interpretation are built into this approach in principle: variables such as personal experience, which may prime accessibility of certain assumptions, or more generally, personal preferences or
tendencies can all be accommodated. For example, if A utters (4), B may interpret it differently depending on whether or not he has just accompanied A to see the film ‘Independence Day’ (involving extremely hostile aliens who want to destroy the Earth).

(4) I feel like an alien.

C might overhear the conversation and immediately think of the song ‘Englishman in New York’ by Sting, thereby missing the point that A intended to make with reference to the film (but B might just think that because the film is still on his mind, but not on A’s), and interpret A’s utterance as having to do with being a marginalised immigrant in a foreign country. A has weakly implicated all of these interpretations and cannot be held fully accountable for any of them, as she has not implicated them strongly. This is because in this case, as in many, optimal relevance can be achieved in different ways.

An important variable pertaining to the investment of as little cognitive effort as possible is accessibility: an addressee always considers the most accessible interpretation first and looks beyond this only if it seems unlikely in the given context. Therefore, the stereotypical or default interpretation of an expression is usually the most likely to be accessed first, due to its strong position in memory as a result of frequent use. It undergoes conceptual modification as required by the context and considerations of relevance (i.e. balancing effort and effect), which takes its course as a mutual adjustment process between encoded and communicated content. In cases of metaphorical expressions, the adjustment can be expected to move further away from the encoded content, and hence can be assumed to be somewhat more severe than in cases of approximation or near-literal uses. This will be exemplified in section 2.2.

Relevance Theory deals with metaphor interpretation by assuming the derivation of a wide array of weak implicatures in comprehension. These can account for the rich impressions the hearer might get upon comprehension, which is combined with a relative freedom on part of the hearer in choosing among the

---

2 Such aspects play also an important role in comparing communication between strangers to communication between people who share a great deal of experiences.
implicatures, releasing the speaker from taking full responsibility of having communicated any specific one of them. In particular, it can account for the indefiniteness of implicatures the interpretation of metaphorical expressions affords, with a focus on potential meaning outcomes, rather than, as is the case with many other approaches, a definition of what the meaning of a metaphor is supposed to be, which tends to result in mere paraphrases. But if a paraphrase would be sufficient to account for the metaphorical meaning, why would we need to use a metaphor in the first place?

In more recent years, the account involving the derivation of weak implicatures has been complemented by a more detailed account of inferential processing at the conceptual level. The reason for this is that the notion of weak implicatures by itself cannot account for so-called ‘emergent properties’, which are features that characterise the end-product of the comprehension process of metaphors of the category-crossing type such as human/machine. Such properties are not standardly associated with the metaphor vehicle, nor with the metaphor topic, and seem just to appear or ‘emerge’ (e.g. Wilson & Carston 2006). For example, the comprehension of ‘Robert is a bulldozer’ has as outcome e.g. that Robert is a very stubborn and persistent person, characteristics which we find neither in the encyclopaedic entry of ‘bulldozer’, which is a machine that obviously does not have any human traits, nor are they initially associated with the person ‘Robert’ in any way. Hence it would not be sufficient to say that a metaphor refers to those features that the topic and vehicle have in common, but rather it is necessary to explain the kind of inferential processing at work which brings about these emergent properties. This is thought to take place in the form of sense modifications, the so-called ‘ad hoc concept construction’, explained in the following.

### 2.2 Ad hoc concept construction

The ad hoc concept construction account has received particular attention in recent developments in Relevance Theory (Vega Moreno 2005, Wilson & Carston 2006, 2007). The mechanisms involve the on-line manipulation of the
encyclopaedic information associated with existing concepts in order to accommodate the situation-specific metaphorical meaning at hand. In Relevance Theory, a concept is understood as consisting of an address in memory which provides access to lexical, logical and encyclopaedic entries. These contain information of the respective kinds: the lexical entry contains linguistic information, such as phonological and other grammatical information about the word or string of words used to express the concept in a specific language. The logical entry provides the rules necessary for truth-preserving inferential computing of the concept (described in section 2.3, and further discussed in section 2.5.2). The encyclopaedic information is about the objects, events, or properties that instantiate the concept (e.g. Sperber & Wilson 1987: 702). It is relatively open-ended and varies between individuals. Although it is not explicitly described as such in Relevance Theory, it is possible, for the sake of simplicity and current purposes, to assume a relatively stable set of encyclopaedic features associated with a concept, i.e. those which are most frequently accessed. They can be understood to provide the core, or default meaning, which is content-constitutive when it features in the stereotypical interpretation of a word in a ‘non-biased context’.³

Sperber & Wilson (1986/95) explain that the distinction between the logical and the encyclopaedic entries is characterised by a distinction between the content-constraining function of the logical entries, and a context-constraining function of the encyclopaedic entries. Thus, the definition of content seems here to be a narrow one, while it might seem that the definition of context here includes elements that intuitively seem to belong to the content, since at least the core part of an encyclopaedic entry can be thought to be the content associated with a concept that is shared by many people. It seems that for Sperber & Wilson, however, there is a clear difference between logical content and encyclopaedic content:

Intuitively, there are clear-enough differences between encyclopaedic and logical entries. Encyclopaedic entries typically vary across speakers and times: we do not all have the same assumptions about Napoleon or about

³ However, I doubt that the term ‘non-biased context’ can refer to real discourse situations, as I consider any context to be relatively biased in one way or another.
cats. They are open-ended: new information is being added to them all the
time. There is no point at which an encyclopaedic entry can be said to be
complete, and no essential minimum without which one would not say that
a concept had been mastered at all. Logical entries, by contrast, are small,
finite and relatively constant across speakers and times. There is a point at
which a logical entry for a concept is complete, and before which one
would not say that the concept had been mastered at all. Suppose, for
example, that a child has not yet realised that \( X \) knows that \( P \) implies \( P \),
and so uses know interchangeably with believe. We would say that he had
not yet mastered the concept.

Sperber&Wilson 1986/95: 88

However, for many concepts I would consider the boundaries between logical and
encyclopaedic entries to be less clear-cut and rather more gradual, depending on
the elements that people consider to be the stable content of a concept. This has
been referred to as the difference between ‘core’ and ‘non-core’ encyclopaedic
features (e.g. Rubio 2008), which will be further discussed in section 4.2.2. Here,
the role of the encoded (or literal) meaning, which is associated with the concept
by default, is understood to be a mere starting point for spontaneous, subconscious
and automatic construction of the appropriate occasion-specific meaning required.
In general terms, the ‘core meaning’ assumed here is usually understood as the
literal meaning. For example, the concept CHAMELEON provides access to various
core features, which, in no particular order, may include the following:

\[
(5) \quad \text{Core features of CHAMELEON:}
\]

- is a lizard
- changes colour
- has separately mobile eyes
- etc.
The concept also provides access to non-core features, which can often be highly idiosyncratic and are likely to change over time, so a few tentative suggestions of what they may include are:

(6) **Non-core features of CHAMELEON:**

- lives in hot countries
- moves slowly
- is always hidden away at the zoo
- is a charismatic creature
- etc.

Ad hoc concept construction is based on the idea that features can change their status from merely vaguely associated with a concept to being content-constitutive in a specific context, and vice versa. Thus, features that can be taken as part of the core meaning by default can become non-core features in the process. For example, if I talk about a mutual friend of ours and say ‘she is a chameleon’, the core-feature ‘is a lizard’ becomes redundant with regard to the communicated meaning, as there is the strong contextual assumption that our friend is a human being and not a lizard. One of the basic relevance-theoretic principles is that a communicator, by way of the very act of communication, makes clear to the addressee that the utterance is relevant to them in this very context. Therefore, due to this expectation of relevance, the addressee takes the inferential steps required to make those features that are the most likely candidates fit the context. The inferential steps employed in ad hoc concept construction are lexical adjustment mechanisms, which involve picking out subsets of encyclopaedic features to arrive at the intended meaning. This involves two different but complementary sub-processes. There is the mechanism of broadening, on the one hand, which requires the dropping of certain encoded features to include denotations of the word that are not part of its encoded meaning. Narrowing, on the other hand, involves promoting certain non-core features to the status of content-constitutive in order to exclude denotations that are associated with the
encoded meaning but not appropriate in the given context. In some cases only broadening or narrowing is involved, in others both.

The established notation for lexical concepts consists of small capital letters: CONCEPT, and for ad hoc concepts of small capital letters and a star: CONCEPT*. In order to make a notational distinction between different ad hoc concepts that originated from the same encoded concept, we vary the number of stars attached: CONCEPT**, CONCEPT*** etc. An example of broadening would occur in many interpretations of the above mentioned concept CHAMELEON, of which (7) is an example:

(7) ‘Sally is a chameleon.’
    (where Sally is a human being)

According to the relevance-theoretic comprehension heuristic, the stereotypical interpretation of the word ‘chameleon’ is the most accessible due to frequent use, and is hence accessed with the most ease. So the core features are the most activated, even if not all of them are contextually adequate. Those features that also fit the context receive an additional activation boost and yield the most cognitive effects (in this case e.g. CHANGES APPEARANCE). In this example, the feature ‘is a lizard’ is dropped so that also human beings are included in the denotation. It is hence a superordinate concept. Figure 1 illustrates the result of broadening: the dots in the small circle stand for the denotations of the lexical concept CHAMELEON, and the dots in the large circle stand for the denotations of the ad hoc concept CHAMELEON* (including the lexical concept).

Figure 1

(taken from Carston 2002: 343)
Further, the feature ‘changes colour’ that is activated as a core feature of CHAMELEON is broadened by way of going from the concept CHANGES COLOUR to its entailment CHANGES APPEARANCE to arrive at the ad hoc concept CHANGES COLOUR*, which includes changes in human behaviour as well as changes in lizards’ skin colour.

Once the addressee has arrived at this point in interpretation, he has achieved optimal relevance: rich cognitive effects, to do with assumptions about Sally’s chameleon-like changes in behaviour (and many more that this metaphor potentially communicates in this context), for a relatively small investment of cognitive effort, as the ad hoc concept has been built up from the most accessible conceptual information, in conjunction with contextual assumptions. For an example of both broadening and narrowing, consider the metaphorical (and often rather rude) use of ‘thick’, as in (8).

(8) ‘Don’t call him thick just because he hasn’t read as many books as you!’

As with the ‘chameleon’ example above, it is helpful for analysis to list just a few of the possible core and non-core features. The concept THICK may give access to the following core features:

(9) Core features of THICK:

- much volume
- deep in diameter
- etc.

Its non-core features may possibly include the following (randomly chosen):

(10) Non-core features of THICK:

- good for insulation
- impenetrable
- won’t break
- heavy
- etc.

Here, a feature such as ‘much volume’, and any feature that denotes a physical quality is dropped to broaden the concept, so that its denotation includes things that are not physically thick. The resulting superordinate concept THICK* then allows for the psychological or mental sense. However, this sense does not yet stretch as far as to refer to a lack of intelligence, and the concept also needs to be narrowed by promoting a feature such as ‘is impenetrable’ which, in its broadened psychological sense, alludes to a mass that makes it difficult for information to be processed, to make its way through. It is thus made content-constitutive to exclude any psychological senses of THICK* that allow for the characteristic of being, say, a quick thinker. Thus we find that the lexical concept THICK and the final ad hoc concept THICK** merely overlap as concerns common features, as illustrated in figure 2:

![Figure 2](taken from Carston 2002: 343)

This is a very rough sketch of the processes at work, but it demonstrates how the concept THICK thus broadened and narrowed is adjusted to denote something similar, but not equivalent, to ‘stupid’.

The ad hoc concept construction account demonstrates how metaphorical meaning is derived at the level of explicatures, since the encyclopaedic information is manipulated in a way that it fits the context at hand. At the same time, a lot of cognitive effects are evoked by an array of weak implicatures, as described above. Thus, the older account of loose use in combination with the newer account of ad hoc concept construction can explain different aspects of the
mechanisms involved in metaphor interpretation. In describing how this is supposed to work, Sperber & Wilson use the following example:

(11) \textit{Woman to uncouth suitor: Keep your paws off me!}

\text{\small (Sperber & Wilson, 2006: 28)}

The encyclopaedic features of the concept PAW are being adjusted by way of broadening to include a human’s hand in the denotation, and by narrowing to exclude paws which are used in a soft and careful way (such as cats’ paws when they are carefully stalking a mouse), and the weakly implicated assumptions are to do with stereotypical features of paws, i.e. that they are used clumsily, grossly, etc. The strongly implicated assumption is that the hearer remove his hands (2006: 28-29). Thus, both types of process contribute to the overall meaning of the metaphorical expression.

While it is likely that ad hoc concept construction takes place in comprehension of novel metaphorical expressions, it is presumably not employed with more conventionalised metaphorical expressions that are used frequently in a language community. Here, the metaphorical sense of the expression will have become entrenched in memory to a certain extent, due to routinisation, so that the concept does not need to be constructed from scratch, but can simply be accessed. Since cognition is geared towards least processing effort for most possible effect, this route is to be preferred as it is less effortful than ad hoc concept construction. This mechanism of easy access due to frequent use, leading to well-worn processing routes, also explains semantic change. An expression that is used frequently enough in the same way in a speech community eventually becomes lexicalised, resulting in polysemy (Sperber & Wilson 1998; Vega Moreno 2005; Wilson & Carston 2007; Wilson 2009). For example, ‘cool’ in English, which has now several senses, such as a moderately cold temperature, and a state of mind and/or appearance, is a classical example of lexicalisation. In section 4.2.3 I run through a few examples of English-German translation asymmetries showing different degrees of lexicalisation, i.e. expressions that are metaphorical in one language but have lexicalised to the status of polysemy in the other.
As I have mentioned in the introduction, it is possible that some metaphors originate in language, while others originate in thought. The kind of metaphor just described appears to have its origin in language use: the metaphorical pattern gets entrenched in memory because of its communicative merit. It has proven a useful tool to express a thought, so it gets used again and again in this way in conversation. But it does not seem to have been present as a cognitive pattern before it got used in language. This would be the case with the other type of metaphor, arising in thought, which will be described in section 3.3.4.

2.3 Adoptions from other approaches: Fodor and Barsalou

This section is concerned with how Relevance Theory has incorporated influences from diverging directions: ideas from the cognitive science tradition, in particular atomic concepts and the Language of Thought by Jerry Fodor (1975; 1981), and from psychology with Lawrence Barsalou’s (1983; 1987) work on ad hoc categories and schematic concepts.

The relevance-theoretic model of a concept is based on Fodor’s notion of ‘atomic concepts’. Atomic concepts are representations that are not analysable, which means that one cannot break them down into constituents. They are primitive mental representations, the most basic elements of the computational system in which our thinking is thought to take place (Fodor 1981). While the relevance-theoretic model of a concept is based on Fodor’s notion of concept atomicity, relevance theorists do not agree with Fodor in every respect. As Carston (2002) points out, Fodor seems to agree with the code model of communication which claims a one-to-one mapping from the concept encoded in a word to the communicated concept (2002: 214). Relevance theorists, on the other hand, maintain that the concept encoded by a word undergoes inferential processing in communication before it becomes the concept that finally features in the interpretation outcome. Thus, an encoded concept serves as a starting point for inference, and takes on the required occasion-specific shape (which may turn out to be quite different from the encoded concept). The main argument in favour of this view is that we can entertain many more concepts in our minds than there are
words in a language to encode them, but we are still able to communicate all (or most) of them (cf. Sperber & Wilson 1998).

However, Relevance Theory shares some basic assumptions with Fodor, in particular the idea that mental representation and computation is distinct from any other form of experience. Cognition is here understood as the processing and manipulation of conceptual representations in the medium of the ‘Language of Thought’ (henceforth LoT). Fodor claims that we do not think in a natural language, such as English, but in LoT, which is a different medium altogether. It is structured in a way that is analogical to natural languages, with syntactic constituents that can be combined recursively. Carston explains that ‘the full formulas are truth-conditional, so have truth values as determined by the way the world is, and they bear logical relations to each other, such as entailment.’ (Carston 2006: 4).

Fodor’s notion of LoT belongs to the strand of cognitive science that is based on the idea of computationalism, which assumes that it is possible to empirically model cognitive processes computationally. Such computational models are thought to represent adequately the physical processes instantiated in the brain during thinking, i.e. cognitive states. Computational realisation of cognitive processes, as in LoT, requires a representational system which has two types of property: formal syntactical, and semantic. The syntactic properties are thought to be the ones that model physical brain states. They follow the laws of logic, which are the inferential rules according to which propositions are connected. That is, in a thought process, the transition from one thought to the next is a matter of operations driven by the purely formal (or syntactic) properties of the symbolic structures. This is only possible with conceptual constituents, which have the necessary logical properties (cf. Preti & Velarde-Mayol 2001).

In Relevance Theory, inferential activity is based on the model of LoT. Sperber & Wilson (1987; 1986/95) describe a deductive device as a model for human deductive ability, which performs the inferences that are necessary to draw conclusions from assumptions. This is seen as an integral part of cognition, and by
extension of communication⁴. Thus, in a given assumption, a conceptual address appears as part of a logical structure, which is detected by the deductive device by means of its symbolic form in the syntax of LoT. This way, access to the three types of conceptual entry of this particular concept is given. So a concept performs two different functions: its computational function, which is merely syntactic as required for LoT, and its function as an address in memory. The logical rules the deductive processes are governed by are very similar to entailment, but entailment is a semantic rather than a syntactic relation. Sperber & Wilson provide the following example:

(12) Apples grow in orchards and grapes grow in vineyards.
(13) Apples grow in orchards.

Here, (12) entails (13). The relation is semantic because it works with reference to a state of affairs. In deductive processing, the entailment relation, as it were, holds by virtue of formal properties only, rather than reference to a state of affairs, but in principle the dynamics of the relation are the same. In syntactic terms, it is a logical implication, where the form \((P \text{ and } Q)\) yields conclusions of the form \(P\) or of the form \(Q\). So this is an example of the kinds of rules the deductive device functions with. Importantly, such a deductive rule is truth-preserving: applied to an assumption, it becomes a semantic relation, as states of affairs in the world are involved in an assumption. In such a semantic relation, the conclusion stands in an entailment relation to the premise. This means that the conclusion must as truly correspond to the states of affairs as the premise.

Sperber & Wilson provide convincing arguments in favour of the idea that human cognition operates according to such a set of deductive rules: for an organism, it would be the most economic way of dealing with representations in terms of storage in memory, as single conclusions from sets of assumptions do not need to be stored separately, but rather can be deduced by the rules whenever required. Further, the truth-preservation that comes with the rules guarantees accuracy of the organism’s representations of states of affairs in the world. Thus,

⁴ In the latter it is understood as a sub-part of non-demonstrative inferential mechanisms (Sperber & Wilson 1986/95).
they argue that it is a system that is economical as well as reliable (1986/95: 84-85).

However, it is not certain that human cognition really operates based on an abstract system that is removed from any experiential modalities. After all, it would mean that perceptual experiences would need to be transformed into abstract symbols so that they can be computed in LoT, and it may seem counter-intuitive that cognitive activity would be so removed from real experience. Barsalou (1999; 2005; 2009) denies the existence of an abstract device such as LoT. He claims that concept formation and cognition are based on modality-specific representations that are direct simulations of the real thing rather than abstractions. In the light of the preceding description of LoT with its emphasis on logic, this raises the question of truth values and reliability. However, Barsalou has compelling arguments in favour of the simulation view, which will be outlined in section 3.3.3.

Relevance theorists would presumably not agree with Barsalou’s claims, as the approaches oppose each other in several respects (outlined below). However, relevance theorists have been inspired by other aspects of Barsalou’s work. The idea of ad hoc concept construction in Relevance Theory has its origin in his work on categorisation (e.g. Wilson 2003; Carston 2002). He conducted experimental studies, based on typicality judgements, to test whether people access ready-made prototypical categories stored in memory. He found that, in fact, people create new categories on the spot, as they could not have had stored in advance all the categories that Barsalou tested. These included very unusual categories, e.g. GOOD THINGS TO STAND ON TO CHANGE A LIGHTBULB, and categories that the subjects might expect specific groups of people, other than their own, to come up with, e.g. a category from the point of view of a housewife, a Chinese person, etc. (Barsalou 1987). Carston emphasises that Barsalou's experiments focused on goal-oriented ad hoc categories rather than ad hoc concepts, as in Relevance Theory (2002: 367). He has, however, drawn attention to the fact that people generally tend to create contents to fit the current situation, and relevance theorists regularly refer to his work. Interestingly, his general views
on cognitive representation seem rather different from relevance-theoretic convictions, in parts even diametrically opposed to them.

To conclude, the relevance-theoretic approach borrows from Fodor the notion of atomic concepts and the idea that conceptual representation functions according to LoT rules of computational manipulation. But relevance theorists do not agree with Fodor on the idea that there is a one-to-one mapping between words and concepts, since there are many more concepts in our minds than readily encoded in words. However, the logical rules according to which inferential completion of propositional utterance outcome is thought to proceed are based on Fodor’s LoT (further elaborated on in section 2.5.2).

Barsalou’s work also plays an important role in Relevance Theory, in that he has sparked the development of the ad hoc concept construction account. But here we find a contradiction between Barsalou’s cognitive-linguistic leanings regarding a holistic experiential view of representation, and the relevance-theoretic modular view that embraces LoT. Thus, the story of the relevance-theoretic account’s development, as it has been influenced by sources of diverging convictions, can demonstrate that it is possible, and may be fruitful, to combine aspects of different schools of thought without having to incorporate or even to agree with all aspects of every model involved. This is the case with the model proposed in chapter 6.

2.4 Modularity
Relevance Theory and Cognitive Linguistics approaches differ in terms of their underlying principles regarding how cognition is thought to work in general, i.e. the mental architecture assumed in the respective models. While Cognitive Linguistics works with the underlying assumption that cognition takes place in a holistic fashion, based on general reasoning and embodiment, an important principle in Relevance Theory is the idea that cognition is modular. In his hybrid model, Tendahl (2006) adopts the cognitive-linguistic position; he prefers to work with the assumption that communication proceeds in a fashion that employs a general and nonmodular cognition. The discussion in section 2.4.2 explains why a
modular architecture seems more advantageous for a hybrid model, and how Tendahl’s arguments against modularity might be influenced by a misconception of Sperber’s (1994) notion of modularity.

The relevance-theoretic account is within the framework of modularity in the sense of Sperber (1994), who has developed a picture of a mental architecture that is based on the notion of a so-called ‘Massive Modularity’. He explains how he comes to the conclusion that cognition must have evolved in this way. Its features go back to Fodor’s model of ‘The Modularity of Mind’ (1983), which influenced relevance-theoretic ideas early on, before Sperber developed Fodor’s original notion of modularity to such an extent that the two do not seem to have very much in common anymore. In particular, Fodor describes a mind that is modular only at the periphery, with central systems that are nonmodular, while for Sperber modularisation goes for all aspects of cognition. This will be explained in the following, beginning with Fodor’s (1983) approach.

2.4.1 Fodor: The Modularity of Mind

Fodor’s (1983) model is based on the distinction between perceptual and conceptual processes. Sperber (1994) provides a simple but clear illustration of the difference between the two: ‘...seeing a cloud and thinking ‘Here is a cloud’ is a perceptual process. Inferring from this perception ‘It might rain’ is a conceptual process’ (1994: 120). Fodor (1983) proposed that perceptual and conceptual processes are to be seen as carrying out entirely different functions, localised in different areas in his model of the mind. This stood in contradiction to the belief held by many, that perceptual and conceptual information is exchanged freely and multidirectionally within the mind, a view which, incidentally, most (if not all) cognitive-linguistics approaches are based on.

Fodor’s proposal is designed after computational models, and he makes an analogy with the Turing machine. Barsalou (2005, discussed in section 3.3.4) suggests that there is a trend in cognitive science to model the architecture of the human mind on computation, and it appears that Fodor is in agreement with this.

---

5 For further discussion of massive modularity see Sperber 2001; Carruthers 2006; Hirschfeld & Gelman 1994; Samuels 1998. For arguments against it, e.g. Fodor (2000).
trend. It is not clear why the analogy is with a Turing machine in particular. It seems to follow a general philosophers’ tendency to refer to Turing machines whenever they draw an analogy between minds and computers. He emphasises that Turing machines (like any computer) are ‘essentially symbol-manipulating devices’, and that this is also the main characteristic of the human mind. However, computers are closed systems, so it is important to take into consideration that the human mind, like Turing machines, is to be seen within ‘a matrix of subsidiary systems which affect their [its] computations in ways that are responsive to the flow of environmental events’ (1983: 38-39). Therefore, there must be a way by which information from the outside world, i.e. sensory information, enters the thinking process. This is how the essential structure of Fodor’s model of the mental architecture comes about, by an analogy with Turing machines within a wider system of information in various forms. Thus, he proposes the mind to have ‘central systems’, which carry out the actual cognitive activity of manipulating symbolic representations (i.e. thinking), and peripheral systems, which receive and mediate the ‘flow of environmental events’. For Fodor, these processes localised at the periphery are modular, carried out by so-called ‘transducers’ and ‘input systems’. The transducers transform incoming information from the sense organs into perceptual representations that can be computed by the input systems, whose function it is to make these representations fit for computation proper (i.e. thinking). Thus, the input systems themselves are not understood as carrying out computational cognition as such, they just ‘translate’ incoming information into a format that can be used in thought in the central systems (1983: 40-46).

Importantly, the peripheral systems are thought to be modular, while the central systems are nonmodular. Sperber (1994) points out that Fodor’s model is not actually modular in a strict sense; he calls ‘Modularity of Mind’ a paradoxical title for it, since the modularity is located only at the perceptual periphery of the mind, while conceptual processes are taking place in the nonmodular centre of the mind (1994: 119).

Mental modules have particular characteristics which clearly distinguish their functions from those of the Fodorian central thought processes. The main ones are the following: mental modules are domain-specific, which means that
they take only specific types of information as their input, such as those pertaining to sensual modalities, and they have developed specific ways of processing them. Their outputs (or conclusions) are also restricted to a small group of properties. Such specific domains and functions include, for example in vision, mechanisms such as shape analysis, spatial relations, colour perception, or face recognition; in audition, grammatical decoding in language, but also mechanisms that cater, for example, for the recognition of other humans’ voices (Fodor 1984: 47). Another important characteristic of mental modules is their informational encapsulation. This means that they are not only specialised regarding their domains, but they are also unable to access any other information present in the overall cognitive system than that contained in their own domain. In vision, this phenomenon can be consciously experienced with optical illusions such as the Mueller-Lyer Illusion: we know that both lines have exactly the same length, due to information other than visual perception present in our mind, such as the outcome of measuring both lines with a ruler, but still our eyes try to make us believe, as it were, that one line is shorter than the other (Figure 3 below).

![Figure 3: The Mueller-Lyer Illusion](image)

The fact that we cannot help ourselves but see a difference in length despite our better knowledge is also an indicator that modular processes are mandatory, we have no control over them in that we cannot stop them from running their course. The high degree of specialisation, the automatic functioning without recourse to any neighbouring types of information, or involvement of decision making of any kind, makes mental modules very fast and efficient (cf. Carston 1996).

However, it should be noted that regarding mandatoriness, Sperber (2005) has an interesting point for discussion. He objects, to a certain degree, to the idea that a module’s reaction to its input is mandatory. There is an ingredient missing on this picture, and this ingredient is attention. That is, if the organism is focusing
on some other incoming stimulus, even when the input in question is within the range of its perceptual radius, it might still not be processed at all. Sperber (2005) refers to experimental evidence demonstrating this phenomenon. For example, subjects were watching a video of people playing basketball, and they were instructed to focus on the movements of the ball. Then a gorilla appeared in the video amidst the players and in full sight. 50% of the subjects reported not to have noticed the gorilla, and they were genuinely surprised when they re-watched the video afterwards (cf. Simons & Chabris 1999). Thus, mandatoriness of processing as a characteristic of mental modules can only be claimed if one supplements the claim with the necessity that the organism pays attention to the stimulus. However, this does not mean that the attention, when it is afforded, is under voluntary control of the organism. Once you notice the gorilla, you cannot prevent yourself from seeing it.

Fodor assumes that the central systems are non-modular. This is the area where computational activity takes place, i.e. the actual thinking involving inference, decision making, belief formation, and anything that is thought to involve manipulation of symbolic representations. According to Fodor, the mind has evolved in this way from an initially entirely modular architecture. Then a process of demodularisation took place to free the central systems from their modular constraints and thus enable them to carry out reasoning proper (1983: 43). Sperber (1994) doubts this evolutionary hypothesis. He argues that it is highly unlikely that an organism would reverse the process of specialisation in order to render itself more general, since this would mean a loss in speed and accuracy. Such a loss could be life threatening for the organism, e.g. when a fast and unmediated reaction to a threat is required. One could argue that Fodor’s model still caters for these kinds of actions, since the peripheral regions have remained modular, and thus carry out such reflex-like activities. However, Sperber’s argument pertains to general evolutionary tendencies, which have not in mind in advance, as it were, how the end product might function, but rather react to environmental pressures in small short-sighted steps. From this perspective, a step away from specialisation would be self-defeating. To illustrate, Sperber gives a ‘toy example’, involving a made-up species that undergoes some evolutionary
developments. The point just highlighted becomes very clear in the context of a species that has a module which functions to analyse and react to, automatically and at high speed, danger, e.g. in the form of a predator. In a mutant exemplar of that species, this very module would have undergone a de-specialisation in favour of the ability to access all kinds of other information to reflect on. In the event of a threat by a predator, its flight behaviour would not automatically be triggered, as the module had been de-specialised, but instead it would reason about the pros and cons of fleeing. This would take too long and the organism would fall victim to the threat, and thus be excluded from natural selection (1994: 126-127). Of course, one could argue that modules specialised on functions required for survival would not undergo demodularisation, while other, less important modules would. However, Sperber points out that in terms of general tendencies of evolutionary movements this seems implausible, as de-specialisation as such is a backwards direction of development which contradicts the rules of evolution.

2.4.2 Sperber: The Massive Modularity Hypothesis
Sperber (1994) defends the idea that the mind is modular all the way through, rather than partly modular. This is called ‘The Massive Modularity Hypothesis’. On this picture, also reflective thinking is carried out by mental modules. However, the modularity that Sperber describes displays much more flexibility than becomes apparent in Fodor’s model, and this is the point which may get overlooked, and due to which it is easy to misinterpret Sperber’s account by thinking that it is too inflexible to be able to accommodate human reasoning. It seems that this has been the case with Tendahl’s (2006) arguments against massive modularity. He seems to have taken Fodor’s description of mental modules and inserted these into the central systems in order to get massive modularity, rather than taking into account the ability of interaction and information exchange between modules, which is an important feature of Sperber’s conception of modularity. This will be outlined below and contrasted with Tendahl’s arguments against the idea of massive modularity. This discussion
will make clear how he seems to have misconstrued Sperber’s ideas to a certain extent.

Tendahl explains that massive modularity and cognitive-linguistic principles, such as embodiment (i.e. the idea that cognition is based on bodily experience in the physical world), are notions which are far too disparate in principle to be combined. So if one intends to endeavour a combination of Cognitive Linguistics and Relevance Theory, one has to make a decision in favour of one or the other underlying principle. Tendahl has made a firm decision in favour of cognitive-linguistic notions of embodiment (2006: 223). I shall argue, however, that it is possible to assume massive modularity in Sperber’s sense and at the same time integrate elements inherited from embodiment ideas. This issue will be looked into in chapter 6. According to Tendahl, the mind cannot be modular all the way through, because he seems to assume the design of the modules in Sperber’s account to be the same as that of Fodor’s peripheral modules. As if such Fodorian modules would then thought to be capable of reflective reasoning just like the central systems in the Fodorian model are. I agree that this seems highly unlikely. Their informational encapsulation would prevent information from sources other than their specific domain to be integrated. With language, this would mean that an informationally encapsulated language decoding module would first decode the linguistic input, regardless of any contextual or other information, and then, afterwards, pragmatic processes would set in to carry out enrichment processes and relevance-driven interpretation. He is right in saying, of course, that this contradicts the relevance-theoretic basic idea of underdeterminacy, that ‘what is said’ is being developed in a process of mutual adjustment of explicatures and implicatures, processes which all run in parallel. Thus, Tendahl seems to think that relevance theorists, Sperber in particular, are contradicting themselves when they believe in modularity with its informational encapsulation, on the one hand, and in linguistic underdeterminacy and the relevance-theoretic comprehension procedure, on the other (2006: 221-224). I agree that these two assumptions are incompatible. However, they are only incompatible if we assume that informational integration among mental modules is not possible. If it were possible, modularity would work in perfect harmony
with the other relevance-theoretic principles. And this is precisely what is the case, as Tendahl seems to have missed out on Sperber’s (1994) perspective of informational integration in massive modularity. Sperber explains that mental modules can exchange information in spite of being informationally encapsulated. While Fodorian peripheral modules get their specialised input from the sensory organs via the transducers, Sperberian modules that perform more central tasks involving inference get their specialised input from other modules. That is, if the input that comes from another module has already been inferentially worked on by the previous module, and perhaps by the one before it, and so on, information that may seem too disparate and not belonging to the specific information that our module is designed to take, may reach it nevertheless via other modules. In Sperber’s words,

If, ..., the output of one conceptual module can serve as input to another one, modules can each be informationally encapsulated, while chains of inference can take a conceptual premise from one module to the next, and therefore integrate the contribution of each in some final conclusion. A holistic effect need not be the outcome of a holistic procedure.

Sperber 1994:133

This can also be envisaged for the language module, so that pragmatic inference may well get integrated in parallel to decoding processes, if pragmatic information coming from a pragmatics module functions as one kind of input to the language module. So Tendahl’s argument against modularity does not hold, since it seems based on an inaccurate conception of massive modularity. Furthermore, Tendahl’s rejection of modularity also entails a rejection of the ‘fast and frugal heuristic’ that, for example, a module dedicated to inferential processes has on offer (e.g. Sperber & Wilson 2002: 11). If Tendahl wants to adopt the relevance-theoretic comprehension heuristic in his model (as we shall see in section 5.2), he better adopt the means by which this heuristic is capable of running its course in actual communication. If these inferential procedures were to run on the basis of a general all-encompassing thinking device, the lack of automaticity would render
them far too slow to be able to function properly in real-time online communication.

However, Sperber points out that there is a fatally weak point, characteristic of Fodor’s nonmodular central systems, that equally holds for an integrative massive modularity: ‘the risk of computational explosion’ (1994: 133). If all incoming information would always be processed directly and without any restrictive selection, it would multiply at increasing speed and all that integration would be harmful rather than useful. But while Fodor’s central systems don’t seem able to impose any restrictive mechanism, as they are just open for information to freely flow (as seems to be the case also with other holistic systems such as those in Cognitive Linguistics), Sperber’s integrative modules can. Their information flow is constrained by competition for computational resources, such that there is a ‘temporary buffer’ governing allocation of resources. It is a device that allows only those representations to be processed which satisfy the requirements of optimal relevance at the time. The device consists in what is ordinarily understood as ‘attention’, which only those representations receive that are relevant enough to get activated. This neatly goes back to the communicative principle of relevance, which I have explained in section 2.1.

My discussion of Sperber’s model of the modularity of thought so far focused on informational encapsulation and integration. Another important characteristic of mental modules is their domain-specificity, the Sperberian version of which makes a difference between different types of domains: actual domains and proper domains. The actual domain of a module is everything that the module is capable of taking as input, independent from the function that the module has evolved to fulfil. Such a function could be, to stay with the example above, to avoid danger by fleeing predators. So it could be the case that an event occurs that is similar to a threat by a predator, e.g. the appearance of an animal that looks dangerous due to its size but is actually a vegetarian. The module might still run its course of processes, i.e. fleeing behaviour, and so the dangerous looking but harmless animal is part of the actual domain of the module. However, it is not part of its proper domain. Proper domains consists of inputs whose processing it is the function of the module. This function contributes to the
survival of the species. In our example, the proper domain consists only of real predators that are a threat to the life of the organism. However, the actual domain is what is actually processed by the module.

The fact that humans, in comparison to other animals, deal with vast amounts of information, has led to a situation where the actual domain, in many cases, has by far outsized the proper domain. This is because there is so much more information around that humans process than there is input that fits proper domains, since humans’ intellectual activities have a much greater role in our lives than maintaining our basic survival. In addition, the large amounts of information in the actual domains, being consciously organised by humans themselves, socially distributed and consumed for various purposes, has led Sperber to hypothesise the existence of a further type of domain: the cultural domain (1994: 136-139). Since all this information is competing for cognitive resources to be processed, as explained above, the regulating force is, again, relevance. However, a lot of information is often not actually relevant in the immediate context at hand. But in the human social environment, there is a wider and more stable, cultural context of beliefs and expectations. The kinds of information that are relevant in this enduring context, i.e. that enrich or contradict socially stable beliefs etc., are the ones that persist in a culture. Further, in human culture, information is not just found but it is produced as well, and Sperber speculates that a module ‘stimulates in every culture the production and distribution of a wide array of information that meets its input conditions’. Thus, the cultural domain accumulates information that ‘mimics the module’s proper domain’ (1994: 141). This obviously has not much to do anymore with biological functions to ensure the species’ survival, it has rather become a means to an end in itself: culture. Often the proper domain of a module has lost its function and ceased to exist, so that there are only the actual and cultural domain left, or even only the cultural domain. In other cases, the proper domain might still be fully functioning, but has been enriched by culturally transmitted information. It might often be the case that someone has never in fact interacted with the input represented (say, that of an animal) in person, but has only been told about it, or read about it, or seen it on television, etc. Sperber suggests that such cultural
information is, of course, useful in communication with other human beings, and he gives as an example metaphorical communication (1994: 143). That is, you might not have encountered a living pig ever in your life (as you grew up in the city), but you are able to represent one with your culturally informed module. If you are, say, eating with your hands while letting half of the food drop on the floor, you will understand the meaning of the metaphorical expression ‘You are such a pig’, because you have heard of the (alleged) features of pigs that make the expression relevant in this context.

Sperber takes the idea of massive modularity further in explaining the important role of metarepresentational abilities within this picture. The idea is that there is a module whose actual domain are mental representations, so it forms representations of representations, operating at a second-order level, above, as it were, representations of things in the world. One might wonder what such a module might be good for, which function it fulfils, i.e. what its proper domain is. If one can represent beliefs and desires towards contents (which follows from the ability to metarepresent), one can represent the mental states of other people as well, which is an essential ability in human cooperation, and most importantly, communication. This module is also called the Theory of Mind module (ToM), and the notion is one of the main principles that underlie Relevance Theory. Inferring the states of mind of others is an important ingredient in guiding a hearer’s expectations towards the intended communicated content. A core idea in Relevance Theory is that ‘a communicator, by means of her communicative behaviour, is deliberately and overtly helping her addressee to infer the content of the mental representation she wants him to adopt’ (1994: 147). Sperber & Wilson (2002) suggest that one could argue that the activities of the ToM module could equally well be carried out with general reasoning functions, which is conceivable, as intuitively one would agree that a rational being, capable of empathy, wouldn’t need a special module for that. However, they explain that the vast amounts of information that speakers’ meanings carry, compared to other kinds of intentions that people may entertain, would just be too much for a general procedure, and that it is, therefore, necessary to have a special device in place which can cope with such a massive input (2002: 13). Thus, contents that enter
people’s minds via communication enter via the metarepresentational module.
What is special about the metarepresentational module is that it enables a person
to critically evaluate beliefs, to question information with view to its source, or to
reject information which does not seem true. If accepted, the new information
may then be added to the appropriate first-order module, like in the 'pig' example
above. This means that we have information about entities in the world available
in two distinct formats: that in the first-order module which specialises in the
conceptual domain in question, plus that in the second-order module. This,
according to Sperber, amounts to two distinct kinds of belief: intuitive beliefs and
reflective beliefs. The latter belong to the realm of the metarepresentational
module and can be detached from any information belonging to first-order
modules. Thus, they often are culture-specific or religious beliefs.

Further, the metarepresentational module enables us to be imaginative, in
ways which the first-order modules would not allow, due to their restricted
domains. The metarepresentational module’s domain is representation as such, so
it has the means to manipulate the contents of the representations it represents,
provided that it remains in its second-order realm. That is, when we imagine
something, say, a pig growing wings and flying out of the window, we are aware,
metarepresentationally, that we are not entertaining true facts of the world, but
rather that we are playfully manipulating representations thereof. Such
imaginative abilities facilitated by the metarepresentational module play an
important role in the use of figurative language. We have the ability to distinguish
between representations of true facts in the world and representations of
representations, which we know are not of true facts in the world. But they can be
used to express a perspective on specific concepts, which we may not have the
literal words to express, and we may thus use blatantly imagined facts instead to
express them, such as metaphors.

Sperber developed the idea of metarepresentation within a modular
architecture further to include specific modules thought to have evolved as sub-
modules of the Theory of Mind module. While the latter’s function is to attribute
other people’s mental states in general (hence it is also called the ‘mindreading
module’), he postulates there also to be a comprehension module whose function
it is specifically to attribute speakers’ meanings in language understanding, and a module whose function it is to consider logical and evidential relations among propositions in the abstract. This further differentiation of the metarepresentational level accounts for the sophistication and speed at which humans are capable of understanding utterances and evaluating their truth and consistency (e.g. Sperber 2000; Sperber et al. 2010).

Seen as a whole, these metarepresentational functions may be the key for explaining metaphor understanding in particular, to do with the role in cognition of the literal meaning of a metaphorical expression. This will become apparent in later chapters, in particular in section 4.2. and in developing the main proposal of this thesis in chapter 6. Metarepresentational functions are a helpful notion in explaining how we perceive and represent these literal meanings in spite of them not being part of the propositional meaning outcome in utterance understanding. This is possible if we are able to represent them as mere representations, rather than true facts in the world, at the metarepresentational level. This is one of the key ideas the new hybrid model is based on (chapter 6).

2.5 Issues with Relevance Theory

In this section, I discuss some issues with the relevance-theoretic account which the model proposed here is going to tackle in later chapters. I address the criticism that the relevance-theoretic ad hoc concept construction process can account only for short metaphors, by attempting to carry out an analysis of a complex metaphor using said framework (section 2.5.1). This attempt does not seem to lead to a successful outcome, and it suggests that work needs to be done in order to make the account fit for complex metaphors also. Beginnings of this appear in Carston (2009; 2010), which will be reported on in sections 4.2.1 and 6.2. Section 2.5.2 discusses the notions of inference as opposed to association in online processing and describes how Carston (2008) discusses the semantics-pragmatics distinction, which is closely related to the notion of propositionality.
2.5.1 Ad hoc concepts in poetry: Issues with extended metaphors

The relevance-theoretic ad hoc concept construction account fares well with short nominal metaphors, which might be the reason why it is usually presented with short and rather simple examples. However, it comes up against serious difficulties in accounting for more complex metaphorical expressions. It then appears that analysis tends to amount to accumulations of single processing steps which eventually turn out difficult to disentangle. Furthermore, the connections between different yet related metaphors within longer stretches of text do not seem to be taken into account sufficiently in this analysis, a problem which will be discussed below. The following is a sketch of an analysis of part of a poem employing the ad hoc concept construction account. It demonstrates how it appears unlikely that comprehension of complex metaphors takes place employing ad hoc concept construction mechanisms as they have been proposed thus far. Example (14) below is taken from the poem ‘A valediction forbidding mourning’ by John Donne, reproduced here:

As virtuous men pass mildly away,  
And whisper to their souls to go,  
Whilst some of their sad friends do say,  
"Now his breath goes," and some say, "No."

So let us melt, and make no noise,  
No tear-floods, nor sigh-tempests move ;
'Twere profanation of our joys  
To tell the laity our love.

Moving of th' earth brings harms and fears ;  
Men reckon what it did, and meant ;
But trepidation of the spheres,  
Though greater far, is innocent.

Dull sublunary lovers' love  
—Whose soul is sense—cannot admit
Of absence, 'cause it doth remove  
The thing which elemented it.

But we by a love so much refined,  
That ourselves know not what it is,  
Inter-assurèd of the mind,  
Care less, eyes, lips and hands to miss.
Our two souls therefore, which are one,
    Though I must go, endure not yet
A breach, but an expansion,
    Like gold to aery thinness beat.

If they be two, they are two so
    As stiff twin compasses are two ;
Thy soul, the fix'd foot, makes no show
    To move, but doth, if th' other do.

And though it in the centre sit,
    Yet, when the other far doth roam,
It leans, and hearkens after it,
    And grows erect, as that comes home.

Such wilt thou be to me, who must,
    Like th' other foot, obliquely run ;
Thy firmness makes my circle just,
    And makes me end where I begun.

Here I focus on lines 22-24. While this is not even a particularly long stretch of a metaphorical expressions, analysis already shows which kinds of difficulties the framework comes up against, as will become apparent in the following.

(14) Though I must go, endure not yet
    A breach, but an expansion,
    Like gold to aery thinness beat.

In describing the relevance-theoretic comprehension procedure presumably employed in interpreting the figurative expression in line 3, I ignore that it is a simile rather than a metaphor, as, arguably, cognitively the difference between the two is of no importance here.\(^6\)

(15) Like gold to aery thinness beat.

Generally speaking, an ad hoc concept construction process consists in choosing among the encyclopaedic information associated with an encoded concept, with

---

\(^6\) I am aware that this is a controversial claim, with view to the relevance-theoretic ad hoc concept construction account in particular, since it is assumed that the processing of similes does not require ad hoc concept construction in the way that processing of metaphors does. However, for ease of exposition I leave this point aside here.
the denotation of the resulting ad hoc concept including a more or less different set of items than that of the encoded concept. In order to capture how ad hoc concept construction might go in this case, it seems necessary to divide up the complex metaphor into its conceptual constituents, which involves spelling out what each part of the metaphor may stand for. This way, it might be possible to model the process as it is thought to function with single lexical items, as a tool that enables me to model the analysis of this example after the relevance-theoretic analyses we have seen so far. These are all of the form ‘X is a Y’. Furthermore, in these examples X is invariably a human being, e.g. ‘Robert is a bulldozer’, ‘Caroline is a princess’, ‘Sally is a chameleon’, etc. So let’s see how the procedure can be thought to apply with a metaphor of a rather different form. In the course of this, we will see which hurdles we come up against.

As above, I use the established notation for lexical concepts (CONCEPT), and for ad hoc concepts (CONCEPT*, CONCEPT**, etc.) First of all, it is essential to take into account the discourse context, which is the theme of the poem: the separation of two people who are very close. This is alluded to in lines 1 and 2 of this excerpt, whereby expressions such as ‘breach’ and ‘expansion’ are, of course, part of the rather complex overall metaphor, but I shall focus on line 3, to begin with. The person who is going to leave (or more precisely, who is going to die) is trying to comfort their partner by expressing that the separation is not absolute, roughly paraphrased in (16):

(16) I am leaving you, but I can assure you that we won’t be separated completely, as we will have a divine bond.

Based on this context, the constituents of the metaphor in (15) can be spelled out as in (17) and (18), respectively:

(17) Our relationship is gold.
(18) A divine bond is gold beaten to thinness.
The context of the poem activates the hearer’s encyclopaedic assumptions about relationships, which, on the whole and for our purposes here should be straightforward enough and thus won’t need spelling out. The utterance itself activates the hearer’s encyclopaedic assumptions about gold, some of which might be as follows:

(19) GOLD: **Encyclopaedic assumptions**

- IS A TYPE OF METAL
- IS THE MOST PRECIOUS MATERIAL
- ITS POSSESSION ENHANCES ONE’S LIFE
- IS SOFT AND FLEXIBLE
- ETC.

Among others, the feature ENHANCES ONE’S LIFE can also apply to relationships, and in particular, to the very close kind referred to in the poem. Contextual assumptions and the presumption of relevance include that the expression should tell us something about the relationship, and so the most plausible inference is to apply this feature (among others) to the relationship just mentioned. This results in broadening the concept GOLD, by way of dropping the logical feature IS A TYPE OF METAL. We then arrive at the ad hoc concept GOLD*, which is associated with the denotation of GOLD as well as other things that have the feature ENHANCES ONE’S LIFE, including non-material notions such as relationships, i.e. it is a superordinate concept. The ad hoc concept construction does not end here, as in this case broadening results in a concept that is too large; it includes too many denotations to pick out the specific one our contextual assumptions lead us to. Hence, in order to make the concept more specific, we can assume a process of lexical narrowing to take place, such that the material aspect of GOLD* is excluded and the denotation that is left only denotes abstract entities, such as human relationships, love, etc. Thus we find that the lexical concept GOLD and the final ad hoc concept GOLD** merely overlap as concerns common features.

The examples of lexical broadening and narrowing just described take place in a parallel mutual adjustment process between implicature and explicature, i.e. the
encoded content, together with contextual information and encyclopaedic assumptions, and the presumption of relevance, functions as a guide in determining the meaning of the expression. I have now outlined how ad hoc concept construction in the case of (17) might go, which is only the first suggested constituent of the more complex metaphor in (15) (which in turn is only a small subpart of the metaphorical complex which runs through the poem as a whole). In order to provide the second metaphorical constituent, of which the content, together with (17), should provide for an understanding of the comprehension procedure of the complete metaphor (15) 'Like gold to aery thinness beat', I now try to describe the ad hoc concept construction that might be at work in the understanding of (20).

(20) A divine bond is gold beaten to thinness.

This attempt of analysis is based on rather constructed lexical concepts, which are probably too complex to be seen as realistic psychological entities. However, for the sake of demonstration I carry this through, since my aim is here to demonstrate how ad hoc concept construction is faced with problems when it comes to complex metaphors. Let’s first spell out the encyclopaedic assumptions for the topic of the metaphor, which is not an explicit part of the poem, but just provided to facilitate analysis:

(21) A DIVINE BOND: **Encyclopaedic assumptions**

- CONNECTS TWO PEOPLE
- HAS NO PHYSICAL REALITY
- REQUIRES TRANSCENDENT BELIEFS
- ETC.

The encyclopaedic assumptions for the metaphor vehicle include all those listed in (19) plus the following:
(22) GOLD BEATEN TO THINNESS: Encyclopaedic assumptions

- HAS BEEN TRANSFORMED
- HAS UNDERGONE A LOT OF SPECIALISED WORK (CRAFTSMANSHIP)
- IS OFTEN USED TO DECORATE OBJECTS
- ETC.

Among others, the feature HAS BEEN TRANSFORMED can also apply to the divine bond alluded to immediately preceding the utterance of the metaphor, which is line 2 of the poem: ‘not a breach but an expansion’. This is an obvious contextual assumption, especially because the poem is about the transition from the state of being together to the state of being apart. So, just as with the broadening of the concept GOLD above, a broadening of the lexical concept GOLD BEATEN TO THINNESS leads to a dropping of the logical feature IS A TYPE OF METAL. We then arrive at the superordinate ad hoc concept GOLD BEATEN TO THINNESS*, which is associated with the denotation of GOLD BEATEN TO THINNESS as well as other things that have the feature HAS BEEN TRANSFORMED, including the relationship the poem is about. So again we have arrived at a concept that has too many denotations to be meaningful in the given context, and to make it specific enough it undergoes narrowing, similar to the above. The concept we then arrive at, GOLD BEATEN TO THINNESS** has in its denotation a transformed relationship, but not the precious metal gold that is the denotation of the lexical concept GOLD. The resulting ad hoc concept carries a wide array of weak implicatures to do with a relationship that has transcended to another level of consciousness, and the like. It should be noted that the processes of broadening and narrowing described here are just examples of the manipulation of many more potential conceptual elements of the poem, and that it does not just hinge on one or two features, but many more.

Obviously, this description, while clumsy due to all sorts of technical difficulties, is rather idealised. The main goal of this exercise was to demonstrate the degree of complication that ad hoc concept construction as a means of analysis of complex metaphors comes up against, even if we take just a small part of the larger picture, i.e. the whole poem. It seems unlikely that this way of simulating online processing can reflect psychological processing in any realistic fashion,
which has also been mentioned by Carston (2010: 12). In referring to the same poem, she focuses on the metaphor of the two feet of a mathematical compass, lines 25-32, which can be analysed to stand for the souls of the two lovers in the poem, whereby the middle of the compass keeps the two together:

If they be two, they are two so
As stiff twin compasses are two;
Thy soul, the fix'd foot, makes no show
To move, but doth, if th' other do.

And though it in the centre sit,
Yet, when the other far doth roam,
It leans, and hearkens after it,
And grows erect, as that comes home.

Given the central theme of the poem, the ‘gold’ metaphor and the ‘compass’ metaphor can be understood to share the same underlying target: the connectedness between the two lovers beyond doubt (and beyond the grave). In the ‘gold’ metaphor this is expressed by the apparently indefinite length to which gold can be stretched, which can be interpreted to stand for distance not being able to drive them apart. In the ‘compass’ metaphor it is expressed by the connecting middle, which keeps the feet of the compass together even if one moves away from the other. Thus, we have a connection between two metaphors which have different sources, but which both allude to the same state of affairs. Ad hoc concept construction does not seem able to account for this kind of relationship between different source domains. By comparison, it might be possible that Conceptual Metaphor Theory can account for this kind of phenomenon, if we assume there to be a very schematic underlying experiential pattern of, say, physical entities and how they hold together or come apart, and this underlying pattern is schematic enough to accommodate all the related analogical connections, which are then creatively elaborated on. However, while it is an interesting point to be made here, the notion of conceptual metaphors and schematic patterns will be discussed and elaborated on in later chapters (especially sections 3.3.3 and 3.3.4).
Another important issue is that of extended metaphors which continue along a longer stretch of text within one conceptual domain. Ad hoc concept construction does not involve conceptual domains as a notion, which could explain how semantically related expressions form different yet connected metaphors. Thus, within this framework it does not seem possible to account for these relations within a text. They can be assumed, however, to play an important role in cognitive processing, as semantic priming effects impact on comprehension. I have discussed the issue in my MA dissertation (Stöver 2006), where I exemplified this phenomenon with a monologue from Shakespeare’s ‘Macbeth’, reproduced here:

**LADY MACBETH**
Give him tending;
He brings great news.

*Exit Messenger*

The raven himself is hoarse
That croaks the fatal entrance of Duncan
Under my battlements. Come, you spirits
That tend on mortal thoughts, unsex me here,
And fill me from the crown to the toe top-full
Of direst cruelty! make thick my blood;
Stop up the access and passage to remorse,
That no compunctious visitings of nature
Shake my fell purpose, nor keep peace between
The effect and it! Come to my woman's breasts,
And take my milk for gall, you murdering ministers,
Wherever in your sightless substances
You wait on nature's mischief! Come, thick night,
And pall thee in the dunnest smoke of hell,
That my keen knife see not the wound it makes,
Nor heaven peep through the blanket of the dark,
To cry 'Hold, hold!'

The general theme of the monologue is Lady Macbeth’s wish to alter her personality to enable her morally to commit the crime of killing the King (or convince her husband to do so). Here, the body functions as a metaphor of her personality, and what follows is a host of interrelated bodily metaphors, which reinforce one another and thus have a semantic priming effect. For example,
conjuring up the evil spirits, she asks them to ‘make thick my blood’ (line 8). This being preceded by the metaphor of cruelty as a liquid the body can be filled up with (lines 7-8), the figurative meaning extends here to the image of cruelty mixing with blood to make it thick, which is reinforced by the metaphor that follows it, ‘Stop up the access and passage to remorse’ (line 9), evoking the idea that thickened blood (still within the extended metaphor) stops the flow of any virtuous characteristics, the veins in the body here also seen as the metaphorical residence of personality. This is related to the traditional metaphor of the heart as the seat of love and kindness, and once its activity of pumping blood through the veins is slowed down, these positive character traits are weakened. These bodily images prime the following, ‘milk’ being exchanged for ‘gall’ (line 13), which will be very easily accessed in comprehension after all the semantic priming that has gone on before. But I would suggest that it is not only accessibility of encyclopaedic features that is increased by several metaphors sharing the same source domain. This is something that can be explained within Relevance Theory, employing the notions of contextual assumptions and activation levels according to the presumption of optimal relevance. However, it seems that in addition, it is also the case that the imagery is intensified by creating this complex picture of interrelated metaphors, and it is experienced in a much livelier way than would have been the case with metaphors from unrelated source domains. These effects seem difficult to account for with standard relevance-theoretic explanations. Therefore, at least for complex metaphors, it can be assumed that there are other kinds of processes at work in comprehension. To account for these, Carston (2009, 2010) proposes a model of dual processing, where one processing route involves metarepresenting the literal meaning of complex metaphors as a whole. This will be explained in sections 4.2.1 and 6.2.

2.5.2 The nature of propositions: The ‘inferential fitness’ issue
The new hybrid model developed in chapter 6 has as one of its main characteristics a clear-cut distinction between propositional and nonpropositional representations, which are here thought to each undergo different kinds of
cognitive processing in utterance understanding. The reasons for assuming this cognitive structure are based on the relevance-theoretic notion of conceptual processing as computation within the medium of LoT (cf. section 2.3). And while the present proposal is influenced to a large part by Tendahl’s (2006) ‘Hybrid Theory’ (cf. section 5.2), what distinguishes it from Tendahl’s model is this characteristic in particular. Tendahl does not employ the notion of LoT, nor does he seem to assume different processing mechanisms for the respective types of representation. Therefore, it is of immediate pertinence, and a theoretical precondition for the describing the development of the new hybrid model, to explain how this type of inference, which takes propositional input only, is thought to work. A substantial part of chapter 6 builds on this definition of ‘inferential fitness’, to be explained in the following.

One of the major differences between Relevance Theory and Conceptual Metaphor Theory is that in the former, conceptual processing is assumed to involve a significant amount of deductive and non-demonstrative inference, a characteristic which underlies the main principles of the theory. In the latter, if it is considered at all, conceptual processing is thought to be based on association. Relevance theorists reject the idea of associationism on the following grounds: if utterance interpretation were to proceed via association, the processes involved would lack directionality, on the one hand, and lead to an overgeneration of conceptual features, on the other, since association is not constrained in any obvious way. Inference, by contrast, guarantees a logical or evidential relation between processing input and output, which has been explained in section 2.3. In the following, I discuss the notion of propositionality and its importance in an inferential account of communication.

Carston (2008) explains her notion of a minimal linguistically encoded meaning (LEM), which is to be understood as a schematic pre-propositional entity. The discussion involves the notion of propositional content as understood in traditional semantics, and whether, or how, this notion works for an account of communication, since it has originally been developed for accounts of thought. An important point here is that traditional truth-conditional semantics cannot be the same as LEM, since a truth-conditional semantics includes too much content
that, according to Carston, can only be pragmatically derived. Hence LEM can only be pre-propositional, that is, non-truth-conditional. Any other proposals would blur the semantics-pragmatics distinction. Carston emphasises that the idea that semantics ought to be truth-conditional is traditionally rooted, and that we could do very well without it in the domain of linguistic communication and work with a notion such as LEM instead. Ultimately, the discussion points towards the idea that linguistic codes are just a guide or an instruction to create concepts from available mental resources, as argued e.g. by Carston (2002) and Pietroski (2005).

Based on the idea of a pre-propositionality of LEM, the issue of ‘inferential fitness’, the established idea that inference has to have propositions as premises, seems to become a problematic issue. Below, I briefly explain how relevance theorists deal with this, given that in Relevance Theory, inference plays a major role in comprehension (although perception and activation also play a role). An important point here is that in the relevance-theoretic framework, representations of the kind traditionally understood as non-propositional are used as input for inference, i.e. sub-propositional conceptual content (LEM). That is, since in Relevance Theory, sentence meanings are thought to be fragmentary and incomplete, and achieve propositional status only through inferential processing (Sperber & Wilson 1986/95, and elsewhere; Carston 2002, and elsewhere), it is not farfetched to ask how fragmentary sentence meanings can undergo inferential processing in the first place if we take seriously the rule that inferential processes take propositional content only. The relevance-theoretic solution to this seems to be that once something sub-propositional is inserted into a metarepresentational propositional schema, it becomes propositional. This will be explained below.

Recanati (2002; 2004) disagrees with this specific point. He therefore arrives at the assumption that primary pragmatic processes are associative, and so, in his framework, inference sets in only after conceptual content has become propositional, in his so-called secondary pragmatic processes (described in the following section).

There is an ongoing discussion in the philosophy of language about how much of the meaning communicated by an utterance exists prior to the utterance, and how much is derived by pragmatic processing in context. This is an important
issue in the analysis of the comprehension of figurative expressions, as deriving their meaning in particular depends heavily on contextual information. Carston (2008) discusses some of the current approaches and compares them to her own favoured position, which is minimalist in the sense that there is a minimal linguistically encoded meaning (LEM), and everything else is pragmatically derived. She discusses several other views concerning the semantics-pragmatics distinction, which I will not follow up here. The main concern in this discussion is how the notion of semantics can be seen as playing a role in a theory concerned with linguistic communication (as opposed to a theory of thought), which in particular depends on how the term is defined. Traditionally, semantics involves truth-evaluable propositions, which is a notion that has its origin in the realm of thought rather than communication. Hence the application of a so defined semantics to linguistic codes, which belong to an entirely different domain, requires some scrutiny. Propositions are either true or false, whereas utterances are representations of thoughts that contain propositions. Utterances are made up of linguistic items, which in turn are given propositional meaning via pragmatic processes in context. So linguistic items are pre-propositional, as they have not yet undergone the pragmatic processes that make them propositional. Therefore, if there is a semantics to play a role in linguistic communication, it needs to be construed as pre-propositional (or be given a label other than ‘semantics’), especially because, if a distinction is to be drawn, semantics should be autonomous from pragmatics, which it cannot be if it relies on pragmatic processes to become fully propositional. And this would blur the distinction – it would be of no use to theorise a separation of one from the other under the condition that the existence of former depends on the latter (i.e. a semantics that requires pragmatic processes for completion). Thus, on this view, a sentence is just ‘a schema or template for building propositions’ (2008: 7). Ultimately, it seems that the controversy about which part of the whole belongs to which, either semantics or pragmatics, might be solved by looking at the issue in a different way altogether, namely by conceiving of semantics as residing at an entirely different level of theorising. Semantics is concerned with unspecified sentence types, and pragmatics with specific tokenings of the former, i.e. with individual
occurrences of utterances. This way of looking at it keeps the two neatly apart, and it justifies Carston’s insistence on keeping the semantics of an expression as unspecified as possible. Thus, in this framework, we can conceive of three layers that make up the meaning of an utterance: 1. the pre-propositional LEM (minimal semantic content) 2. the proposition expressed/explicature (which is LEM that has undergone saturation and free enrichment) and 3. the full meaning of the utterance which includes implicatures.

Another important term in this discussion is the expression ‘what is said’, coined by Grice (1989). Recanati (2004) understands ‘what is said’ to be the meaning of an utterance that exists prior to implicatures, which are being worked out after ‘what is said’ has been computed. Depending on one’s conception of how comprehension works, ‘what is said’ can be construed as existing at different levels of processing (or pre-processing): at the level of the linguistically encoded meaning or at the level of saturation, such as disambiguation and reference assignment; or ‘what is said’ might be determined at the later stage of free enrichment. These stages are described further below. Ultimately, this discussion is about where the line is to be drawn between semantics and pragmatics, how far ‘literal meaning’ can be conceived of as being semantically encoded or pragmatically derived. Clarification of the issue, or at least a description of it, is an important preliminary for working out how the interpretation of figurative language might go in comprehension, if one wishes to specify the different types of processing involved, which process might be employed at which stage, how memory is drawn on, etc.

Recanati explains that in the minimalist tradition (which he rejects), ‘what is said’ may go beyond the linguistically encoded meaning of a sentence only insofar as saturation is concerned. This means that only linguistically mandated processes can contribute to ‘what is said’, i.e. the filling in of empty slots in the structure, such as given in deictic expressions such as personal pronouns and locations relating to the interlocutors’ spatial locations, e.g. ‘here’ or ‘there’, etc. So, ‘what is said’ is conceived of as something existing to a large extent outside and independently of speech situation and discourse context, with the exception of contextual elements that must be attended to as prescribed by the constituents of
the sentence (saturation). Recanati describes the latter as bottom-up processes, which follows logically from the idea that the content of ‘what is said’ is built exclusively out of the given linguistic form, insensitive to any external influences that are not linguistically mandated. Which is why this approach is called minimalist, it keeps the involvement of the discourse context in deriving the content of the literal meaning minimal. ‘What is said’ is here equivalent to the proposition expressed, the minimal semantic entity that is truth-evaluable. So in the minimalist framework the distinction is made between literal meaning and speaker’s meaning: the literal meaning is made up of the sentence meaning (linguistically encoded) and ‘what is said’ (the proposition expressed, resulting from saturation). The speaker’s meaning comes about when the literal meaning undergoes free enrichment, which is a specification of the literal meaning with respect to the discourse context that is not linguistically mandated, but does not yet fall under the label implicature (in this framework anyway). For example, someone replies to the question whether they are hungry ‘I’ve had breakfast’. Due to free enrichment, this utterance is understood to mean that the speaker has had breakfast on the day of the utterance rather than, say, a year ago. There is no empty slot or any other linguistic motivation that provides for this information to be filled in, but it is not a case of implicature either. The implicature here, in the Gricean sense, would be that the speaker is not hungry.

Whereas Carston distinguishes between layer one, on the one hand (LEM), and layers two and three, on the other, Recanati draws the line between layers one and two versus layer three, i.e. he draws the distinction between the intuitive proposition expressed and speaker meaning, derived by primary and secondary pragmatic processes, respectively. Carston acknowledges this as a valid distinction, with the caveat that intuitions vary, and so are not reliable entities that could be utilised to draw clear distinctions. But she points out that, since an intuitive proposition expressed relies on pragmatic processes, this is not a distinction between semantics and pragmatics, but rather between different stages of pragmatic processing (or different levels of speaker meaning). Furthermore, her distinction has the advantage that it corresponds to a difference between the psychological processes employed: LEM is derived via decoding processes,
whereas the rest of the meaning is inferentially derived. However, as Recanati sees it, the distinction that he draws is also along the lines of distinct psychological processes. In his framework, which will be outlined in more detail in the next chapter, there are primary pragmatic processes which are associative, and secondary pragmatic processes which are inferential. There is a specific reason for this conception of primary pragmatic processes, i.e. the derivation of the explicature, or ‘what is said’, as non-inferential: inference presupposes propositional content as premises, which LEM by itself cannot provide. This is why in his and similar frameworks the divide is between a semantics that must include a minimal proposition expressed and the rest of speaker meaning. Here, the minimal proposition expressed has to be partly context-dependent, as it has been shown that there cannot be enough shared content across contexts to make encoded meaning fully propositional. This is, in Carston’s view, at odds with an actual distinction between semantics and pragmatics: in order to achieve a minimal proposition expressed, pragmatic processes, such as reference assignment etc., need to be employed, which again results in an undesirable intrusion of pragmatics into semantics.

The issue of ‘inferential fitness’, i.e. the notion that premises for inference are necessarily propositional is something that needs to be addressed, in particular with view to the relevance-theoretic account of utterance comprehension, according to which the derivation of speaker meaning is entirely inferential, once LEM has been decoded. How can this be? Seeing that LEM is just like an empty shell to be filled in, a mere conceptual sketch that is by no means propositional but rather functions as a guide for pragmatic meaning completion, the claim that it could function as a premise for inference seems like an obvious contradiction. The solution that relevance theorists have come up with goes as follows: upon hearing an utterance consisting of underdetermined, non-propositional constituents, the hearer embeds it in a schema of some sort which renders the utterance propositional. It looks something like this: ‘The speaker has said that [non-propositional, logical form]’. Thus, the hearer works with the proposition that consists in the fact that the speaker has expressed something, which is a representational form that can be used as a premise for inference. So this schema
functions as a kind of inferential fitness generator, by means of shifting the linguistic expression to a different representational level. This is related to the hearer’s presumption of relevance, i.e. that the hearer can expect that what the speaker has said is relevant in the specific discourse context for the hearer to be processed, which in turn guides the derivation of speaker meaning. This proceeds with the relevance-theoretic comprehension heuristic, as described above in section 2.1. The mere fact that an utterance is intended to communicate something sets in motion the process of a propositional embedding in the form of ‘The speaker has said that [...]’, so that inferential processing can run its course. Thus, we can say that the embedding of pre-propositional content in a schema goes hand in hand with the presumption of relevance. This means that we have a metarepresentational level at which inference can be employed, and where the mutual adjustment process between implicature and explicature kicks off: with this initial proposition as input (‘The speaker has said that [...]’), inferential processes, including implicatures at an early stage at the conceptual level, lead to determining the explicature. This is an issue that plays an important role in the new hybrid model developed in this thesis, and it will be taken up again in chapter 6.

The idea just outlined suggests that a central relevance-theoretic notion, namely that in utterance comprehension explicatures are derived via inferential pragmatic processes, seems to rest entirely on the assumption that the hearer puts the pre-propositional material in ‘some schema’, which makes it fit for inference. This interplay between the level of decoding and the metarepresentational level at which the hearer employs the schema, i.e. the getting from the proposition ‘The speaker has said that [...]’ to determining the content of [...] is equivalent to what has been described as the ‘parallel mutual adjustment process between explicature and implicature’, which seems a very insightful conception as such. However, the employment of ‘a schema of some sort’ that renders something non-propositional into something propositional seems to be a controversial idea. Recanati (2002), for example, disagrees with the idea, he argues that just embedding non-propositional material in a propositional schema still doesn’t make the actual content itself propositional (2002: 26). Accordingly, in his framework of utterance
comprehension, processes other than inference must be at work in deriving the content of ‘what is said’, processes that do not presuppose propositional content. For him, these are associations, and we find them in his associative primary pragmatic processes, which will be described at the beginning of the next chapter (section 3.1).
Chapter 3
Alternative Approaches

After an outline of the relevance-theoretic approach in the last chapter, this chapter discusses alternative approaches to communication in general, and to metaphor in particular. This is ‘to set the scene’, to provide some theoretical contextualisation of the present proposal and to outline some of the approaches that have influenced and inspired it. It shows alternative points of view in order to emphasise that, while it is important to observe relevant constraints on theorising (an issue that will be discussed in more depth in chapter 5, in particular), it is equally important to keep an open mind in assessing the theoretical value of alternative explanations of the phenomena under investigation.

The main focus of this chapter is on Conceptual Metaphor Theory (section 3.3), since it is one of the main building blocks of the hybrid theory proposed. That is, while Relevance Theory takes here the role mainly of accounting for inferential processing of propositions in utterance understanding, processes to do with nonpropositional, or more impressionistic representations, are based largely on ideas from Conceptual Metaphor Theory. Therefore the main principles of the theory, and some of its developmental history, will be explained, since integral parts of the proposal in chapter 6 are based on this.

The chapter begins, however, with resuming the discussion from the previous chapter in describing an approach that is relatively close to Relevance Theory (Recanati 2004, section 3.1). The main difference is that it incorporates association processes in addition to inferential processes. Thus, it demonstrates the possibility of an account that integrates different types of processing, and is therefore an example case for the proposal developed in chapter 6, in that it also assumes an interplay of inference and association at different levels (here the ‘personal’ and the ‘subpersonal’ level). And while Recanati’s approach is not employed to play a role in the proposed model directly, it serves the purpose of theoretical contextualisation.

In addition, in order to give an idea of what is possible in terms of envisaging approaches that deal with propositional models of representation,
section 3.1 outlines an approach that introduces ‘gappy’ propositions, i.e. propositions that are not fully fledged, yet still fulfil the functions of full propositions (Corazza & Dokic 2007). Again, this is an example case of an alternative approach to provide a backdrop of varied theoretical approaches to contextualise the present proposal. In particular, it demonstrates a flexible view of what propositions as such can be understood to represent, and while the present approach does not incorporate Corazza & Dokic’s notion of propositions directly, the latter offers inspiration not least pertaining to what may be possible when refining the present model in future research.

Section 3.3.1 gives an overview of Conceptual Metaphor Theory (e.g. Lakoff & Johnson 1980), and section 3.3.2 discusses how it has been developed further (Grady 1997). Section 3.3.3. outlines Barsalou’s (1999; 2005; 2009) model of situated conceptualization. The version of conceptual metaphors employed in the proposal in chapter 6 is based on it to a large extent. It is defined in section 3.3.4, with an orientation towards what seems psychologically plausible for such a notion, based on the criticisms discussed.

In order to demonstrate practical applications of Conceptual Metaphor Theory (if not empirical evidence), the chapter concludes with examples of how the notion of conceptual metaphor has been used productively in second language acquisition and related research.

3.1 Recanati: Primary and secondary pragmatic processes
Recanati (2004) discards the notion of literal meaning as described in section 2.5.2 above, and instead makes a distinction between sentence meaning and speaker’s meaning. Here, sentence meaning is just the linguistically encoded meaning, like a skeleton that needs fleshing out, and everything else of ‘what is said’ is pragmatically derived. So ‘what is said’ appears as the result of saturation as well as free enrichment and other pragmatic processes such as loosening.
Recanati labels all processes that are involved in the determination of ‘what is said’ primary pragmatic processes. Hence the proposition itself is to a considerable extent derived by pragmatic processes, as opposed to the traditional
minimalist idea that pragmatics comes into play after the proposition has been determined. Processes that take the fully fledged proposition as input are what Recanati labels secondary pragmatic processes, which are implicatures in the Gricean (1989) sense. Accordingly, primary pragmatic processes are pre-propositional, whereas secondary pragmatic processes are post-propositional. Recanati arrives at this conception of utterance interpretation by rejecting the traditional idea that there is a meaning of an utterance that exists outside or prior to the speech situation, and by defining utterance meaning with view to the actual processing outcome, i.e. what interlocutors would actually report to have said or understood. So on this view, ‘what is said’ must be intuitively accessible to the conversational participants, rather than a semantic construct that would be difficult to grasp for a non-linguist (2004: 20). He emphasises that this psychological conception of ‘what is said’ is much more compatible with the Gricean maxims than the minimalist notion of ‘what is said’. This is because in the non-minimalist framework, ‘what is said’ is sufficiently fleshed out to allow for implicatures to arise, whereas the minimalist ‘what is said’ needs to undergo several further processing stages, such as free enrichment, before it can provide a basis for implicatures.

According to Recanati, local processes of lexical access are necessarily associative. His reason for this claim is that inferential processing presupposes computation of the whole sentence, that is, of the proposition expressed, and the derivation of the meaning of constituents takes place before arriving at the global sense of the sentence. Accordingly, he assumes that only secondary pragmatic processes, which are post-propositional, can be inferential. This calls for a comparison with the relevance-theoretic account, according to which all processes, including those at the pre-propositional lexical level, are strictly inferential. Wilson & Carston (2007) argue that their approach, i.e. Relevance Theory, is to be preferred for theoretical reasons, as purely inferential processing is much more constrained than other processing (cf. section 2.3). They reject such a mixed-type approach, as well as purely associative approaches, on the grounds that the activation of more representations than necessary requires a further processing stage of ‘filtering out’ the inappropriate representations (2007: 35). It
would be worth investigating in the future how approaches that are not strictly inferential, including Recanati’s, might construe these further processes of suppressing unwanted representational material, what they might look like and whether they really do constitute unnecessary processing effort.

Recanati draws attention to the traditional view, where it is thought necessary for the literal meaning of the sentence, i.e. the full proposition, to be computed before any other interpretive steps can be taken, such as an apparent ‘move’ from the literal meaning to the figurative meaning of an expression. This idea is not compatible with Recanati’s model, or any other contextualist account; however he does assume it to hold for constituents, of which the literal meaning is thought to be always activated first in comprehension, independently of the proposition expressed. The latter is thought to come about at a later stage in the overall interpretation process. The determination of the meaning of constituents as they feature eventually in the proposition expressed is here described as follows: the activation of the literal meaning of a constituent initially triggers by association the activation of all kinds of related representations, which are processed in parallel. All such activated meanings, the derived as well as the literal, compete on equal footing. Contextual clues, either pertaining to the linguistic or the discourse context or both, then increase or decrease the activation of the representations involved, until one has succeeded (probably by exceeding a threshold), and this process may quite possibly lead to a suppression of the literal candidate. Recanati calls this a ‘temporal shift in accessibility’ (2004: 30), which he demonstrates with the classic example ‘The ham sandwich has left without paying’: The utterance activates a representation of an actual ham sandwich, which triggers associatively related representations, including the ham-sandwich-orderer. Once the second part of the utterance ‘has left without paying’ influences the competition between all sorts of associatively activated representations, the ham-sandwich-orderer succeeds as the only plausible candidate, as its activation is now the strongest by way of fulfilling the status of argument in ‘...has left without paying’.

The idea that all related meanings are triggered once the encoded sense is activated seems plausible, in particular because there have been psycholinguistic
experiments that investigate the processing stages of comprehension in cases of polysemy, showing that all meanings of ambiguous words are being automatically activated upon hearing the word, with contextually inadequate meanings being suppressed at a later processing stage. A classic example is the cross-modal priming study carried out by Swinney (1979). Swinney’s participants listened to sentences containing an ambiguous word, e.g. ‘He swam across the far side of the river and scrambled up the bank before running off.’ Immediately after they heard the ambiguous word, with varying onset times to test the different processing stages, the participants had to make a lexical decision task concerning the meaning of a visually presented word, which supported either of the two meanings (e.g. ‘river’ or ‘money’). He found that at an early stage of processing all meanings of the polysemous word were activated. This is because lexical access happens automatically without regard to context at the earlier stage (1979: 656-8). So would it be too farfetched to assume that not only polysemous meanings but also all related meanings are triggered resulting from activation of the encoded meaning? Although relevance theorists do not seem to share this idea, as such associative triggering is thought to lead to undesirable overgeneration of features, it is worth thinking about whether their treatment of polysemy, described below, suggests a generalisation that accounts for polysemous meanings as well as for related meanings, as an expression might undergo different stages of figurativeness and encoded polysemous meanings, due to semantic change: through frequent use, a figurative meaning that has been merely related to the linguistically encoded meaning might catch on in a speech community, so that in time the related meaning will become an encoded meaning, having created a new case of polysemy. This will be different for different people at different times. But even if for one individual the occasion-specific meaning has not yet become an encoded sense of the word and for another individual it already has, with respect to the speech situation, and expectations of relevance, etc., they will arrive at the same understanding of the intended meaning, albeit in different ways (with the usual risk of misunderstanding) (Wilson & Carston, 2007). Hence the differences between polysemous and encoded meanings can be seen as lying on a continuum, as is the case with nearly all ‘categories’ in Relevance Theory. Here, the
underlying rationale is to achieve the widest possible generalisation that can still explain individual cases adequately and sufficiently. However, Wilson & Carston (2007) do not assume the same processes to be involved in the understanding of polysemous and figurative meanings, respectively: in their view, understanding a polysemous expression goes via disambiguation, i.e. choosing between the two or more encoded meanings, whereas understanding a figurative expression goes via ad hoc concept construction.

3.2 Corazza & Dokic: Incomplete propositions
The discussion I have outlined so far rests on the assumption that propositions are fully determined in their content. Just like thoughts that we are consciously aware of, they cannot be ‘gappy’ or incomplete, and even if the views described here are incompatible in other respects, there definitely is agreement on this point. But surprisingly, there seem to be a few thinkers nowadays who advocate an entirely different view. They envisage the existence of underdetermined thoughts which, similar to linguistic expressions, have a partly non-propositional or underdetermined character, an idea of which the recognition would presumably change the arguments presented here considerably, in particular as concerns inferential fitness, i.e. that a premise for inference must be fully propositional. So the question would be, if we had some kind of semi-propositional content, would it qualify as input for inferential processing? Perhaps as input to a different kind of inference? Or a combination of inference and a different kind of processing? The possibility that comprehension comprises different types of processing that complement each other is discussed in chapter 6.

An example of propositional form that is not fully determined has been proposed by Corazza & Dokic (2007). They disagree with contextualists like relevance theorists and Recanati, who assume that the content of a proposition must be fully explicit, which is usually construed as the outcome of free enrichment. For example, on the contextualist view the expression ‘Jane is too tall’ is seen as not fully propositional, as it contains an unarticulated constituent. For it to be fully propositional, it would need to be enriched by way of taking
contextual information and the speaker’s intentions into account, as in ‘Jane is too tall [for a jockey]’. Corazza & Dokic take the expression to be propositional without enrichment, because ‘Jane is too tall’ can be understood by people across contexts, regardless of speaker intention. This would be different with an expression containing indexicals, of which the referents need to be known so it can be understood, i.e. reference assignment is a mandatory requirement for rendering an expression propositional. For example, a context-independent meaning of ‘He was ready to go then’ is difficult to imagine. Expressions containing unarticulated constituents, on the other hand, present us with a minimal proposition, which can have different truth values in different situations. This is why Corazza & Dokic call their position situationalism or situated minimalism, as the truth of a minimal proposition is relative to the situation it occurs in: ‘Jane is too tall’ has a positive truth value in the situation of her wanting to be a jockey, but has a negative truth value in the context of playing basketball, i.e. ‘Jane is too tall [for a basketball player]’ would not be true. This shows that Corazza & Dokic have a fundamentally different conception of thoughts, such that thoughts can have parts that are not explicit and don’t need to be made explicit, either. They assume that incomplete propositions such as ‘Jane is too tall [situational truth value]’ can be cognitively maintained. To be clear, by this they do not mean that there is some sort of default interpretation for expressions with unarticulated constituents, which would presumably look like this: ‘Jane is too tall [compared to the average height of the population]’. Rather, they assume ‘Jane is too tall’ to be represented as ‘Jane is too tall [gap]’ and nothing more, i.e. they assume a ‘gappy’ proposition. Intuitively, this doesn’t feel wrong to me, at least I can imagine thoughts of which parts are just left blank, given that the parts in question are not needed for the specific computations at hand, which is why we still understand utterances like this, even if no context for free enrichment is provided. Contextualists would disagree with this, claiming that minimal propositions of this form have no psychological reality. The argument Corazza & Dokic put forward against this is that it is unlikely in general that speaker and hearer arrive at entertaining exactly the same representations; at best there is a similarity between entertained propositions (cf. interpretive resemblance), which is
sufficient for communication to be successful. To support their claim, they quote Sperber & Wilson who originally put forward this point:

It seems to us neither paradoxical nor counterintuitive to say that there are thoughts[propositions] that we cannot exactly share, and that communication can be successful without resulting in an exact duplication of thoughts in communicator and audience.

Sperber & Wilson 1986/95: 193

It follows that, if a similarity of representations is enough, one can argue that minimal, i.e. non-enriched propositions similarly suffice to reach understanding. But is this correct? This is difficult to decide, as there seem to be no other means but introspection to find out what thoughts might look like. Be it as it may, this conception of minimal propositional content definitely deserves further investigation in the future, in particular with respect to how it could be utilised in a cognitive account of comprehension procedures.

3.3 Conceptual Metaphor Theory

This section describes the main characteristics of Conceptual Metaphor Theory as originated in Lakoff & Johnson (1980) and discusses criticisms of it. Section 3.3.2 outlines Grady’s (1997) proposal of a more structured definition of conceptual metaphors, which involves a compositional approach. Section 3.3.4 offers a new definition of conceptual metaphors, which is the one adopted in the model proposed here. It is based on the notion of situated conceptualization developed by Barsalou (1999; 2005; 2009), which is therefore outlined first (section 3.3.3). The idea of conceptual metaphors as underlying thought is then revisited from a multilingual perspective (section 3.3.5): first a brief outline of a study in second language acquisition (Boers 2000), which demonstrates that Conceptual Metaphor Theory can have beneficial practical applications, regardless of whether the theory is supported by empirical evidence, then a proposal for a second language teaching technique based on Conceptual Metaphor Theory by Sabine deKnop
(2008), and a study that works with the assumption that learning prepositions in a second language might require reorganisation of (supposed) underlying conceptual structures (Ijaz 1986).

3.3.1 Lakoff & Johnson: ‘Metaphors We Live By’
Cognitive Linguistics is a wide field with various strands of approaches to language and cognition. Here I focus on Conceptual Metaphor Theory (e.g. Lakoff & Johnson 1980; Lakoff 1987; Johnson 1991), since it is the most popular and most influential approach within the field. It seems to have been there right from the beginning of the cognitive-linguistic movement, and many researchers build their work on its principles without questioning them (McGlone 2007: 110). At the same time, it is a fairly controversial theory that has been fiercely criticised by theorists outside Cognitive Linguistics (McGlone 2007; Vervaeke & Kennedy 1996; Vervaeke & Green 1997; Keysar et al 2000). While it does not always seem theoretically conclusive, as will be outlined below, it has strong intuitive appeal and provides explanations concerning a perceived connectedness of representations in cognition by way of stating a metaphorical interrelatedness between concepts. Some might say that conceptual metaphor theorists have gone too far in claiming to have found the key to how cognition is organised and to know, even, the lines along which people think. I would agree with them, but I suggest that this is an issue that one should not take as a cause to reject the notion of conceptual metaphor per se. It may be worth to develop a moderate version of conceptual metaphor which does not seem overly patronising in terms of claiming to be able to predict the ways in which people think. It could, after all, help explain one way in which our understanding of things in the world (abstract and concrete) might be facilitated to a certain degree.

Whereas Relevance Theory views metaphor as originating primarily in language, Conceptual Metaphor Theory, like Cognitive Linguistics in general, conceives of metaphor as originating in thought. According to Lakoff & Johnson (1980), most of our thinking rests on metaphorical mappings, involving schemas such as PHYSICAL FORCE IS PSYCHOLOGICAL FORCE, which puts elements of the
psychological domain in correspondence with elements of the physical domain, and this is understood to be reflected in most of our everyday language. These domain mappings are thought to be fixed entities in long-term memory, called conceptual metaphors. For example, the above conceptual metaphor would give rise to expressions like ‘I’m depressed’. Our thoughts about the psychological domain are conceived of as being modelled on the physical domain, such that a forced downwards movement represents negative feelings: psychological forces are weighing us down, so to speak. Accordingly, our entire conceptual system is seen as metaphorically structured in this fashion, which is thought to be based on recurrent patterns of our bodily interaction with the world. Often these are orientational metaphors, grounded in our bodily orientation in physical space (cf. Gibbs 2006). The most obvious examples are the metaphorical conceptualizations HAPPY IS UP and SAD IS DOWN, which are derived from the bodily postures expressing the respective emotions (Lakoff & Johnson 1980: 14-15). This UP/DOWN orientation gives rise to an ‘overall external systematicity’, in the sense that following the conceptual metaphor HAPPY IS UP, a wide range of positive experiences are conceived of as UP, e.g. HEALTH IS UP, CONTROL IS UP etc. (1980: 18). Most conventional expressions are thought to have their origin in conceptual metaphors. For example, the image schema BALANCE appears in all kinds of arenas of experience which are otherwise not related, such as the psychological, mathematical, legal, etc. (Johnson 1991). Lakoff & Johnson take the idea that cognition is structured in this way a step further and claim that our cognitive system has to rely on conceptual metaphors in order to understand language at all:

‘...many aspects of our experience cannot be clearly delineated in terms of the naturally emergent dimensions of our experience. This is typically the case for human emotions, abstract concepts, mental activity... Though most of these can be experienced, none of them can be fully comprehended on their own terms. Instead, we must understand them in terms of other entities and experiences, typically other kinds of entities and experiences.’

Lakoff & Johnson 1980: 177
So for Lakoff & Johnson it follows that an expression like ‘You’re running out of time’ (1980: 8) reflects that time is conceived of as a physical entity, and so it is seen to reflect a metaphorical thought. Since they prefer their conceptual metaphors to be at the most concrete level, the expression is thought to go back to the conceptual metaphor TIME IS MONEY (the issue of metaphor levels and primary metaphors will be discussed below), even though an untutored person might perceive the expression as straightforwardly literal. It follows that to Lakoff & Johnson most, if not all, ordinary language can be traced back to specific conceptual metaphors. Consequently, they dispose of the category boundary between the literal and the metaphorical altogether. One could argue that this is something they have in common with relevance theorists, who claim there to be a continuum with no sharp boundaries between categories of expression types (i.e. literal or figurative), as explained in section 2.1. However, the approaches disagree on the question whether metaphor originates in thought, as in Conceptual Metaphor Theory, or whether metaphors originate in utterances designed to communicate the speaker’s thoughts to others, as in Relevance Theory. In general, Lakoff & Johnson focus on conventional language rather than poetic use of expressions, and it does not seem to be of interest to them that the latter more obviously displays metaphoricality. They define many expressions as metaphorical that would traditionally be seen as literal.

Lakoff & Johnson claim that our conceptual system is ‘fundamentally metaphorical’ (1980: 3). This claim is supported entirely by the observation of semantic regularities in the linguistic expression of concepts. This goes back to the conviction that the existence of conceptual metaphors as cognitive entities can be evidenced by grouping linguistic expressions according to their (apparent) source domains. For example, the conceptual metaphor ARGUMENT IS WAR is thought to be reflected in expressions like ‘Your claims are indefensible’, ‘He attacked every weak point in my argument’, ‘His criticisms were right on target’, etc. (Lakoff and Johnson 1980: 4). The idea that this kind of lexical structure counts as evidence for a cognitive structure has been criticised for its circularity, as it lacks evidence on the cognitive side. The claim that conceptual metaphors govern our thinking cannot be founded on linguistic evidence alone (see e.g.
This problem of circularity is also well known in the debate about linguistic relativity, the idea that the language we speak shapes the way we think. It is discussed in section 4.1. There is an inherent difficulty when it comes to providing evidence of this kind: in order to prove that linguistic behaviour reflects cognitive behaviour, one needs non-verbal evidence of the thought patterns claimed, as otherwise there is no way of telling whether the language specifics are due just to linguistic form, or whether they tell us something about thought (for a critical discussion of the role of intuition in Cognitive Linguistics, see Gibbs 2006a).

An important notion in Conceptual Metaphor Theory (and in Cognitive Linguistics in general) is that of ‘embodiment’, or ‘experiential grounding’. Lakoff & Johnson (1980a) state that the only nonmetaphorical concepts we have are very basic and arise from experience early in life. They divide these types of early experiences into three categories: spatial orientation, ontological concepts (such as ENTITY, CONTAINER, etc.), and structured experiences and activities (such as EATING, MOVING, etc.). Ontological metaphors are those which involve a projection of substance onto an abstract concept which does not have substance inherently, for example an expression like ‘I have no time’ would be understood via the conceptual metaphor TIME IS A PHYSICAL ENTITY. A structural metaphor, concerns the structuring of one kind of experience or activity in terms of another kind of experience or activity. A conceptual metaphor such as THINKING IS EATING would accordingly be the basis for an expression such as ‘She digested all the ideas’ (1980a: 195-200). Basic metaphors like these are thought to be bases of a vast array of linguistic expressions, and a large part of Lakoff & Johnson’s canonical book ‘Metaphors we live by’ (1980) consists of long lists of examples that show how these single instantiations of underlying theses are thought to be interconnected.

Vervaeke and Kennedy (1996) point out that, according to Lakoff & Johnson’s (1980) proposal, there is an inherent flaw in the theory in that it seems impossible to determine which particular conceptual metaphor could be thought to have evoked which particular linguistic expression, and in particular, at which
level a conceptual metaphor can be thought to be the most ‘basic’, i.e. the actual metaphor that does the job of the cognitive structuring. For example, the conceptual metaphor often cited as a typical case is\textit{ARGUMENT IS WAR}. Couldn’t it be that the underlying metaphor is actually\textit{ARGUMENT IS COMPETITION}, and if not, what are the reasons to choose the former over the latter as the better candidate? They go on in suggesting that perhaps the correct level of generality should be \textit{ARGUMENT HAS A GOAL}, which should be understood in connection with\textit{ARGUMENT IS A BODY}, since the goal of a body is \textit{HEALTH} (1996: 276). This little thought experiment demonstrates how Lakoff & Johnson’s theory seems often a bit too intuitive and rather arbitrary, suggesting that it suffers from a lack of rules to define what conceptual metaphors actually are. Vervaeke and Kennedy could not be any clearer in making their point as follows:

As a result, the very flexibility of the implicit-metaphor theory, we suggest, may be its undoing in the long run when it aims to be an empirical, testable theory rather than a brilliant feat of hunting, gathering, and interpreting.

Vervaeke and Kennedy 1996: 287

3.3.2 Grady: Primary conceptual metaphors and metaphorical compositionality

Joseph Grady (1997) did some important work to give the notion of conceptual metaphor more of a structure. He questioned Lakoff & Johnson’s practice of postulating conceptual metaphors at the most specific, rather than most general level (e.g. \textit{TIME IS MONEY} instead of e.g. the more general \textit{TIME IS A VALUABLE COMMODITY}), a practice that Lakoff & Johnson give no particular reasons of adopting (1980: 9). In order to show how conceptual metaphors seem to make much more sense at a more general and experiential level, on the one hand, and that more specific metaphors arise from combinations of the more general ones, on the other hand, Grady analyses in detail the conceptual metaphor \textit{THEORIES ARE BUILDINGS}, proposed as a typical example in ‘Metaphors we live by’. Grady
draws attention to the fact that Lakoff & Johnson (1980) emphasise how our conceptual system is shaped on an experiential basis, that is, our early bodily experiences in the world. This may work for conceptual metaphors of the type MORE IS UP, but, by contrast, there is no experiential basis that we can draw on to explain the association between theories and buildings. Thus, there is reason to doubt the idea that the latter could be an instance of a conceptual metaphor, since it does not seem to fit the definition. So Grady suggests that restricting conceptual metaphors to experiential concepts might be more adequate. These may then be more general, and in order to make them more specific, they could be combined with others in a compositional fashion. He looks at other ways in which theories are frequently conceptualised, i.e. by employing other physical structures as source domain, such as fabric. It can be found in examples like ‘He tore the theory to shreds’. Accordingly, based on a few more examples of this kind, Grady proposes the more general metaphor THEORIES ARE PHYSICAL STRUCTURES (1997: 270-272). However, Grady reasons that this is a metaphor that is too broad to be considered as invoking specific experiential domains, and he explains that he comes to this conclusion on the basis of comparison with the conceptual domains discussed in the literature so far (1997: 273). I am not convinced by the suggestion that this conceptual metaphor is too general. Rather, I would like to suggest that conceptual metaphors are cognitive entities which are at a very schematic, pre-propositional level and function as bases for more specific ones which take shape in a situation-specific way. It is possible that such processes are employed in a fashion as proposed by Barsalou (e.g. 2005, cf. section 3.3.3). Thus, on the contrary, I would rather assume that THEORIES ARE PHYSICAL STRUCTURES may still be too specific as a source of the many meanings observed in the linguistic examples, e.g. ‘The theory shook in its foundations’, ‘These are solid facts’, etc. These could all be built out of a much less concrete schematic pattern.

Be that as it may, Grady then goes on to propose to add an additional conceptual metaphor, and to demonstrate that these two metaphors together form a more specific compound metaphor. This is then thought to be much more adequate as a source for these meanings than the original THEORIES ARE
BUILDINGS metaphor. He adds PERSISTENCE IS ERECTNESS/VERTICALITY, where ‘persistence’ is to be understood as shorthand for continued function. This conceptual metaphor, in combination with THEORIES ARE PHYSICAL STRUCTURES, accounts nicely for the many metaphorical meanings to do with stability, coherence, etc., evoked in so many expressions to do with theories. Note that the source domains of both metaphorical constituents are identifiable as experiential, which suggests that they are authentic candidates for conceptual metaphors. Grady points out that the constituents in such a compositional metaphor should be understood as still independent of each other. For a metaphorical expression involving, for example, fabric, as exemplified above, the second constituent would not be appropriate, since ERECTNESS is not something concerning fabric. Grady calls this practice of combining basic constituents into more complex larger units ‘unification’, a term which he explains is borrowed from grammar. Here, syntactic constituents, such as a determiner and an unspecified nominal structure, combine to yield a phrase that carries more specific information than the constituents on their own (1997: 275). Grady’s proposal is an achievement that develops Conceptual Metaphor Theory into a more sophisticated theory, or, to say it in the spirit of Vervaeke and Kennedy, it lifts something that would more adequately be described as a collection of linguistic examples sorted into categories enriched with some wise intuitions, into the realm of actual theories. The move of explaining more complex metaphors by way of decomposing them into more basic components has brought us a step closer towards delineating the nature of conceptual metaphors by reducing their characteristics to those that are experiential.

Lakoff & Johnson (1980) do not seem at all sure as to what it is that their (alleged) conceptual metaphors are based on, and it is unclear whether they see any difference between physical (experiential) and cultural motivations for conceptual metaphors to arise in the first place. For example, they make the following two claims: first, that the vastness of physical and cultural experience provides the grounds for many different possibilities of basic conceptual metaphors to arise, and that this may vary between different cultures, and second that ‘it is hard to distinguish the physical from the cultural basis of a metaphor,
since the choice of one physical basis from among many possible ones has to do with cultural coherence’ (1980: 19). Grady’s treatment of the issue seems to have straightened out this kind of vagueness, as in his proposal it is clear that the basic metaphors are the experiential ones. Further, he explains that bodily experience is shared by all humans, and so the basic metaphors can be assumed to be the universal ones, while combinations into more complex metaphors may result in the more culturally specific metaphors (1997: 288). However, the question remains as to how these combinatorial mechanisms, if they exist, would proceed. The issue is not clear, and, on the face of it, an account that limits itself just to the basic conceptual metaphors seems psychologically more plausible. Such an account can explain metaphors that arise in thought, based on embodiment, while it seems likely that the more complex metaphors have their origin in language use. This issue will be elaborated on at the end of chapter 6.

Now that I have described the idea of conceptual metaphors in some detail, I would like to emphasise that I am in principle in favour of maintaining the notion, but only under the condition that we define conceptual metaphors as a very basic and schematic pattern which has a low-level function of guiding our thoughts, based on our own experiences in the world. It seems that conceptual metaphor theorists, in some cases, have gone too far regarding the power that conceptual metaphors apparently have over our thinking. Therefore I wish to distance myself from these views and, for the model developed here, use a more moderate definition of conceptual metaphors, based on Lawrence Barsalou’s work, described in the next section.

### 3.3.3 Barsalou’s mental simulations

The hybrid model that will be developed in chapter 6 includes a notion of conceptual metaphors in the more basic fashion alluded to above. It is inspired by Barsalou’s work (1999; 2005; 2009) on situated conceptualizations, outlined in the following.

Barsalou (e.g. 2005) develops a system of conceptual representation that contrasts with the relevance-theoretic view that conceptual processing takes place
within the medium of the Language of Thought, as outlined in section 2.3. Representations in Barsalou’s model are modality-specific (i.e. based on sensory perceptions), as opposed to the modality-free abstract symbols in the Language of Thought (LoT). Further, his model is nonmodular, as opposed to the modular view in Relevance Theory (outlined in section 2.4). Barsalou writes that ‘modular amodal and nonmodular modal theories make different predictions about the roles of modality-specific systems in conceptual processing’ (2005: 630). That is, we are dealing with theories based on different principles, so that it is perhaps remarkable that one of them can draw on work from the other. It demonstrates how productive it can be to combine elements of different schools of thought, and it is possible that these different theories may eventually be seen to show different angles of the same phenomena. In chapter 6, I offer a version of such a view. On Barsalou’s view, a concept is not a representation of a whole entity, but rather an incomplete pattern to be filled in on occasion, each part of the pattern corresponding to a component of an experience. This process of ‘filling in’ is nonmodular and modality-specific. This is in contrast to what Barsalou calls ‘the semantic memory view’, the dominant theory since ‘the cognitive revolution’ (2005: 621). Here, semantic memory is thought to be modular, with the representations contained in it being amodal, i.e. abstract symbols that are fundamentally different from Barsalou’s modality-specific representations (described below), and which are to be understood as ‘redescriptions’ or ‘transductions’ of the latter. A version of this is Fodor’s Language of Thought (see above), a system of amodal symbols, which are representations that fulfil the criteria for inferential fitness, i.e. that have the characteristics necessary for computational processing. For Fodor, these representations are computed in the central systems and are to be distinguished from his modular input systems and transducers. In Relevance Theory, LoT is modular, too, as Fodor’s notion of central systems has been further developed into an all-modular model in Relevance Theory (see chapter 4). Barsalou describes these symbolic representations as decontextualised prototypes, functioning as an encyclopaedic database. By contrast, Barsalou conceives of conceptual representations as

\[\text{Note that the different schools of thought operate with different terminologies, depending on}\]
analogous to modality-specific representations: the conceptual system re-enacts, or simulates, modality-specific representations, rather than re-describing them in the form of amodal symbols. Furthermore, these simulations are understood as situated, i.e. highly context-sensitive and non-modular.

So what we have here is quite a contrast, and Barsalou himself mentions that there has been done very little in terms of a comparison. He argues that this is not surprising, since the modular amodal view now represents the mainstream of cognitive science. He explains that this strong tendency reflects the ascendance of logic, statistics and computer science in the twentieth century, and it is so popular not least due to its symbolic representation languages which can be formalised and easily be implemented in computer hardware (2005: 630). It should be clear though that these facts alone are not sufficient as arguments in favour of the computational view. More compelling seem the above mentioned arguments of memory storage economy due to LoT’s recursive power and a guaranteed accuracy gained through the truth-preserving properties of the inference rules in LoT. However, one may wonder whether the experiential character of Barsalou’s model does not intuitively seem more natural and hence more likely for human beings, as it seems much more inclusive not least with a view to nonpropositional, such as emotional or sensual, aspects of experience.

So how does this conceptual re-enactment work? Barsalou explains that ‘when a physical entity is experienced, conjunctive neurons in association areas capture the pattern for later cognitive use’ (2005: 623). This is then reactivated in subsequent encounters, which is thought also to be possible in the absence of bottom-up stimulation. That is, only parts of the experience are then being simulated, as a mere guide to assist occasion-specific conceptualisations. If attention selects components of experience repeatedly, a corresponding simulator develops. An important point is that a representation never gets activated as a whole; the pattern is envisaged as very schematic, thus being responsive to the dynamics of context-sensitivity. However, when a component of the pattern is recognised as matching a situation, the rest of the pattern also becomes active in memory in the form of inferences that are justified by frequent occurrence. This
leads to entrenched situated conceptualisation (further described below). Simulators provide the agent with the ability to interpret entities it encounters as tokens of a type. Simulations include sensory and introspective (e.g. emotional, evaluative, etc.) representations. Thus, several simulators work together when reacting to a pattern that becomes active in the feature system. In terms of neural location, simulators develop near perceptual modalities, and then become integrated at a higher level in anterior regions. It is important to note that this is not the same as the idea of a prototype, because a prototype is a holistic image. Here, the type is just a very schematic pattern. On this view, a concept is seen as ‘an agent-dependent instruction manual that delivers specialized packages of inferences to guide an agent’s interactions with particular category members in specific situations’ (2005: 626). This process is often unconscious. It may be conscious, though, to a certain extent, when we are making a deliberate attempt at constructing mental imagery, for example. Barsalou states that situated conceptualisations support four types of situated inference:

1. Inferences about goal-relevant properties of the focal category;
2. Inferences about the background setting;
3. Inferences about likely actions that the agent could take in order to achieve an associated goal
4. Inferences about likely introspective states that the agent might have while interacting with the category, such as evaluations, emotions, goals, and cognitive operations

Barsalou 2005: 639

When people find themselves frequently in similar situations (such as social interactions, institutions, etc.) they develop entrenched multimodal knowledge patterns which are activated as a whole automatically on minimal cuing, enabling the agent to act skilfully. The accumulation of such processes amounts to what is commonly understood by ‘life experience’. Even if these entrenched knowledge patterns don’t always fit the situation perfectly, they still provide useful inferences which enable the agent to anticipate what is going to happen and to make
predictions. This is particularly important in communication, and thought of as an automatic process, to the extent that it is entrenched in memory.\(^8\)

Barsalou provides a wealth of empirical evidence which support the hypothesis that conceptual processing is based on modal systems, as opposed to a system of amodal abstracts symbols (as in LoT). He argues that a great extent of these processes occurs subconsciously, so it is useful to investigate neural representations of the processes, rather than introspection. When subjects produced or verified properties of categories, they did so by mentally simulating the categories and checking the resulting simulations for the relevant properties. This could be shown, for example, by testing the subjects’ responses in cases of visual occlusion: [LAWN] generated different properties than did [ROLLED-UP LAWN], where the accessibility of properties like ROOTS and DIRT varied (Wu & Barsalou 2009). Similar examples of experiments are given involving e.g. the size of properties: the greater the size, the longer processing took, as apparently the region that is simulated becomes larger corresponding to perceived size (Solomon & Barsalou 2004). Barsalou also reports experiments involving brain scanning (fMRI), with supporting evidence suggesting that brain regions used in mental imagery and object perception were active during these tasks (Kan et al. 2003).

Other experiments show that simulating visual shape also plays a crucial role in property verification. For example, MANE for [LION] did not facilitate MANE for [PONY], because of its different shape, whereas MANE for [HORSE] did. On the other hand, BELLY for [LION] facilitated BELLY for [PONY] as much as did BELLY for [HORSE] (Solomon & Barsalou 2001). Further evidence for the simulation view is provided by experiments involving modality-switching during property verification. Participants were faster at verifying a conceptual property if the words labelling it were presented relating to the same modality as the one before, e.g. LOUD for [BLENDER] was verified faster when following RUSTLING for [LEAVES] than when following TART for [CRANBERRIES] (Pecher et al. 2003).

\(^8\) This idea seems well worth a comparison with the relevance-theoretic comprehension heuristic, which is based on a similar principle of anticipation and prediction, such as the presumption of relevance, and the notion of mind reading (or ToM), i.e. the ability to see a situation from another person’s point of view.
Furthermore, it has been shown that in cases of modal deficits, i.e. lesions in a modality-specific system, there is an increased risk of losing category knowledge tied to modality-specific representations, e.g. damage to the motor areas might entail losing knowledge about categories such as [TOOLS] (e.g. Damasio & Damasio 1994; Gainotti et al. 1995; Humphreys & Forde 2001). This has also been supported by neuroimaging studies showing that the processing of different categories activates the corresponding modality-specific systems (e.g. Chao & Martin 2000). Other studies have shown that just seeing a picture of a manipulable object in isolation activates the motor system, which was also the case upon hearing specific words (e.g. Hauk, Johnsrude and Pulvermüller 2004) (Barsalou 2005: 631-638).

If conceptual representations are conceived of as having the form of simulations rather than amodal symbols, the question to ask is whether such simulations can have the classic symbolic functions, such as being computationally manipulable and thus providing premises for inference. The latter is the case in the computational processes as assumed in Fodor’s Language of Thought and in Relevance Theory. According to Barsalou, a simulation-based system can implement these functions, but he concedes that the issue requires further research. He speculates that the sets of conjunctive neurons in association areas might have some amodal triggering functions, but that, however, they cannot count as actual representations (2005: 645). I doubt that this comes anywhere near to the inferential fitness provided by LoT. More compelling is the question of how abstract concepts, i.e. abstract thought, could be realised with situated simulations. Here, Barsalou makes clear that he sees abstract thought as central to human cognition. A possibility he offers is that abstractions could be realised through constructed configurations of simulations, which raises the question of how different such configurations may be from Fodorian transductions. However, Barsalou insists that the simulations of the content of abstract concepts are very similar to those of the content of concrete concepts, they only differ in terms of their content. He claims that the content of abstract concepts is drawn from events and introspections, and explains that this could be simulated in the relevant modality-specific systems just as well. Barsalou &
Wiener-Hastings (2005) tested this hypothesis and report that in experiments involving property generation, subjects generated situated properties for both concrete and abstract concepts. Although the content differed in terms of concreteness, the properties generated for both types of concept seemed to have been guided by the four types of inference involved in simulation (see above). However, this seems to be a tentative result requiring further research in the future. Note that there is an important parallel here to the cognitive-linguistic view that our ability to cognize the abstract in terms of the concrete, i.e. metaphorical thinking, has its basis in concrete bodily experience. For Barsalou, this ability would be implemented in re-enactments and simulations. Nevertheless, in terms of parallels to ideas of symbolic representation, such as LoT, he considers the activity of verifying a property as belonging to a category to be based on simulation, according to the experimental evidence he provides. Hence, according to Barsalou, simulations must play a role of some sort in cognition (2005: 646).

Note also that Barsalou’s cognitive system is based on a nonmodular view of the mind. Thus, it may not seem possible to implement it in a modular model, which is ultimately what the account proposed here is going to amount to. However, Barsalou’s simulations can be understood to be carried out by a dedicated module which specialises in processing experiential representations. In chapter 6, I propose a way in which simulation processes can be integrated in an otherwise proposition-based model, by developing a new modular hybrid model of communication. The way in which conceptual metaphors are conceived in the model, based on Barsalou’s mental simulation just described, is the focus of the next section.

3.3.4 Refining the definition: Conceptual metaphors as schematic patterns
Conceptual metaphors, as originally seen in some (or most) of the Cognitive Linguistics literature (e.g. Lakoff & Johnson 1980; Grady 1997) seem too specific and concrete concerning the representations involved, if they are intended to function within cognitive processing mechanisms, such as those proposed in Relevance Theory. Since the hybrid model outlined in chapter 6 includes
conceptual metaphors as well as relevance-theoretic processing mechanisms, the former need to be modelled in a psychologically plausible way. I propose conceptual metaphors to be schematic patterns in memory that become more concrete in situation-specific instantiations. This idea is based on the processing model of situated conceptualizations, or mental re-enactments, as proposed by Barsalou (1999; 2005, above), which strikes me as the psychologically most plausible. He describes inconcrete patterns in memory to have developed from our experiences in the world. These are thought to be partially reactivated when the agent experiences similar events or objects repeatedly. Thus, a conceptual metaphor can be understood as a low-level schematic pattern that has developed as an entrenchment of repeated bodily experiences in the physical world. It results in representations which, although very basic and unspecific, are thought to motivate analogical thinking, which may in turn help cognitive orientation. This idea has much in common with a notion proposed by Vega Moreno (2005), suggesting the existence of memorised representations that have a more moderate role than Lakoff & Johnson’s conceptual metaphors in terms of their assumed salience in cognition, and which could indeed be involved in utterance understanding. Vega Moreno emphasises that entrenched analogies may well direct us towards specific inferential routes in comprehension, while it seems unlikely that they actually structure and dominate our thinking in the way that some cognitive linguists suggest. She further suggests that representations which cognitive linguists assume to be conceptual metaphors, such as ANGER IS HEATED FLUID IN A CONTAINER, do not necessarily need to be instantiations of conceptual domain mappings. Rather, they may just as well be understood in terms of a high accessibility of information pertaining to human nature, since, in this example, anger leads to changes in the human body, such as rising heat and temporary higher blood pressure (2005: 263). However, this can equally be seen as a description of the experiential basis of conceptual metaphors (i.e. the body itself), and it could be argued, instead, that the labelling of conceptual metaphors, in this case as ANGER IS HEATED FLUID IN A CONTAINER, is just not sufficiently accurate. It is far too specific and concrete, while the experiential bases of conceptual metaphors could be more realistically assumed to be rather schematic and
inconcrete. Thus, while Vega Moreno intends this as an argument against the notion of conceptual metaphors, I would understand it as an argument in favour of a more refined definition of them. The conceptual schematic pattern I have in mind cannot be labelled in a way comparable to the one above, as this, again, would presuppose an initial meaning that is too concrete. Rather, I would suggest that such underdetermined patterns that get fleshed out by re-enactment in situated use (cf. Barsalou 2005), leads to creating analogies between current and former experiences. In this sense, we can say that a mapping takes place between experiences that are similar but not identical. They share a common structural pattern.

The kinds of conceptual metaphor proposed in Cognitive Linguistics that are properly rooted in bodily experiences with the world, such HAPPY IS UP or MORE IS UP, are likely candidates that fall under the definition of conceptual metaphors proposed here. By contrast, an unlikely candidate would be a conceptual metaphor like IDEAS ARE CUTTING INSTRUMENTS, as proposed by Lakoff & Johnson based on linguistic expressions like ‘That was a cutting remark’ and ‘He’s sharp’ (1980: 48). Of course, the claim that this apparent conceptual metaphor seems unrealistic is based on intuition and would require empirical evidence of some kind to support it. However, trying to explain how IDEAS ARE CUTTING INSTRUMENTS may have arisen from a basic experiential pattern just doesn’t appear to lead anywhere near a plausible cognitive picture.

Alternatively, in situated simulations (or re-enactments), not the whole memory of an experience gets reactivated, only the parts that fit the current situation. They are very basic structures which are thought to generally support cognitive orientation in the world. While this corresponds to the claim made by conceptual metaphor theorists, that conceptual metaphors govern our thinking, the notion of conceptual metaphors proposed here is much more moderate, as the processes involved are thought to proceed at a low and inconcrete level of cognition. So, rather than combining several general conceptual metaphors to get a compound that is more specific, as proposed by Grady (1997, above), I go in the opposite direction in favour of less specificity and assume that repeated situation-specific simulations of the same schematic patterns take place. This is supported
by the cognitive ability to interpret entities in the world as tokens of a type (Barsalou 2005: 624ff.). Further, I understand such a ‘type’ to be the schematic pattern just described. And this is what I call ‘conceptual metaphor’ in the proposal outlined here. The tokens, i.e. the realisations of that pattern, linguistic or otherwise, can take various forms and shapes (metaphorically speaking), since the type seems to be merely a general orientational guide in cognition, i.e. it is vastly underdetermined. Thus, the notion proposed here is a much more moderate version of conceptual metaphors compared to those presented in the prevailing cognitive-linguistic literature.

At the end of section 2.2 I described how routinised metaphorical usages may arise in language use through lexicalisation. In a complementary fashion, I propose here that other metaphors arise in thought. These may be the ones that have their origin in the basic cognitive patterns described above. For example, an expression like ‘She was so angry that she nearly exploded’, may have its origin in entrenched patterns of bodily experience which are repeatedly perceived also in analogy with entities outside the body. This means that, for example, we may experience biological reactions to anger in an analogical fashion to observing how things explode. Due to some analogical relations between them, such experiences can be assumed to be processed each time by way of simulation, or re-enactment, of the corresponding schematic pattern. Such a pattern, in this example, may be activated upon perception of one’s rising body temperature caused by anger, as well as when seeing a container filled with petrol about to explode. Thus, while metaphors that arise in language use come about by repeated communicative means, metaphors that arise in thought may come about by repeated instantiations of the same underlying cognitive patterns, which we may call ‘conceptual metaphors’ here.

Evidence supporting the idea that an underlying cognitive pattern may find expression in different concrete forms is provided e.g. in Cienki & Müller (2008), with an example of data showing how varying realisations of a (presumably) underlying conceptual metaphor materialise. In a study on meaningful gestures in conjunction with speech, they observed a multimodal verbo-gestural metaphorical expression of the same general idea, which, roughly, was about the difference
between right and wrong. Two different metaphors were used in parallel in the two modalities: the verbal metaphor employed black vs. white as a source, and the gestural one employed clearly separated spaces. A conceptual metaphor underlying this multi-modal expression could only have been a very unspecific one, which nevertheless helped to structure the thought that led to the utterance. Cienki & Müller’s research will be described in more detail in section 4.1.4. It can be seen to support the notion of conceptual metaphors as maintained in the present proposal.

However, the criticism expressed by many still stands, namely that the notion of conceptual metaphor lacks empirical support. In particular, the idea that conceptual metaphors are evoked in online reasoning whenever linguistic metaphors of any type are used, remains open to charge. Especially when it comes to those expressions that to Lakoff & Johnson are metaphors, yet elsewhere would be described as polysemous items with metaphorical origin, in etymological terms. McGlone (2007) puts it as follows:

> It is presumptuous to infer that a spouse who confesses that she has ‘‘fallen out of love’’ with her partner has mentally invoked (let alone embraced) the schema RELATIONSHIPS ARE CONTAINERS. Evidence independent from the mere occurrence of idioms in conversation is necessary to demonstrate the conscious or unconscious deployment of a conceptual metaphor.

McGlone 2007: 123

Furthermore, there is a terminological issue with the way conceptual metaphor theorists use the term ‘metaphor’. They use it synonymously for two different things: the conceptual entity they assume underlies and structures thinking, on the one hand, and metaphorical expressions as they appear in language use, on the other. This often leads to a conflation of the two, and ultimately to the assumption that linguistic metaphorical expressions inevitably are instantiations of very specific conceptual metaphors, which often seem to be understood as one-to-one congruent with the expression. Accordingly, there seems to be the implicit
assumption that even lexicalised metaphorical expressions are cognitively processed as if they were meant to denote their original literal, but now distinct, meaning, as in McGlone’s example of the spouse quoted above. It seems that this practice of equating (presumably) metaphorically motivated polysemy with expressions that are actually used metaphorically leads to a picture of language use that does not reflect psychological reality in any way. It seems much more adequate to define metaphoricity, i.e. metaphorical expressions which are experienced as metaphors, in terms of an awareness of it in communication, which will be discussed in section 4.2 (e.g. Carston 2009; Camp 2008), and picked up again in section 5.1.1 (Steen 2008).

So far, this chapter has described the notion of conceptual metaphor, criticisms and a redefinition of it, which has prepared its employment, in a psychologically plausible way, in the account developed in chapter 6. It has thus explained how some metaphors may arise in thought, as compared to the relevance-theoretic account which explains how some metaphors arise in language use (section 2.1. and 2.2). The next section is devoted to describing how conceptual metaphors have been fruitfully used in a practical way in second language learning and second language acquisition research. The purpose of this is to show that, even though to date we have very little empirical evidence of their existence in cognition (but see section 4.1.4), conceptual metaphors are a useful tool, seen from a multilingual perspective. To a certain degree, this justifies postulating them in the new hybrid model, bringing the so far relatively abstract notion of conceptual metaphors ‘down to earth’.

3.3.5 The multilingual perspective: Applications in second language acquisition

It may seem that cognition involves conceptual metaphors in one way or another, even though it does not seem entirely clear yet how this involvement is implemented. There is much research to be done in the future. For a start, the notion seems to play a useful role in explaining conceptualisation, at least in second language acquisition (deKnop 2008, below). While evidence of the
apparent workings of conceptual metaphors in ordinary communication seems hard to come by, it has been shown that employing the notion of conceptual metaphor as a practical means to an end in the didactics of second language acquisition can improve performance. Frank Boers (2000) carried out a study which investigated conceptual metaphor as a facilitator of vocabulary acquisition. Students were asked to memorise lists of expressions describing emotions, which were either arranged according to various metaphoric themes (or conceptual metaphors), or according to some other groupings, such as type of function. (23) is an example of a metaphorically organised list, and (24) of another type of list:

(23)  *anger as fire*

  an inflammatory remark  
she was breathing fire  
she was all steamed up  
she erupted  

  *angry people as dangerous animals*

  he has a ferocious temper  
don’t snap at me!  
he unleashed his anger  
don’t bite my head off  

(24)  *to describe acute and sudden anger*

she exploded  
he unleashed his anger  
she erupted  
she flipped her lid  

  *to describe the way angry people speak*

she blew at me  
don’t snap at me  
don’t bite my head off  
she was breathing fire  

Boers 2000: 555-556

Students that had memorised the metaphorically organised lists performed much better at recalling the vocabulary than those who had memorised the other types of list. This facilitation through metaphor may suggest a cognitive predisposition towards metaphorical structuring. The study cannot demonstrate how such a predisposition might have arisen, and which cognitive processes may have been at
work that took advantage of the metaphorical relationships between the expressions. However, generally speaking, it could be understood to support the assumption that conceptual metaphors might, in one way or another, play some role in cognition.

Another piece of research that is concerned with the employment of conceptual metaphors in second language acquisition is a theoretical investigation carried out by Sabine deKnop (2008). While it does not involve performance tests, the study makes sophisticated suggestions for language teaching practices which seem to be promising in terms of improving French students’ performance in learning German as regards speed and depth of acquisition. It is based on explaining the meaning of German prepositions and their cases (in German, prepositions are case-marked) using the notion of conceptual metaphor. The idea is that ‘experiential learning’, i.e. learning based on bodily experiences in the world, is promising as the proposed rationalisations of lexical and syntactic meaning in the second language seem likely to facilitate acquisition. In order to exemplify this, I first need to explain what deKnop’s proposal looks like regarding the didactics of prepositional case marking of the non-abstract prepositions, i.e. those that are not metaphorically motivated, as this is the basis for explaining her heuristic of acquiring the abstract, i.e. metaphorically motivated prepositions. To give an example, ‘in einen Apfel beißen’ (to bite into an apple) requires accusative case. In order to make the reason for this case graspable to the language learner in some way, it may be helpful to explain that it is accusative because it is a dynamic location, it involves a ‘dynamic movement from the teeth to the apple’ (2008: 56). DeKnop, using the terminology of Cognitive Linguistics, explains that at the concept level, the apple is a visualised schema, experienced as a CONTAINER (trajectory), and the teeth are the trajector moving into it. The same goes for ‘an die Tür klopfen’ (to knock on the door), which also requires the accusative. In conceptual terms, the door is experienced as a SURFACE (trajectory) and the hand dynamically moving towards it is experienced as the trajector. Teeth and hand are to be understood as implicit objects which are in a metonymic relationship with the action designated by the verb. Making these conceptual connections which underlie the grammatical designation explicit helps the learner
to understand the use of the accusative; it does not feel like a random feature anymore, which helps them understand its function and facilitates a reliable use of it in communication (2008: 57). Next, deKnop gives examples of abstract expressions in German which seem metaphorically motivated, such as ‘auf ein Problem eingehen’ (to go into a problem) and ‘ich bin an den Vertrag gebunden’ (I am bound to the contract). In order to give learners of German a tool to understand the connection between the use of the accusative in these examples and that of the concrete examples above, deKnop invokes the conceptual metaphor ABSTRACT ACTIVITY IS MOTION. Carrying out the relevant domain mapping in their mind, the learners can now follow the same procedure as they did with the concrete expressions in order to apply the correct case, here again the accusative. Thus, the implicit trajector in ‘auf ein Problem eingehen’ (to go into a problem) is the brain, the problem is the trajectory, and the dynamic movement, in an abstract sense, goes from the brain to the problem. In ‘ich bin an den Vertrag gebunden’ (I am bound to the contract), the implicit trajector are, again, the hands, and their boundedness to the contract can be understood as an abstract movement from the trajector to the trajectory (2008: 59). Thus, according to this approach, conceptual metaphors have the potential to facilitate language acquisition, due to their experiential basis, which seems to ease cognitive processing as it is based on former experience. Although deKnop has not devised any tests to prove the success of her teaching practice, I am convinced that it is efficient. This conviction is based on my own experience of second language learning, which confirms that it helps to consciously construct an explanation of an otherwise opaque meaning of an expression in the new language. It aids memory and the process of building a lexical structure, and a system of conceptual metaphors is a useful tool for this as it can be easily grasped while having a holistic appeal. Therefore, this study supports the notion of (some version of) conceptual metaphor (albeit merely in terms of its practical applicability).

DeKnop refers to the often expressed intuition that one needs ‘to learn to think in the foreign language’ in order to master it (2008: 61), which is based on the idea that different languages categorise entities in the world differently and thus give rise to different conceptual structures. This position is rooted in
linguistic relativism, which will be discussed in more depth in section 4.1. In the same vein, Ijaz (1986) investigated the meaning boundaries of prepositions by analysing participants’ responses to several semantic tests such as sentence completion tests and cloze tests. She then compared the responses of native speakers of English to those of advanced second language speakers who lived in either an English or non-English speaking setting, and found that each group ascribed different semantic boundaries to the words. According to her analysis, this was due to different semantic dimensions in the respective native languages. She understands second language acquisition to involve the mapping of two conceptual systems, whereby some concepts do not match exactly but overlap only partially. She claims that in order to accommodate second language lexical meanings that are not congruent with the first language translations, second language learners need to either restructure their first language conceptual system or develop entirely new concepts to be able ‘to think in’ the second language.

This chapter has explained the notion of conceptual metaphor as it is conceived of in Cognitive Linguistics, followed by a description of how the notion should be defined in order to be usefully employed in the new hybrid model developed here. This is based on Barsalou’s view of cognition, i.e. situated conceptualisations, which has been outlined in section 3.3.3. Even though Barsalou works within a nonmodular framework, his simulation processes can be fruitfully adopted into a modular approach, since, instead of dominating cognition as a whole, they can instead be seen as implemented within a module dedicated to re-enactments, which in turn contributes, in parallel with other modules, to a holistic overall cognition, as will be outlined in chapter 6. Conceptual metaphor is a notion that seems popular within second language acquisition research, since it invites practical applications in the field. We have seen a few examples of this in this chapter. The question is whether these practical applications could be entertained as evidence of the claim that we actually do have conceptual metaphors in our minds. To begin answering this and related questions about language and thought, the next chapter is concerned with research that aims to show how it might be possible to gain non-linguistic evidence of the interplay between language and cognitive structure.
Chapter 4
Language and Thought

This chapter investigates, broadly speaking, the relationship between language and thought. It touches on important issues that are pertinent to the proposal developed in chapter 6, since the latter aims to explain how metaphor arises in thought, on the one hand, and in language use, on the other hand, and how the two are related. The representational formats and their specific processing mechanisms, and their mutual interaction as envisaged in this model are based on an understanding of the interplay between language and thought drawn from the insights outlined here. Thus, this chapter offers an outlook aimed at providing arguments in favour of the present theoretical stance, at the same time as giving an orientation to locate the present model in the wider research field of language and thought.

It begins with a description of the notion of linguistic relativity, and how it has been researched. It is an important issue with a view to Conceptual Metaphor Theory, which has been criticised for the claim that linguistic evidence alone can conclusively inform us about cognitive structures. The linguistic structures observed e.g. by Lakoff & Johnson (1980) could just as well be evidence for the linguistic structures themselves, since there is a lack of evidence on the cognitive side. However, later in this chapter, in section 4.1.4, we shall see that studies of spontaneous gestures accompanying speech might offer some insight in that respect, as gesturing is a non-verbal behaviour and may thus count as non-linguistic, and therefore cognitive, evidence of metaphorical thinking (Cienki & Müller 2008). But first, an overview is offered of the most salient findings within the field of linguistic relativity studies, in sections 4.1.1 to 4.1.3.

Section 4.2 then returns to the issue of metaphorical utterance interpretation, here in the context of the mental representations that we may entertain in comprehension. Thus the focus is on the role of the literal meaning of metaphorical expressions in comprehension, and how it contributes to an awareness of metaphoricity (Stöver 2011). As we have seen in chapter 2, in Relevance Theory the literal meaning is understood as a mere starting point for
conceptual adjustment processes in context, whereby the resulting ad hoc concepts and the weak implicatures derived both contribute to the communicated meaning. However, the literal meaning is perceived as well during the course of processing, and somehow seems to remain present in addition to the proposition expressed (Carston 2009, 2010; Camp 2008). This phenomenon is discussed in section 4.2.1 and supplemented with a report on a psycholinguistic study that provides evidence of the sustained role of the literal meaning in metaphorical interpretation (Rubio 2008), in section 4.2.2. To illustrate the discussion by offering an insightful multilingual perspective of the issue, section 4.2.3 provides examples of translation asymmetries, where an expression might be perceived as metaphorical in one language, but not with its equivalent, in another language.

4.1 Linguistic Relativity

The hypothesis of linguistic relativity, also called the ‘Sapir/Whorf Hypothesis’, states that the language one speaks influences one’s way of thinking, or, in extremer versions, how one perceives the world. This is a controversial idea, in cognitive science in particular, and it raises many questions. The discussion seems emotionally charged, and often, more than in other areas, driven by opinion at the cost of empirical facts. Levinson (2003) describes the situation as follows: ‘It is as if the topic of ‘Whorfianism’ is a domain where anybody can let off steam, go on mental holiday, or pounce upon an ideological enemy’ (2003: 25).

I first give a brief overview of the history of linguistic relativism in the next section, sketching its erratic changes in popularity over the course of the last century, and explain why nowadays it is still a problematic issue many scientists prefer to avoid. I then focus on experimental work on spatial reasoning, carried out by two different groups of researchers who argue the direct opposite of each other (Pederson et al. 1998; Levinson et al. 2002; Li & Gleitman 2002). This is followed by an introduction of a mild version of linguistic relativism: Slobin’s (2003) ‘Thinking for Speaking’. This approach has been drawn on in a study involving gesture, speech, and thought reported on in section 4.1.4 (Cienki 2000).

---

9 Part of this chapter is based on a research project carried out during my MA (Stöver 2006a).
This study highlights some important ideas concerning the interplay between different modalities in communication, which may have the potential to inform accounts of the interplay between different levels of cognitive utterance processing. Since the new hybrid model proposed in this thesis involves different levels of processing, the issue is of great interest here.

4.1.1 The history of linguistic relativism

In the first half of the 20th century the idea of linguistic relativity was nowhere near as controversial as it is today. The then predominant belief was that differences in the grammar of languages create cognitive differences. According to the contemporary empiricist epistemology, all learning takes place through experience: the child is seen as an ‘empty vessel’ to be filled with knowledge during development. It was also the time of structuralism, with its assumption that formal distinctions reflect and influence meaning distinctions. Structuralists assume that the structure of cultural practices, including language, produces and reproduces structures of meaning in the human mind. In addition, it was thought that there is an unconscious mental life, with linguistic effects below the level of awareness, so that the individual has no control over the ways in which particular languages affect ways of thinking (cf. Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996). These ideas were the perfect basis for linguistic relativism, and in this intellectual landscape Whorf's proposal seemed unquestionable. The following quote from his work summarises well what lies at the heart of the idea of linguistic relativity:

... what I have called the “linguistic relativity principle,” which means, in informal terms, that users of markedly different grammars are pointed by their grammars toward different types of observations and different evaluations of externally similar acts of observation, and hence are not equivalent as observers, but must arrive at somewhat different views of the world.

Whorf 1956: 221
In his own research, Whorf investigated grammatical differences between English and Hopi, a Native American language, and how these differences might influence the respective native speakers’ ways of thinking. In English, time units are encoded as count nouns (hours, years, etc.), whereas in Hopi they are encoded as recurrent events. From this, Whorf drew the conclusion that speakers of Hopi think about time differently than speakers of English do. But he did not provide any actual evidence for these differences in thinking; he mainly drew this conclusions on the basis of the linguistic evidence itself (Lucy 1996: 42-3). This is a circular argument. One cannot show that a particular language creates particular ways of thinking by merely pointing to the language itself, as this fails to provide evidence of a correlation with the apparently corresponding cognitive processes. Unfortunately, this problem of circularity, which has also been called ‘linguacentrism’ (Lucy 1996: 44), reappears in later studies in the Whorfian tradition, and, as mentioned in the last chapter, seems also to be a symptom that Conceptual Metaphor Theory still suffers today. However, from the 1960s onwards, when the intellectual climate changed, it became increasingly difficult to make claims based on linguistic relativism.

In the second half of the 20th century, Whorfian ideas became less popular in academia as cognitive science developed in the 1960s. Linguistic relativism became part of popular folk-wisdom, which in turn reinforced the negative attitudes of scientists, and scientific interest in it came to a halt. The predominant ideas in cognitive science are radically opposed to linguistic relativism: a salient one is universalism, according to which conceptual structure in its core features is constant across cultures. On this basis, it is impossible to claim that different languages encode radically different concepts. Accordingly, it is assumed that concepts come first, and words are merely labels for them. Probably the most influential idea in cognitive science is nativism, according to which semantics is given in an innate ‘Language of Thought’, sometimes also called ‘mentalese’, which is completely independent of language. This was described in chapter 2. It follows that a specific language cannot have any serious influence on its speakers’ cognition (cf. Levinson 2003).
In spite of these dominant ideas, linguistic relativism seems to have attracted new interest in contemporary cognitive science, presumably due to an increase in knowledge about cognition and new research techniques, which allow for more diversity of approaches. Nowadays there is a wide range of positions, ranging from the extremes to a variety of intermediate positions, some of which are outlined below. On the whole, there has not been very much empirical research so far (although theoretical ideas on universalism and variation have developed considerably). This is presumably due to the bad reputation linguistic relativism has acquired over the last decades (cf. Lucy 1996: 37). Another reason why people shy away from the topic might be that it is very difficult to devise tests which give us reliable conclusions, as will become apparent below.

Cognitive processes are hard to capture, and in many cases there is a problem with discriminating linguistic from non-linguistic processes. Accordingly, studies so far have had positive as well as negative outcomes with regard to linguistic influence on cognition, which seems partly to depend on the choice of the cognitive aspects focused on. Experiments seem feasible only if one picks out a small range of aspects pertaining to cognition, on the one hand, and to language, on the other.

A classic example of experimentation is colour perception, which is a strand of research that had its greatest popularity in the 1960s and 70s (e.g. Berlin & Kay 1969). This strand of research investigates whether different lexicalization of colours – for instance different numbers of words for colours and/or different colour boundaries – does lead to different conceptualisation of colour. For example, in English the colours purple, orange and red are each referred to by a different name, whereas Shona, an African language, has only one term for all three colours (Glucksberg 1988). But does that mean that the Shona do not conceptualise or even perceive these colour differences? Reviews of studies in the field give different answers to this question: Glucksberg reports that the majority of evidence suggests that there is no correlation between numbers of words and numbers of concepts and/or perceptual discriminations (1988: 232), whereas according to Levinson, the state of affairs is quite the opposite: on the whole it seems that there is a tendency for linguistic colour coding to influence colour
perception (2003: 37-38). By contrast, Papafragou et al. express an insightful objection to research on relativism in colour perception: ‘... the last place we might expect malleability of human thought is in domains involving sensory representations shared with many non-human species’ (2002: 7). I think that this is a valuable comment: language is specific to humans, so it seems sensible to investigate its relationship to human-specific cognition, as influence of a specific language on thought would, if it takes place at all, presumably take place in connection with more abstract or variable thinking. In particular, the question of the plausibility of a role for conceptual metaphor in cognition, and whether its (apparent) manifestations in the language in turn influence cognition, is more usefully discussed with referral to abstract contents, where abstract-to-concrete mappings seem to be more easily discernible. However, before I describe research that begins to approach issues related to abstract reasoning, I describe a series of studies that focused on spatial reasoning, with contradicting outcomes (Pederson et al. 1998; Li and Gleitman 2002; Levinson et al. 2002). They reveal important methodological issues, as well as demonstrating how testing non-linguistic behaviour may allows us, depending on experimental design, to draw conclusions that pertain to linguistic variation as having an influence on reasoning, at least in terms of physical orientation.

---

10 For example, in a study comparing English and Zuni speakers (Native American), the discrimination between ‘yellow’ vs. ‘orange’ helped English speakers to memorize colours, as compared to speakers of Zuni, which doesn’t have this discrimination (Lenneberg & Roberts, 1956). Similar findings were reported by Lucy (1981) for Yucatec Maya vs. Spanish vs. English speakers, and by Davidoff, Davies and Roberson (1999) for English vs. Bernimo (Papua New Guinea). Kay and Kempton (1984) found that English speakers exaggerate the perceptual differences between ‘blue’ and ‘green’, as compared to speakers of Tarahumara (Mexico) which doesn’t have this discrimination (Levinson 2003).
4.1.2 Spatial Reasoning

Spatial orientation seems to be a cognitive domain which lends itself well to research into linguistic relativity, since it allows for an experimental design that can demonstrate aspects of non-linguistic cognition in a relatively straightforward way, as described below. First, it is necessary to explain the three frames of reference that play a major role in spatial orientation, and that are used in different languages, respectively: relative, intrinsic and absolute. Levinson et al. (2002) explain that spatial orientation in the relative frame of reference, also called ‘egocentric’, is relative to the viewer’s body. Here, terms such as left, right, front and back are used. For example, if you stand opposite to me facing a model of, say, a car and a house, from your point of view the car may be on the left of the house, but from my point of view it is on the right. In the intrinsic frame of reference the coordinates are object-centered, i.e. no matter where I stand, the car is at the front of the house. In the absolute frame of reference orientation depends on overall geographic directions, hence it is also called ‘geocentric’, and one would say ‘the car is south of the house’ (2002: 158).

According to Levinson et al., the particular frame of reference one uses in spatial reasoning depends on how it is linguistically encoded in the particular language one speaks. For instance, in English the relative frame of reference is dominant, but English also makes available the intrinsic one. Accordingly, English speakers use the relative and intrinsic frames in their spatial cognition, and spatial reasoning in the absolute frame is hard to compute (2002: 157). For example, being told ‘Give me the spoon that’s northeast of your teacup’, would make it difficult for us to mentally picture the location of the requested spoon. In order to test this apparent correlation between the linguistic and the orientational frame of reference, Levinson et al. developed the experimental technique of rotation. Participants were turned to face a different direction while carrying out non-linguistic tasks, as a way of revealing which frame of reference the participants use in spatial cognition. There is, for example, the so-called ‘Animals-in-a-row’ test: three toy animals are placed in a row on the presentation table, all facing in the same direction. After the participant had a look at them, she is rotated by 180 degrees. Then she is given four toy animals and is asked to pick
the same three animals out of the four, and to rebuild the same array she had seen on the presentation table. If she places the animals facing the direction relative to her point of view (e.g. left), she is using the relative frame of reference. If she places them facing the same geographic direction as before (e.g. east), she is using the absolute frame. The procedure of choosing the three animals out of four is to distract the subjects from what is being tested, and to increase the memory load, which makes it more likely that they will use the form of spatial reasoning they are most used to (2002: 166-7).

Levinson et al. tested, among others, speakers of Japanese and Dutch, which are languages that use the relative frame of reference (54 participants), compared to speakers of Tzeltal, Arrernte and Longgu, languages that use the absolute frame of reference (51 participants). The results suggest a correlation between the participants’ linguistic frames of reference and their non-linguistic spatial orientation: the average number of absolute responses out of 5 trials was 0.50 for the linguistically relative participants and 4.02 for the linguistically absolute participants (Pederson et al. 1998: 579-80).

Li and Gleitman (2002) expressed doubts about this experimental outcome. Being opposed to the idea of linguistic relativity to start with, they tried to invalidate Levinson et al.’s results by claiming that the experimental design did not control for particular settings influencing the participants, as each group had apparently been tested in their natural environment, so that people who use languages with an allocentric (or absolute) frame of reference were tested outdoors, and people with egocentric (or relative) languages indoors (2002: 273). According to Levinson et al. this was simply not the case, as they carried out the experiments in settings that did not always correlate with participants’ natural environment (2002: 162). However, based on this criticism, Li and Gleitman argued as follows: in order to demonstrate that choice of a particular frame of reference was not due to a particular linguistic usage, but could be manipulated by the experimental setting, they set up an ‘Animals-in-a-row’ experiment with only speakers of English as participants, varying outdoor/indoor settings and orientational cues. In this way, they aimed to show that English speakers would change their frame of reference from relative to absolute, depending on the
environment (2002: 272-3). The results suggest that this was indeed the case. For example, in the indoors condition with the blinds down (excluding orientational cues from the outside) the majority of subjects responded in a relative frame of reference, whereas indoors, with the blinds up, and in the outdoors condition some subjects were using relative, some absolute frames of reference (2002: 277-8). Consequently, Li and Gleitman conclude that in general, differences in spatial cognition are independent of linguistic frames of reference, but depend solely on environmental cues.

Levinson et al. replied to Li and Gleitman’s attack as follows: First of all, they objected that Li and Gleitman had simplified the original experiment to an extent that influenced the results. When they asked the participants to recall the array of animals, they gave them the same three animals, so that the participants only needed to recall their order. But in the original experiment it was crucial that they had to pick the three animals out of four, as this way they had to concentrate not only on the order and direction of the array but also on the identity of the animals. Thus they were deflected from guessing that the experiment was all about direction. Furthermore, it added to the memory load involved in the task, which was of major importance, because it made it much more likely that participants would fall back on their habitual forms of spatial reasoning (2002: 163). Li and Gleitman’s version of the experiment was far too transparent, increasing the likelihood that participants would second-guess the experimenters’ intentions. Levinson et al. point out that this became apparent in the fact that the participants seemed to be aware of the ambiguity of the task: 70% asked for clarification as to what was meant by: ‘Make it as it was’, wanting to know whether they should rearrange according to the left/right axis or otherwise (2002:172). If participants focus on the main aspect to be tested in this way, while the experiment is designed to bring out unreflective behaviour, the results are unlikely to be genuine.

But most importantly, Levinson et al. point out the major flaw in the study: Li and Gleitman ignore the difference between absolute and intrinsic frames of reference. They operate with the term ‘allocentric’, which is a higher-order classification that includes both intrinsic and absolute frames of reference
In their experiments, the allocentric frames of reference were in fact really always intrinsic, i.e. array-internal relationships. A reason for this was that Li and Gleitman seem to think that in absolute frames landmarks are used as orientational cues. This is not correct, as ‘absolute’ is to be understood literally, i.e. people using this frame of reference maintain a ‘mental compass’, independent of landmarks (2002: 172-3). But here, whenever a participant used a frame of reference that was not relative, the assumption was that it was absolute. For example, a small toy resembling a duck pond was used in the experiment to function as a landmark to manipulate orientation. So whenever a participant rebuilt the array on the recall table in a relation to the duck pond that was the reverse of the left-right axis on the presentation table, it was assumed they were using an absolute frame of reference. But in fact it must have been an intrinsic one if the duck pond was perceived as a landmark. And it was even more likely that it was perceived by the participants as part of the array to be rebuilt after rotation. Levinson et al. conclude that Li and Gleitman’s participants were not being manipulated into behaving like users of an absolute frame of reference, which would have been alien to them (as explained above), but rather that they switched between the two frames available in their language: relative and intrinsic (2002: 173-4).

To demonstrate that this must have been the case, Levinson et al. carried out the ‘Animals-in-a-row’ test with a slight variation: they rotated subjects only by 90 degrees instead of 180, in order to show that the Dutch subjects (who use relative and intrinsic frames of reference, just like Li and Gleitman’s American subjects) only used the orientational frames made available by their language. They explained this as follows: the difference between absolute and intrinsic frames of reference does not show after a rotation below 180 degrees. An intrinsic frame of reference stays the same, independently of the geographic directions in which an array is laid out. So if a participant sets up the arrangement on the recall table after a 180 degree rotation, both the absolute and the intrinsic orientation can be inferred: she could have thought either of an object facing the end of the table (intrinsic) or facing north (absolute). But if the participant is rotated by 90 degrees instead, the difference becomes apparent: now the object faces the end of the table.
when placed on a different geographical axis than before. So if the participant uses the absolute frame of reference, she will put the object facing the side of the table rather than the end. The results suggest that the great majority of subjects used the intrinsic frame of reference, and not the absolute one. Furthermore, when the memory load was increased by adding more animals to pick from, the subjects tended to use their habitual frame of reference, which was the relative one (2002: 174). Hence Li and Gleitman’s claim that speakers of any language use all frames of reference available, depending only on environmental cues and not on language, cannot be right (2002: 177-9).

The debate between these two groups of researchers shows how small differences in experimental design can have major effects on the results. Most interestingly, it shows how it is possible to design experiments according to the desired outcome: it seems that, in order to support their ideological convictions that thought is independent of language, Li and Gleitman were able to provide counter-evidence against Levinson et al.’s results by overlooking an apparently insignificant detail: the difference between an intrinsic and an absolute frame of reference. But, apart from the controversies which reveal meaningful methodological issues, what is most interesting is that Levinson et al.’s experiments convincingly demonstrate that participants do indeed display non-linguistic behaviour that corresponds to the linguistic frame of reference in their language. Thus, linguistic relativity cannot that easily be denied, at least when it comes to spatial reasoning.

However, the studies on spatial reasoning can only give insight regarding a very small subdivision of language, and one could argue that the difference in cognition between a person who has an intrinsic frame of reference and a person who has an absolute one does not make any significant difference with a view to language, and further, it may not be a sufficiently informative finding if we consider the overall organisation of cognition. Yet if we consider how spatial reasoning might relate to other types of reasoning, in particular when it comes to metaphorical reasoning of the type that draws on an analogy between space and time, findings of this kind might be of significance. This will be elaborated on further in section 4.1.4., which also relates to research to do with the intuitive
impression many learners of foreign languages have, namely that speaking another language requires you to think in a different way. This is addressed by the research presented in the next section.

4.1.3 ‘Thinking for Speaking’
To date, most research into linguistic relativity has focused on the influence that specific languages might have on non-linguistic cognition. Slobin (1996; 2003) disagrees with the idea that investigation of cognitive behaviour that is disconnected from language is most likely to enlighten the issue. He considers it most important to look at cognition that is tied to language in the first place, as humans spend a large amount of their cognitive activity on linguistic behaviour. Thus, the starting point for his research is online cognition in the course of language use. He calls his approach ‘Thinking for Speaking’. The rationale behind it is that in order to talk about something, an event for example, one has to ‘package’ it in language, and hence think accordingly in order to prepare the appropriate utterances. This idea does not only pertain to actual events of communication. This type of cognitive behaviour is thought to also occur in situations when an individual is not actively communicating, since, even if one does not actually talk about particular experiences or events, there is always the potential that one might. Thus, cognition is geared towards communication in the sense that one is always ready to speak. This leads the speaker of a particular language to pay selective attention to aspects of cognition that are obligatorily marked in the grammar of that language (2003: 157-9). For example, some languages encode manner of motion but not path in their verbs, they are so-called ‘manner languages’. They may also be called ‘satellite-framed’ languages, as the path of motion is encoded in a ‘satellite’, usually in prepositions such as ‘in’ or ‘out’. To illustrate, manner verbs include examples such as: *walk*, *run*, *crawl* and *bounce*. Speakers of such languages are seen as habitually paying more attention to motion. Accordingly, Slobin hypothesizes that they have ‘a rich mental imagery of manner of motion’ (2003: 164). By contrast, speakers of languages that do not encode manner in motion verbs, or encode it to a lesser extent, are
called ‘path languages’ (or ‘verb-framed languages’, with the path of motion framed within the verb), are seen as habitually paying less attention to motion, and as paying more attention to path or the direction in which motion proceeds. Examples of path verbs are: *cross, ascend, exit* and *turn*.

Note, however, that much of the evidence that Slobin provides is purely linguistic, and this opens his hypothesis to the charge of ‘lingua-centrism’ (see definition above). In these cases he just emphasizes certain aspects of the grammar of a language, and claims that this provides evidence for certain thought patterns. Thus, he presents evidence showing that speakers and writers of manner languages encode manner in their verbs more than speakers and writers of path languages. For example, he compares pieces of fiction writing in Spanish, French, Turkish and Hebrew (path languages) with those in English, German and Russian (manner languages), counting the numbers of manner versus path verbs (2003: 166). This of course doesn’t show anything more than the differences between the languages, rather than differences in cognition. He does, however, suggest that speakers/writers of path languages use adverbs of manner, descriptions of motor behaviour and body condition, etc. to compensate for the lack of manner verbs, but that apparently overall they still convey less information about manner of motion (2003: 167). Incidentally, Papafragou et al. (2002) suggest the possibility that the fact that speakers of path languages have to rely on additional means to convey manner of motion, and so have to ‘foreground’ it, has the opposite cognitive effect, i.e. that this leads them to paying more attention to manner of motion than speakers of manner languages (2002: 22).

Slobin also provides more convincing evidence. In a small-scale study, he focuses on the mental imagery reported by participants when retelling stories or describing video-clips. For example, speakers of English and Spanish were asked to read passages from novels in their native language. The English versions were literal translations of the Spanish original, so that the same small number of manner verbs were used in both versions. It was thought that this way the habitual ways of paying attention to manner or path of motion, respectively, could not be affected by the number of manner verbs that appeared in the stories. The English speakers reported a rich mental imagery for the manner in which the protagonist
moved, whereas the Spanish speakers reported seeing a series of static pictures which were more like photographs. Thus he provided evidence that speakers of the manner language English pay more attention to motion than speakers of the path language Spanish.

Slobin reports another study in which speakers of Korean and speakers of English watched video-clips that showed an actor who carried out a series of activities. In a surprise memory test the participants were shown clips with the same actor carrying out the same activities, but moving in slightly different ways. The English speakers reported having spotted the differences in manner of movement, whereas the Korean speakers did not report having memorized any such differences (2003: 173-4). Although this seems to be quite convincing evidence of how a focus on manner of motion in a language might influence perception, we still cannot be sure of this, as it might be the case that the Korean speakers just did not report the differences in motion, although they still perceived them, possibly due to some cultural norm. On a merely speculative note, it might just be that Koreans perceive it as impolite to comment on the way someone moves.

However, Slobin’s analysis of these results concludes that some influence of language on cognition, depending on whether the language encodes manner or path, must have taken place. This is not implausible from a cognitive point of view. But it seems that his evidence leaves too much room for speculation: for one thing, these studies were based on the subject’s verbal reporting of impressions, which is linguistic behaviour in itself and might not accurately reflect cognitive processes. To control for this, a more sophisticated experimental setup might be required, which leaves some exciting work to be done in the future.

This section has discussed research relating to the question of if or how language might influence cognition. It has shown the difficulties in developing experimental designs and methodologies that researchers of this sensitive issue come up against in the endeavour to deliver conclusive evidence. A particular problem is here to break though the ‘lingua-centric cycle’, which is not only an issue with approaches that are concerned with linguistic relativity, but also for
Conceptual Metaphor Theory, which is equally open to charge on the grounds of ‘lingua-centrism’. Since the model developed in chapter 6 includes conceptual metaphors, this is an issue that needs to be sufficiently addressed here. Therefore, in the next section, I describe an approach which has the potential to deliver convincing non-linguistic evidence in favour of the notion of conceptual metaphor.

4.1.4 Metaphor and Gesture

There is a type of non-linguistic behaviour that might be more informative with regard to the study of language and cognition than those reported on so far: gestures. Gesturing is a form of behaviour that is intimately intertwined with language and accompanies it in usage, but equally it is possible for gestures to carry meaning that is independent, to a certain degree, of that of the verbal utterance at hand. Research into gesture and conceptual metaphor suggests that detailed observation of speakers’ spontaneous gesturing can provide non-linguistic evidence for metaphorical reasoning (Cienki & Müller, 2008; Cienki, 2000). This means that the most convincing argument against the existence of conceptual metaphors, the argument of circularity resulting from exclusively linguistic evidence (discussed above), could possibly be refuted.

Alan Cienki and Cornelia Müller (2008) report that spontaneous gestures have been observed to be affected by spatial frame of reference. For example, when talking about location or motion relating to an event, speakers of the Australian language Guugu Yimithirr reliably point in the relevant compass directions, irrespectively of their own position while talking, which demonstrates their absolute frame of spatial reference (Haviland, 2000). This also applies to the ways in which time is expressed metaphorically as space. Cultural differences in this can be observed, for example in populations that talk about the future as behind them rather than ahead, such as speakers of the South American language Aymara. In particular, they not only talk about the future as behind them, they also gesture in the backwards direction when they refer to the future (Núñez & Sweetser, 2006). Thus, it can be demonstrated how culture-specific ways of
spatial reasoning are reflected in spontaneous gestures, as well as reasoning involving, presumably, the conceptual metaphor TIME IS SPACE, and the more specific culture-dependent one THE FUTURE IS BEHIND. It is reasonable to assume that other kinds of metaphorical reasoning are reflected in gestures also, and it is then not just an intuitive assumption but can be evidenced by these cross-cultural observations. Cienki & Müller (2008) call the kind of gesture which communicates meaning suggesting an underlying conceptual metaphor ‘metaphoric gesture’. This is elaborated on in more detail below.

McGlone (2007), making a case for the necessity of non-verbal behaviour as evidence for conceptual metaphor, argues that conceptual metaphors could only be taken seriously as an aspect of cognition if it were to be tested in a procedure involving at least three steps: first, one would need to identify a conceptual metaphor which in a particular culture seems to underlie a number of idiomatic expressions used to describe a specific abstract concept, in our example, TIME or THE FUTURE. Then, according to McGlone, one would need to find another culture that talks about the same concept using a different metaphor, to be able to compare the two. In our case, this would be THE FUTURE IS BEHIND versus THE FUTURE IS AHEAD. The third, and crucial, step would be to demonstrate in the population’s performance of non-verbal reasoning that the conceptual metaphors are not just a linguistic phenomenon but are used in reasoning (2007: 114). The metaphoric gestures employed by the Aymara to talk about the future clearly fulfil McGlone’s third criterion, in addition to the other two having already been established. Thus, metaphoric gestures are shown to deliver the desired non-linguistic evidence in favour of conceptual metaphors, evidence that so many people have called for for decades, and that has been all around us all along.

Cienki & Müller (2008) have made systematic observations of how speech and gestures relate to one another in different ways according to the communicated content. They observed four different ways in which this may occur: 1. speech and gesture expressing the same metaphor, 2. a gesture expressing a metaphor that does not co-occur in speech, 3. a gesture that expresses one metaphor in parallel to another metaphor that is expressed in the co-occurring speech, and 4. a gesture expressing a metaphor which is never used in the
language. Of the data Cienki recorded in student interviews, an example of the first type is described with the utterance in (25) accompanied by a gesture of a fist pushing forward against a boundary indicated by the other hand.\textsuperscript{11}

(25) And I think they’re willing to push their moral limits

As an example of the second type Cienki & Müller describe an utterance from the data in which the student talks about the topic of honesty and truth in a non-metaphorical way (without any metaphors for truth in the surrounding co-text), accompanied by ‘a flat-hand gesture in the vertical plane, fingers pointing away from his body’ (2008: 488). The timing is such that the gesture occurs exactly on the word ‘truth’. This is interpreted to mean metaphorically STRAIGHT and involving the conceptual metaphor HONESTY IS STRAIGHT, a for English speakers conventional conceptualisation of the truth, or honesty. This is an example that suggests the possibility of a conceptual metaphor being cognized independently of verbal behaviour. The third type can demonstrate the multi-modality of metaphors, highlighting different aspects of the source domain, or possibly several source domains. For example, a student metaphorically talks about black and white to characterise two abstract moral categories, to do with students’ behaviour at exams, while simultaneously making a chopping gesture in front of her and thus using a spatial metaphor to set the two categories apart. So here she uses two different metaphors in parallel to express the same content. Another, and perhaps more plausible interpretation could be that the same underlying metaphor is expressed in different ways, whereby the setting apart of two ideas referred to could alternatively be expressed by contrasting colours, or by parting space. This accords with my favoured definition of conceptual metaphors, which I have described in section 3.3.4: A conceptual metaphor may be just a very schematic representation, akin to Barsalou’s entrenched situated conceptualisations, of which the various realisations cannot be expressed in sum at a concrete level. They may have, however, a structuring function guiding cognition. Here, the

\textsuperscript{11} The descriptions of the gestures are here reproduced in an incomplete and simplified way, and the notations used are left out. The reproduction of a more detailed example follows in the next section.
difference between black and white, and the difference between the two spaces can be seen as both arising from the same schematic representation.

Cienki & Müller give a particularly interesting example of type four: a student talks about exam preparations while gesturing from left to right along a metaphorical time line. The conceptual metaphor presumably employed is TIME IS SPACE, while the movement from left to right to indicate temporal progression reflects a cultural practice having to do with the left-right orientation of our writing system. However, this left-right orientation does not occur in the language in this way; we do not use expressions like ‘I got up left of having coffee’ to mean ‘I got up before having coffee’. Hence this example can be understood to reflect aspects of visuo-spatial thinking, rather than linguistic aspects, since this concept does not appear in the language (2008: 487-492).

The idea that primary metaphors are the basic experiential ones (cf. the discussion in section 3.3.4) can be seen as related to the consideration that these are the only ones that can be expressed with gestures, due to the constraining practicalities of gesturing. A less basic conceptual metaphor, such as THEORIES ARE BUILDINGS cannot be expressed with a gesture, as it is far too complex, whereas PERSISTENCE IS ERECTNESS/VERTICALITY can, for example by holding one’s forearm in a vertical position. By this I do not intend to imply that only those conceptual metaphors that can be expressed with gestures can be such by definition. However, the observation goes in favour of the idea that an assumption of the existence of conceptual metaphors requires, as a preliminary, that we define precisely what a conceptual metaphor is in terms of its potential psychological reality. Those conceptual metaphors that can be observed to be expressed in gestures certainly seem to be the right kind of candidate, but perhaps not the only one. The issue calls for thorough empirical research in the future. However, I do not agree with the idea that we can infer from mere collections of semantically related lexical items, i.e. belonging to the same conceptual domain, that these are governed by a conceptual entity as suggested by Lakoff & Johnson (1980) (cf. Vervaeke & Kennedy 1996). Spontaneous gestures, however, could give us an indication of how to define basic conceptual metaphors, since, arguably, they
seem to have the function to contribute to our thoughts’ structuring. This will be explained in the following.

Alan Cienki (2000) reports on some specific observations concerning the temporal flow of gestures in conjunction with speech. His analyses invite suggestions about the nature of the interplay between gestures, speech, and thought. He draws on Slobin’s (1987) notion of Thinking for Speaking (2003, described above) and McNeill’s (1992, 1997) ideas involving multi-directional interactions between speech, gesture, and thought. As mentioned above, findings suggest that gestures reflect metaphorical thought. However, if we consider all the possibilities that a multi-directionality of interaction between speech, gesture and thought may afford, we may ask whether it might also be the other way round, that is, whether gestures inspire metaphorical thought. The idea behind this is that, in thinking for speaking, which usually occurs simultaneously to utterances while cognitively preparing the next utterance, metaphoric gestures might subconsciously (or consciously) inspire thoughts, which then get expressed in the following utterances.

Cienki (2000) proposes the tentative hypothesis that conceptual metaphors, which are expressed in gestures, are taken up and elaborated on in speech, which in turn makes their metaphorical meaning more dominant in overall cognition. Thus it may be that conceptual metaphors, via gesturing, have an influence on the thought processes of the speaker on an imagistic-experiential level. The temporal aspect is important here, as a speaker builds up their ideas in the course of speaking: in thinking for speaking, on the search for appropriate expressions to capture an idea in the course of developing it, gestures can facilitate the use of metaphors which otherwise would have been less pronounced, or would perhaps not even have come up. Figure 4 shows a diagram illustrating the multi-directional interplay between thought, gesture, and speech.
The arrows show how thoughts are expressed in speech and gestures, which occur at the same time, and which can be understood to constitute the main activity. The arrows consisting of dotted lines indicate the opposite movement of speech, as well as gestures, functioning in turn as input for thought. If it were shown on a timeline, we would see that speech and gestures influence thought during an utterance and those that follow it. Cienki describes the flow of these mutual influences as a ‘cycle’ (2000: 7). Importantly, gestures as well as speech are here seen to have an impact on thought processes. Thus, gestures are understood to inherently carry meaning of their own, rather than merely being movements that support the meaning communicated in speech.

Cienki observed two distinct patterns of gesturing behaviour that might be relevant to the formation and/or elaboration of metaphors in thought. The first consists of a series of preceding gestures, which initially may not even carry any inherent meaning at all, but just serve to emphasise the verbal content. They then turn into some schematic image which is subsequently used as a source domain of a metaphor. The other pattern involves setting up a ‘scene’ which then gets repeatedly elaborated on, either by just the speaker who initiated it, or by both (or all) interlocutors. I discuss an example of each in turn. The transcripts with the
notations, as well as the descriptions of the video materials and gesture analyses, are taken from Cienki (2000). It should be noted that the analyses are speculative in that they are entirely based on observations of naturally occurring behaviour, rather than psychological testing. However, they appear plausible, and it is worth following them up with experimental testing in the future.

Notation conventions:
{ } = gesture stroke
___ = verbal metaphoric expression of interest in the example
^ = move outward
˘ = move toward self

Example (26)
The speaker’s hands were raised and spread, making (non-metaphoric) beats for emphasis as she explained how students study for exams. At this point she said, “For example I know that I’ve .. y’know, I’m hugely drawn to procrastination of any type,” and exaggerated the motion before and while saying “hugely.” The gesture afforded by the given hand position and previous motion appears to have guided the speaker into thinking in terms of a given metaphor (INTENSITY IS SIZE), which was expressed in a rather unconventional way (hugely).\(^{12}\)

A  Umm {...}
See there's {such a wide variation/@
^ ^ [left hand]
^ [right hand]
of how people,
^ ^
^ y'know,
prepare for exa=ms?
^ ^
^ I -- ..
˘ ^
^ For example I know that I've, ..

\(^{12}\) The ‘unconventionality’ of the expression ‘hugely’ in this context seems to pertain to speakers of American English, as I believe it is quite common in the UK.
y'know I'm -- ..

^ ^

^ hugely drawn to procrastination of any type,

^ ^ ^

^ an' it would be absolutely .. `atypical for me} ta, ..

^ ^ ^

y'know,

even, ..

{`m mildly prepare in adva=}nce for an exa=m unless it's an exam of, {.. really} great astronomical im`portance.

Cienki 2000: 4-5

In this example, it is suggested that the hand movements, which preceeded the utterance, guided the speaker into focusing on the concept of size and using that as a metaphor, as it seems to have created an image in the speaker’s mind, inspiring her to draw on it. Whether or not this process really occurred in the speaker’s mind in this way remains open to charge. However, I would suggest that this first pattern Cienki observed, involving gestures that only in the course of ‘thinking for speaking’ take on a meaning, is a particularly difficult one to pin down in observation of naturally occurring speech behaviour. Nevertheless, this specific example can give us some insight into how the interaction of gestures, thought, and speech might proceed.

The second pattern, involving setting up an imagistic scene for further elaboration, lends itself much better to analysis, and I would also assume that it occurs more frequently in communication. In example (27), the gestures are numbered and described shorthand below the transcript, including analysis. Part of this example, including the main theme of the metaphor, that of ‘pushing the boundaries’, has already been introduced above. Here, the speaker sets up a picture before herself, and keeps referring back to it, which helps her lexicalise her ideas while she draws on this one image from the same source domain (or frame) for a series of metaphoric expressions. It is like ‘filling in the slots of the frame with various elaborations’ (2000: 7). These elaborations seem to reinforce
one another, and the speaker is drawn into the imagined scene by her own gestures, as it were.

Example (27)

Gesture groups are numbered. See key below.

B I don't necessarily think that the majority of American students cheat?
I don't.
I would say I would keep that in a minority?
A They look for the [easy way out.]
B [However,]
I think they look for the easy way out,
very excellent point XXX,
A @@
B and I think that they're willing to push their moral limits,
(1a)                     (1b)        (1c)     (2a) (2b)
or y'know,
their judgment of,
(2c) y'know,
what is right or wrong,
(2d) (2e)
to the extent that it can or cannot be `labelled cheating.
(3a)         (3b>)       (3c<--) (4a<)     (4b<)
And `this you can't really `label an' y'know deem them `cheaters or deem like a `cheater an',
y'know, con- -- ...
con`denn them for that.
So he's taken it to the point where he could benefit from it,
(3d)
yet still rationalize himself,
an' rationalize his situation as not being unethical.
And I think that,
y'know,
really almost any American student,
y'know,
who could,
y'know,
will take it to its limit,
and reap the [benefits to the extent] that he can.
A [I don't mind his rationalization.]
The speaker sets up the scene with “push their moral limits”:

(1) formation of fist, with beats (PUSH)
(2) flat C-shaped hand (LIMIT) [she keeps showing the limit]
(3) flat-O drawing line, R to L (EXTENT)
(4) open hand, palm up (presentation gesture)
and then goes back to (3): “taken it to the point”

Note the cycle: speaker’s mental image of the source domain of a conceptual metaphor leads to gesture, which traces an image in space of the source domain of the metaphor that the speaker has in mind (and has also expressed verbally) [reification of the image], which can then facilitate the creation of additional metaphoric mappings using that source domain, which are then expressed in more verbal and gestural metaphoric expressions.

Cienki 2000: 7

These observations show how imagistic thinking, which is expressed in metaphoric gestures, may get reinforced by the gestures themselves. This mutual influence also involves verbal expression, as when the conceptual metaphor thus strengthened appears to guide further thinking, which in turn partly functions to prepare verbal communicative acts (cf. Slobin’s Thinking for Speaking). The observations give rise to the tentative suggestion that cognition proceeds along two different paths in parallel, an idea that will be elaborated on further in sections 6.4 to 6.6. One path is imagistic, including spatial reasoning, as well as metaphorical spatial reasoning. This can partly be inferred to be the case from observing the spontaneous metaphoric gestures that occurred independently from, and prior to, verbally uttered metaphors. Furthermore, it may suggest that the imagistic processing route also involves conceptual metaphors (in the sense that they are restricted to experiential bases). The other processing route is propositional, and it seems likely that the latter gets influenced and informed by the former during online processing. Obviously, this assumption needs theoretical support, which will be offered in later chapters, and ultimately it will need empirical evidence to be gathered experimentally in the future.

However, first assumptions can be drawn from the observations of the interplay between metaphoric gestures and verbal utterances, described above, which is an important point in the context of my own proposal, which will be outlined in chapter 6. The observation that spontaneous and unreflective gestures can prompt subsequent verbal utterances that employ the same metaphorical content, and that these jointly lead to a propositional statement, supports the
following assumption: These two ways of thinking, i.e. imagistic metaphorical (as in gesture), and propositional (as in speech), are both at work in parallel, with each fulfilling their own purpose while mutually reinforcing each other. It might be that either of its own would hardly be sufficient for coherence in thinking, let alone in communication. The proposal of the new hybrid model outlined in this thesis runs along very similar lines as the analyses of speech-gesture behaviour just outlined. Hence, in the future, a potent way of testing the model empirically, in particular pertaining to the interplay of different representational formats, will be found in gesture research.

4.2 Awareness of metaphoricity

An important aspect, which many people intuitively feel to be an integral part of metaphor, is the impression made, or the image mentally evoked, of the literal meaning of a metaphorical expression. Think, for example, of an expression like cliff-hanger to describe the ending of an episode of a TV series, which remains open and leaves us waiting in tension for the next episode to deliver the solution. At the very least, people who have seen the film ‘The Italian Job’ can’t help but access the image of a bus hanging halfway over a cliff (at least I can’t help it), and this image facilitates understanding of the expression cliff-hanger and enriches it with a mental experience much more vivid than if the ending of the series had just been called ‘unresolved’, or the like. This surely must play an important role in the explanation of metaphor understanding (cf. the definition of ‘metaphoricity’ in the introduction p. 10-12). However, Relevance Theory does not seem to cater for this sufficiently, at least not explicitly, since the focus here is on propositional meaning. How we experience an expression in all kinds of imagistic or otherwise sensual ways seems to be more of an accompanying side effect. But is that really so? Couldn’t it be that such ‘side effects’ have a much more substantial influence on meaning-making? In a recent paper (Stöver, submitted) I consider to what extent theoretical approaches should take account of intuitions like this, and whether awareness of metaphoricity is an aspect that needs to be included in a comprehensive account of metaphor understanding. Within Relevance Theory
(e.g. Sperber & Wilson 1986/95, 2006; Wilson & Carston 2006, 2007) the focus so far seems to be more on the analysis of subconscious representations and/or processing, while the contribution of conscious and/or reflective processes to pragmatics has not received much explicit attention. However, more recent discussion has begun to look more closely at this, e.g. Carston (2009; 2010) considers the role of awareness and metarepresentation in the pragmatics of metaphor understanding. Further useful input comes from work by Camp (2008), who emphasises that we feel a tension between the literal and the figurative meaning, and Rubio (2008), who, from a relevance-theoretic point of view, provides experimental evidence which supports such intuitions by investigating the processing role of so-called ‘core features’ (section 4.2.2). The conception of core features can be understood as relatively close to what are generally understood as literal meanings. The results from this study can be seen as bringing us a little closer to an account of how an awareness of the literal meaning, as opposed to the figurative meaning of an expression, might influence comprehension. To illustrate this further, section 4.2.3 provides English-German cross-linguistic examples which show how awareness of metaphoricity clearly affects comprehension. Translation asymmetries encountered by bilinguals can be explained by taking the issue into account, in particular when it comes to expressions that are novel metaphors in one language but conventional in the other. Here, experience of awareness of metaphoricity may vary, depending on the meaning in the native language of the particular expression, which might be more or less lexicalised than in the second language.

4.2.1 The role of the literal meaning

Being aware of the literal meaning while interpreting a metaphorical expression often amounts to being aware that the expression is a metaphor, to a sense of its metaphoricity. Carston (2009) has dubbed this phenomenon ‘the lingering of the literal’. This is a consciously accessible process, which may, in particular with creative metaphors, involve reflective thinking leading to the interpretation outcome. For example, a poetic utterance such as ‘His ink is pale’, used by
Flaubert to comment on another poet, Leconte de Lisle, is likely to prompt a mulling over what the author intends to mean by the utterance (example from Sperber & Wilson, 1986/95: 237). This leads to the tentative hypothesis that differences in processing are relative to awareness of metaphoricity: the more a communicator is aware of the tension between the literal meaning and the figurative meaning of an expression, the ‘intuitively felt gap’ (cf. Camp 2008, see below), the more conscious effort is invested in the interpretation process; the less a communicator is aware of an expression’s metaphoricity, the more automatically and unreflectively it is processed. As already mentioned in the introduction (p. 10-12), Camp (2008) describes this phenomenon as the ‘intuitively felt gap between literal and intended meaning, where the first provides the perspective for constructing the second.’ (2008: 14). It can be argued, to follow Camp’s line of thought, that this reliance of metaphorical meaning on the difference between the literal and the intended meaning shows that the relevance-theoretic continuity account may need adjusting, for it does not seem sufficiently to take account of this ‘felt gap’ in terms of comprehension. It assumes conceptual meaning modification to consist of subconscious and automatic inferential steps, which seem to provide a smooth transition from the literal to the intended meaning by way of broadening and/or narrowing, as described in section 2.2. Intuitively, this seems to lack an account of why we feel that metaphor is ‘special’, as it sees metaphor as an expression type that is not processed in different ways than other types of expression. Thus, the continuity account cannot explain the strong intuition most of us have because of this gap. By contrast, Camp’s gap is consciously accessible via introspection and prompts an awareness of metaphoricity, which leads us in the first place to view the expression at hand in a different way from its literal counterpart. The reason for this is that in many cases we bear the literal meaning in mind while interpreting the figurative. Carston (2009; 2010) explores the idea that with sustained creative metaphors, processing might involve a metarepresentational level at which the literal meaning is retained but understood to describe another, imagined world, and the intended meaning is arrived at within this world, as it were. Thus she departs from the ad hoc concept construction account as capable of catering for all kinds of metaphor.
and proposes a new processing route for extended metaphors. Therefore, the problems arising with the account and complex metaphor, as I have demonstrated in section 2.5.1, seem to have been tackled. Furthermore, Carston (2010) proposes a dual processing account based on this new approach, giving a more dominant role to the literal meaning of short and simple metaphors also. This will be described in chapter 6. At first glance, the idea of holding the literal meaning in mind while processing the propositional meaning outcome, does not seem too far away from the notion of source and target domains in Cognitive Linguistics (e.g. Lakoff & Johnson 1980; Gibbs 1994; Steen 2007), since it is reminiscent of the idea that two domains, the source domain and the target domain, are mapped against each other to derive the metaphorical meaning. Carston’s new account seems to suggest a similar juxtaposition. Comparable to the description of metaphor offered by Camp, the difference between the literal and the intended figurative meaning is taken here as the focal force that brings out the metaphor as such.

In order to investigate how conscious awareness might influence the processing of metaphor, one should be clear about the differences between conscious and subconscious processing. Carston (2009) suggests that it may be worth paying explicit attention to the differences between reflective thought processes and subconscious inferential computations. Recanati (2002) points out that in the literature, two senses of ‘inferential’ are being used, depending on the author’s outlook and choice. In some cases ‘inferential’ refers to reflective thinking (or reasoning proper), and in others, in modern psychology in particular, it refers to automatic and subconscious computational processes, where the agent is consciously aware only of the conclusion of the inference but not of how it has been derived (2002: 16). Therefore, it is not possible to judge via introspection whether an interpretation is arrived at by associative or by inferential steps (for further discussion see Recanati 2004, Mercier & Sperber 2009). Relevance theorists tend to use the term ‘inferential’ in this latter sense, and I suggest that the subconscious processes explicated should be complemented with an explanation of how reflection at the conscious level may impact on the former, since the aspect of awareness of metaphoricity appears to play a role in the description of
the mechanisms involved. Ad hoc concept construction, as part of the relevance-theoretic comprehension heuristic, is an inferential process that is thought to take place subconsciously, fast and automatically. That is, according to this approach we are aware only of the overall interpretation outcome of a metaphorical expression, because the inferential steps involved in broadening and narrowing take their course at a level below consciousness. However, this does not seem to be the whole story, since we have strong intuitions about many metaphors being understood by consciously and reflectively taking into consideration the literal meaning and how it relates to the intended figurative meaning. The literal meaning does not just ‘disappear’, it seems to linger on in the mind, in parallel with and related to the propositional meaning the metaphor is used to express. This has been shown in processing terms in an experimental study described in the next section.

4.2.2 The literal meaning in metaphor processing: A psycholinguistic study
Carston (2009) emphasised the immediate pertinence of a study by Rubio (2008) which provides psycholinguistic evidence for the ‘lingering of the literal’. The study employed cross-modal priming experiments, and the results suggest that dominant encyclopaedic features remain active in the process even in cases when they are not computed as part of the meaning outcome. Rubio’s aim was to test the hypothesis that concepts have core features and non-core features, a distinction inspired by Barsalou’s (1982) definition of context-independent and context-dependent properties. It corresponds (but is not equivalent) to the relevance-theoretic idea that features are either strongly or weakly associated with a concept. For example, ROUND is understood to be a core feature of the concept BASKETBALL, while FLOATS is a non-core feature of BASKETBALL, as it is thought only to become dominant in a biasing context, in this case possibly involving a lake or a swimming pool.

The results provide supporting evidence for the relevance-theoretic account in terms of ad hoc concept construction, but also suggest additional insight into the comprehension of figurative language: conceptual features can be activated
without necessarily being computed as part of the meaning. This distinction between feature activation, and computation in addition to activation has not been made in previous studies, where it was held that mere activation of features is evidence for meaning-constitutive computation (e.g. Barsalou 1982; Whitney et al. 1985; Greenspan 1986; Tabossi 1988). However, it can be argued that features that are activated but not computed as parts of the propositional meaning may still be consciously accessible and play a role specific to the comprehension of figurative language.

In preparing the experiments, Rubio attempted to establish a consensus as to which features are regarded as stable meanings. The determination of which features count as dominant and which as non-dominant was carried out with the use of questionnaires, guided by the literature on prototypes (Barsalou 1983; 1985; 1987; Rosch 1973; Rosch & Mervis 1975). Cross-modal priming was employed to test the idea that all properties associated with a concept are activated on word recognition, but only some of them remain as parts of the comprehension outcome. This is in tune with the relevance-theoretic notion of narrowing, where specific encyclopaedic properties are thought to gain content-constitutive status, and thus dominate the interpretation process, whereas other features associated with the concept play no active role in computing the meaning at hand. The dominant properties could also be non-core features, depending on occasion-specific use. This point is of particular interest with regard to figurative language, which, according to this approach, is characterised specifically by changes in the status of conceptual features. It is possible that core features do not end up as parts of the comprehension outcome; this happens in cases where they are not computed as part of the word meaning. However, the study suggests that they still remain highly activated and accessible throughout the time span of processing, up to 1000ms, unless they are actively suppressed in cases of contextual inconsistency, as happens with inappropriate meanings of polysemous words (e.g. bank referring to the financial institution vs. bank referring to a riverside). By contrast, non-core features in general become deactivated after 300-400ms, unless they are contextually relevant and thus become content-constitutive. For example, ROUND as a core feature of BASKETBALL would remain activated, while its non-
core feature FLOATS would passively decay after 300-400ms if contextually irrelevant, as in a situation without a body of water around.

Rubio shows that there is a difference in processing between active suppression of inconsistent conceptual features and passive decay of irrelevant features. As concerns active suppression, different researchers have developed models that vary in how specific a suppression mechanism might be: one possibility is a small-scope suppression mechanism which is confined to cases of disambiguation of polysemy, i.e. a mechanism focused on inconsistent meaning features (e.g. Simpson & Kang 1994, Simpson & Adamopoulos 2002). Another is a more general, wide-scope kind of mechanism, which includes the suppression of merely less relevant features in addition to inconsistent ones (e.g. Gernsbacher 1990). It turns out that the small-scope view seems to be supported, since core features remain active after 400ms even if irrelevant. To suppress core features, competition between property candidates is necessary, which is the case with disambiguation only. Non-core features lose activation even if they are just irrelevant, rather than contextually inconsistent, but remain active in biasing contexts where they are relevant. The study thus indicates that there is a clear difference in processing between core and non-core features. It can be concluded that, in investigating lexical processing, attention needs to be paid to the difference between levels of activation on the one hand, and whether a conceptual feature is being computed as meaning constitutive, on the other.

Overall, then, core features function as default meaning components in ‘neutral’, non-biasing contexts and appear in the stereotypical interpretation, unless prompted otherwise. If such a stable element of a concept is currently meaning-irrelevant, it is not computed but could still play some role, due to its activation throughout the comprehension process. Since Rubio tested subconscious automatic pragmatic processes, she does not discuss the extent to which the sustained activation of a meaning-irrelevant core feature is available to consciousness. Generally, the observed behaviour of core features can be understood in relation to the tension between literal and figurative meaning, which we intuitively perceive and which may prompt reflective processes in metaphor interpretation.
4.2.3 The multilingual perspective: Metaphorical translation asymmetries

It should be fruitful to look at translation equivalents that are fully metaphorical in an individual’s first language due to being novel, but conventionally metaphorical or with a lexicalised metaphorical sense in their second language. Presumably, second language learners need to deal with such translation asymmetries in a reflective way, in order to get the intended meaning right in the given context. In cases where an expression’s metaphorical meaning is novel to the second language learner, they will have a much higher awareness of its metaphoricity than the native speaker. According to Steen (2008, cf. section 5.1.1), the novel/conventional distinction, which applies here, should not be conflated with the deliberate/nondeliberate distinction. However, it seems likely that when second language learners come across such translation asymmetries, they perceive the expression as if it were a deliberately used metaphor, i.e. it would be a case of a mild misunderstanding, but which, if communication is successful, would be resolved by employing metaphorical processing mechanisms.

For example, the adjective thick, analysed in relevance-theoretic terms in section 2, is frequently used in an idiomatic way in English, paraphrasable with ‘stupid, slow on the uptake’. In German, it is just used in its basic, physical sense ‘large or dense’ (German: dick, dicht), but not ‘stupid’. A German learner of English can be assumed not to be familiar with this specific metaphorical sense upon first encounter, and to perceive it as deliberately metaphorically used, and so to consciously reflect on it, hopefully to arrive at the intended meaning. That is, we can assume that they would think about the basic physical meaning of thick and how it may relate to the domain of character traits of a person (rather than, say, body size), as required by the context. It is clear that awareness of metaphoricity plays a crucial part here, as compared to the processes involved in merely subconsciously accessing the intended meaning. Another apt example is provided by the idiomatic phrase This song has grown on me, which can be paraphrased thus: ‘I got to like this song more and more, even though I may not have liked it at first.’ Literally translated into German it makes no sense (German: *Dieses Lied ist auf mich/mir gewachsen), as German does not have a

13 It is, incidentally, metaphorically used in some contexts in German, for example ‘dicke Freunde’ means ‘close friends’.
conventional metaphorical equivalent, so there is only the physical sense, which sounds rather bizarre. Similarly, the translation equivalent of *to mince one’s words*, paraphrasable as ‘to speak carefully, to speak in affected way’, has not lexicalised in this way in German, and only evokes the basic physical sense of the words (German: *seine Worte hacken*). Other examples are idiomatic expressions like *pick someone’s brains, getting the hang of it*, and many more.

While these expressions are nonsensical in German, their use is conventional in English. When a native speaker of German (or another language), who is learning English, comes across them in English discourse, they will presumably be fully aware of their metaphoricity and arrive at the intended meaning, if all goes well, via conscious reflection, at least to a certain degree. For example, when I first heard the expression *don’t mince your words*, I mentally pictured a mincing machine with meat going through it, and I had to reflect on how that might relate to the current context and to *your words*, before I was able to grasp the intended meaning. Thus, my awareness of metaphoricity was alerted, presumably more than would have been the case if English had been my native language and I would have accessed the lexicalised metaphorical meaning of *mince* readily stored in memory. In English, these and other examples are cases of polysemy, which from an etymological point of view are metaphorically motivated. So an average native speaker of English (and non-expert in linguistics) might detect no metaphoricity at all there, and we can expect them to process the expressions automatically without any conscious reflection. Investigating how speakers with different linguistic backgrounds process such asymmetric translation equivalents in fundamentally different ways may cast light on the question of how metaphor awareness influences processing. What the interplay between the different kinds of cognitive processes may look like, however, still needs to be explored; whether conscious reflection triggers, or runs in parallel with, the automatic lexical adjustment processes, or whether the intended meaning is derived via mapping processes as in Cognitive Linguistics, or perhaps via a combination of processing types, are still open questions. In particular, how imagistic or experiential representations might influence propositional processing is an issue of interest concerning the hybrid model developed in chapter 6. But
first, let’s have a look at other, yet related attempts at merging different approaches to create a new account of metaphor comprehension, in the next chapter.
Chapter 5  
Combining approaches

This chapter describes attempts to unite different approaches to metaphor understanding that have informed the development of the model proposed in this thesis. Section 5.1 reports on a discussion aiming to combine accounts of metaphor analysis (Steen 2008; 2009; 2010) with the processing mechanisms as proposed in Relevance Theory (sections 5.1.1 to 5.1.4). This exercise came up against various difficulties, and its description provided here is intended to draw attention to the importance of a careful analysis of the differences between the accounts of metaphor in question, in particular concerning their analytic perspectives and aims, before embarking on merging them. These observations in turn can inform us on how to develop a successfully combined account of metaphor, provided that we consider which elements and principles of the respective frameworks are suitable for combination, and which, in principle, rule each other out. This ensures an awareness of constraints on theorising. Gerard Steen (2007) describes some of the different perspectives and aims which ultimately render some approaches to metaphor incompatible with others. This can be, for example, due to a focus on language use versus a view of language as a ‘closed system’, potentially making the respective approaches mutually exclusive. This is discussed in section 5.1. Section 5.2 describes Tendahl’s (2006) hybrid model of metaphor understanding in some detail, as it has significantly influenced the new hybrid model developed here.

5.1 Steen: Symbolic and behavioural approaches

Steen (2007) paints a picture of a range of approaches to metaphor that can, broadly speaking, be classified into two main categories: symbolic and behavioural (e.g. 2007: 10-13). Analysis of metaphor in a symbolic framework treats language as an entity in itself, i.e. it assumes that language exists as if it were an object that can be looked at independently and detached from its use. Such a point of view does not take into account cognitive aspects of language use...
and what happens in production and comprehension psychologically. Its purpose is to describe semantic relations and their origins within a language, many of which do not play any role in online processing. That is, for a communicator in a discourse situation there is no need to consider whether a polysemous sense is, from an etymological point of view, metaphorically motivated or not. Hence we can say that, generally, symbolic analysis does not correspond to a psychological reality. It can, however, have merits of its own, depending on a researcher’s goals and intentions.

Behavioural approaches, by contrast, are aimed at investigating the processes that take place in actual language use. Here, expressions that might have been metaphorical in the past but have now lexicalised (e.g. ‘cool’) are not understood or processed as metaphorical, simply because their comprehension does not employ metaphorical mappings (or, depending on the theorist’s conviction, some other form of involvement of the original literal meaning). Such cases are then treated as polysemous, regardless of whether the polysemy is metaphorically motivated on etymological grounds (Steen 2007; 2008).

Steen emphasises that each type of approach is based on different perspectives, and accordingly leads to different outcomes of analysis, and while each is useful in their own way, conflation of the two should be avoided. Furthermore, whenever interaction between the two takes place, it should be clearly spelled out. Many expressions are, by many theorists, interpreted as metaphorical on linguistic grounds, such that from a diachronic perspective they have metaphorical origin. Thus they are properly metaphorical in a symbolic (or structural) framework. In Cognitive Linguistics however, as Steen observes, the model of metaphoricity as defined within a symbolic analysis is also expected to lend itself to the analysis of cognitive processing, i.e. an account’s symbolic orientation is interpreted as behavioural (2008: 215), presumably without the analysts making any distinction between the two. As a result, expressions that can be seen as metaphorical on merely etymological grounds, but are not psychologically experienced as such, are assumed to be cognitively processed as metaphors. There are symbolic approaches, such as Conceptual Metaphor Theory, that tend to assume that conventional and even fully lexicalised metaphorical
expressions are processed via cross-domain mappings, due to their metaphorical origin (Steen 2007; 2008). However, psycholinguistic evidence suggests that this is not the case, and many expressions that are understood as metaphorical from a symbolic perspective are not actually processed metaphorically (2008, see below).

5.1.1 The three-dimensional model
Steen (2008) postulates three dimensions of metaphor, namely language, thought, and communication. He first describes two of them, the dimension of language and that of thought, and explains that most (if not all) approaches in Cognitive Linguistics are based on the idea of such two-dimensionality. He draws particular attention to the Career of Metaphor Theory (Bowdle & Gentner 2005). Here it is assumed that a metaphor has a cognitive representation in the dimension of thought, independent of language, on the one hand. The language dimension, on the other hand, refers to linguistic form, i.e. how a metaphor is presented: either as a metaphor (Juliet is the sun) or as a simile (Juliet is like the sun). Thus, there is no distinction made between metaphor and simile in terms of their cognitive representation, as the conceptual entity is thought to be the same in either case, independent of its realization in the linguistic dimension. This accords with the view generally held in Cognitive Linguistics, and Steen bases his own theory on the principle that metaphor is primarily a matter of thought.14

However, in the Career of Metaphor Theory a distinction is made concerning the processing of a metaphorical representation: is it expressed as a simile, it is thought to be processed by comparison, whereas a conventional metaphor, if linguistically expressed as a metaphor, leads to a processing by categorization. The latter involves a superordinate category which includes the conventionalised metaphorical sense, an idea that originated in Class Inclusion Theory (e.g. Glucksberg 2001).

It should be noted that the term ‘metaphor’ is being used here with two slightly different senses: one sense refers to a cognitive representation, the other refers to a linguistic form. This is because a metaphor in the dimension of thought

14 Relevance theorists, by comparison, do make a processing difference between metaphor and simile.
can be expressed either as a metaphor or as a simile in the dimension of language (2008: 216). In the Career of Metaphor Theory, the other factor that determines how a metaphor is processed is its relative novelty or conventionality. This is understood as a matter of cognitive structure, i.e. residing in the dimension of thought. Thus, whether a metaphor is processed by comparison or by conventionalization depends on the interplay between these two factors: linguistic form and conceptual structure. While novel metaphors are thought to be processed by comparison, i.e. cross-domain mappings, conventional metaphors, such as ‘faith is an anchor’, would be processed by categorization due to their conventionality. However, a conventional metaphor that is presented as a simile (‘faith is like an anchor’), would be processed by comparison, which is thought to be prompted by its linguistic form (2008: 216). In short, the dimension of thought covers what in Cognitive Linguistics is understood by ‘cognitive structure’, here in concrete terms whether a metaphor is novel or conventional. The dimension of language covers the linguistic form of a metaphor (its packaging, so to speak), here in concrete terms whether a metaphor is presented as a metaphor or as a simile.

These two dimensions are the starting point for Steen’s (2008) three-dimensional model of metaphor, with the addition of a third dimension: communication. While communication plays a central role in Relevance Theory, Steen emphasizes that in contemporary mainstream metaphor studies, including the Career of Metaphor Theory, the focus has been on language and thought, but not on communication. The Career of Metaphor Theory can demonstrate how the dimensions of language and thought interact, and it can account for a lot of cases. But, according to Steen, it does not address the various purposes that metaphorical usages in discourse are intended for (2008: 217). So communication is an aspect of metaphor which deserves more attention than it has received so far in Cognitive Linguistics, as it impacts on whether an expression is processed metaphorically or not. Regarding the dimensions of language and thought, however, Steen follows the Career of Metaphor Theory in the definition of metaphorical processing as involving some form of comparison, i.e. a metaphorical mapping, and nonmetaphorical processing as involving categorization.
The communicative dimension can be usefully employed for analysis of a lot of aspects. For now, Steen focuses on the distinction between deliberate and nondeliberate use of metaphor, as this is the aspect that determines whether a metaphor is processed metaphorically or not. Deliberate use is intimately connected to an awareness of metaphoricity, which is thought to prompt a comparison process. Steen describes this process along the following lines: the addressee is expected to view a conceptual domain that is part of the discourse topic at hand from the perspective of another conceptual domain, and to set up a cross-domain mapping from the latter (the source) to the former (the target). Awareness of the fact that the source domain has a disruptive effect on the referential coherence of the target domain’s conceptual structure, which is the topic of the discourse, sets this process in motion. When a metaphor is used nondeliberately, it goes relatively unnoticed in discourse.

In Cognitive Linguistics, all metaphors are thought to be interpreted via domain mappings, but Steen points out that psycholinguistic research suggests that many metaphors are processed via categorisation (or class inclusion) instead of cross-domain mappings (or comparison) (e.g. Glucksberg 2001, Glucksberg & Haught 2006). Using the metaphor of ‘a psycholinguistic crack in the cognitive-linguistic mirror’ (2008: 214), he argues that the divergence between the cognitive-linguistic sweeping generalisation of all-applicable domain mappings and the psycholinguistic research results mentioned above represents a paradox that needs to be remedied. The solution he offers is based on the type of use of metaphor: it is only deliberate use that invites the addressee to set up a domain mapping. This characteristic is not to be confused with the distinction made between novel versus conventional metaphor, which is another aspect determining metaphoricity. The latter relies on the degree of frequency and familiarity of the expression in question, which, according to Steen, is defined within the conceptual dimension of metaphor. That is, it is a matter of where a metaphor stands within the dimension of thought, rather than its use in discourse (dimension of communication). Crucially, deliberate use is characterised by its rhetorical purpose: when a metaphor invites the addressee to change their perspective to a

\[\text{15 and, according to the definition used in this thesis, whether there is metaphoricity or not.}\]
different conceptual domain, it is used deliberately. Its relative novelty is not always relevant to its deliberateness, since such an effect can also be achieved with a conventional metaphor, for example by means of word play. An example of a conventional metaphor used deliberately for rhetoric effect is this one involving the word ‘hard’:

(28) Old bread isn’t hard – no bread, that’s hard!

‘Hard’ in the psychological sense is a conventional metaphor, i.e. its figurative sense has lexicalised. However, the play on words here draws attention to its metaphoricity. We would not normally expect the word ‘hard’ to be used as a deliberate metaphor. Its metaphorical origin usually goes unnoticed, and the appropriate sense is then just accessed rather than metaphorically processed.16

Steen identifies his three dimensions with their main characteristics which he labels naming, framing and changing, and describes them as follows:

1. Naming: to fill linguistic gaps for existing conceptual forms (labelling a concept with an existing word from a different domain)
2. Framing: to place a metaphor in a conceptual framework (conceptual metaphor)
3. Changing: changing perspective to a different conceptual domain as a communicative purpose (deliberate use)

Steen 2008: 231

An interesting question to be asked is how Steen’s threefold model compares to Relevance Theory as concerns the ‘cognitive dimension’, given that Relevance Theory is not least a theory of cognition. In Relevance Theory, language is more or less loosely used to convey thoughts, i.e. departing from the encoded meaning to a greater or lesser extent. The thoughts conveyed ‘may be vague, but need not themselves be metaphorical’ (Wilson 2009: 2). This does indeed match Steen’s idea of a ‘linguistic dimension’, which he defines as ‘filling linguistic gaps’ (2008: 231), i.e. labelling a thought or conceptual representation that has no

16 Another example of this type is this one: ‘You are so cool it’s snowing behind you!’
existing linguistic expression associated with it, by way of using a word that is 
originally used for another conceptual representation. In Relevance Theory, this is 
done via lexical adjustment (ad hoc concept construction). However, Steen’s 
‘cognitive dimension’ clearly corresponds to Conceptual Metaphor Theory, and is 
in Relevance Theory not assumed to exist in this way.

Steen emphasises that it is important to maintain these three dimensions 
and to keep them apart in analysis. Nondeliberate metaphor, for instance, even 
though in the communicative dimension it is in contrast with deliberate metaphor, 
in the conceptual dimension is no different from deliberate metaphor with its 
relative degree of novelty versus conventionality, etc. (2008: 228). And symbolic 
analysis in the language dimension, for example, allows for the identification of 
metaphorically motivated polysemy. The latter may not be present in behavioural 
analysis which involves the processing of cognitive representations, in the 
dimension of thought and/or communication. Thus the three-dimensional model 
allows for a view on the language system ‘as such’ and analysing semantic 
relations across the lexicon, on the one hand, and for an analysis of actual 
language use, including rhetoric effects and cognitive processing, on the other. 
However, Steen does not seem to have much to say about how the three 
dimensions interact and leaves the issue to be tackled in future research (2008: 
228)\textsuperscript{17}. But it can be argued that an interaction of the dimensions should be 
considered to play the most important role, given the question of the model’s 
merits. As long as we are not able to see the connections, the three-dimensional 
model bears no difference to three separate models, one per dimension, that stand 
independently of each other. Thus, we can analyse symbolically the conceptual 
structures as is done e.g. in Conceptual Metaphor Theory (dimension of thought), 
we can do a behavioural discourse analysis of the rhetoric use (dimension of 
communication), and in parallel we can look at the linguistic form (dimension of 
language). It does not seem to make a difference whether different outcomes of 
analysis, such that a conventional metaphor is classified as metaphorical on 
structural grounds but nonmetaphorical according to processing considerations,

\textsuperscript{17}He does, however, tentatively begin to integrate the dimensions in his five-step model (2010), 
see below, but I am not convinced due to the mismatch between symbolic and behavioural 
analysis, to be discussed below.
are due to different models of analysis, or due to different dimensions of analysis within one model. Therefore the three-dimensional model remains to be shown to be favourable, in which case the relations between the dimensions should tell us something more about metaphor than other models can.

5.1.2 Combining the three-dimensional model with Relevance Theory?
Leaving the criticisms outlined in the previous section aside, I thought it worth investigating whether there could be role for Relevance Theory to play in an exercise of combining approaches. That is, since Relevance Theory has a lot to offer in terms of cognitive processing mechanisms, which Steen’s three-dimensional model (henceforth 3D model) seems to fall short of, the former could be envisaged to complement the latter. In order to investigate the theoretical implications of such an endeavour, I tried to ignore the seemingly mutually exclusive principles for the time being, to enable this kind of thought experiment. Such principles included, for example, that Steen views metaphoricity as a binary matter: an expression is either used metaphorically, or not. In Relevance Theory, metaphor as such is seen as an artificial construct, while the conceptual distance between the encoded concept and the communicated concept (i.e. figurativeness) is relative and gradual, so that metaphor lies on a continuum with other types of expression (cf. section 2.1). Steen follows the idea of mainstream Cognitive Linguistics that metaphor resides in thought (as explained above), while in Relevance Theory, metaphor is a linguistic means to convey thoughts of all kinds. Steen’s perspective, however, which is a cognitive-linguistic one with an awareness of the advantages that a pragmatic model such as Relevance Theory has to offer, allows for a merging of the two approaches. Here, the anticipated advantages of employing parts of a different framework that has the useful features of flexible processing on offer, such as Relevance Theory, initially seem to outweigh the disadvantages of a potential theoretical mismatch.

Within the 3D model, the area most suitable to accommodate relevance-theoretical elements seemed the dimension of communication, since this is where utterance processing can be described (and it is Relevance Theory’s natural
habitat, so to speak). The main reason to set the focus here was the observation that the 3D model could not account for the indeterminacy and uncertainty of intended meanings that deliberate metaphors can evoke, in particular when they involve, according to the cognitive-linguistic view, novel cross-domain mappings. So far, the 3D model just stated that deliberate use of metaphor invites the addressee to set up such a cross-domain mapping, but there was no account of the pragmatic mechanisms by which occasion-specific meanings are chosen over other potential meanings a domain mapping is thought to produce. This is something that Relevance Theory could cater for. Furthermore, a relevance-theoretic analysis can account for how the addressee is engaged more actively in ‘meaning making’ as compared to comprehension of literal language, due to their individual input as regards the choice of interpretation. The more weakly implicated weak implicatures are communicated, the more responsibility the addressee has to take concerning their understanding of the metaphorical expression being used (as described in section 2.1).

Thus, the idea was to leave the dimensions of language and thought in the 3D model as they were, and to integrate in the dimension of communication the relevance-theoretic notion of weak implicatures as they unfold in online processing, according to the communicative principle of relevance. So, cross-domain mappings in metaphorical interpretation were taken as a given, and weak implicatures were rehearsed to operate on top of them. Intuitively, this can be explained as follows: As introspection tells us, in understanding deliberately used metaphor we perceive a tension between the literal and the intended figurative meaning, on the one hand, and a potential for various simultaneous interpretations (i.e. weak implicatures), creating a certain vagueness, on the other. Thus the 3D model should incorporate both types of processing, a comparison between the source and the target representation, i.e. domain mappings, as well as weak implicatures that lead to a situation-specific array of meanings.

To illustrate, the metaphor in example (29) can be assumed to be used deliberately for rhetoric effect in a discourse situation of someone describing their work surroundings (the person not being a beekeeper by profession).
(29) Our office is a beehive.

The basic (or literal) sense of the word ‘beehive’ obviously not being part of the discourse topic at hand, triggers activation of the concept to do with insects’ habitat, and the addressee is invited to take the word ‘beehive’ as referring to the workplace. This is done by way of a metaphorical mapping between the domains of work (target domain) and nature (source domain). The comparison thus undertaken draws attention to what the referents have in common, e.g. beehives and offices are both places where agents systematically follow tasks in an organised way. Furthermore, features of the source domain are projected onto the target domain: it is typical of beehives that they display frantic activity, they are crowded and the bees never seem to rest. As a result of the domain mapping, these typical traits are viewed as features of the workplace, and the metaphor is interpreted to mean that the office in question is a very busy place. But this is not the only meaning that the metaphor gets across. If that were the case, the speaker then might just as well have said ‘Our office is busy’. But it evokes an abundance of meanings, a few examples of which are listed here:

(30) 
- Beehives buzz with frantic activity  
→ employees work very hard and are energetic and motivated.  
- Bees are all the same  
→ employees lack individuality.  
- Bees never rest and mechanically carry on regardless  
→ employees are forced to do overtime without breaks.  
- Beehives are crowded  
→ employees must make do with squeezing into tiny workstations.  
... and so on.

The 3D model seems to allow for any or all of these interpretations to occur regardless, as long as they satisfy the speaker’s goal of changing the addressee’s
perspective, and the addressee’s aim at restoring coherence within a framework that relates to the characteristics of the genre within which the metaphor occurs. We can assume that the choice (voluntary or not) of features that are in effect projected in comprehension, i.e. are psychologically real, depends on a) the specific discourse situation, including further contextual information, and the information the addressee has about the speaker, e.g. whether they prefer a busy work environment or not, b) the addressee’s individual preferences or tendencies, e.g. depending on their general outlook or mood they might attend more to the positive features of an office compared to a beehive, or more to the negatives. Crucially, the interpretation that the addressee entertains might not be the same as the one intended by the speaker.

So the 3D model needs to account for differences in interpretation between individuals, and other occasion-specific differences. It also needs to account for the effects of vagueness caused by the array of meanings, all of which are potentially intended but none of them with determinacy, more or less, apart from the most obvious ones, e.g. in (29) that the office is busy. So far it cannot do this, and the idea was to remedy this lack by adding elements from Relevance Theory. This can be illustrated with example (29) ‘Our office is a beehive’. Again, we imagine this metaphor to be used deliberately in a situation of someone describing their workplace. Now, if we merely consider a domain mapping to take place in comprehension, (leaving aside, for the purpose of exposition, constraints that pertain to feature accessibility due to familiarity or frequency), we suggest that either all features possibly made available by the domain mapping contribute to the interpretation outcome, or merely a choice of features that seems rather arbitrary. For example, in the latter case, how do we explain, in terms of processing, an interpretation outcome of (29) that favours a reading where the office is an overcrowded place where it is annoying to work? Or, alternatively, an inspiring workplace where one feels invigorated and motivated due to the high level of activity? Here, relevance-theoretic mechanisms can be employed to accommodate a) the fact that there are choices to be made by the addressee, because the figurativeness of the utterance leaves it relatively open which interpretations are intended and b) how the addressee may arrive at these choices.
Thus, with (29) interpretation might proceed as follows: The speaker invites the addressee to compare their office with a beehive, i.e. to activate the conceptual domain of nature in order to view the domain of work in a different light. The domain mapping may evoke representations that lead to any of the potential interpretations exemplified in (29) above (of which in reality the range would be much broader than this list, of course). The domain mapping provides the basis for a derivation of the potential meanings, the choice of which is eventually arrived at via implicatures. That is, only those features that comply with the contextual information both speaker and addressee have, and with both of their individual preferences etc., are being computed, according to the relevance-theoretic comprehension heuristic. Let’s consider two different scenarios to illustrate this possibility in the case of (29):

a) A knows that B is very noise sensitive, needs lots of space and finds it difficult to concentrate in a busy environment. In addition, A can tell by B’s sarcastic tone of voice that B is not pleased with their workplace. Based on the domain mapping A set up prompted by B’s deliberate use of the ‘beehive’ metaphor, A’s choice of features thereby opened up will fall on the negative ones, such as ‘Employees must make do with squeezing into tiny workstations’.

b) A knows that B works best with lots of people around. So the implicatures derived from the domain mapping, this knowledge about B and other aspects, will lead to interpretations on the positive side, such as ‘Employees work very hard and are energetic and motivated’.

In addition to the directionality of interpretation that the relevance-theoretic comprehension heuristic provides (i.e. it is not arbitrary anymore which interpretations are chosen from all the possible ones), also the indeterminacy of meaning is catered for: in both cases, none of the interpretations that would be inappropriate on the specific occasion are discarded entirely; they are still weakly implicated due to the kind of interpretation that the idea of weak implicatures affords.
Thus, at first glance, my exercise of enhancing Steen’s 3D model with Relevance Theory’s flexibility and psychologically plausible computations resulted in an interesting outcome. However, I was trying to reinterpret the three-dimensional model from a processing perspective in order to be able to accommodate relevance-theoretic elements, even though the 3D model is mainly based on a symbolic approach that views cross-domain mappings from a structural perspective, that is, its purpose is to analyse semantic relations between lexical items regardless of use. The only exception to this is the distinction between deliberate and nondeliberate use in the communicative dimension. This means that I was not able to explain how, from a processing point of view, relevance-theoretic implicatures would arise ‘on top of’ domain mappings. There was no link to be construed that could have enabled this jump between the distinct frameworks, i.e. between the symbolic domain mappings and the relevance-theoretic comprehension heuristic, which focuses on language use.

This mismatch became much clearer when I extended the exercise towards Steen’s five-step method of metaphor identification. This model explicates in detail how the analyst can identify cross-domain mappings, and how metaphor analyses, which are intended to define the meaning of the metaphors in question, can be drawn from them. Initially, it seemed that an insertion of the notion of relevance-theoretic weak implicatures into this model seemed suitable. The reason for this was that the model seemed to offer a clear point of where to create the link to Relevance Theory, namely step five of the five-step method which, again, seemed to include a communicative aspect. The model will be described next.

5.1.3 The five-step method

Steen (2009) has developed a method for metaphor identification and analysis within a symbolic framework, which consists of five clearly defined steps. It is intended for longer texts with metaphorical elements, poetry in particular. Steen characterises the procedure proposed as ‘a form of text analysis which should be seen as relatively independent of the analysis of people’s text processing and its products’ (2009: 3), which means that it does not aim to model the computational
interplay of representations assumed in cognitive processing of poetry. Rather, the idea is for the method to provide tools for the analyst to identify the potential underlying domain mappings of metaphorical expressions, as assumed in Cognitive Linguistics. Steen explains that he ‘refer[s] to the abstract or symbolic structures which may be discerned by linguists and discourse analysts without necessarily having to make assumptions about their cognitive representation...’ (2009: 3), whereby it is not clear what the relationship could be between the mentioned structures, and actual cognitive representations and computations in discourse. But since the proposed method is not thought to be dealing with cognition, we would only be required to establish such a relationship if we were to claim that we simultaneously intended to look at behavioural aspects by using the same model. This issue is discussed further below, in the context of rehearsing a combination of the five-step method with elements taken from Relevance Theory.

To be sure, the latter is not investing any effort in explaining symbolic structures of language that do not play a role in communication. Thus, there are reasons to doubt that the endeavour of merging the approaches is feasible in this way, even though at first glance it seems an interesting approach. This will be discussed below. But first, I describe Steen’s five-step method. It aims to systematically show the connections between a metaphor’s linguistic form and the (assumed) corresponding conceptual structures as derived by the analyst.

According to Steen’s view, metaporphicity is a binary matter, i.e. a linguistic expression is seen as either metaphorical or nonmetaphorical. This is in contrast to the relevance-theoretic assumption that there is no clear cut-off point between metaphorical and nonmetaphorical uses and that there actually is no need for the notions of metaphor and metaporphicity, since all expressions are interpreted using essentially the same mechanisms (e.g. Wilson & Carston 2006; Sperber & Wilson 2006; see also section 2.1, page 32 ‘continuity account’).

Steen’s assumption, on the other hand, is that if the meaning of an expression is derived by a cross-domain mapping in the conceptual structure, we have a metaphor. The five-step method attempts to reveal (or reconstruct) how the linguistic form is related to these conceptual structures, and how these relations can be identified analytically. Steen argues that in Cognitive Linguistics, although
descriptions of the relations between domains have been central, so far it has not been attempted to show how these relations can be retraced systematically (2009: 4-5).

Steen (2009) exemplifies the method with a poem, ‘Now sleeps the crimson petal’ by Lord Alfred Tennyson. He begins with the first half of the first line, which is the same as the title. Step 1 involves identification of the linguistic expressions of metaphor in the sentence to be analysed, which is based on the assumption that in a longer stretch of metaphorical text there is always one or a small number of metaphorically used words in a sentence that render the sentence metaphorical. This is done by the Pragglejaz method (Pragglejaz Group 2007) which has been developed for the purpose of dealing with large corpus data. In principle, using this method, one defines lexical meaning to be metaphorical whenever there is a more ‘basic’ meaning of the word to be found in other contexts that can be compared with the metaphorically used version. The method is defined as follows:

1. Read the entire text–discourse to establish a general understanding of the meaning.
2. Determine the lexical units in the text–discourse
3. (a) For each lexical unit in the text, establish its meaning in context, that is, how it applies to an entity, relation, or attribute in the situation evoked by the text (contextual meaning). Take into account what comes before and after the lexical unit.
   (b) For each lexical unit, determine if it has a more basic contemporary meaning in other contexts than the one in the given context. For our purposes, basic meanings tend to be
   —More concrete [what they evoke is easier to imagine, see, hear, feel, smell, and taste];

---

18 It should be noted that, other than in pragmatic or discourse analytical frameworks, the analytical units we are dealing with I call ‘sentences’ and not ‘utterances’, since this is symbolic analysis, i.e. we are treating language as an entity outside use, as it were. For now I leave aside the question of whether this kind of treatment is not an artificial construct that cannot tell us anything about real life language, or even communication.
—Related to bodily action;
—More precise (as opposed to vague);
—Historically older;

Basic meanings are not necessarily the most frequent meanings of the lexical unit.

(c) If the lexical unit has a more basic current–contemporary meaning in other contexts than the given context, decide whether the contextual meaning contrasts with the basic meaning but can be understood in comparison with it.

4. If yes, mark the lexical unit as metaphorical.

Praglejaz Group 2007

Within the five-step analysis, an important characteristic of the metaphorically used word(s) is that it indicates the source domain of a metaphor (2009: 8). The analysis includes the identification of ‘metaphor-related words’ (not only metaphorically used words), i.e. words that may not be linguistically interpretable as metaphors, but that somehow express metaphor in thought. This means that also other linguistic forms, such as simile, are included. The theoretical basis for this is that in Cognitive Linguistics metaphor is primarily a conceptual matter, i.e. metaphor resides in thought and can be expressed in various ways linguistically. This also includes conceptual metaphor that is expressed by means of longer stretches of text, such as allegory, but I will leave a discussion of these cases aside for now and focus on smaller lexical units, for the sake of simplicity.

In the example, in the line ‘Now sleeps the crimson petal’, ‘sleeps’ is identified as the metaphorically used word, as it is the one from another domain, that of human or animal activity (rather than plants). The basic meaning of ‘sleeps’ is not an action that can be done by a plant, hence the word is being used in a ‘non-basic’, i.e. metaphorical way. Steen suggests a cross-domain mapping between plants and animate beings, with a plausible action corresponding to the basic meaning of the word, for a plant, as ‘being inactive’.

In step 2 the linguistic expressions identified as metaphorical are reformulated as propositions, which, according to Steen, are the same as conceptual structures. Again, this is clearly based on the idea that metaphors are
conceptual matters, rather than linguistic. Hence the propositionalization in step 2 is seen as making the metaphorical elements explicit, with a labelling of the concepts as either belonging to the source or to the target domain (2009: 8-9). In this example, only SLEEP is labelled as belonging to the source domain. In my view, labelling the concepts in question as parts of source and target domains comes already very close to spelling out a cross-domain mapping, or even consists of it, as the notions of ‘source’ and ‘target’ already imply a specific mapping relationship between the two. However, the method is intended to explicate in detail how source and target relate, that is, how this relationship may be constructed in a structural analysis. This is what the next steps are meant to do.

Step 3 serves to express the proposition established in step 2 as an open comparison between two propositions, i.e. it explicitly draws apart the source and target domain in order to juxtapose them. Thus there is the incomplete proposition of the crimson petal (the target) on the one hand, and the incomplete proposition of the action of sleeping (the source) on the other. The propositions are seen as incomplete, on this view, and the comparison as open, because the crimson petal is performing some action that, in this step of analysis, remains an undefined variable, and the action of sleeping is being performed by some agent, which also remains an undefined variable. This is because the domains and their contents are here expressed as separate entities. In step 4 the gaps created in the comparison of step 3 are filled in, and a fully fledged analogy is articulated, an ‘analogical structure’. Now the analyst needs to make decisions of how to fill in the gaps, for example the crimson petal as performing the action of being inactive. Another decision is required to determine the agent of the sleeping action, which in this case falls on a human being. ‘BE-INACTIVE’ and ‘HUMAN’ constitute additional conceptual content that is obviously not part of the original text, and Steen makes it clear that these are options that may vary depending on the analyst’s choice.

Note that making the analyst’s choice an integral part of analysis in this way, i.e. making a suggestion for a possible paraphrase, misses the point that metaphors carry more meaning than their paraphrase (or even more than the sum of their paraphrases). Although Steen does indeed suggest that the paraphrase chosen by the analyst is only one of several possible options, he is making this
point as an aside comment that is not part of the analysis itself. Thus, this analysis does not include the indeterminacy that I consider characteristic of metaphorical meaning. Indeterminacy of figurative meaning is, by comparison, an integral part of the relevance-theoretic analysis involving many weak implicatures that evoke poetic effects. The question is whether a structural, or symbolic, analysis would ever be able to capture the fact that an expression can suggest a wide array of meanings that are all potentially intended at the same time, and that a decision for one of them is not required, as the propositional meaning thereby created remains flexible. Of course, these are considerations that characterise analysis of pragmatic processing, i.e. behavioural analysis, rather than a structural ‘language outside use’ view.

Returning to the technical description of the five-step model, in step 4 the assumed correspondences between the domains that are being mapped are made explicit by filling in the propositional gaps. Step 5 completes the spelling out of the mapping. It is meant to make explicit the precise correspondences between the propositional elements juxtaposed in step 3 and analogically connected in step 4. Thus, the concept SLEEP is directly put in correspondence with the concept BE-INACTIVE, and HUMAN with CRIMSON PETAL. Furthermore, inferences that can be drawn from these correspondences are identified. For example, by projecting elements from source to target\(^{19}\), that is from the sleeping schema onto the crimson petal schema, may be inferred that the petal is tired, as the goal of sleeping as well as the goal of being inactive may be rest (2009: 10). Step 5 is the step that Steen admits to be the least clear in terms of how the inferences are arrived at, and which mechanisms to employ in analysis (2009: 22). This may be due to the fact that making inferences as such seems to have more to do with communicative behaviour rather than structural analysis, and explication of this seems to belong in a behavioural framework rather than a symbolic one. Furthermore, it remains unclear what the merits may be of explicating a derivation of meaning that is not intended to model cognitive processing in actual language use and follows different principles than those of communication.

\(^{19}\) Unfortunately, by which mechanisms the technique of ‘projecting’ is carried out remains unclear.
5.1.4 Combining the five-step method with Relevance Theory?

In a later seminar presentation of his five-step method (Steen 2010), Steen offers an outline of how it relates to his 3D model. Although the five-step method was originally created as a kind of manual providing the analyst with the tools for metaphor identification within a structuralist realm, it seems now to be reinterpreted as a more general model that claims to integrate symbolic as well as behavioural elements. The five steps of analysis are matched with the three dimensions of the 3D model, a relation that had not been spelled out before. Thus, step 1, the identification of metaphor-related words, is located in the dimension of language, and steps 2 to 4, covering the identification of propositions, open comparison, and analogical structure, are described as belonging to the dimension of thought. Step 5, which in the earlier version consisted of fleshing out the cross-domain mapping and then drawing inferences, has now been reallocated to cater for the identification of implicatures, here seen as including the fleshing out of cross-domain mappings. Clearly, this belongs to the dimension of communication, and Steen describes this dimension to be about how an expression functions between participants. Thus, the new version invites the accommodation of relevance-theoretic implicatures. However, apart from labelling them implicatures, Steen has not made any changes to the actual procedures proposed in step 5. That is, schemas are thought to be projected onto each other, and thus inferences are thought to come about. In spite of the relabeling of step 5 and the clear reference to the dimension of communication, I am again confronted with the question of how to find a basis for a context-sensitive communicative process, such as the derivation of weak implicatures, on top of a relatively fixed two-domain analytic construct that bears no reference to real-world cognitive processing.

This and the preceding sections have illustrated how it is not always possible to combine approaches. Nevertheless, this has proved to be an informative endeavour, since we can learn a lot from these examples of how some forms of theorising might be refuted due to constraints that clearly mark out boundaries when it comes to combining theoretical approaches that are all too
disparate in principle. These insights have informed the development of the new hybrid model developed in chapter 6.

The next section describes the hybrid model developed by Tendahl (2006), a combination of approaches that seem more compatible. It has been the starting point for my own model, which will be developed in the next chapter.

5.2 Tendahl’s Hybrid Theory of Metaphor
In this section I describe a proposal by Markus Tendahl (2006) in which he attempts to combine Cognitive Linguistics approaches with elements from Relevance Theory. It was the model case for the hybrid model developed here: while Tendahl combines Relevance Theory and Cognitive Linguistic approaches from a cognitive-linguistic point of view, the present thesis critically evaluates the approach from a relevance-theoretic perspective and accordingly sets up a new hybrid that aims to observe relevance-theoretic principles to a greater degree than the former.

However, I consider Tendahl’s merger of theories to be more successful than the one I described in the last section, because it aims to draw a picture that focuses mainly on cognitive processing, rather than trying to encompass a language-outside-use view as well as processing aspects. It does, however, attempt to unite different schools of thought at the cost of relevance-theoretic principles and in favour of cognitive-linguistic ideas. This, strictly speaking, only works because it seems to overlook some of the conflicts between the different underlying principles. On the other hand, it does not take on two different perspectives simultaneously, which the merging attempt of Steen’s five-step method with Relevance Theory does, as critically observed above. Therefore, it deserves some in-depth investigation, and the next chapter attempts to solve some of the conflicts that will become apparent in this chapter, by proposing the new model.

Tendahl (2006) generally has a cognitive-linguistics orientation. In particular, he supports the main ideas of Conceptual Metaphor Theory, e.g. that metaphors in thought are cross-domain mappings that with repeated use get entrenched in memory. With entrenchment they become ‘conceptual metaphors’
which are stable and enduring parts of cognitive networks in long-term knowledge (e.g. PEOPLE ARE PLANTS, LIFE IS A JOURNEY, etc.) (e.g. 2006: 251). With regard to Relevance Theory, he is all in favour of the cognitive mechanisms guided by the relevance-theoretic comprehension heuristic, while he argues against a mental architecture based on modularity, and the idea that cognition takes place in an amodal Language of Thought. For Tendahl, it is more than likely that cognition involves several different types of representation, such as image schemas alongside propositional representations, because he feels that the complexity of human experience and, in connection to that, the diversity and expressivity of metaphors, can only be captured in such a multi-modal fashion (2006: 184). He contends that such combinations of representational formats would allow for conceptual projections, as envisaged in Cognitive Linguistics with domain mappings, as well as cognitive processes based on inference as employed in Relevance Theory.

On his view, both Cognitive Linguistics and Relevance Theory appear to gain in a project of merging the two. Cognitive Linguistics, on the one hand, lacks explanations of online processing including constraining mechanisms for feature selection in interpretation. These are aspects the relevance-theoretic comprehension heuristic is most useful for. Relevance Theory, on the other hand, could gain from looking at how different metaphors seem to be connected in memory, which is the focus in Conceptual Metaphor Theory. In section 2.5.1 I have demonstrated the difficulties that can arise in relevance-theoretic analysis of poetry involving complex metaphorical interconnections.

Explaining the mechanisms of online processing, as focused on in Relevance Theory, so far does not seem to belong to the goals of cognitive-linguistic research, which aims to explore the connectedness between domain mappings in metaphorical thinking. Conceptual Metaphor Theory in particular is looking to systematise structures of conceptual metaphors and their hierarchies. Thus, the aim is to determine entrenched and even static entities in long-term memory, and how these can be discovered in language use. Therefore, pragmatic considerations are not so important here. There is just the assumption that a word expresses a concept. This concept is part of a conceptual schema and evokes a
conceptual metaphor, which facilitates understanding a metaphorical expression. But there is no step-by-step explanation of exactly how the conceptual metaphor comes into play, of what lies in between hearing a word and accessing the corresponding conceptual metaphor, or by which criteria a particular conceptual metaphor is selected (2006: 189-193). In Relevance Theory, by contrast, the main focus is on context-dependent utterance processing, whereby assumptions drawn from long-term memory play a role as one type of input to the inferential comprehension procedure. However, relevance theorists do not assume the existence of conceptual metaphors in the form of domain mappings. There may be entrenchments of linguistic metaphors due to frequent use (cf. Vega Moreno 2005, see section 3.3.4), but metaphorical thinking conceptual-metaphor-style is not part of the relevance-theoretic picture. This comes down to the fact that in Relevance Theory, conceptual representations are in the format of the Language of Thought, where inference needs to take input of facts that are true in the real world. The issue was discussed in some detail in sections 2.3 and 2.5.2. So Tendahl explains that merging Cognitive Linguistics with Relevance Theory, given that we allow for different types of representation and thus ignore specific relevance-theoretic principles, has the advantage that Relevance Theory can cater for processing aspects that Cognitive Linguistics does not explain.

Further, he argues that Relevance Theory does not provide an explanation of how metaphors are related to each other, as is one major focus in Cognitive Linguistics. In domain mappings, which are thought to be facilitated by conceptual metaphors, an important aspect is that not just single features are projected between source and target domain, but relational structures between those elements, which gives rise to inferences. That is, the topology of the source domain, for instance, equally holds for the target domain. Tendahl gives the example ‘Ruud is a tree’. The inference that this must be a very tall person is here explained as arising from the conceptual metaphor PEOPLE ARE PLANTS, and in the domain of plants, trees are very tall exemplars. Projecting this topological relation that holds in the world of plants onto people leads to the inference that Ruud is tall. This kind of cross-domain relational projection is based on a principle known as the ‘invariance hypothesis’, according to which the topology
of the source domain is carried over to the target domain (e.g. 2006: 154). The explanatory power that this and other relations that hold between metaphors provide, is not made use of in Relevance Theory, where pragmatic explanation of single instances of metaphorical expressions in specific communicative situations is the goal. Context-dependence plays a major role here, so naturally, this perspective does not pay too much attention to the idea of statically fixed networks in long-term memory. Again, Tendahl is convinced that one approach can benefit from an aspect which it does not have but which the other approach can provide.

He develops a model of metaphor understanding which aims to combine insights from Cognitive Linguistics such as domain mappings, with relevance-theoretic pragmatics, in particular the relevance-theoretic comprehension heuristic. He also incorporates the idea that comprehension proceeds via ad hoc concept construction. As a starting point, he takes inspiration from a speculative idea Carston (2002) discusses. Here, she is beginning to question the position that words encode atomic concepts which function as a basis for pragmatic processing. She considers that it might be possible that what a word encodes is rather more schematic, and that it might be the case that pragmatic processing in occasion-specific use starts off with an entirely underdetermined concept, rather than an atomic concept, which is much more specific. It might be the case that something very sketchy becomes a full concept, with specifically defined content, only through pragmatic processing. She rehearses this idea using the word ‘happy’ as an example. Like most words, it takes on many different occasion-specific meanings, and it is questionable whether there is a specific general concept that the word ‘happy’ provides access to. Its various senses, which Carston exemplifies as ‘a steady state of low-key well-being’, ‘a momentary experience of intense joy’, and ‘the sense of satisfaction that accompanies a successful negotiation’ etc., could well be established starting from a much less specific and more abstract concept which can accommodate all of the senses of the word ‘happy’, but which itself has no such level of specification (2002: 359-60). For Tendahl, the notion of such an underdetermined cognitive entity is a welcome addition to the collection of different kinds of knowledge structures he makes use
of in his ‘hybrid theory’, as it offers a basic structure, or an environment which can accommodate the interplay between the different knowledge formats that Tendahl proposes. He begins with Carston’s tentative and slightly vague definition of it:

Could it be that the word ‘happy’ does not encode a concept, but rather ‘points’ to a conceptual region, or maps to an address (or node, or gateway, or whatever) in memory? This pointing or mapping provides access to certain bundles of information from which the relevance-constrained processes of pragmatic inference extract or construct the conceptual unit which features in the speaker’s thought.

Carston 2002: 360-361

Tendahl adopts the term ‘conceptual region’ as a proper technical term to feature in his model, where it is understood as the representational basis, or starting point, from which the process of comprehension unfolds. In principle, this was the role that Carston (2002) speculatively intended for it, while she was not employing it in a concrete manner, i.e. as a technical term with a defined role in a model of comprehension. So Tendahl set out to creatively use this element for his model. While Carston does not specify what a conceptual region can be assumed to consist of or how it may be structured, Tendahl uses the notion as a first building block for his model.

At the onset of outlining his model, he defines the differences between several representational entities, or knowledge structures, and formats he employs. These come from the different theories he combines to form his hybrid theory, namely Relevance Theory, Conceptual Metaphor Theory, and Blending Theory. The latter two both come from Cognitive Linguistics and share a fair amount of assumptions. The formats he uses are ad hoc concepts from Relevance Theory, conceptual domains (source and target domains) from Conceptual Metaphor Theory, and mental spaces from Blending Theory. It is important to note that it is not the case that we have here the same phenomena which carry different labels in the respective theories, but due to the (partly) fundamental differences between
the mental architectures assumed, we are dealing with different conceptual formats altogether. Nevertheless, since the different approaches try to tackle the same, or at least similar, phenomena, that is, how we use and understand metaphors, the respective notions of the knowledge structures sometimes partially overlap, and Tendahl aims to make use of this. The way he does this will be explained in the following.

First of all, he explains that ad hoc concepts and conceptual domains are different types of representation in particular due to ad hoc concepts being temporary and highly context-dependent, as opposed to conceptual domains, which are fairly stable entities in long-term memory that are accessed rather than constructed. Mental spaces are similar to ad hoc concepts in that they are also temporary context-dependent constructions. However, while an ad hoc concept construction amounts to one single concept, mental spaces involve several concepts. With regard to frequent use, Tendahl explains that entrenchment ultimately leads to stable lexical concepts if their origin is in ad hoc concept construction, but amounts to conceptual domains if arising from mental spaces.

From the point of view of the current investigation, which focuses on metaphor at the lexical level, it seems that the input that comes from Blending Theory is not as important here and will remain in the background when it comes to illustrating and discussing the outline of the model in this chapter, even though Tendahl himself discusses it at some length. The reason for this is that the recruitment of Blending Theory mainly serves explaining metaphor at the sentential level, which is not the focus here. Here, the main priority is to draw attention to the (very few) weak points of this otherwise sophisticated model that successfully incorporates the strengths of the different schools of thought, and to prepare the ground for the development of my own model, which has more relevance-theoretic leanings than Tendahl’s cognitive-linguistics based model. This will be outlined in chapter 6. Due to the scope of this work I restrict myself to tackling these issues at the level of single metaphorically used words in context, to avoid unnecessary complication and to focus on conflicts related to underlying theoretical principles and their solutions. Hence Blending Theory, the other Cognitive Linguistics approach that Tendahl uses alongside Conceptual Metaphor
Theory, will only be mentioned in the overview (without being able to give the complexity of Tendahl’s model full justice).

As concerns the conceptual region, an important point is that it is the very first thing accessed upon hearing a word, so all processes start from there. Tendahl proposes to equate the conceptual region with lexical semantics, in the sense that it provides the content that is already present in memory prior to any contextually dependent modification or enrichment. However, importantly, the conceptual region does not contain nearly enough content to be understood as ‘the meaning’ of a word, as it is very schematic, as described above. It is thought to merely function to give access to the various representational formats thought to be involved in comprehension (2006: 236-7). This does, however, include central meaning elements that can be assumed to be stable across contexts, as Tendahl describes it, and he refers to this as the ‘encoded content of the lexical concept’ (2006: 238), which seems to be similar to what Rubio (2008) has called the ‘core features’ of a concept (section 4.2.2).

Tendahl emphasises that his model can incorporate all of the different representational notions. What makes this position tenable is that he does away with the relevance-theoretic principle that cognitive processing takes place in the medium of the Language of Thought, which takes as input only specific formats (cf. sections 2.3 and 2.5.2). From his point of view, the relevance-theoretic comprehension heuristic can be adopted without the logical format, and even without the mental architecture of modularity of mind, which plays an important role in Relevance Theory: ‘The resulting, slightly truncated, version of relevance theory would still be a very rewarding completion of cognitive linguistic approaches to communication and cognition. Such a combination should be possible for the simple reason that the defining assumptions of relevance theory are not logically dependent on the modularity of mind’ (2008: 229). I have explained my disagreement with this statement in section 2.4.2, where I pointed out the reasons why the relevance-theoretic comprehension heuristic needs to be based on a modular mental architecture, which allows speedy and constrained processing of vast amounts of information. Thus, Tendahl’s hybrid model has been constructed as an enhancement of cognitive-linguistic approaches. Whether a
similar construct can be envisaged from the opposite perspective will become clear in due course, i.e. enhancing Relevance Theory with elements from Cognitive Linguistics while adhering to relevance-theoretic principles. First, I give an overview of the model proposed by Tendahl.

A conceptual region is here thought to have a specific structure which accommodates ‘free slots’ which can be understood as gateways to different types of knowledge structure, and as tokens for the gaps in meaning that the schematic conceptual region leaves open to be fleshed out. Filling in these free slots proceeds through so-called ‘connectors’ to external knowledge structures. It remains to be defined what such connectors may look like, as it were, and how they may function exactly in terms of cognitive processing. The knowledge structures that thus become part of the occasion-specific meaning are different from those that are already part of the encoded content, that is, internal parts of the conceptual region. Tendahl calls the former ‘external knowledge structures’, as the content in question is not initially part of the conceptual region, but becomes part of the ad hoc concept in the process. This is facilitated by the free slots in the conceptual region.

Tendahl explains that a free slot is either completely ad hoc, or it is an ‘entrenched free slot’. He describes the role of entrenched free slots as follows: ‘Entrenched free slots are part of our long-term knowledge associated with a particular word, but they are not part of the lexical concept, because they are enriched via recourse to external knowledge structures.’ (2006: 238). Thus, one can conceive of entrenched free slots as comparable to the encyclopaedic entry in Relevance Theory, which is similar in the sense that it is not to be understood as inherently encoded by the concept, but rather that it is content associated with the concept, and the concept gives access to it (Sperber & Wilson 1986/95). However, Tendahl seems to understand Sperber and Wilson to mean that encyclopaedic knowledge is a fixed part of the encoded content of a word (2006: 238). For Sperber and Wilson, however, it seems that this is a slightly different matter in terms of how stable this type of knowledge can be expected to be. But I will leave the issue aside for the moment, as shortly it will become clear that Tendahl’s
notion of ‘entrenched free slots’ is still different from either definition of encyclopaedic knowledge.

So, if we accept Tendahl’s definition, knowledge structures that get incorporated into the meaning of a word through connectors via entrenched free slots are not part of the encyclopaedic knowledge but can be understood to have a more distant connection to the conceptual region. Such knowledge structures are then not associated with a concept, in the sense of the relevance-theoretic definition of encyclopaedic information, but it seems to be the case that this type of information is still further removed from the conceptual region, as it were, while there still is a connection on occasion. It is called external knowledge structure. When the connection to the conceptual region gets established more frequently, it becomes part of the encyclopaedic information, and the knowledge structure becomes internal. Tendahl contends that conceptual metaphors can be understood as one type of external knowledge structure. However, given that conceptual metaphors are thought to be frequently used in understanding, they should be considered to become internal knowledge structures over time. This amounts to a picture where conceptual metaphors become part of the encyclopaedic entry of a concept, in relevance-theoretic terms. Such a construct runs into problems with Relevance Theory, if adhering to the relevance-theoretic representational format (cf. section 2.5.2). However, since Tendahl makes it clear that he does not see a necessity to stick to the full relevance-theoretic picture which includes LoT compatible truth-evaluable formats, this is not an issue within his model. Furthermore, as Tendahl explains, it actually is not the case here that conceptual metaphors do become parts of the encyclopaedic entry, or in his terms, internal knowledge structures. This is because it depends entirely on the context of an utterance whether or how a domain mapping is used, i.e. which features from which domain get incorporated into the meaning. Thus, conceptual metaphors remain in the realm of external knowledge structures, even if their connection to the conceptual region gets entrenched over time. So they seem to be somewhere in between contextual assumptions (non-entrenched external knowledge structures) and knowledge that is properly associated with a concept (internal knowledge structures). In an earlier chapter of his thesis (2006: 201),
where he discusses the role of context in Cognitive Linguistics, Tendahl mentions that here conceptual metaphors are understood for people to be used as parts of the context. Taking this into account, in addition to the assumed entrenched connection to the conceptual region, results in this intermediate position of conceptual metaphors in this model.

Free slots of the non-entrenched kind get filled under entirely context-dependent conditions, so assumptions that are not derived from the conceptual region (i.e. not parts of internal knowledge structures), and not from associated (or entrenched) external knowledge structures either, get connected to the free slot ad hoc and from scratch. Tendahl describes it like this: ‘When, for example, the preceding discourse has made more manifest certain assumptions in the cognitive environment and thereby activated a conceptual domain that is usually not linked to the conceptual region of the word, then we can construct an ad hoc concept that contains a slot which receives structure from such an external domain’ (2006: 238/39). This raises the question whether it is the conceptual region that contains the free slots in advance, as Tendahl described it just above this quote, or whether it is single ad hoc concepts that contain the free slots which have been created in the process, or both. I interpret Tendahl’s diagram of a conceptual region (reproduced below, figure 5) to show that the filling of empty slots of the non-entrenched kind is a way of creating an ad hoc concept, i.e. the free slots have been there in order to facilitate the creation of an ad hoc concept by way of facilitating the connections to external knowledge structures. But it seems also to be the case that these types of free slot can be created ad hoc themselves. However, since this is not too important an issue with a view to the overall picture of the model, I’ll leave this question aside for now.
Tendahl calls the conceptual region a ‘blueprint’ for the construction of an ad hoc concept. The conceptual region provides the context-independent information, which is then contextually adapted to varying degrees. The processes employed are accommodated by entrenched and non-entrenched ‘free slots’, which are filled with external knowledge structures by ‘connectors’. By comparison, in Relevance Theory, these contextual adaptations are called ‘broadening’ and ‘narrowing’ which can do without a machinery of ‘free slots’ and ‘connectors’, but are simply constrained by the relevance-theoretic comprehension heuristic. For Tendahl, one reason for introducing this machinery may, possibly, be the difference between internal knowledge structures as parts of the conceptual region, on the one hand, and external knowledge structures that are connected to the conceptual region via entrenched free slots, on the other. In addition, there are external knowledge structures which get incorporated completely ad hoc. Thus, there is a distinction between two types of external knowledge structure, with a third type of knowledge structure being the internal. So we have a kind of three-way distinction between types of knowledge structure. By comparison, Rubio (2008) describes a two-way distinction between core (or default) and non-core encyclopaedic
features of a concept. She explains that this distinction is inspired by Barsalou’s (1982) distinction between context-independent and context-dependent properties. For the current purposes, it can be seen in the light of the relevance-theoretic idea that features are either strongly or weakly associated with a concept. For example, for Rubio (2008, referring to Barsalou 1982), ROUND is a core feature of the concept BASKETBALL, while FLOATS is a non-core feature of BASKETBALL, as it is thought only to become dominant in a biasing context, in this case possibly involving some sort of water mass, like a lake or a bath tub. It can be argued that non-core features are associated with a concept to varying degrees, with variation between individuals and times etc., but for the present discussion it is helpful to accept a simple two-way distinction to simplify the matter for the sake of exposition. The three-way distinction that Tendahl proposes has the advantage that it can accommodate conceptual metaphors as connected to a word’s meaning. Rather than assuming a connection to a conceptual region to a different degree of strength, compared to other types of knowledge, it is connected to a conceptual region in a different way. This way of seeing it could refute the issue of inferential fitness, explained in section 2.5.2, or a lack thereof in the case of conceptual metaphors. While inferential fitness is not an issue anyway from Tendahl’s point of view, as his model enjoys the luxury of freedom as concerns representational formats (2006: 237), it is important from a relevance-theoretic point of view. The reason for this is the importance of inference and its constraints in Relevance Theory.

Since most of the description so far has been concerned with the filling in of slots and adding external knowledge features, one might wonder how Tendahl envisages the process that in Relevance Theory is known as ‘narrowing’. What if parts of the encoded content get dropped in the comprehension process? Does Tendahl adopt the relevance-theoretic notion of narrowing, or does he have some other process in mind for feature demotion? He mentions that not all elements of the lexical concept necessarily enter the ad hoc concept, and that the choice of those that do enter the ad hoc concept is constrained by the usual relevance-theoretic comprehension heuristic (2006: 239). The conceptual region is thought of as very schematic, and the intended meaning of an expression is arrived at by
using all contextual assumptions as well as increased activation of features through entrenchment. So, lexically encoded features, though part of the conceptual region, are not accessed in the first place if there is no contextual pressure facilitating this, hence a feature demotion is not necessary. It would only be necessary in a model that assumes atomic concepts, where the encoded content is not as schematic and includes specific features that may not be part of the intended meaning on specific occasions. This is not the case in Tendahl’s picture. Here, whether a connector to a free slot gets activated depends on contextual pressure and relevance considerations, on the one hand, and on degree of accessibility through frequent prior use (entrenchment), on the other. But even if a connector is entrenched, there must be sufficient activation from the context, otherwise it will not activate. For free slots that are not entrenched through prior use, connectors to knowledge structures that have not been incorporated before in this way must be created ad hoc.

In the diagram of Tendahl’s model of a conceptual region above (figure 5), the shaded elements are those that get incorporated into the ad hoc concept, which does not include all elements that are part of the lexical concept, i.e. that potentially could have been accessed in the conceptual region on a different occasion. At present, it is not entirely clear whether this corresponds to the notion of narrowing after all, or whether it would be more in Tendahl’s sense to say that a conceptual region’s potential is never fully exploited in comprehension, and so there will always be some encoded elements that do not get picked out in the process. The diagram also shows that different kinds of representation and knowledge type are made use of, including image schemas, conceptual metaphors and metonymies. What is not shown in the diagram is whether a free slot is entrenched or not. A way of showing the difference between entrenched and non-entrenched slots in the diagram could be spatially. Entrenched empty slots could be pictured somewhat closer to the lexical concept (but not inside it, as these are still connections to external knowledge structures), and non-entrenched empty slots could be somewhat further removed. This, however, would not capture the binary distinction between entrenched and non-entrenched, it would rather appear as a relative difference in distance to the lexical concept.
Another question that remains open, in the diagram as well as the verbal description of the model, is what ‘connectors’ actually are and how they work. Do we need them in our terminology? It might turn out to be sufficient to talk of knowledge structures that get incorporated in the process, it might not be necessary to assign an agent such as a ‘connector’, to be doing the action in the process. Further, what is the significance of the term ‘free slot’? It might be a useful term in that it signifies the fact that a conceptual region is incomplete in terms of meaning and needs to be fleshed out. But since it appears to be the case that some free slots get filled in, others don’t, and yet others are being created ad hoc, the term does not seem to clarify the issue. Hence it might be sufficient to just talk of underdeterminacy of the conceptual region. However, in order to give my description of Tendahl’s model some substance by way of exercising it with an example, I will for now stick to his terminology. I use example (29) again, which has been discussed in section 5.1.2.

(29) Our office is a beehive.

Upon hearing the word ‘beehive’, the hearer automatically accesses the conceptual region of the lexeme. The conceptual region contains the encoded content of the word, which itself is rather sparse but gives access to information that can be assumed to be more or less stable across contexts. From my point of view, in Tendahl’s work it does not become sufficiently clear as to what degree the context-invariant meaning in the conceptual region is schematic or unspecified. He seems slightly contradictory concerning this specific point. For example, he states that ‘a conceptual region is usually much too unspecified in order to call it ‘the meaning’ of a word’ (2006: 237). But only a few pages later, he assumes rather clearly defined conceptual features that are stable across contexts: ‘for trees this probably includes the assumption that they are predominantly green, it might contain an image of the general form consisting of a trunk plus branches and, depending on the time of the year, also leaves, the assumption that the trunk consists of wood...’ (2006: 242). This sounds quite specified to me. So it might be the case that, at least at that point in time, he had
not made up his mind yet as to the degree of specification of the encoded concept. However, what surely does seem unambiguous is the fact that in his model, independently of degree of specification of internal knowledge structures, the larger part of the meaning outcome is constrained by the context and considerations of relevance, which determine the occasion-specific filling in of free slots with external knowledge structures.

For now, let’s assume that with ‘beehive’, we can employ examples of encoded information that is relatively specified, for the sake of exposition. This may be information like ‘accommodation/centre of living of bees’, ‘brown or beige structure with holes and tunnels bees move in and out of’, ‘occurs in nature’, ‘is constructed by bees’, etc. The conceptual region further contains free slots, the filling in of which contributes to the fleshing out of (parts of) the encoded content, and adds further content from external knowledge structures. One type of content that it is connected to is propositional and belongs to various domains the concept may occur in. These may differ due to individual variation pertaining to experience, cultural differences, etc. Tendahl refers to Langacker (1987) who calls this kind of knowledge structure accumulation ‘domain matrix’ (2006: 243). In the case of ‘beehive’, this may include connectors to domains like INSECTS, NATURAL STRUCTURES, HONEY PRODUCTION, etc. I assume that this information is likely to overlap with the encoded content, and that the boundaries between the two are not that rigid. This picture lends itself well to a comparison with the relevance-theoretic notion of encyclopaedic entry, which also may be more or less close to the encoded content of a concept, and with some of which it is hard to decide whether it should be assigned to the group of core or to that of non-core features. So, both knowledge types can be thought of as belonging to the type of content that traditionally would be understood as the ‘literal meaning’. In the case of example (29), however, the word ‘beehive’ is used metaphorically, which is obvious because of the context it is uttered in, which has nothing to do with real bees in nature. This discursive state of affairs activates the connector to a specific conceptual metaphor. The assumption in this model is that connections from a conceptual region to a conceptual metaphor are activated via a concept in the conceptual region that also features either in the source domain or in the target
domain of the conceptual metaphor. In metaphorical use, the role of the encoded content is to activate the connection to the conceptual metaphor, where the encoded content is part of the source domain. The context in which it is uttered, here in particular the mention of an office, further determines which conceptual metaphor should be the most relevant. That is, the immediate linguistic context of the metaphorically used word, here ‘beehive’, is an indicator of the target domain of the conceptual metaphor, which in this case is the workplace. Thus the connector to the conceptual metaphor *A WORK PLACE IS THE ANIMAL WORLD* gets activated. This, in addition to relevance-driven selection processes, provides the basis for the hearer to draw the inferences the speaker most likely intends, such as the ones I suggested in the discussion of this example in section 5.1.2, reproduced here:

\[(30)\]

- Beehives buzz with frantic activity
  \[\rightarrow\] employees work very hard and are energetic and motivated.
- Bees are all the same
  \[\rightarrow\] employees lack individuality.
- Bees never rest and mechanically carry on regardless
  \[\rightarrow\] employees are forced to do overtime without breaks.
- Beehives are crowded
  \[\rightarrow\] employees must make do with squeezing into tiny workstations.

... and so on.

This way, meaning features of ‘beehive’ that do not play any role in the interpretation of (29) do not even enter the process, even though connectors to them are part of the conceptual region. Thus, we do not assume, for example, that the office in question is brown or has lots of holes and tunnels.

I have now described Tendahl’s model, which is very insightful and can account for many aspects of metaphor understanding that none of the models involved can account for on its own. However, from a relevance-theoretic perspective, there are a few open questions with this approach as concerns, for example, the requirements for the relevance-theoretic comprehension procedure in
particular, and for inferential processing in general. These issues will be looked into in the next chapter, and in the course of solving them, a new hybrid model of metaphor understanding will be developed.
Chapter 6
The new hybrid model: A relevance-theoretic perspective

In this chapter, I redefine Tendahl’s (2006) model from a relevance-theoretic point of view, with the aim of retaining, or reintroducing, relevance-theoretic principles, such as a modular mental architecture. The reason for this reintroduction is not only that the relevance-theoretic comprehension mechanisms would not function as efficiently in a nonmodular mental architecture, as explained in section 2.4.2. It is also because it seems unlikely that different representational formats are all processed in the same way in cognition, which would be the case if they were all dealt with employing a general reasoning capacity. In that case, how would we able to tell the difference between, for example, logical facts and subjective intuitions? Section 6.1 explains the advantages of developing a hybrid model within a modular framework, and outlines how it is done in the present proposal.

It seems fruitful to incorporate nonpropositional representations into the model, as exemplified in Tendahl (2006), as it is a way of accounting for levels of comprehension beyond the propositional. This, intuitively, seems necessary in particular with regard to metaphor understanding and the sensory impressions it evokes. According to the relevance-theoretic account, metaphor is not a special type of expression, due to the view that all types of expression lie on a continuum between literal and figurative (cf. section 2.1). This, however, seems to go against the intuition that metaphors seem to be experienced in a special way, as explained in the introduction (p. 10-12). This phenomenon is addressed by Carston (2009; 2010), who proposes a dual processing route involving a special role of the literal meaning of metaphorical expressions at a metarepresentational level. Her idea represents one of the building blocks of the proposal outlined in this chapter and will be described in section 6.2. The other building blocks are Tendahl’s (2006) hybrid model of metaphor (section 5.2), and Barsalou’s proposal of mental simulation (section 3.3.3). A combination of relevance-theoretic and cognitive-linguistic ideas seems the right move in the desired direction of a comprehensive account of metaphor understanding, since it combines inferential logic with
intuitive impressions such that the latter are allocated an appropriately pronounced role, which can account for the ‘specialness’ of metaphor. Uniting such apparently opposing elements in a complementary fashion seems promising. However, thorough scrutiny of Tendahl’s model led to the identification of some mismatches with regard to relevance-theoretic principles. These, and suggestions of they might be addressed, are explained in section 6.3. Section 6.4 discusses how different representational formats might be best organised in an inclusive model, so that each format can be processed in its specific way. One of the formats introduced is that of conceptual metaphors (as defined in section 3.3.4), and section 6.5 is devoted to discussing how it might most usefully feature in the present proposal. One advantage of including conceptual metaphors is that they provide the grounds for explaining how it might be cognitively conceived of that some metaphors arise in language and others in thought. The issue will be elaborated on further at the end of this chapter. First, section 6.6 describes and explains the different processing mechanisms involved in this account, and how they are organised in a way that results in a fine-tuned interplay between three different levels of processing. Hence, this model can be called a triple processing account. Its strength is that it adheres to relevance-theoretic principles, which includes a characterisation of metaphorical expressions as equal in kind with other expressions, at the level of processing the proposition expressed. At the same time, the model accounts for the feeling that metaphor is ‘special’, thanks to the introduction of a) a dedicated nonpropositional level and b) the particular interplay of that level with the metarepresentational level. How this is supposed to work will be described and illustrated in detail in section 6.6.

6.1 A hybrid model within a modular mental architecture

In Relevance Theory, comprehension mechanisms are thought to be based on non-demonstrative inference processes, within a modular architecture of the mind. This ensures a specialisation at the propositional level, ascertaining accuracy and directionality of thinking, kept apart from purely associative processes. With associative processes that are not as constrained, e.g. by deductive rules, it is hard
to avoid an overgeneration of features, and true premises do not guarantee true conclusions (cf. sections 2.3. and 2.5.2). Hence in the model proposed here, different types of representation, which are processed in different ways, are kept apart from each other. In section 2.4.2, I explained how modules have their own dedicated procedures with their specific input, which can come from other modules, while each have their different specialisations. Thus, we can retain the relevance-theoretic picture of a communication module, working according to the relevance-theoretic comprehension procedure based on propositional representation. The conceptual features computed in this fashion are those that we find in the encyclopaedic entry; they are represented as true in the world. At the same time, it is proposed here to assume another kind of knowledge base in a different module, which specialises in imagistic impressions and experiential analogies such as conceptual metaphors. These representations are here proposed to be processed in a different way, involving mental simulations, as proposed by Barsalou (1999, 2005; see section 3.3.3).

Thus, the present account introduces a further module into this modular picture, which might prompt questions about what theoretical constraints there are on postulating new modules to deal with particular phenomena. Fodor has referred to approaches which postulate a large number of modules as ‘modularity gone mad’ (Fodor 1987), a description which Sperber (2001) quotes in defending his version of the ‘massive modularity’ hypothesis. In section 2.4.2 I have outlined the main reasons why, according to Sperber, it seems sensible to assume that evolutionary pressures have resulted in the development of a modular mental architecture comprising various highly specialised modules. In particular, modules allow for speed and efficiency in processing due to their automaticity, and they produce reliable output due to their specialisation. Thus, the fast and frugal heuristic, which is assumed in Relevance Theory in particular for language comprehension, balancing out effort and effects (cf. section 2.1), becomes psychologically feasible. General reasoning processes, by comparison, would, on a psychologically plausible picture, take into account far too many aspects in processing and would not be able to deal with the vast amounts of information humans process in communication in real time. These are the reasons why the
model proposed here assumes a modular mental architecture. However, there are also phenomena which are explained in Cognitive Linguistics but not in Relevance Theory, such as processes of mental simulation and conceptual metaphor. These phenomena are explained within nonmodular approaches, and an effort is made here to integrate them into a modular picture, in an attempt to strive for overall coherence. Of course, we would not want to rush to postulate the existence of a module if there was no direct evidence or other justification for this. The account proposed here provides both: while a modular account is to be preferred for the theoretical reasons just outlined, there is indeed strong empirical evidence in favour of assuming cognition to also involve simulation processes, some of which I have outlined in section 3.3.3 (e.g. Barsalou 2005). Therefore, it makes perfect sense to assume a module that specialises in simulation within a modular framework. Another way to account for the specific phenomenon of metaphor understanding and metaphoricity with its accompanying effects (outlined e.g. in the introduction above) could be, for example, cross-domain mappings as described in the Cognitive Linguistics literature. This would, however, not account for the specific cognitive mechanisms employed in communication, which are exhaustively and convincingly explained within the relevance-theoretic framework. Therefore, with the aim of developing a coherent explication of processing mechanisms, the overall framework assumed here is based on the principles of Relevance Theory. The constraints afforded by a requirement of overall theoretical coherence, as well as the arguments pertaining to the empirical evidence presented in section 3.3.3, are taken to justify the assumption of a module dedicated to mental simulation processes. The following discussion aims to explain how this works in more detail.

In the spirit of Relevance Theory, the assumption here is that comprehension requires a clear pragmatic explanation, based on rationality (in particular non-demonstrative inference and deductive rules). This is one of the reasons why, unlike Tendahl, I assume a modular architecture as proposed by Sperber (1994; 2000). This has the added advantage of allowing for a communication module specialising in inferential processing of propositional representations. At the same time, such a mental architecture can allow for a
module that specialises in mental simulation and thus include experiential processing and embodiment as envisaged in Cognitive Linguistics. Seen in this way, the latter processes as such do not interfere with the relevance-theoretic communication module. It is sensible to include experiential mental simulation processes in a model of metaphor interpretation, as they can account for this further level we experience with metaphor understanding. This, in particular, involves experiencing a ‘gap’ between the literal and the figurative meaning of a metaphorical expression, with a special weight given to the nonpropositional representations evoked by the literal meaning. The latter are not included as such in the proposition expressed. In Cognitive Linguistics, this process usually involves a domain mapping. It should be noted, however, that the approach outlined here is not intended to imply that the original relevance-theoretic account does not provide for the processing of sensory experiential impressions, including those that are evoked by language. On the contrary, on the relevance-theoretic view, there are perceptual modules which process external sights, sounds, smells, touches, etc. and internal modules that process sensations such as hunger or thirst, feelings such as sadness or happiness, etc. These will be sub-attentively monitored until they become relevant enough to be worth attentional processing. Further, imagery that is evoked by utterances is also seen to play a role in the relevance-theoretic picture of utterance understanding. So the claim is not that Relevance Theory lacks an account of these. However, the model proposed here is different from Relevance Theory in that it aims to describe and explain: a) the tension caused by the apparent ‘gap’ between literal and figurative meaning in metaphor understanding and the associated sensory experiences (which is a view more akin to Cognitive Linguistics than Relevance Theory) and b) the feelings and sensations involved in metaphor understanding that go beyond the computation of the propositional meaning outcome and the associated cognitive effects. That is, the present account aims to include in the description an account of the experience of metaphoricity (as defined in the introduction p. 10-12). This does not feature in Relevance Theory where the focus is primarily on propositional utterance meaning. Thus, the present account aims at a more holistic picture involving the description of a broader scope of communicative experience.
In the model developed here, the representations resulting from simulation processes mentioned above are not thought to function as direct input to comprehension processes, but rather are mediated at the metarepresentational level. Here they are either embedded into a propositional schema so that they can function as input to the communication module (in literal interpretation), or they are not embedded into a propositional schema, so that they are experienced as impressions that are not part of the proposition expressed in metaphorical interpretation, but rather feature as nonpropositional experiences (which are felt to be ‘special’ on this account).

This mental design ensures rational thinking based on reality in the world, while holistic simulations, such as imagistic impressions, can contribute to the overall meaning in their own way. These two channels are kept separate in the process by the domain specificity of the modules and metarepresentational monitoring. We are aware, for example, of the difference between facts represented as true and imagistic impressions. Thus, while imagistic-experiential representations featuring in (the literal meaning of) a metaphorical expression are not computed as part of the proposition expressed, i.e. are not themselves understood as part of the ‘factual meaning’, they serve to influence the proposition expressed only indirectly, through accompanying feelings and sensations. The processing of these representations is here proposed to occur at a different level, in a mental module that specialises in the processing of imagistic-experiential representations by way of simulation. This is closely related to the idea that the literal meaning of a metaphorical expression is metarepresented in utterance interpretation to aid understanding, involving a specific processing route. This is described in the next section.

6.2 Carston: The metarepresentational level
Carston (2010) draws attention to the fact that the main aim of the relevance-theoretic account so far has been to explain how propositional content is communicated, and that the imagistic quality of metaphors is something that does not play a constitutive role here and might require further explanation. This was
explained in section 4.2.1. That is, the meaning of metaphors frequently comes about through nonpropositional aspects, i.e. through images and other sensory experiences which cannot be captured in propositional terms. Therefore, the relevance-theoretic account does not quite seem to capture what intuitively makes metaphor ‘special’ compared to other forms of expression (2010: 9).

Carston paints a picture of two camps of theories, where theories that explain metaphor at a propositional level (such as Relevance Theory), are on one side and theories that view metaphor in terms of their imagistic properties (e.g. Davidson 1978; Black 1979; Camp 2008), on the other (2010: 4). She thus describes a dichotomy of proposition-based theories versus image-based theories. Note, however, that Carston focuses largely on the philosophy of language and refers mainly to theories with a philosophical outlook. Although approaches based in Cognitive Linguistics traditionally have a strong focus on explanations of metaphor, Carston ignores them in this picture. This might be due to the fact that, as Tendahl (2006) explains, relevance theorists and cognitive linguists tend to keep apart from each other, due to their principled theoretical disparities. He suggests that this lack of mutual acknowledgement may be disadvantageous for all parties, as they consequently are missing out on ways in which they could otherwise benefit from each other. Tendahl has made a start changing this state of affairs, and I follow his example. If approaches within Cognitive Linguistics were included in the dichotomy Carston suggests, they may be more suitably allocated to the group of image theories, as they work with image schemas. Domain mappings, which play a major role in cognitive-linguistic approaches, are image-driven, and more associative than inferential, which also better fits with image theories than with proposition theories. This might be taken to support the idea that cognitive-linguistic approaches have the potential to complement Relevance Theory, as they might be able to account for imagistic representations in a way that Relevance Theory, as a proposition-based theory, cannot. This idea will be taken up again in due course.

Carston tentatively sketches a new version of the relevance-theoretic account that is still primarily propositional but allows for imagistic representations to play a role in comprehension. Here, images evoked by the literal meaning of a
metaphorical expression might well have an influence on the proposition communicated. But it is important to note that the account remains propositional in principle, not least due to the fact that in Relevance Theory, images as such cannot be communicated. This is because it is a model of communication that is based on the notion of the Language of Thought which computes only propositional representations (cf. section 2.3.). Imagistic and other sensory representations belong to the sensory modalities and would need to be ‘repackaged’ in order to function as input to Language of Thought computations. This was discussed in section 2.5.2. Thus, only propositions can be communicated, which has to do with the current definition of ostensive communication where deductive inference and representations that are truth-evaluable play an important role. Carston explains that images are not truth-evaluable, i.e. they are not propositions. Only the content of images may be truth-evaluable, which is a different representational format, i.e. propositional (2010: 21). The nonpropositional character of images and other sensory representations might be usefully explained in relation to what Fabb (1995) calls ‘density’, in conjunction with its counterpart, ‘articulateness’. These terms have been coined by Goodman (1968), as Fabb explains, to characterise two different types of descriptive systems (or schemas): ‘Articulateness and density are complementary potential characteristics of schemas, which relate to the distinctness vs. continuity between the signs in the schema’ (1995: 145). Musical notation, for example, is an articulate system, as it consists of clear and definable units, each of which signifies one entity, i.e. one note stands for one tone, no more and no less. Images, on the other hand, are dense. They can be appreciated only as a whole, as they do not have constituent structure, as sentences do, so it is not clear where one element of an image ends and the other begins. An example which makes the difference instantly clear is the case of two different systems made up of numbers. One contains only whole numbers, so between 1 and 2 there is a gap. The advantage of this system is that it is always unambiguously clear which number is intended. This is called ‘finite differentiation’, because of the sharp boundaries between units. Such a system is labelled ‘articulate’. The other system of numbers in our example does not just contain whole numbers, but all possible numbers,
which entails that there can’t be any gaps between them, as there will always be a further differentiation: between 1 and 2 there is 1.1 and 1.2 etc., and between 1.1 and 1.2 there is 1.11 and 1.12, and so forth, ad infinitum. This is an example of a dense system. Articulate descriptive systems are effortlessly applied e.g. in describing figurative language, such as poetry, but they will always be restricted to the description of elements such as rhythmic notation or semantic content, as they cannot cover imagistic meanings, which are continuous and not definable in terms of distinctive units (1995: 143-45). These opposing terms are strongly related to what we call ‘analogical’ and ‘digital’, terms that are best illustrated by referring to the corresponding different types of clocks: the hands of an analogue clock, by way of analogy, reflect the passing of time in a dense fashion, as the way the hands move is continuous. By contrast, a digital clock shows the time in discrete units, with the seconds usually being the smallest (1995: 146). Because anything dense is inherently difficult to describe (which is where the everyday meaning of the word ‘articulate’ comes in – density cannot be articulated), literary critics, for example, usually use metaphorical expressions when it comes to describing sensory aspects of a text, for example by referring to the ‘movement of words in verse’ or by saying that a text ‘must be read with the body as well as with the eye’ (1995: 149). Correspondingly, Pugmire, discussing the nature of feelings, stresses that metaphorical language is the most suitable to describe feelings (1998: 102), and one might want to add that this presumably applies to any ‘dense’ phenomena. In the Language of Thought, the medium that in Relevance Theory is assumed to accommodate the processes applied in communication, conceptual representations consist of entities that are described as discrete units; hence LoT is an articulate system (1995: 151). This explains in part why there seems no way of incorporating the communication of images into the relevance-theoretic framework. However, intuitively, it seems that, in one way or another, metaphors do communicate images.\footnote{And so do, arguably, other types of figurative, and even literal, language. Due to the scope of this thesis, however, I focus on the imagistic evocativeness of metaphorical expressions.} For this reason it is proposed in the following how this might be accommodated in a model metaphor understanding, in particular. Carston (2010) makes a promising move in this direction, and her
approach plays an important role in the new hybrid model. It is therefore described further in the following.

She begins with the observation that with extended metaphors, such as metaphorical conceits or other metaphorical forms where the meaning is communicated through an accumulation of metaphorical expressions, it seems unlikely that for each small part of the whole an ad hoc concept construction process takes place. I have illustrated this with example analyses in section 2.5.1. Rather, according to Carston, it seems to be the case that the literal meaning of each constituent is held in working memory, where the entire construction is represented as a whole, i.e. it is interpreted literally. It is then scrutinised in front of the ‘inner eye’, as it were, ‘within the mental equivalent of scare-quotes’ (2010: 13), and the relevant implications are then inferentially derived from the literal scenario as a whole. Thus, she proposes an account of dual processing, where processes of automatic ad hoc concept construction take place for short lexical metaphors, and slower and reflective processes that involve the above described metarepresentations of the literal meaning, for longer connected stretches of metaphorical expressions (2010: 12-14). This allocation of a role for the literal in metaphorical interpretation can be taken a step further, given the notion of the ‘lingering of the literal’, as described in section 4.2.1. That is, it seems that the literal meaning is mentally held, or metarepresented, in interpretation of short metaphors also, at least with some types of metaphor (e.g. poetic ones) (2010: 16). This is not only to do with the difference between the literal and the metaphorical meaning and their processing, but in particular with the way in which metaphors are experienced: often, the intended meaning comes across by the way the images, feelings, or other sensory perceptions (or more precisely, memories thereof) are evoked by the metaphor vehicle. These nonpropositional, phenomenological experiences are difficult to account for in a theory that is dedicated to explaining the communication of propositions, such as Relevance Theory. However, it seems clear that Carston is now making a case for including mental imagery (and possibly other sensory experiences) in a description of comprehension processes:
Suppose it is right that mental imagery is triggered by (descriptive) words quite generally and requires no particularly effortful processing, no special attention beyond that required for accessing encoded concepts, then it is there for speakers to exploit, to highlight and induce hearers/readers to pay extra attention to, to scrutinize with their mind’s eye.

Carston 2010: 19

If this line of thinking is adopted, then mental imagery will be seen as playing a more significant role in relevance-theoretic accounts of utterance interpretation in general, and not only in understanding metaphor. Furthermore, the process of metarepresenting imagery of the literal meaning is now not only considered as a, rather exceptional, comprehension route for extended metaphors, but also for short and more conventional (though not lexicalised) ones, such as ‘my lawyer is a shark’. Thus, processing that involves mental imagery seems to take place with all, or perhaps nearly all, metaphorical expressions, and Carston explains that it might just be a matter of the degree to which we scrutinise the literal image in front of our ‘inner eye’ (2010: 20-21).

So Carston sketches a model of metaphor understanding that ascribes a prominent role to the metarepresentational capacity in cognition, a role that seems more specific, or more pronounced, than traditionally assumed in Relevance Theory. To recall, on Sperber’s (2000) model of modularity, there are three metarepresentational modules: one for attributing mental states in general, one for attributing speakers’ meanings, and one for considering logical and evidential relations among propositions in the abstract. Now, if we were to locate the function of metarepresenting the literal meaning for further inspection in figurative interpretation, as Carston describes it, in one or more of the metarepresentational modules that Sperber postulates, which one(s) would it be? Carston does not specify this, given that her discussion focuses on the potential processes involved, without aiming to implement them in a concrete model:
All I mean here by ‘metarepresenting’ the conceptual representation which comprises the literal meaning is that it is neither taken as descriptive/factual itself nor adjusted into another descriptive representation, but is held for a further process—of inspection, as it were, of its conceptual properties (implications, associations).

Carston 2010: 13, n.17

It is not clear whether she intends eventually to explicitly integrate these ideas into the relevance-theoretic model, or whether they are going to remain a temporary philosophical excursion. In the latter case, it would indeed not be necessary to describe them in relation to the relevance-theoretic modular architecture of the mind. However, since the model developed in this thesis is going to take Carston’s idea of metarepresenting the literal meaning in metaphorical interpretation as an integral part of accounting for metaphor, it ought to be required to implement this function in an overall modular picture (given that the model developed here is a modular one). Unfortunately, this requirement exceeds the scope of this thesis, such that for now we will need to make do with a more general notion of a ‘metarepresentational module’ or a ‘metarepresentational level’, in the sense that it has been described previously by Sperber (1994), and leave a more fine-grained specification of the module(s) carrying out this metarepresentational function for future research. Carston describes it as involving more reflective processes, where ‘the literal meaning of metaphorically used language is maintained, framed or metarepresented, and subjected to slower, more reflective interpretive inferences’ (2010: 13), so one way of envisaging it is that it is not carried out by a specific module, but rather involves an interplay of several modules, in the sense of semi-conscious reasoning (and in some cases fully conscious).

Coming back to the discussion of Carston (2010), conceiving of imagistic information as constituting a part of, or some other kind of involvement in, the propositional meaning outcome may pose a problem within Relevance Theory in its current version. According to the relevance-theoretic understanding of communication, we cannot communicate images or other sensory perceptions,
only propositions (which may represent the content of the former in a truth-evaluable fashion). Thus, any phenomenological experience, as opposed to pure facts, cannot be true or false in a logical sense, and can therefore not be the object of ostensive communication, on this definition. Carston, in explaining the issue, speculates on whether it may be in order to reformulate the definition of ostensive communication, so that it includes communication of the content of imagery, in cases where the latter is mentally intended. Carston emphasises that this does not mean that images are the same as concepts, and that they therefore cannot be communicated in the same way as concepts can. But they can be evoked through the use of language (2010: 21-22). It seems to follow that, even if it may be the case that imagery is, strictly speaking, not communicated in relevance-theoretic terms, its role in metaphor interpretation is so much foregrounded that it typically ends up determining the propositional content in most cases, and it seems that this is indeed a way of actually communicating it. I would think that imagistic and propositional representations often get extensively intertwined in communication, so why should we keep imagistic representations out of the content that is thought to get ostensively communicated? The relevance-theoretic argument against such a move is that inference needs propositional input so that logical computations can take place. However, given the strong intuition that imagistic representations are involved in the online processes of communication, it should be possible to have an account that includes them, while having truth-preserving inferential processes at the same time. Tendahl’s model outlined in the last chapter seems to be pointing in the right direction here, although I do not completely agree with his all-embracing account. Nevertheless, it seems promising to take his approach as a model case to be developed further. How this can be done will be outlined further below. First, I discuss some issues to do with Tendahl’s proceedings, in order to prepare the ground for developing a model with comparatively more relevance-theoretic leanings, which also includes Carston’s metarepresentational level of metaphor understanding.
6.3 Resolving issues with Tendahl’s model

In Cognitive Linguistics approaches, there often seem to be ingredients that remain vague and unspecific, such as the mechanisms by which we, apparently, get from a conceptual metaphor to an interpretation of a metaphorically used expression. Tendahl, aware of these shortcomings, does make remarkable progress at remedying them. However, he does this at the cost of valuable advantages that Relevance Theory has to offer, such as precision in balancing processing effort and effect. The reason for this is that this precision relies on the automaticity and speed of modular processes, which do not get distracted, as it were, by various intersecting other processes, as would be the case with general reasoning. But Tendahl prefers his approach to be nonmodular, as I have explained in section 2.4.2. Further, the accuracy of inference, i.e. the derivation of warranted conclusions, is compromised, as such processes are seen on a par with association and impressionistic intuitions. In Tendahl’s model, factual information, image schemas, and conceptual metaphors and metonymies are all seen as potential knowledge structures used to fill the free slots to construct the meaning outcome, with no apparent differentiation between them as concerns the type of representation. They are only differentiated by the kind of connection to the conceptual region, i.e. whether the knowledge structure is internal or external, and for external knowledge structures, whether the free slot is entrenched or not.

From a relevance-theoretic point of view, however, it makes a great difference whether a representational format used in comprehension is assumed to be propositional or not.

While Carston emphasises the difference between imagistic and propositional representations, Tendahl’s account seems to represent just the opposite view when it comes to considering different representational formats. According to his rather more holistic perspective, there are many different representational formats included in metaphor comprehension. This is an interesting approach, but it seems that Tendahl’s way of giving all formats equal status in processing does not seem all that useful when it comes to processing considerations. As we have seen above, on his picture of a conceptual region, the lexically encoded content consists of propositional as well as image schematic
knowledge, and the external knowledge structures that the empty slots can be filed with include all formats possible. Now, while it may be that different formats are indeed involved in processing, I doubt that they could all be used employing the same kind of processing mechanism. Tendahl remains unspecific when it comes to such a mechanism. He assumes relevance-theoretic ad hoc concept construction, in his version of ‘filling in free slots’, but how this ‘filling in’ actually proceeds remains unclear. Consequently, it can be argued that a model where the different representational formats have all the same processing status can only be upheld as long as the details of processing mechanisms remain unspecified. As soon as we specify the mechanisms, we run into problems regarding, for example, a lack of inferential fitness of nonpropositional representations, an issue that I have explained above. This is not a desirable situation, and Tendahl’s model therefore requires to undergo changes that make it fit for the processes that are involved in communication, at least from a relevance-theoretic point of view. In general, I agree with Relevance Theory that inference plays the dominant role in communication. In this way, my proposal is distinct from Tendahl’s in that I propose a hierarchy among the different types of knowledge structures that enter the comprehension process, where imagistic components at one level are seen as subservient to the propositional ones, which are at another level. How this is supposed to work will be outlined in the next section.

6.4 Representational formats
In spite of lacking the propositional requirements for inferential processing, according to computational principles, it seems important that imagistic and other experiential representations be included as meaning constituting knowledge structures, because of the obvious role they seem to play in comprehension, as explained above. They should, however, be seen as belonging to a different category, or type, than that of factual encyclopaedic information in the relevance-theoretic sense. Carston (2010: 19, n. 23; and elsewhere) suggests that encyclopaedic entries may include nonpropositional representations in some way,
and I would like to add to this that it may be beneficial to clearly define how their storage in memory is kept apart from that of propositional representations, as it is important that propositional content can still remain valid input for inference without interferences from the nonpropositional content. A mental architecture that makes this possible is modularity, since each module works in its domain-specific way, and exclusively takes its allocated type of input. Therefore, the model proposed here assumes a modular framework.

An important aspect that seems to be characteristic of imagistic-experiential representations is that they are experienced in a direct and ‘unfiltered’ fashion, as it were; rather than having analytic and objective quality, they are best described as sensory and subjective. Thus, we can conceive of them as a different type of representation altogether, which by implication is processed in a different way. Since imagistic-experiential information should be understood as also undergoing online processing, as does propositional information, it is sensible to assume two processing mechanisms, one for each type of information. This will be outlined in section 6.6. So we now have a picture of two types of information structure associated with a concept: propositional information, on the one hand, and imagistic-experiential information, on the other.

The approach proposed here is modelled after Tendahl’s (2006) incorporation of nonpropositional knowledge structures into the conceptual region (section 5.2; cf. Tendahl’s diagram figure 5 on page 175), albeit in a more ordered and hierarchical way. Here, nonpropositional\textsuperscript{21} formats are subservient to propositional formats (in computational terms). This is shown in the diagram below (figure 6), depicting the conceptual region which is associated with a concept. It shows both the propositional and the imagistic-experiential parts of potential utterance meaning in the conceptual region, and how the imagistic-experiential representations might inform propositions. I use the word ‘potential’, because it depends on the context and considerations of relevance which representational features eventually get incorporated into the meaning communicated. The diagram shows, in an approximate and simplified way, what

\textsuperscript{21} I prefer to positively call them imagistic-experiential formats, since they are more than just not propositional.
the connections between the two categories of knowledge structure are within a conceptual region.

![Diagram of conceptual region](image)

**Figure 6:** Different types of representation in the conceptual region.

The propositional representations are assumed to be dominant (hence the bold font), as it is facts in the world which can be evaluated as true or false that eventually make up the proposition expressed. In other words, comprehension is still assumed to take place within the medium of the Language of Thought, employing inference as in the original relevance-theoretic picture. Thus, propositional representations can be understood as the medium of rational thinking, and they have the last word, so to speak, in evaluating what is true or false, in decision making, and in communication within its dedicated module. The processing of the imagistic-experiential representations takes place in a different module, which informs the communication module indirectly. Hence the arrows depicting the influence of imagistic-experiential representations on the propositional ones are made up of dotted lines, to indicate that they are not taking over the derivation of communicated meaning. However, it should be noted that with metaphorical expressions, imagistic representations play quite a special role, involving a metarepresentational level as suggested by Carston (2010, section
6.2). This it will be defined in more detail, in pragmatic processing terms, in section 6.6.

For now, let’s clarify further what this mental architecture may be like, as depicted in a simplified way in figure 6 above. Note that, since in a modular framework, each module has its own dedicated processing mechanisms, and takes only those representations as input that belong to its allocated domain, the different knowledge structures are kept neatly apart. Further, the influence of the imagistic-experiential representations on the propositional representations is here thought to be ‘mediated’ at the metarepresentational level. By this I mean that nonpropositional material only finds its way into propositional representation if it is intended to feature as meaning-constitutive in the proposition expressed, and this differentiation happens via the metarepresentational level. In the diagram above, it is depicted with the blue circle between the two boxes containing the knowledge structures, called ‘metarepresentational mediator’, and the arrows that depict the influence on the propositional interpretation outcome have to go through it. In section 2.4.2, it was explained that on Sperber’s (1994) view of modularity, there is a metarepresentational module which has representations of representations as its domain. Thus, it represents propositional representations as well as imagistic-experiential representations as conceived of here (i.e. nonpropositional), and deals with the relations between the different representational formats. In section 2.5.2 it was explained how sub-propositional material is thought to be made ‘fit for inference’ by way of embedding it in a propositional schema (‘the speaker has said that [...]’). This is here conceived to work in a similar way for imagistic-experiential impressions to make them processable as computational input, as part of the propositional meaning outcome, mediated by the metarepresentational module (hence in the diagram it is called a ‘mediator’). This notion of integrating nonpropositional impressions has been mentioned before in the relevance-theoretic literature (Wilson & Carston 2006), where it was discussed whether, and in what way, nonpropositional representations might get incorporated as parts of encyclopaedic entries. Thus, using the example ‘That surgeon is a butcher’, it was suggested that, for example,

\[22\] Note that this is not meant to reflect actual brain structures, but rather just abstract relationships between representations.
with the embedding of kinaesthetic experiences, the relevant propositional representation might look as follows: ‘cuts meat like this xxx’, where xxx was thought to stand for the kinaesthetic representation (2006: 28, n.11). This process of embedding content in a schema for inferential fitness was also explained in section 2.5.2. It fits well with the account proposed here, whereby additionally, the present account offers a way of explaining how such a kinaesthetic experience would have been processed in the first place, before it got embedded into a schema to fit the propositional format. Namely, nonpropositional representations are here thought to be processed by way of mental simulation, as proposed by Barsalou (section 3.3.3), within a dedicated module. If these representations are required to feature in the interpretation outcome, i.e. contextually relevant, they are propositionally embedded to be computed in the language module. The latter functions with the Language of Thought taking amodal symbols only – hence the embedding (cf. sections 2.3 and 2.5.2).

Recall that in section 6.2, I described Carston’s (2010) proposal of metaphor interpretation as involving a process of metarepresenting the literal meaning. I explained that further research is needed to define exactly which kind of modular interplay at the representational level would do this job, so that for now, we must make do with describing these processes more generally as taking place at the metarepresentational level. Thus, we have two sub-processes of metaphor interpretation which take place at the metarepresentational level: a) embedding nonpropositional material in a propositional schema, if required for the context-dependent meaning, and b) holding the literal meaning in metaphorical interpretation for further inspection. In the course of describing the model developed here, I am going to claim the following: nonpropositional elements of the literal meaning (vivid imagistic impressions), which are not part of the proposition expressed in metaphorical interpretation but nevertheless feature as a defining (sensory) part of the metaphor, are those representations which are held at the metarepresentational level and not propositionalised. These nonpropositional representations (resulting from simulation processes) create an intense cognitive effect and are responsible for making metaphor feel ‘special’. This will be elaborated on further below (section 6.5).
In order to further illustrate the functions carried out at the metarepresentational level, let’s look again at example 29, which was last used in section 5.2:

(29) Our office is a beehive.

An intuitive impression evoked by this metaphoric expression, i.e. an imagistic-experiential representation, might be the buzzing noise that beehives make, and the feeling of how such a noise, if it were to be heard at an office, would get on one’s nerves. The impression would be evoked by a simulation process: by cognitively re-enacting the perception of a beehive, the memory of the noise it makes is activated. It is as if we heard the noise with our ‘inner ear’, and we re-enact the experience of how it feels. In literal interpretation, this experience is then turned into a proposition by way of embedding it in a truth-evaluable schema such as in example (31) below, where the constituent in front of the square brackets represents the propositional embedding, and the content inside the square brackets consists of the imagistic-experiential representation.

(31) “My impression is this [ANNOYING BUZZ NOISE].”

Thus, although imagistic-experiential representations may play a content-defining role in comprehension in some way, they are still mediated at the metarepresentational level prior to entering inferential processing, by embedding them in truth-evaluable propositions. With metaphorical expressions, the sensory perceptions (or memories thereof) evoked by the literal meaning of a metaphorical expression are not taken as part of the propositional meaning, and hence are not embedded in a propositional schema. Instead, they held at the metarepresentational level and just ‘felt’ or somehow perceived and experienced. In this way they contribute to the often intensely felt experience of the metaphor. This is here thought to occur at the metarepresentational level, which not only has the function of re-packaging nonpropositional representations to make them fit for inference, but also plays this important role particular to metaphor understanding. In the model proposed here, this idea is seen in connection to metarepresenting the
literal meaning in metaphorical interpretation as rehearsed by Carston (2010). This will be explained further in section 6.6. First, the discussion in the next section is aimed at clarifying the role of conceptual metaphors in this picture.

### 6.5 Conceptual metaphors in the new hybrid model

While there have been attempts at integrating conceptual metaphors as parts of comprehension procedures, it does not seem clear what their position in the overall picture is supposed to be (Tendahl 2006, Gibbs & Tendahl 2008). Gibbs and Tendahl mention that conceptual metaphors could ‘perhaps’ be assumed to be located in the encyclopaedic entry of a concept (2008: 14). But they do not discuss the idea any further. Tendahl (2006) does not appear to see conceptual metaphors as encyclopaedic information. He seems to propose that they are a knowledge format somewhere intermediate between encyclopaedic knowledge and contextual assumption, as I outlined in section 5.2. On this view, conceptual metaphors, through regular use, become a type of concept-external knowledge residing in long-term memory which is connected to a free slot in an entrenched fashion, but they do not become part of the conceptual region as internal knowledge. Tendahl does not explain why, on his picture, they do not become internal knowledge and thus part of the encyclopaedic entry. The position they currently have in his model seems to be rather vague: they are connected to the concept but somehow, at the same time, they are not. Perhaps we could say that they are part of the context, in a wider sense. However, they also seem too closely associated with the concept itself for this to be a satisfying definition. I doubt that it is helpful to have such an underdefined role of conceptual metaphors in comprehension, given that it is such a central notion in Cognitive Linguistics. Therefore it might be necessary to be more specific about its position in a cognitive model.

The position of conceptual metaphors in a hybrid model is problematic from a relevance-theoretic point of view, as it clashes with the idea of ‘inferential fitness’. If we were to use a conceptual metaphor as encyclopaedic input for non-demonstrative inference in comprehension, it would violate the rule that premises
for inference need to be represented as true in the real world to yield a true outcome. Since a conceptual metaphor, like PEOPLE ARE MACHINES, is obviously not factually true, the system would presumably still take it as such and the interpretation would result in a misunderstanding, because the metaphor in question would have been taken literally. Clearly, this is not how metaphor understanding works. Therefore, it seems a suitable alternative to introduce conceptual metaphors into a model where types of knowledge structures are clearly distinguishable, so that it is clear which type they belong to and, by extension, how they are processed. Conceptual metaphors are not propositional representations, as should have become clear in my definition in section 3.3.4, but rather, they are of an experiential nature and have arisen in repeated situated conceptualisations. Thus they should be understood as part of the imagistic-experiential group of representations. So, according to the picture painted thus far, conceptual metaphors do not constitute propositional content, but they are allocated a role of merely informing the propositional content in question. This ensures that inference in the spirit of Relevance Theory can take place, while conceptual metaphors can still be involved in the comprehension of metaphors whose origin is in thought, i.e. expressions which may be based on a conceptual metaphor. It might even be the case that also comprehension of metaphorical expressions which have their origin in language is aided by accessing conceptual metaphors, but trying to answer this question would exceed the scope of this project. Note that an advantage of Tendahl’s model is that it does, however, incorporate conceptual metaphors in a model of metaphor understanding, albeit in a seemingly underdefined way. Intuitively, conceptual metaphors do seem to play a role both in cognition in general, and in language use in particular, as discussed in sections 3.3 and 4.1.4.

Another argument that might also go in favour of involving conceptual metaphors in a relevance-theoretic approach to comprehension is the following: the metarepresentations involved in metaphor understanding which Carston (2010) describes, i.e. the scrutinising of the literal meaning in front of the ‘inner eye’, seem comparable to cross-domain mappings as featuring in Cognitive Linguistics. They also involve a juxtaposition between the literal and the
figurative meaning, and we can equally describe this, using the cognitive-linguistic terminology, as mentally holding the source domain in order to define the target domain. Conceptual metaphors, which arise as a result of frequently entertaining similar or related cross-domain mappings (or, to use a more processing-specific terminology, analogical re-enactments), are thought to be stored in long-term memory and reactivated in the course of metaphor understanding. If a conceptual metaphor were to be processed as a different type of information, such that it functioned to merely inform the propositional content, rather than be taken as propositional content itself, it could be treated as nonpropositional information that is processed by simulation, much like I have sketched above, pertaining to imagistic-experiential content. Thus, we can extend the imagistic-experiential group of representations in this model to include conceptual metaphors. Importantly, the kind of conceptual metaphor I have in mind is very schematic and underdetermined, as I have described in section 3.3.4. Conceptual metaphors are thought to arise from experiential simulations, which in Cognitive Linguists is called ‘embodiment’. This type of experience is best explained with Barsalou’s (2005) situated conceptualisations (section 3.3.3), an approach which offers a psychologically plausible version of conceptual metaphors.

6.6 Processing routes
If we go along with a definition of conceptual metaphors akin to Barsalou’s simulation patterns, it is adequate to also incorporate his view of processing and adopt it for our imagistic-experiential level of representation. Thus, we can conceive of two processing routes that the two different types of representation undergo. The propositional representations undergo inferential processing, and the imagistic-experiential representations undergo processes of mental simulation (or situated conceptualisation). The latter consist in accessing schematic patterns which have arisen from past bodily experiences with the world (cf. sections 3.3.3 and 3.3.4). Each type of representation, along with their specific processing routes, are accommodated in their own module, respectively, ensuring that they do
not interfere with each other, while some informational interplay is mediated by
the metarepresentational level, as described in section 6.4. Importantly, the
inferential route is still understood here as playing the dominant role in
comprehension, while the simulation processes, running in parallel to it in
comprehension, inform it by feeding it with imagistic-experiential impressions via
the metarepresentational module which embeds these impressions in a
propositional format, if contextually required, as exemplified in section 6.4.

Conceptual metaphors, as defined here, are underdetermined and
schematic, and thus, just as the meaning of linguistic expressions is thought to be
underdetermined and then fleshed out conceptually in comprehension through
pragmatic processes in context, conceptual metaphors are fleshed out in the
parallel processing route of holistic situated simulation, as proposed by Barsalou.
So, in terms of running the course of processing, completion of both processing
routes can be envisaged to neatly run in parallel.

It should be noted that Barsalou’s model presupposes a nonmodular
mental architecture, as the processes of situated conceptualisation are thought to
be implemented in a holistic fashion of general reasoning. However, the processes
he proposes can equally be seen to run within a modular model, where the
simulation processes are carried out by a module that gets its input from modules
that process sensory perception. Thus, there is nothing that stands in the way of
integrating these processes into a modular framework (cf. section 6.1). Figure 7
shows, in a simplified fashion, how this parallel processing might go for utterance
comprehension in general.
Figure 7: Literal interpretation.

The diagram shows inferential processes, indicated by the long thick arrow. These include the relevance-theoretic comprehension procedure. Based on logical rules, they are the dominant ones that hold comprehension together due to being rooted in rationality. The long arrow below indicates the simulation processes which imagistic representations, including conceptual metaphors, undergo. It is a dotted line to indicate that these processes, involving the imagistic-experiential representations, usually have a weaker and only indirect influence on the interpretation process than the processes involving the propositional representations. The two processes run in parallel, which is to be understood in accordance with the mutual adjustment of explicatures and implicatures in the process, i.e. the more an ad hoc concept is nearing its completion, at the inferential level, the more specified the re-enacted pattern becomes, at the simulation level. This will be shown with an example further below. The comprehension process stops once optimal relevance is achieved and the hearer has finished interpreting the utterance, which is indicated here with a horizontal line labelled ‘communicated meaning’. The small dotted arrows show how the simulation processes influence the inferential processes by way of manipulating contextual assumptions (to a mild degree).

As we have seen in the diagram above (figure 6), the way in which imagistic-experiential representations inform propositions is mediated at the metarepresentational level, which by extension also applies to the processing
mechanisms involved in comprehension. This model of dual processing is intended to show how comprehension might go with more or less literally used expressions, i.e. with utterance meanings that are relatively close to the linguistically encoded content. So the imagistic-experiential representations are evoked by uttering expressions in a discourse context that suggests denotations at least closely related to the literal meaning of the expression, in other words, the core features of the encoded concept. Here, the simulation processes that inform the propositional meaning by getting embedded into propositional schemas can be understood to have the function of accompanying the overall meaning outcome. The function of the metarepresentational module, in the diagram in figure 6 depicted by a blue circle, is left out in figure 7 merely for ease of exposition. However, this depiction is not entirely correct, since it makes it look as if the inferential level would receive direct input from the simulation level, which is here thought not to be the case. As explained in section 6.4, the metarepresentational level has the function to mediate the information flow from the imagistic-experiential representations to the propositional representations by way of embedding the former in propositional schemas so that they can function as input to the inferential module. Thus, in addition to the two processing levels of inference and simulation, we have a third level of processing: the mediating function at the metarepresentational level. So what we have here is triple processing, rather than dual processing. In order to provide a more complete picture, the diagram in figure 8 includes the mediation processes at the metarepresentational level.
Figure 8: Literal interpretation revisited.

With the interpretation of literal utterances, i.e. utterances that are not accompanied by a perceived metaphoricity (in processing terms\textsuperscript{23}), the metarepresentational module has merely the function in comprehension to make imagistic-experiential representations, which are processed at the lower level, fit for inference by embedding them in propositional schemas. The long blue dotted arrow in the middle depicts the metarepresentational processing level. The input from the simulation level is here shown correctly not to impact on the inferential level directly, but rather its output reaches the metarepresentational level first, which is depicted with the small black arrows pointing upwards. After being processed at the metarepresentational level (i.e. made fit for inference), the information is ready to enter the inferential level at the top, which is here shown with the small blue dotted arrows going from the metarepresentational level upwards to the inferential level.

By comparison, imagistic-experiential representations play a much more salient role in the comprehension of metaphorical expressions, that is, expressions whose metaphoricity we consciously perceive (cf. section 4.2), or, to use Steen’s terminology, deliberately used metaphors (section 5.1.1). As Carston (2010) suggests, comprehension of this type of metaphorical expression involves

\textsuperscript{23} i.e. where the perceived metaphorical ‘gap’ between the literal and the figurative meaning plays a role; this excludes metaphorically motivated polysemy, which is not thought to be processed metaphorically; cf. section 4.2
perceiving the literal meaning in front of the ‘inner eye’, and scrutinising it there, so to speak. By extension, it is here proposed that this involves processing it in a metaphor-specific way at the metarepresentational level. The literal meaning of the expression as such does not enter the inferential level, since it is not computed as the communicated propositional meaning. Or, more precisely, at the inferential level, it merely features as a starting point for ad hoc concept construction (explained in section 2.2). However, the literal meaning still appears in processing, and is consciously perceivable, which occurs at the metarepresentational level. This is because, just as in literal interpretation, the simulation module sends its imagistic-inferential representations to the metarepresentational level for further processing. But, other than in literal interpretation, the metarepresentational level has, in the case of a metaphorical expression, not the function of preparing the imagistic-experiential representations for entering the inferential processes (cf. the blue circle in figure 6), since in metaphorical interpretation they do not feature in the proposition communicated. Instead, they feature as imagistic impressions that the metaphor imparts, and the metarepresentational level has the function of representing them, just as they are: nonpropositional. This leads to the effect of metaphor as a ‘special experience’, since these imagistic impressions, which are not propositionalised, and thus not ‘rationalised’, as it were, are experienced in an intense way. This conception of the effects that such nonpropositional representations may have at the metarepresentational level can be understood as an elaboration on what Carston has called the ‘lingering of the literal’, as described in section 4.2.1. Figure 9 shows, in a simplified and abstract way, how this might go.
Here, the imagistic-experiential representations that are evoked upon hearing the expression take the same route as they do in literal interpretation: they reach the metarepresentational level, here again depicted with the long blue dotted line. However, the discourse context makes it clear that literal interpretation would not achieve relevance (i.e. it would be wrong), and therefore the imagistic-experiential representations of the literal meaning have to be prevented from entering the inferential level in the form of propositions. Thus, they are being held at the metarepresentational level, which is shown in the diagram with the large shaded dots at the ends of the black dotted arrows coming from the simulation level. Further processing, as would happen with literal interpretation, is blocked. This is to prevent the imagistic-experiential representations from getting embedded in a propositional schema and thus taken to be part of the proposition expressed. If that happened, the interpretation would lead to a misunderstanding, because a metaphorical utterance would have been taken literally. Thus, in order to prevent such misinterpretation, the output from the simulation processes (the lowest level in the diagram in figure 9) is not turned into a propositional format. Since it is not processed any further (the ‘lingering of the literal’), it is perceived in an ‘unfiltered’ way and experienced in an intense experiential fashion. This explains why the comprehension of metaphors with a high degree of metaphoricity (poetic metaphors in particular), give us the feeling of a special, often strongly felt
experience. To illustrate how this triple processing might go, let’s take another look at example (29).

(29) Our office is a beehive.

On the inferential processing route, ad hoc concept construction takes place along the following lines: of the many encyclopaedic features associated with the concept BEEHIVE, the feature IS A HOME FOR BEES is dropped in the broadening process, so that the denotation of the word also includes other structures such as buildings, which could be inhabited by humans. Thus we arrive at the concept BEEHIVE*. Further, a narrowing process is required in order to exclude features of beehives which are not intended in this context, for example that beehives can be very quiet and cosy places in corners where the eggs are kept until the queen’s offspring hatches. Thus, the feature IS QUIET AND COSY is dropped, and we arrive at the concept BEEHIVE**, which can capture the situation of a busy office that is filled with hard working people. While there is the strong implicature that the office is very busy, the metaphorical expression also weakly communicates a wide array of weak implicatures to do with lack of individuality, mechanical monotony perhaps, and other stereotypical features of beehives that fit the context here.

While this analysis tells us how the propositional outcome of the metaphorical interpretation might be arrived at, it does not really account for the imagistic impression of a beehive which we see in front of our ‘inner eye’, and the sensory experience triggered by imagining the buzzing noise that the bees make, and how this imagined experience makes us feel. These are the processes of mental simulation that go on in parallel at the level of imagistic-experiential representations. At that level, we may also represent a conceptual metaphor, which at first might just be a schematic pattern, but in the course of processing get fleshed out to become a representation such as this one: WORK PLACE IS ANIMAL WORLD. This representation might come about by a juxtaposition we mentally carry out, of images of an actual beehive, as the source domain, and an actual office, as the target domain, resulting in a domain mapping. These are just imagistic impressions, which as such do not influence the propositional content
communicated. They do, however, have a high impact at the metarepresentational level, where they are being held for inner inspection, as it were, instead of being repackaged by way of embedding them in a propositional schema. The latter would only be the case in literal, or near-literal interpretation. Thus, the sensory impressions evoked by imagining a beehive, and the domain mapping which is also felt experientially, make understanding this metaphorical expression in discourse a more intense overall experience. Arguably, this might in turn indirectly influence the propositional representations to a small degree, by way of reinforcement of the literal meaning, which might draw attention to specific weak implicatures. For example, the sensory impression of the imagined buzzing noise may draw attention to the weak implicature that the busy atmosphere at the office gets on one’s nerves, since a constant buzzing noise has this very effect. However, we remain aware that the noise is not real, since at the metarepresentational level this representation is processed as a simulation, and not as a representation of a true fact (the latter being a propositional representation). Thus, all three processing routes inform and reinforce each other, where at the metarepresentational level it is ensured that the inferential route remains the one that guarantees the factually correct outcome, adhering to considerations of optimal relevance, while at the same time we can enjoy the expressiveness that figurative language has to offer, via representing the imagistic-experiential representations at the metarepresentational level. Most notably, this framework accommodates the way in which the impact of the imagistic-experiential representations (or impressions) gets intensified as metarepresentational processes prevent the propositional embedding of the former - their mere lingering allows them to unfold their intensity to the fullest.

Tendahl calls his approach ‘The Hybrid Theory of Metaphor’ (2006), as it incorporates elements from both Cognitive Linguistics and Relevance Theory. The model proposed here, by comparison, can be understood to be even more inclusive: not only does it incorporate representational elements from both schools of thought, it also offers hybrid processing proper, i.e. relevance-theoretic inference in parallel to cognitive-linguistic simulation processes. Tendahl’s model, by contrast, seems to presuppose that the relevance-theoretic
comprehension procedure can deal with all types of representation; and it is not clear how he envisages the former to be implemented in his holistic amodular model. My model is hybrid all the way through, organised along the two different types of representation and the respective type-specific processing mechanisms. This is made possible by assuming a mental architecture of massive modularity, which has the advantage of a great flexibility, allowing for different modes of representation, on the one hand, and of an organisation in a modular fashion, allowing the different modes to be kept apart, on the other hand. This model has the advantage that it provides the logical precision of Relevance Theory, catering for psychologically plausible communication processes, while at the same time benefiting from cognitive-linguistic insights concerning a holistic understanding of the world and how things relate to each other. Unlike approaches in Cognitive Linguistics, including Tendahl’s, it does not assume communication to rest on general reasoning, and thus, by employing a massively modular view of cognition, makes sure that holistic impressions based on embodiment, such as conceptual metaphors, are neatly kept apart from rational processes based on propositions. Still, the former retain their deserved place in the explanation of human behaviour and reasoning, which is clearly not only based on propositional representation but is much more inclusive of all kinds of experience that life throws at us.

Furthermore, when it comes to metaphor understanding, we have here a special case of triple processing: at the most basic level, experiential simulation processes take place, which enter the metarepresentational level, and the imagistic aspects of the literal content are being held there to provide us with an intensely felt experience. By comparison, interpreting literal expressions involves further processing of these imagistic aspects as they eventually get incorporated in inferential processing. Thus, they are not ‘lingering on’ at the metarepresentational level. This ‘lingering’, on the other hand, is the phenomenon that accounts for the gap we feel between the vehicle and the tenor of a metaphor (or the source and the target). Due to the mechanics involved, it does not interfere with processing the proposition expressed, at the inferential communication level.

In addition, this model can also account for the idea that some conventionalised metaphors originate in language and others originate in thought.
While we do not yet have empirical evidence of this difference in origin, it seems plausible and worth exploring in the future. Thus, the comprehension module, which works with propositional representations, takes advantage of the high accessibility which conventional linguistic metaphors have due to their frequent use. Using them again reduces the effort invested for a maximum of effects, which is achieved by the expressive power that metaphors have. Thus, the presumption of optimal relevance generates use and re-use of metaphors, which become entrenched in the memory of people from the same speech community. An example of this might be the expression in example (32).

(32) He is Marmite.

British readers might be familiar with the advertising campaign that presented this spread as something ‘you either love or hate’. Thus, the brand name took on the linguistic role as shorthand to describe a person that is perceived only in extreme ways (with either love or hate), as it is quicker and easier to use just one word than describing this state of affairs. If such an expression catches on and is used widely in a speech community, the metaphor can be understood to have its origin in language use. Conceptual metaphors, on the other hand, might give rise to linguistic expressions which seem to have their origin in thought. For example, an expression like (33) might have its origin in a repeatedly entertained conceptual metaphor such as THOUGHTS ARE FOOD.

(33) ‘I first have to digest your proposal.’

This can be explained by the reoccurring patterns of experience, which get accessed by way of holistic, modality-specific re-enactments, or simulations. Thus, the process of thinking about something is repeatedly experienced in analogy with eating and digestion. This and other ways of analogical thinking can be understood as basic tools for cognitive orientation by way of categorising events and objects. This in turn might get reflected in language and other ways of
expression which have been observed, for example, in gesture studies (section 4.1.4).
Chapter 7
Conclusion

This thesis develops a hybrid model of metaphor that incorporates elements from both Relevance Theory (Sperber & Wilson; Carston) and Cognitive Linguistics, in particular Conceptual Metaphor Theory (Lakoff & Johnson) and Situated Conceptualization (Barsalou). While it has its starting point in the hybrid model proposed by Tendahl, it differs from Tendahl’s proposal in taking a relevance-theoretic perspective rather than a cognitive-linguistic one. It assumes, for example, predominantly inferential processing in understanding communication. One advantage of preserving relevance-theoretic principles is that this model can explain accuracy in rational thinking, while approaches based on purely associative processes cannot guarantee that only true conclusions follow from true premises (or, more generally, truth-preserving processes). Another advantage is that the assumption of a massively modular architecture, with the speed and automaticity of a modular processing device, explains the processing of the vast amounts of information that human communication deals with. The approach developed here also offers an account of the nonpropositional effects that come about in metaphor understanding in particular. This is in part made possible by adopting the idea proposed by Carston, that during metaphor interpretation the literal meaning is entertained at a metarepresentational level, which can account for the imagistic quality of metaphors. A comprehensive account of metaphor understanding, and by extension of utterance understanding in general, should include analogical thinking arising from experiential simulation, as argued in Cognitive Linguistics. This reflects the intuition that propositional inference seems to be informed by this more basic level of processing. Thus, the new hybrid model proposed here incorporates an imagistic-experiential level of processing, which is seen as subservient to the propositional level of processing. Since it proposes a modular framework of cognition, the different types of processes can run in parallel, each within its specialised module, while there are still routes of interaction between them. The input to the propositional module has to be propositional, as this is the domain it specialises in. In understanding literal
expressions, there is a mechanism at the metarepresentational level which processes nonpropositional input and makes it fit for inference in the propositional module. This mechanism takes on a special role when it comes to metaphorical expressions: since the literal meaning is not processed as part of the interpretation outcome, its imagistic-experiential elements are not ‘re-packaged’ for inferential fitness. Thus, they just remain at the metarepresentational level. This distinct processing for metaphor explains how the ‘lingering of the literal’ in metaphor understanding evokes the feeling that metaphor is special. Note that this account preserves the relevance-theoretic continuity account, which states that metaphor is not a special kind as compared to other types of expression. This is because the inferential processes involved in metaphor understanding are not directly influenced by the specific interaction between the imagistic-experiential representations and the metarepresentational module. At the inferential level, ad hoc concept construction proceeds as originally proposed in Relevance Theory, the figurativeness of an expression being a matter of degree in terms of situation-specific broadening and narrowing processes. Thus, the interplay between the three levels of processing (simulation, inference, metarepresentation) results in a propositional communicated meaning that is accompanied by nonpropositional effects. Thus, this model explains the specific pragmatic processes that lead to the ‘special effects’ experienced in metaphor understanding.

The introduction of an imagistic-experiential level within a relevance-theoretic account of metaphor further allows for conceptual metaphors to play a role in understanding. Since imagistic-experiential representations do not interfere directly with inferential representations, due to the modular specialisation of each, conceptual metaphors can be seen as an impressionistic experience based on mental simulation processes. The advantage of including conceptual metaphors in this hybrid model is that we can explain how some metaphors might arise in thought based on experiential structures while others originate in language use. The idea that conceptual metaphors might be involved in cognition is supported by observations reported in gesture studies, indicating that basic metaphorical structures might have a guiding function in the process of developing an idea during conversation. It appears that the interplay between the three levels of
cognition proposed in this thesis have much in common with the observed interplay of gesture, thought, and speech. This apparent parallel in observation does not seem to be a coincidence, but rather it may be assumed that our cognitive processing patterns are reflected in the way we express ourselves. Thus, if a gesture inspires an idea in an imagistic way, a similar pattern might occur with an imagistic cognitive representation that inspires propositional thought. This parallel suggests exciting research to be carried out in the future to empirically test the proposed connections between thought and verbo-gestural expression, on the one hand, and different kinds of thought processes, on the other.

It could be argued that the model developed in this thesis assumes several different processing levels which all, essentially, compute representations of the same content, albeit in different formats. This might be taken to imply a waste of cognitive resources, because such representational duplications are ultimately superfluous. However, while different representational formats may be associated with the same conceptual content, they have entirely different functions, which all make important contributions to an overall understanding of utterances, and of the world. Thus, propositional representations are the ingredients that constitute rational thinking, while imagistic-experiential representations are important in the processing of feelings, non-rational intuition such as instinctively drawing analogies between domains (leading to conceptual metaphors), and generally more holistic impressions. The metarepresentational level has a different function, which is to manage these different kinds of representations. This, then, does not involve a waste of resources but rather a very efficient system that caters for a range of levels of human experience.

In future research, it will be necessary to test these theoretical considerations empirically. A promising starting point is suggested by the data on gesture, speech and thought mentioned above. These observational studies require more rigorous experimental testing in order to count as evidence. We could, for example, compare how people develop ideas in conversation while freely gesturing, as opposed to situations where their hands are kept busy, so that they need to rely on speech alone. Such research could cast light on the extent to which conceptual metaphors are involved in reasoning. The model developed here
further suggests future research investigating the interplay between the different
cognitive processing modes. While there is a long way to go until we can develop
methodologies that can provide clear evidence on this, perhaps involving
technologies employed in brain tomography, the future promises exciting research
in this matter.
References


