From Collision to Collusion

An Optimality approach for Metonymic Type Coercion

Marlies Kluck

From Collision to Collusion

An Optimality Approach for Metonymic Type Coercion

(met een samenvatting in het Nederlands)

Marlies Kluck
9800166

Thesis submitted in partial fulfillment
of the requirements for the degree of Master
of Linguistics. Utrecht University,
Faculty of Arts, Dpt. of Linguistics.

Supervisors:

Prof.dr. H.E. de Swart
Prof.dr. E.J. Reuland

June 12, 2005

'It's three a.m., there's too much noise
Don't you people ever wanna go to bed?
Just 'cause you feel so good, do you have
To drive me out of my head?'

I said, Hey! You! Get off of my cloud
Don't hang around 'cause two's a crowd
On my cloud, baby

Mick Jagger and Keith Richards, *Get Off My Cloud*
(December's Children, 1965)

'But how can you walk away from something and still come back to it?'
'Easy', said the cat. 'Think of somebody walking around the world. You start out walking away from something, and end up coming back to it.'
'Small world', said Coraline.

Neil Gaiman, *Coraline*.

From Collision to Collusion

An Optimality approach for Metonymic Type Coercion

Marlies Kluck

June 12, 2005

Contents

Acknowledgements

It seems ages ago that I started this study. In fact, it *is* ages ago. I remember that I attended a class of Henk Verkuyl and got acquainted with conceptual semantics. I guess that's where 'things' like mind, reality, concepts and language started to intrigue me. The idea for this thesis was born shortly after that, but I still had to do lots of things before I could the actual writing of this MA thesis. The usual stuff, like other courses and yet another side-job because of the permanent state of being broke, but also the experience of being an Erasmus student in Italy[1] and several interesting assisting jobs at university. The last adventure before ending my studies and entering a new phase in life was my trip to the Philippines, which was in one word unforgettable. All made the time I spent as a student at university really... diverse. And long.

A first warm *thank you* goes to my family. My mother who would always be so down-to-earth when I was complaining over the phone about the mountain of work ahead of me. She'd be like '*Well, you just need to start working, Marlies. Stop* talking *and start* doing. *One foot before the other.*' Thanks for always taking the time to talk when I call, mom. Daddy, thanks for the numerous wake-up calls at 7 AM. My long days couldn't have had a better start. I thank the both of them for the special help in *making* this thesis: my mother who in fact realised my idea for the cover (we are both the type of perfectionist person, I am so glad she understands!) and my father who printed several drafts for me and is responsible for the print you are holding right now. There is no better printshop than Thora Copy Service, and I am proud to be the bosses daughter! My sister Rosalinde and I do not differ much in age- in fact she's younger but graduated first- she and I also have a history of long phone calls. Especially in the middle of the night when we both need to vent and rant- about life in general and little annoyances in particular. My two special aunties Joke and Toni, thanks for all the support and the words of comfort in the harder times. And finally my granddad, the only grandparent who lives to see me graduate- who, on one of my worst days, played Debussy on the piano over the phone. There are no words for things like that.

Returning in the post-thesis life, I have a solid circle of friends that I have missed for far too long. I am very thankful to all of them, especially for the fact that they are still my friends, even after the periods of neglect (otherwise seen as Periods of Academic Inspiration). I thank Lotje, my housemate for over 6 years here in Utrecht, for the many cups of coffee and self-made cakes and the infinite times I had to use her oven for a long overdue meal. I

[1] In 2003 I spent half a year studying at the Dipartimento della Communicazione dell'*Università di Bologna*. I was given the opportunity to attend courses of John Perry and a seminar of Umberto Eco, and I travelled to Geneva, Swiss, to participate in the *Generative Lexicon Conference* 2003. I had the pleasure to meet James Pustejovsky, who was willing to share his thoughts on some of the examples that are part of this study in a personal conversation.

owe you big time. I thank Jantine for the many episodes of Friends she borrowed me and our nights just hanging at the couch with a hot cup of coffee and Tuesday-TV. I think things would have been a lot more difficult without the support of Annemarieke, expert in coping with academic and other forms of stress. Jozefien, thank you for always keeping in touch, calling to see if I was still breathing underneath the pile of books and articles I made myself read for this study. The memories of my stay in Italy are kept alive by my friendship with May, who I met there and who will start her thesis in law as I finish mine in linguistics. Grazie tanto bene, bella! My friend Inez, thanks for all the emails and late-night phone calls. And I thank Stefan for the mountain of pancakes and encouraging coffeebreaks at the library. Caroline for taking me out dancing when I had spent days on the inside of four walls. And Marthe, for helping me with the wonderful yet *peculiar* world of LaTeX- so many times I can't even count. Finally Jorn, thanks for just being there for me, no matter how stressed and unreasonable I was. I wouldn't have coped so well with this thesis if it wasn't for all of you guys!

Introduction

> '*If I try to say the joke to myself, making the word mean the three different things at the same time, it is like hearing three different pieces of music at the same time, which is uncomfortable and confusing and not nice like white noise. It is like three people trying to talk to you at the same time about different things.*'[2]

I met my first elephant in a book. At home, we had this book with pictures with the words belonging to them. In fact, I think I have met *most* of my elephants in books, except for maybe the ones I met on my rare visits to the zoo. The amazing thing is, that although I was only acquainted with very minimalistic and two-dimensional elephants, I yelled *Look, an elephant!* in the zoo, and pointed to the correct animal.

There is nothing unique about this little anecdote whatsoever. Not even my elephant. I think I learned the words for most animals from books. Some books even produced the noise that the animal makes when opening the page with the picture of that animal, which, in case of the lion, was quite traumatizing. Next to that, this story doesn't make me a prodigy at all. All children learn to point at books and reproduce cat's noises, followed by the word *cat* and finally calling actual cats *cats*. And finally, I am not the first to wonder about the relation between object and image.

In 1926, the Belgian painter Rene Magritte (1898-1967) surprised both artists and philosophers with his painting that is known as '*Ceci n'est pas une pipe'*, or the *Treachery of Images*. The painting itself is in my opinion not the best he ever made, showing no more than a photographically precise pipe with the subtitle which says it's not a pipe. This apparent contradiction was the cause of quite a debate amongst philosophers. One of the philosophers and a personal acquaintance of Magritte, was Michel Foucault (1926-1984), with whom he exchanged some letters that were later published in his short *This is not a pipe* (1983). The paradox of this painting is that it represents an object that we commonly call a *pipe*, and yet Magritte claims it's not. The underlying thought is that no matter how closely the painting resembles a pipe, as long as it's impossible to actually *smoke* the pipe, it is *not* a pipe.

In early stages of my studies at university, I learned that it is often the case that what is an issue for philosophers, is an issue for linguists as well. After all, we're all curious about the human mind and what goes on in that gray matter that makes us do the things we do, and say the things we say. In the painting of Magritte, it is obvious that there is a difference between an object and its representation- regardless of their external similarities. I assume there are two reasons for a painting like this to be the source of confusion. First of all, there is no representation without an object. Second of all, the human mind can connect an image of an object easily to the object itself.

In natural language, we often switch between image and object, without even noticing

[2]The introductory notes to the chapters of this study are all quotes from Christopher, character in Mark Haddon's brilliant novel 'The curious incident of the dog in the night-time'

it. How many times will we look at photo's we took at holidays, and say *Hey, that's me!*, rather than *Hey, that's a picture of me*, or even more correctly if we'd follow the reasoning of Foucault (1983) Hey, that is a representation with a set of properties that are similar to the set of properties I associate with myself. A photo of me becomes me, when we talk about it. This ability to switch and interpret an image as its object and vice versa is of use in the interpretational phenomenon that is topic to this study. I will focus on examples in language were we are forced to interpret an object in terms of its image. This phenomenon is commonly known as METONYMIC TYPE COERCION (MTC).[3] A good example of this is *stone lion*, where *lion* becomes a statue due to the combination with the adjective *stone*. So far I have discussed only natural kind terms. Although it is the case that this study discusses the coercion of *nouns*, from the data it will be clear that *all* nouns are sensitive for coercion, and that adjectives are not the only cause of MTC. A sentence like *When I crossed the street, I tripped over a truck* shows that verbs can have a similar effect on nouns.

Before I explain the goals of this study, and which questions I aim to answer, there are a couple of related issues that need to be addressed. Those are: what is it that I mean by 'type' and following from that, which type shift is the result of MTC and, most importantly, why exactly is this something a linguist would be worried about.

In this study, the notion type is used to refer to an *ontological* type. As I noted before, MTC can 'happen' to all kinds of nouns. My ontological claims about nouns are based on the TRIPARTITE CONCEPT LATTICE in Pustejovsky (2003:iii). According to this lattice, a word denotes either an ENTITY, an EVENT or a QUALITY. For my purposes, only the first matters, since entities are commonly denoted by nouns in natural language. Those entities can be NATURAL, FUNCTIONAL or COMPLEX. The first can be for example *lion*, although natural kind nouns can also denote abstract concepts. PHYSICAL natural kinds are divided in COUNT and MASS nouns. In this study, I will put Pustejovsky's FUNCTIONAL and COMPLEX type together and call that ARTEFACTS. Nouns that denote artefacts are of the ontological type N_{art}. The other relevant type here is NATURAL, in this study nouns denoting natural kinds are assigned the ontological type N_{nat}.

The shift taking place is to an ontological type that is not present in the lattice of Pustejovsky (2003): IMAGE. This type is what you obtain by mirroring the concepts of the types in the lattice, assuming that for every object there can be some kind of image. The data in this study show that language may force us to interpret objects (denoted by nouns) as images of themselves. The shift is thus from the ontological type they have by default (N_{def}) to the image type N_{ima}.

This shift, although apparently unproblematic for our minds, is rather problematic for theories about natural language interpretation. First of all, we semanticists like to assume there is some kind of meaning that is intrinsically related to a word. This can be seen as a set of properties that at least form necessary conditions for the definition of a concept. The possibility to interpret a living lion as some stone in the shape of a lion, would mean that the property of being alive is not a necessary property of *lion*. Another problem will be clear when we study the examples of MTC, namely that in most of the constructions in which MTC occurs, the construction itself does not account for type coercion. What basically happens in MTC is that there is some kind of collision between some of the properties of the noun that undergoes the coercion and the linguistic element that serves as the MTC functor. In fact, this is not something to worry about, it is something you look at and think *Mental tricks*

[3]The term Type Coercion finds its source in mathematics.

like this are the reason I am obsessed with semantics. Because the most amazing thing is, we have no problem to come to an understanding of utterances that contain colliding material.

The aim of this study is to find a semantic framework that is able to account for MTC. This theory should be explanatory on different levels, accounting for both the linguistic causes of coercion as the relative cognitive ease with which MTC constructions are understood.

The quests and subquests to this semantic grail are described in four chapters. Chapter 1 is a detailed description of MTC. The aim of this chapter is to take a close look at examples containing MTC. We should then see in which constructions MTC can occur, what happens when we combine different data, to what extend this phenomenon is productive and what kind of interpretations are associated with the ontological type N_{ima}. In this chapter I will not describe MTC from the perspective of a particular semantic theory, to avoid a theoretical bias. The knowledge about MTC that we will gain in chapter 1, will form a sort of test for chapter 2, where I will start browsing through current semantic theories that offer some perspective on MTC. I will discuss theories from different directions within semantics. The assumption is that the theories all account for a part of MTC, but that they are either too specific or too broad to give a satisfying account for all of the examples. This is in fact no surprise given the data of MTC, their variety yields that part of them will traditionally be seen as *pragmatic* effects, and the other as *semantic* effects. A complete account of MTC accounts for linguistic effects on interpretation- whether these are contextual or not. Chapter 3 is a proposal for an optimality based approach for MTC. I will argue that OT offers the perspective on natural language interpretation that we need for this phenomenon, and that the strict rules and principles offered by the theories discussed in chapter 2, should be reconsidered in the perspective of optimal interpretations. In the final chapter this study comes to an end, I will give conclusive remarks on the results of my OT proposal for MTC and give some suggestions for extended research.

Preliminaries

> *'This is because I do not always do what I am told. And this is because when people tell you what to do it is usually confusing and it does not make sense.'*

This study makes use of symbols in more or less 'common' ways, however, in some cases I needed to use a symbol to explain something particular to this study, or make an abbreviation because I use a term so many times that reading the full word repeatedly would become annoying. That's why I will add a short list of things that I think will come in handy when reading or before reading this study.

Use of Symbols

Symbol	Description
$\|\|\alpha\|\|$	*semantic value or meaning of* α
x→y	*x becomes y*
*S	*S is semantically unwellformed*
#S	*S does not have target meaning*
?S	*S is not unwellformed, but its meaning is odd*
A∩B	*Intersection of A and B*
A∪B	*Union of A and B*
A⊆B	*A is a subset of B or A is subsumed by B*
☞	*optimal output candidate*
*	*violation of constraint*
!*	*fatal violation of constraint*
a∈A	*a is a member of A*
R > S	*R is higher ranked than S*

Abbreviations

Abbreviation	Description
MTC	*Metonymic Type Coercion*
N_{def}	*N in its default ontological type*
N_{art}	*N of type artefact*
N_{nat}	*N of type natural kind*
N_{ima}	*N of image type*
V/VP	*Verb/Verb Phrase*
N/NP	*Noun/Noun Phrase*

Adj	*Adjective*
Mod	*Modifier*
MST	*Mental Space Theory*
SG	*Sense Generation*
GL	*Generative Lexicon*
DRT	*Discourse Representation Theory*
DRS	*Discourse Representation Structure*
OT	*Optimality Theory*
OTS	*Optimality Theoretic Semantics*
QS/QS$_{def}$	*Qualia Structure/default Qualia Structure*
C	*Concept*
C$_p$	*prototype of C*
U	*Utterance*

Chapter 1

MTC in detail

> *'I said I wasn't clever. I was just noticing how things were, and that wasn't clever. That was just being observant. Being clever was when you looked at how things were and used the evidence to work out something new.'*

Before we start with the search for a suitable semantic framework accounting for MTC, it is necessary to obtain an overall view of what this phenomenon entails. That is, we need to figure out what linguistic factors play a role in MTC and in which constructions MTC occurs. Next to that, we should attempt to make a clear definition of metonymic type coercion. This can only be done if we look at both the domain and the range of MTC. In other words, next to giving possible examples of MTC in natural language, we will need to find the limits of this phenomenon. Last, but certainly not least, it is useful to see if there is a specific meaning associated with the coercion of the nouns in MTC constructions, and if so, how this meaning can be described.

The first section of this chapter will be a listing and explanation of the core examples that are subject to this study. After that, the second section will be an exploration of the limits of MTC, followed by section 1.3 where I will give an overview of the constructions that involve MTC and abstract typical meanings that we can associate with metonymic type coercion.

1.1 The domain of examples

This section shows two kinds of appearances of MTC: in Adj-N constructions and in discourse. The latter means MTC in larger units than the noun phrase of the noun that undergoes the type-shift. The listing will be accompanied by a brief explanation of how the shift of interpretation takes place. Note that these interpretations are based on intuitions, and are not rephrased formally. Later sections of this section each offer their own way of formalising the interpretations. For terminologies and usages of symbols I refer back to the chapter Preliminaries. Note that every example is contrasted by other examples. In case of the Adj-N constructions, I added a counterexample of a equivalent construction that does not undergo a type-shift of N. In case of the discourse examples, I sometimes added more counterexamples, depending on how many elements in the discourse potentially give rise to the type-shift. In general, the (a) sentences are examples of type-shifts; I will indicate if different.

1.1.1 MTC in Adj-N constructions

In current linguistic research, the examples of MTC that are discussed, are Adj-N constructions. In the examples below, the Adj that modifies the N, seems to force the N into another type than its original type. Still, the interpretation of N is not completely different from the original N, the type that N shifts into is somehow related to the type it comes from. Consider (1):

(1) a. a stone lion
b. an Asian lion

The default type of the noun lion is N_{nat}. A stone lion is a representation of a lion made of stone, like a statue. The interpretation of (1b) is thus of another type, the type that I refer to as N_{ima}. Intuitively, the combination of the Adj *stone* with N_{nat}(lion) gives rise to a type-shift. In other words, the meaning of the head in this NP is altered by its modifier. The type-shift is not particular to natural kind terms only. Consider (2):

(2) a. a fake gun
b. a straight gun

In this example, the semantic type of *gun* would normally be N_{art}. However, if you combine it with the Adj *fake*, the gun is interpreted as N_{ima}. A fake gun is not a gun, but still you refer to it as some kind of gun.[1] Intuitively, the interpretation of (2a) is that the gun is a reconstruction of a real gun. A reconstruction of an (original) object is typically what the semantic type N_{ima} means. Fake seems a stronger cause of this shift than stone. Fake will shift any noun that it's combined with into N_{ima}. Stone will do that for any N_{nat}, but not for all N_{art}.[2] The next example seems like (1) only with a N_{art}:

(3) a. a chocolate teapot
b. a porcelain teapot

In this example, the N_{art} is shifted to N_{ima}. Intuitively, a teapot made of chocolate is not a teapot but a chocolate with the shape of a teapot (an image of a teapot). Note that the adjectives in (1) and (3) are both material adjectives, i.e. they can function as an nouns (stone, chocolate) and as adjectives. This is not true for *fake*, which can only be used adjectival. That is, the construction *chocolate N* coerces more N's into N_{ima} than *stone N*. This has to do with differences in meaning between the two adjectives which will be discussed in later sections. In (3), the Adj takes away the functional properties of the noun, whereas *stone* takes away the property of being alive in (1). The adjective *wannabe* seems to correspond to *fake* in some respect:

(4) a. a wannabe punker

[1] I am aware of the fact that the territoria of the semantic type of gun and the conceptual issue of what is a gun are separated by a very thin line. However, this chapter deals with the semantic relation between linguistic items. The conceptual side of MTC will inevitably shine through the analysis of the semantic relation, because conceptual meaning and semantic types are inherently related. The conceptual issue lies beyond the scope of this thesis, but will be shortly discussed in chapter 4.

[2] Consider *stone bridge* that remains N_{art} versus *stone book* that shifts from N_{art} into N_{ima}.

b. a tall punker

The punker that is referred to in (4) is someone who (desperately) wants to be a punker, but is not in the eyes of the person who asserts (4). The adjective is subjective, more than stone or chocolate. Somebody who is a wannabe punker can be a wannabe in the eyes of his classmates but not in the eyes of himself or, say, his grandmother. Note that fake has a similar property: the property of being fake in the case of gun may mean that only the person who uses the fake gun knows that it's fake. While fakeness is not a subjective matter in the same sense as wannabe-ness[3] , both depend on varying perceptions. This example is not as clear as the other examples we have seen so far. In some respect, I doubt that what we see here is the same kind of type-shift as in (1)-(3). This might have to do with the fact that *wannabe* modifies a noun that denotes a human being. Describing the coercion that takes place here could seem somewhat problematic when it comes to assigning the appropriate type to this noun, as we have seen in the preliminaries to this thesis. Since I decided there to keep things as simple as possible, *punker* should be type N_{nat}. It is in this example that we can see the difficulty of subcategories of human kinds: a wannabe punker is still a human being. However, he does not have the properties that the speaker thinks a punker should have. In that respect he is an infelicitous image of a punker. So, in this case the properties of punker are coerced into those of an image of a punker. The resulting type after coercion is N_{ima}, although the properties of the punker being a human are maintained. *A wannabe punker* is assigned type N_{ima} because its underlying subtype of the socially defined 'class of punkers' is coerced, and not its main type N_{nat}. To keep things as simple as possible in this thesis, the coercion is, however, described as the common $N_{nat} \rightarrow N_{ima}$. An even harder case is (5):

(5) a. an imaginary friend
b. a blond friend

The friend in (5a) is a friend that only exists in the mind of the person that has that friend. In other words, this is not just an N_{ima}, the image here is a mental image alone. There is no actual image in reality that is the reference point for this expression. A mental image of a friend is not the same as an imaginary friend though, the adjective *imaginary* has a meaning that contains a form of *intentionality*, which also holds for *wannabe*. There are more adjectives with the effect of *imaginary*: *mythical* and *fictitious*. A noun N_{art} or N_{nat} that is combined with one of these adjectives does not only shift to N_{ima}, it also loses its reference in reality. This will be further discussed in section 1.1.3.

1.1.2 MTC on a larger scale

This paragraph is concerned with examples that show MTC at a bigger level than within the NP. There are of course differences between them, I call them discourse examples, but in fact it will become clear that we can divide discourse into linguistic context (outside the phrasal level) and non-linguistic context (where a type-shift occurs due to context of utterance and supposed mutual knowledge of the speakers). 'In discourse' in this view just means that the shift does not occur through modification of an adjective. The examples are ordered I will start, though, with an example that in fact lies outside this definition of discourse, but is peculiar enough to regard:

[3] *Wannabe-ness* is subjective by definition, *fakeness* is testable in reality, and is subjective in the sense that a person might not see that the gun is fake. This will be discussed extensively in later sections.

(6) a. Jenny chose a dress with roses.

b. Jenny chose a garden with roses.

Roses in (6a) shift from $N_{nat} \rightarrow N_{ima}$. In the examples I have discussed so far, it was the modifier that altered the head. In this example, the modification takes place after the head, by way of a PP construction that is embedded in the NP. Here, the modifying element undergoes the type-shift, namely the roses. It might thus be that on the syntax-semantics level, backward type-shifting is not allowed. I suppose that the type-shift in this example is caused by a mixture of lexical knowledge about *a dress* and *roses* and common-sense that roses on a dress are pictures of roses. In the following example, the shift occurs within a VP:

(7) a. When I walked in my street, I tripped over a truck.

b. When I walked in my street, I collided with a truck.

In (7a) the verb *to trip over* seems to imply that the object that is the THEME of the event of *tripping over*, is not bigger than until the AGENT's knee. This is an intuitive paraphrase of what demands the verb puts onto its θ-roles. In other words, a verb can force a NP to type-shifting if the original meaning of that NP forms a conflict with the demands of the verbs θ-roles. In (7a) the truck means a small truck, i.e. a toy truck that is probably left outside by some kid. A toy that is a reconstruction of a truck is an image of a truck. The type-shift in this sentence is thus $N_{art} \rightarrow N_{ima}$. (7b) offers a counterexample with a different verb. This verb demands THEME to be bigger than until someones knee. In this sense, *to collide with* and *to trip over* are complementary in terms of meaning. Since the default truck is big enough in order for AGENT to collide with it, no coercion of *truck* takes place. For now, my hypothesis for this example is that the type-shift takes place by means of lexical knowledge of the head verb of the sentence. Extending the distance between shifting nouns and elements that seem to be responsible for that shift, let's look at (8):

(8) a. I walked into the room. The elephant had fallen down.

b. I walked into the zoo. The elephant had fallen down.

c. I walked into the room. The goldfish had fallen down.

The type-shift in this example is similar to the one in (1), N_{nat} *elephant* in the second sentence becomes N_{ima}. This seems to happen because of the setting/location of the event, which is presented in the first sentence. If you have an elephant in a *room*, this will probably be a small decorative statue in the shape of an elephant. Look at the contrast sentence (1b). An elephant in another location, the zoo, means an elephant N_{nat}. Next to that, there is no type-shift when we use the same location and another instance of N_{nat}. Intuitively, the setting of the location room makes the noun elephant shifts $N_{nat} \rightarrow N_{ima}$. Note that this can even occur over a bigger distance than a sequence of two sentences. Suppose that the story is: '*I walked into the room. The window was open, the curtains were blown out by the wind. The elephant had fallen down.*' The setting of a story is an important factor in the meaning of the discourse, apparently it's strong enough to cause type-shifts for as long as the location is not changed in that story. Now let's look at an example that has both a difficulties with the location and the verb:

(9) a. The duck lay in the bath.

b. The duck lay in the lake.

c. The duck swam in the bath.

d. The duck swam in the lake.

In (9a) the type of duck is shifted from $N_{nat} \rightarrow N_{ima}$. The ducks usually found in baths are those funny yellow ones that peep when you squeeze them. When we contrast this sentence by changing first the location (9b), the interpretation is N_{nat}, but probably the duck is dead here. Still, a dead duck is a N_{nat}. When we change the verb into *swim*, the location has no influence. The verb to swim needs AGENT to be N_{nat}, so type coercion is not needed, moreover, it seems to be impossible. This might mean that the demands that are put on a nominal argument N are stronger than the demands of the location-setting in which N occurs. The examples (10) and (11) offer a complete different picture, namely AGENT in the following examples seem to lead to type-shift of THEME:

(10) a. The child noticed the jaguar had disappeared from his collection.

b. The millionaire noticed the jaguar had disappeared from his collection.

c. The man noticed the jaguar had disappeared from his collection.

In (10a) the jaguar shifts from $N_{art} \rightarrow N_{ima}$. The subject of the sentence is *the child*. In this case it seems that common sense predicts that children will have a toy jaguar (N_{ima}) rather than real jaguars. In case of the millionaire, *the jaguar* is probably real (N_{art}). In case of (10c), the situation is unclear. I use this example to indicate what I think is at stake here. The factors that cause this type-shift in (10a) are not based on pure lexical information like in the previous cases. In this case, the type-shift occurs because of *common-sense* knowledge. *The man* in (10c) can have an actual jaguar if the man is referred to in a context where owning a jaguar is normal. In sum, the necessity for a type-shift depends on the context of utterance and on conceptual knowledge combined with common sense, rather than on lexical knowledge. Another sort of knowledge is causing shifts in (11):

(11) a. Next season, Steven Spielberg will work with dinosaurs.

b. Next season, Bert and Ernie will work with dinosaurs.

c. Next season, Roy C. Andrews will work with dinosaurs.

In this case, there are two sentences with a shift $N_{nat} \rightarrow N_{ima}$: (11a) and (11b). In (11c), the dinosaurs, probably meaning a heap of bones that once belonged to dinosaurs have the interpretation of the default type N_{nat}.[4] It seems that the subjects of the sentences in (11) force the noun dinosaurs into N_{ima}. The subjects of these sentences are of a special kind, they are all famous people that a lot of people know. Intuitively, the shift does not occur because lexical knowledge but because the hearer is supposed to know that Steven Spielberg is one of the worlds most famous movie-producers, who directed a movie about dinosaurs (in supposed life-size), that Bert and Ernie are famous puppets in Sesame Street, who will probably play with puppet-dinosaurs, and that Roy C. Andrews is a well-known paleontologist who was the first to discover dinosaur eggs (of the later called *ovi*raptor). In sum, knowledge about elements in the discourse might lead to type-shifts.

[4]Note that the dinosaurs in (11a) probably look the most like the concept we have in mind by the word *dinosaur*, and the dinosaurs in (11c), the default type, look least like dinosaurs.

To summarize what we have seen so far, I close this section by giving an informal description of what I have called metonymic type coercion. We have seen that there are two default semantic types for nouns that are relevant here: N_{nat} for nouns referring to natural kinds, and N_{art} for those referring to objects made by mankind. Representations of those kinds are a subclass of artifacts: *images*. The type that I use for this class is N_{ima}. We have seen that nouns of both types can be coerced into N_{ima}, by either an adjective or an MTC operator somewhere else in the discourse that seems to make the interpretation belonging to the default type impossible. So far we can represent MTC in the following way:

METONYMIC TYPE COERCION I
$F_{\text{MTC}}(N_{nat/art}) \rightarrow N_{ima}$

Whether this informal representation is enough to describe the phenomenon is still the question. An attempt to discover the limits of MTC is made in the next section. For now, let me restate that this 'formula' and the types I use to discuss MTC are, theoretically seen, not intended to have any formal value. This chapter is no more than a description of the phenomenon. For the purposes of this thesis it is necessary to maintain at a level of 'theory-neutrality', which is why I refrain from representations that would easily fall into the range of a specific semantic framework.

1.2 The range of examples

This section will be an exploration of the limits of MTC. What I want to see is to which degree MTC occurs when we 'mix' the examples of the previous sections. The underlying motivation for this exploration is theoretical: a semantic framework that accounts for the existing cases of MTC, must also account for the limits and restrictions of MTC. In 1.2.1 I will attempt to derive the meanings of the adjectival cases of MTC in the sentences of section 1.1.2. That is, without expressing the adjective. In 1.2.2 I will embed the adjectival cases into sentences that are similar to example (6).

1.2.1 To a higher leveled phrase

So far, we have seen that MTC can occur in constructions with a modifier and a head (Adj-N constructions) and in discourse with or without a specific syntactic relation. It would be interesting to see if the examples of MTC in Adj-N constructions can also occur in discourse. So, let's look at each example and integrate it in different contexts, without no adjective, but with the same type shift on the noun. For each of the examples (1)-(5) I will try formulate sentences with similar type-shifts, making use of the construction-scale of (6)-(11). Note that sometimes I needed to alter the verbs or other elements. If the sentence is successful for MTC, it can be contrasted with the same sort of sentences I used in (6)-(11):

(12) a. My son has a sweater with a lion (on it).
b. ?Walking to the toilet, I accidentally stepped on the lion.
c. I came into the room. The lion had fallen down.
d. ?The lion lay in the bath.
e. The child missed a lion from his collection.

f. Next season, Bert and Ernie will work with lions.

In the PP-NP construction in (12a), there is a shift $N_{nat} \rightarrow N_{ima}$. Lions on a sweater are like roses on a dress. These kind of constructs we will find continuously in constructions like a dress with, wallpaper with, a sweater with, a pajama with, etc. These things usually have pictures printed on them, so anything that said to be 'on' them, will probably shift into N_{ima}. In (12b) the case is different than in the similar example (7). I chose for the verb *to step on* instead of *to trip over*, because a N_{nat} lion (when lying down) has a size such that a regular human being could trip over. This sentence is not as clearcut as (7), though. If you replace *to trip over* in (7) by *to step on*, the truck will undergo type-shift, but in case of N_{nat} lion, somehow it seems to be possible to interpret to step on a lion as to step of a part of a lion. This sentence could easily mean that I stepped on a lion's tail. In other words, the type-shift in a sentence like this, only seems to take place when it is absolutely impossible to interpret the noun in the default way. This could be an important fact in the analysis of MTC, as we shall see later on. Continuing to (12c), we see the same shift $N_{nat} \rightarrow N_{ima}$ as in (8). The location *room* and the object N_{nat} *lion* seem to conflict, an ornament in the shape of a lion would be more probable in this context. This is not necessarily true for (12d). A lion in a bath doesn't seem to have the same effect as a duck in a bath (9). Intuitively, this is because there are this funny yellow plastic ducks that are designed especially as toys for in the bath. However, a N_{nat} *lion* in a bath is also odd, so it could be the case that there is a type-shift here. Example (12e) has a funny effect that doesn't occur when the object is *jaguar* like in (10). Jaguars and all kinds of (N_{art} or N_{ima}) cars are suitable to be a part of a collection. This is not true for N_{nat} lion. A lion can be part of a group of lions (in wildlife or a zoo), but is not likely to be a collector's item. So, in this example the type-shift occurs, but for a different reason. The lion that is a collector's item is probably some plastic toy that belongs to a series of toys. In this case, man or child as the subject of the sentence (that are contrasted in (10)), do not make a difference. Finally, (12f) has the same type-shift as (11b). Note that the objects in this type are very sensitive to the kinds of subjects. The effect does not arise with Steven Spielberg, because if Steven Spielberg worked with lion is his movies, they would probably be N_{nat} lions. In case of Bert and Ernie, one would think directly of a puppet-show, in which all animals are puppets, i.e. N_{ima}.
Let's see if a gun can be interpreted as N_{ima} without being modified by the adjective *fake*:

(13) a. Bart has wallpaper with guns.

b. When I walked across my street, I got soaked by a gun.

c. #I came into the nursery. The gun had fallen down.

d. #The gun lay in the bath (on the floor, toilet, nursery,...).

e. ?The child noticed a gun missing from his collection.

f. Next season, Bert and Ernie will work with guns.

Recall that the sentences that I marked with #, are not ungrammatical, but do not have the 'target' shifted interpretation of *gun*. I will go over these example briefly, since a lot of explanation is already done in the previous example. As suggested above, anything that is likely to have pictures on it, will easily lead to type-shift of the noun it's combined with by way of a PP-NP construction. This also holds for (13a). In (13b) I used the phrase to get soaked by, which automatically leads to a type-shift to N_{ima} of gun. Note however, that in

this case, you do not only get the interpretation of fake gun, this fake gun is also a toy gun (water pistol). I will come back to that when I discuss (13e). The following examples (13c) and (13d) do not get a shifted interpretation. I suppose that for an object like gun, there are no locations that it cannot be a real gun. This might have to do with the fact that people in possession of a gun, sometimes hide it in the most unlikely places. In (13d) I could not differentiate with a verb like in (9), because a gun is no N_{nat}. This could be important. The shift from $N_{nat} \rightarrow N_{ima}$ might 'outweigh' the shift from $N_{art} \rightarrow N_{ima}$: there are more verbs that will be MTC operators when they have N_{nat} as one of their arguments. Whenever the verb demands one of its arguments to be [-ALIVE], this will lead to an inevitable coercion of nouns of type N_{nat}, but not for N_{art}. In other words, it might take more for N_{art} to be coerced to N_{ima}. In order to trigger a type-shift, the verb probably differentiates in the feature [+/-ALIVE] for its θ-roles. Any noun of N_{art} will remain [-ALIVE], while any verb that demands an argument to be [-ALIVE], will probably shift $N_{nat} \rightarrow N_{ima}$, or end in ungrammaticality. In (13e), I regret to say that it's questionable for the gun that is in possession of a child to be a fake gun. However, if this sentence was uttered in a context where someone tells about Santa Claus, it will most likely be a toy gun. The child as a subject, however, does not necessarily lead to a shift to N_{ima}. In the final example, Bert and Ernie would be doing a show about weapons, and considering the educational purpose of Sesame Street, this will be fake guns that probably look like real guns. So, unlike (13b), this fake gun will not be a toy gun. The difference could be that toy guns that shoot water, are visibly fake, while a fake gun in Bert and Ernie's show can look like a real gun. Let's for now assume that there are two kinds of interpretations for a gun that shifts into N_{ima}; a fake-but-looking-real interpretation, and a discernible kind of fake in the sense of a toy-interpretation. So, what I mentioned before about fakeness and subjective knowledge in (2) vs. (4), only holds for non-toy interpretations of gun.
Moving on to the next example, (14) will show a nice struggle to imply that a teapot is N_{ima} (made of chocolate if possible) without using the word chocolate as modifier:

(14) a. Jenny choose a dress with teapots.
b. When I crossed the street, I chewed on the teapot.
c. *I looked into the box. The teapot was gone.
d. ?The teapot melted in the oven/in the sun.
e. ?The child missed a teapot from her collection.
f. ?Next season, Bert and Ernie will work with teapots.

As was the case before, the PP-NP construction in (14a) directly shifts teapot $N_{art} \rightarrow N_{ima}$. It seems redundant to say that the meaning of adjectival chocolate is very hard to derive in context. But it is interesting to see that it is easier to imply a meaning of stone than one of chocolate. I think this has to do with default interpretations on a different level, i.e. after the type-shift has taken place. Compare (14b) to (7a) and (12c). In the latter, the lion is likely to be made of stone, wood, marble, porcelain or glass. If the interpretation is ornament/statue, these are the substances that are likely to constitute the N_{ima} lion (or elephant). In other words, the default interpretation once you have shifted to an N_{ima} interpretation is that it's made of one of these substances. For (14b) it is not clear if a teapot that you can eat is necessarily made of chocolate, however it must be made of edible substances, in the shape of a teapot. So the shift to N_{ima} does take place here, substance unspecified. Note that the

acceptability of (14b) depends more on the bigger context it occurs in than in case of elephants or lions that fall down in rooms. To arrive at an interpretation of chocolate, you'd have to build a context that includes food, or verbs related to eating, or strong properties of chocolate that are inferable. In (14c) I tried to make a location that is more likely for a chocolate teapot to be in, than for a default teapot. Not very successful, because it's not implausible for a regular teapot to be packed in a box.[5] The effect of location and verb, in (14d) is interesting, a teapot can only melt if made of a substance that can melt, i.e. glass, plastic or chocolate. A teapot that melts in the oven, is most likely made of plastic or glass (if the oven is 950° Celsius). These teapots would not melt in the sun, though, so one could interpret (14d) as chocolate, but it's not very convincing. In (14e), it's not too clear either. It's possible to interpret teapot as N_{ima} here, meaning a toy teapot. This teapot could however be made of the same materials normal teapots are made of, only smaller. The same holds more or less for the interpretation of (14e). It is not that unlikely that if Bert and Ernie play (work) with a teapot, this is a real teapot. Here I see the same as I suggested when discussing example (12), the conflict between the default type of N and the possible trigger for type-shift has to be strong, in order for MTC to take place.
The following two examples are different from what we have seen so far. In the previous examples, we have discussed adjectives that modify nouns that refer to objects in reality. Now we move over to more abstract adjectives, that modify human concepts *punker* and *friend*:

(15) a. William has wallpaper with punkers.
b. I walked in the street, dressed as a punker.
c. I looked into the box. The punker was stolen/?was gone.
d. *The punker prayed in church.
e. The child/The man missed a punker from his collection.
f. Next season, Bert and Ernie will work with punkers.

The problem that we had in trying to make a teapot of N_{ima} chocolate is also at play in this example. It seems that only a few of the adjectival meanings are derivable in context. However, it is still possible to focus on the shift of the noun, and we see that in some cases we have a successful context. Needless to say, a wallpaper (15a) with punkers has pictures of punkers on it. In (15b) the verb *to dress as* might suggest that the person is not a real punker. The same effect can be reached with verbal constructions like to look like, to smell like, to sound like. The adverbs *as* and *like* seem to be responsible for this meaning, *like* in itself suggest an image-interpretation: it looks like N, but it's not a real N.[6] The interpretation of (15c) is not clear. I tried to make up a location that is unusual for a punker to be in, i.e. a box. This by itself does not cause a shift, though. A punker in a box could be a homeless punker living in a box. The verb in the second sentence is able to cause a shift. In the case of *to steal*, the punker will be interpreted as a little doll with the appearance of a punker,

[5]The interpretation I was aiming for was a box of chocolates, which again could be derived in a context where people are talking about food or chocolates in particular. However, that was not the point of this sentence, the challenge was to find a construction with a location alone that made it impossible for *teapot* to be interpreted in a default way.

[6]I will incorporate this meaning as shifted to N_{ima}, although it might seem redundant, exactly this meaning is what is meant by N_{ima}, the noun that is interpreted to be an image, looks *like* its original.

i.e. N_{ima}. With *to go*, the punker has gone himself, so no shift occurs. In (15d) there is no shift. Again, I picked a strange place for a punker. It might be that locations do not have the same kind of strength that verbs have to shift their arguments. Put in a collection, (15e), the punker shifts from $N_{nat} \rightarrow N_{ima}$.[7] The change of subject from *the child* to *the man* does not effect this type-shift, like it did for car. In the final example (15f), the punker is shifted to N_{ima}, because it will probably be puppet-punkers in Sesamestreet.
The last bit of juggling with the data we had, concerns the example of the *imaginary friend*:

(16) a. ?Mary has a t-shirt with a friend on it.
b. *When I walked in my street, I thought about my friend.
c. #When I lay in my bed, my friend wouldn't come to me.
d. ?My friend was present in my head for more hours a day, when I was unhappy.
e. *The child missed a friend from his collection.
f. *Next season, Bert and Ernie will work with (their) friends.

Also here, it's almost impossible to get the interpretation of *imaginary* without modification from the adjective. There is a possibility to get a N_{ima} interpretation, but not in the fake-sense. In (16a), Mary can have a t-shirt with a print of a friend on it. This is a photo of a friend, so $N_{nat} \rightarrow N_{ima}$. This interpretation most resembles the case of stone lion or the (ornament) elephant in the room. In (16b), to think about a friend does certainly not imply that this friend is made-up, the same holds for (16c), but the interpretation of an imaginary friend is more likely here. It should be pointed out that (16c) needs a context in which it is clear that the speaker, and that could possibly include the adjective *imaginary* or verbs like *to make up*. This is more or less the case in (16d), but still it is possible that the person that utters this sentence is thinking about a good friend when he/she is having a hard time. Regardless of the fact that (16e) is a strange sentence, collection this time doesn't cause a type-shift. It might be that the word friend is not easy to shift, so it takes a lot from its context in order to shift. Finally in (16f), I argue that although Bert and Ernie's friend is a puppet, there is no type-shift. We speak here of a change of perspective, when talking about the friend of someone else, it does not need to be the case that this friend is our friend, still it is a friend to someone. So from the perspective of Bert and Ernie the friend is not a fake friend or the image of what a friend is. To sum, it is very difficult to come up with contexts that will shift the noun *friend* into N_{ima}; none of these examples contain successful shifts.

1.2.2 Embedding MTC in MTC

A last thing I want to show in using these examples is that they are productive, i.e. it is possible to make 'recursive' examples with type-shifts that go even beyond the ones in the previous examples. I will do so by using two different examples for each of the adjectival cases of MTC, one with the object as an image on a sweater (corresponding to example (6)) and one with the object as an image on a poster:

(17) a. stone lion

[7] Note that I label *punker* and *friend* with N_{nat}, which in fact is a bit problematic. A human is a natural kind, but how humans label each other is a human property, punkers and friends are not real natural kinds. I will leave them there, though, because they are worse artifacts than natural kinds. Concepts that divide human beings into categories will be topic to chapter 3.

b. John has a sweater with a stone lion on it.

c. Mary likes the poster of/with the stone lion.

d. Paul prefers the sweater/poster with the lion.

Intuitively, (17b) is harder to understand than (17c) and (17d). This could be caused by the difference between the way an object is imaged on a piece of clothing and the way it is on a poster. For *the poster* it's easier to have the interpretation of a picture (like a photo of a statue of a lion printed on the format of a poster). It is more crucial to look at what happens with the type shifts in this case. In the original example (17a) the the adjective *stone* coerces *lion* from N_{nat} to N_{ima}. In (17b) the trigger for the coercion is not only *stone* but also *sweater*. The interpretation of this type of embedding is 'image of an image'. If this is as simple as it sounds, we will still have to figure out. Semantically, the first trigger coerces the complete noun phrase $[stone_A[lion]]_{NP}$ into N_{ima}. Within this phrase the lion is coerced from N_{nat} to N_{ima}. The image type of *lion* is thus coerced by *sweater/poster* into what I will refer to as $(N_{ima})_{ima}$. This type can be seen as the secondary type of N_{ima}, which, in theory, we could (infinitely) extend to any n-ary type of N_{ima}. However, for the provisional definition that I will propose at the end of this section, I will restrict type N_{ima} to the secondary level.

(18) a. fake gun

b. John has a sweater with a fake gun on it.

c. Mary likes the poster of/with the fake gun.

d. Paul prefers the sweater/poster with the gun.

When looking at the examples (18b) and (18c), it seems that the meaning of *fake* doesn't allow this kind of embedding the same way as *stone* does. Intuitively, (18b) and (18c) correspond more to (18d) than (17b) and (17c) to (17d). That is, the meaning of *fake* can be seen as becoming redundant in the examples where *fake gun* is coerced by *sweater* or *poster*. Next to this, *stone* is a substantive adjective whereas *fake* is an abstract adjective. Although a stone lion depicted on a sweater is made of some kind of fabric like cotton or wool, (17b) must mean that this lion doesn't look like a natural lion. The double coercion results in a meaning where *lion* looses its natural kind meaning by *stone* and *stone* in turn looses it's default meaning by *sweater*. However, the lion made out of fabric in (17b) looks like a stone lion and not like a natural kind lion. For (18) we can propose two interpretations. In the first, the adjective *fake* is subsumed by *sweater/poster*. The interpretation of the examples is then 'a poster or sweater with a gun on it'; the gun being made out of some fabric instead of steel is fake by definition. In this case, adding the adjective *fake* explicitly is redundant. The second possibility is for *fake* to coerce *gun* into a gun that is fake by appearance, and not only by function. Remember from section 1.1.1 that a fake gun can look like a real gun in all respects. If *fake* is considered (semantically) redundant, the image on the sweater/poster is a representation of a real gun. If *fake* is taken as a second trigger for MTC, it seems more plausible that the image on the sweater/poster is a picture of a gun that is (visibly) not a real gun, for example a picture of a water pistol. In short, we can say that *fake*, being an abstract adjective, leaves open the possibility to be interpreted as redundant. In that case, coercion takes place only once (triggered by sweater/poster), with *fake* possibly functioning

as 'reinforcement' for the meaning of the coercion.[8] Then, the coercion is performed over gun only: $N_{art} \rightarrow N_{ima}$. In the case of *fake* functioning as a separate operator for MTC, the coercion looks the same as in (17): $N_{art} \rightarrow (N_{ima})_{ima}$.

(19) a. chocolate teapot

b. John has a sweater with a chocolate teapot on it.

c. Mary likes the poster of/with the chocolate teapot.

d. Paul prefers the sweater/poster with the teapot.

In (19) we see more or less the same effects as in the examples of (17). *Chocolate* being an substantive adjective, the teapots in (19b) and (19c) still have to look like teapots made of chocolate even if represented on, and made by paper or some fabric. In other words, in these examples the two possible triggers for MTC (*sweater/poster* and *chocolate*) both function necessarily as actual MTC operators for MTC. The complete (provisional) operation is again $N_{art} \rightarrow (N_{ima})_{ima}$. It could be argued though, that a stone lion depicted on a poster is 'more plausible' than the same sort of representation of a chocolate teapot. In general, it is more likely for representations of stone (often functioning as statues) to be re-represented on pictures, posters, photos etc. *Chocolate* as an substantive adjective has a different type of function. Whereas *stone* is commonly used to build things out of, *chocolate* functions as food. Nonetheless, in a certain context, a price-winning chocolatier could have a picture or image of his design on a poster, sweater, etc. The plausibility of the examples in this thesis is related to issues of *common sense*, which will not be addressed in detail here, but could be subject to more conceptual-semantic research (see chapter 4).

(20) a. wannabe punker

b. John has a sweater with a wannabe punker on it.

c. Mary likes the poster of/with the wannabe punker.

d. Paul prefers the sweater/poster with the punker.

In my opinion, this example illustrates best the difference between the (b) and (c) examples in this section. In general, I would say that when an object is represented on a poster, the representation is more likely to be a photo than when the object is represented on a piece of clothing. But is that really what differentiates between (20b) and (20c)? It is in this example that I can explain why I chose to use a indefinite for the (b) examples and a definite for the (c) examples. If you interchange the articles in (19b) and (20c), it is more likely that the image on the sweater is a resembling representation (ideally a photo) of a wannabe punker. In discussing the adjective *wannabe* in section 1.1.1, I indicated that this adjective only describes the opinion of the person who makes the utterance in which it occurs. In other words, 'wannabe-ness' is in the eyes of the beholder. Let's for the sake of simplicity assume that it doesn't matter whether the object is represented on a poster or sweater. The use of a definite article results in a different meaning than the use of an indefinite article. *The wannabe punker* is uttered when the speaker knows that the listeners have no doubt about

[8] According to Grice (1957) it would be conversationally infelicitous to use words that are semantically redundant in an utterance. Following that line of reasoning, which will be a part of discussion in chapter 3, we will have to assume that if a picture on a sweater/poster is a representation of a real gun, *fake* will only be added when this is demanded by the context of the utterance.

what/who is being referred to. It can be the case that a group of people agrees that someone known presents himself as a punker, and think that this is infelicitous. A representation of that someone on a poster or sweater can in this context be referred to using a definite article. When the speaker uses an indefinite expression here, it is not referring to someone in particular, which in this case means that it is not part of the common knowledge of the people taking part in the conversation. This is a little besides the issue of this particular section, but it should be noted anyway that for adjectives like *wannabe*, the choice for the article is more significant for the meaning of the adjective itself than it is for adjectives like *stone* or *chocolate*.[9]

Time to take a look at the coercions in (20). As assumed in section 1.1.1, the coercion in (20a) results in the type-shift $N_{nat} \rightarrow N_{ima}$. When *a/the wannabe punker* is depicted on a sweater/poster, the coercion based on what I proposed for (20a) would be $N_{nat} \rightarrow (N_{ima})_{ima}$. However, the operators *sweater/poster* have a different coercion effect than we have seen in previous examples. Take (17): in *a sweater with a stone lion on it*, *lion* loses the properties of a living natural kind when it is modified by the adjective *stone*. *Stone lion* loses all its properties except a 'recognizable appearance' when represented on a sweater/poster. For (20) the coercions are not that straightforward. *Punker* loses only the property of being acknowledged as being part of this particular (social) class. The result is in fact that *a wannabe punker* is still of type N_{nat}, but does not meet the conditions that are set for being an actual punker. In this respect he is an image of a punker, but not an image of a natural kind. The actual shift that eliminates N_{nat} as the type of *punker* takes place when it is depicted on a sweater/poster. The problem that this example puts forward could be an argument to allow subtypes as 'input nouns' of the operation of MTC.[10] For example, we could create N_{hum} as a subtype of N_{nat} and sub-subtypes that can either be natural, such as N_{rac} for race, or N_{soc} as a type for categories of N_{hum} that are based on social factors. Not only is this unpractical to have so many 'types', it still does not allow us to really represent the coercion that takes place in *a poster with a wannabe punker*. If punker is N_{soc}, the combination wannabe punker results in $N_{soc} \rightarrow N_{ima}$. So far, so good, as I have indicated early on that the image type is based on the original type. When we move on to the case of the sweater/poster, we can simply use the same coercion as we used for embedded cases before: $N_{nat} \rightarrow (N_{ima})_{ima}$.

We have now left away the tiny issue of the punker becoming inanimate only after the second coercion took place. It seems to me that this is merely due to the meaning of the adjective wannabe. This adjective seems to have strong requirements for the noun it modifies. The noun doesn't necessarily have to denote a human being. It is possible to call a small dog (for example the ones that you can mistake for a cat) '*a wannabe dog*'. In the era of big brands and expensive fashion someone might even say '*that is such a wannabe Louis Vuitton*'. But does *wannabe* in these cases apply to the *owners* of the dog/bag, or to the nouns themselves? I think I will stick to the definition that the utterance *a wannabe N* is any

[9] I could write a complete chapter on this issue, and I am aware of the poor choice of words in 'the meaning of the adjective itself' here. The issue is in fact not the meaning of the adjective, but in a way the 'scope' of the validity of the utterance in the conversation. The interpretation of an indefinite construction is restricted to a (personal) belief, while the definite is able to extend from personal to common opinion.

[10] Note that those subtypes are different for the secondary type we created for N_{ima}: an image of an image is still an image, whereas the subtypes as N_{hum} represent semantic categories more specifically. It is also for this reason that it will be better to restrict the definition of types that we need for the account of MTC to the 'supertypes' N_{nat}, N_{art} and N_{ima}.

N that in they eyes of the speaker wants to look like N but is not successful.[11] So if you were to say *a wannabe Louis Vuitton*, it refers to the owner wanting to appear as part of the rich community that buys ugly bags for an average month's salary. To put it simple: *wannabe* needs its complement to be alive. This is why *punker* becomes inanimate only after he has been represented on a sweater/poster. Although this is not clear from the description of the type shifts above, an image of a natural kind does not need to become inanimate because of coercion. It could be coerced into N_{ima} by taking away other essential properties of the original type. This is illustrated best when we contrast *a sweater/poster with a/the wannabe punker* with *a sweater/poster with a/the stone punker*. In the latter, *stone* coerces *punker* directly into an N_{ima} that is inanimate.

(21) a. imaginary friend
b. *John has a sweater with an imaginary friend on it.
c. ?Mary likes the poster of/with the imaginary friend.
d. Paul prefers the sweater/poster with the friend.

This last example leads to a funny effect. As I have indicated in 1.1.1, the problem of (21a) is that this NP has no physical reference point. The abstract adjective *imaginary* seems to eliminate the physical aspect of the meaning of friend. In my opinion, (21a) is semantically unwellformed. I have marked (21b) with a question mark because the sentence is semantically at the most questionable. The difference can again be explained by the choice of article for *imaginary friend*. The indefinite article leaves an object that is already without reference point, open. My guess is that this is why (21b) is semantically unwellformed. The use of a definite article in (21c) creates the possibility for *imaginary friend* to be actually represented (visibly). This interpretation is, however, restricted: *the imaginary friend* in (21c) can only be Mary's imaginary friend. In this respect, (21c) would be the semantical equivalent of *Mary likes the poster with her imaginary friend*. Still, the sentence is odd. Just to show how the case of *imaginary* has bickerings everywhere, just consider the variations in (22):

(22) a. *Mary liked the painting of an imaginary friend.
b. Mary liked ?the/her painting of the imaginary friend.
c. Mary made a painting of an imaginary friend.
d. Mary made a painting of the imaginary friend.

It is difficult not to discuss this example extensively. I have marked (22a) as semantically unwellformed, for the same reasons as I have argued before in the case of (22b). In (22c) we see that the sentence is questionable like (21c), but it's ok when we replace the definite article with the possessive *her*. This way it's possible for (21c) to be interpreted such that Mary liked the painting that she has (made) of *her* imaginary friend. This has led to the examples (22c) and (22d). As long as we take care that Mary is the creator of the paintings, it does not matter which article we use for *imaginary friend*. In case of (22c) Mary has painted one of her imaginary friends, in (22d) she painted her imaginary friend (implying that she has only one). Note that (22c) is more wellformed than (22d): the latter is, once again, better when we

[11] This is especially true for natural kinds that have the potential to pretend that they are something they are not in the eyes of others. Various (verbal) constructions in language illustrate this capacity for human beings: *to dress like* (a princess), *to look like* (a wanderer), *to pretend*, ...etc.

use the possessive instead of the definite article. So, what do the diverse examples of the embedded *imaginary friend* tell us about MTC? Like in the case of *punker*, we deal here with a subtype of N_{nat}. To keep things simple, I will just refer back to this discussion and assume we can define *friend* as being of type N_{nat}, coerced to type N_{ima} by the adjective *imaginary* and the depiction of the latter on a painting/drawing/picture/poster/etc. resulting in $N_{nat} \rightarrow (N_{ima})_{ima}$. Also in this case we see a difference between indefinite or definite articles; this time the indefinite even creates a semantic unwellformedness. A new phenomenon is that these instances of MTC lead to a 'physical realization' of the objects that the coerced nouns refer to. *Imaginary* first coerces *friend* to an abstract object that only exists in the mind of someone in particular. The depiction of this mental N_{ima} gives it a visible appearance. This is effect is stronger in (23), where the coercing noun *unicorn* is of type N_{ima} in itself:

(23) a. John made a painting of the unicorn.

b. Mary likes a picture of a unicorn.

c. ?I saw a photo of a unicorn.

It is evident that triggers of MTC like painting will in most cases coerce the nouns that represent the objects that are depicted into N_{ima}. We wouldn't expect anything else from things like paintings, statues and all other objects that have the inherent function of representing another object, right? Example (23) just shows how powerful we (as humans) are to represent things we perceive, whether this is from imagination or what we actually see. In language this causes type-shifts that go beyond expectation, we can even shift something that we have never seen into something visible. However, the case of *the unicorn* leads to a complicated bunch of new questions that unfortunately lie beyond the scope of my present study. So I will just restrict myself by a synopsis the questions that are relevant to MTC. In introducing (23) I proposed N_{ima} as the default type for the (non-modified) unicorn. This is problematic, since we have established this type for representations of existing things. It is generally assumed that unicorns do not exist. Still, ask several people to draw one and you will probably get a series of similar drawings. The question is *if* we can call this N_{ima} and *how* we should analyse the coercion that takes place in the examples of (23).[12] When the input of MTC is a N_{ima} that has as referent an object that exists only mentally, is the result of coercion $(N_{ima})_{ima}$? It would be interesting to include these examples in a more elaborated analysis on MTC, since we are now arriving at the borders of semiotics. What holds for the case of unicorn, is also true for a lot of creatures that are part of stories, fairy tales and other forms of art.

This section has given a picture of the possibilities of embedding Adj-N constructions with MTC in contexts with other triggers for MTC. In most cases this led to a successful 'double metonymic type coercions'. I have restricted these examples to cases of *depictions*; posters, paintings, drawings, prints on sweaters, etc. It is obvious why these nouns are so talented in triggering MTC. As stated above, their nature is to represent other objects. I ended section 1.1 by giving an informal description of MTC. The examples of 1.2.1 do not add anything to this definition, although they underline the assumption that (to a certain extent) MTC can be caused by triggers other than adjectives, i.e. other elements than adjectives are MTC operators. It will be clear in the next chapter why this assumption is essential. As for

[12]I have assigned this type to unicorn because of the apparent similarity to imaginary friend. However, unicorn will never have a referent in the actual world except when depicted. I suppose that this is what makes this decision for a type difficult.

the examples I have just discussed, it is clear that any definition of MTC should include the following operation (regardless of semantic representation):

METONYMIC TYPE COERCION II
$F_{\text{MTC}}(N_{ima}) \rightarrow (N_{ima})_{ima}$

1.3 A survey of MTC

Before we start looking at the way present semantic frameworks deal with MTC, it's practical to give structure to the data. In this survey I will approach the examples on two levels. In 1.3.1 the examples will be analysed syntactically, we will look if the distance between MTC operator and coercing N is restricted, and I will formulate a final version of the definition of MTC. In 1.3.2 I will generalize the interpretations of the examples of MTC, and, where possible, divide them in categories. Finally, 1.3.3 will shortly explain which questions will be addressed henceforth in this study.

1.3.1 Different relations

In this section I will evaluate the different syntactic, semantic and pragmatic relations that can lead to metonymic type coercion. I will do so briefly; this section is merely a summary of what has been discussed in the previous section. In section 1.1.1 I discussed several examples of Adj-N constructions that gave rise to type coercion. In section 1.1.2 I walked through some examples in which the noun is not coerced by a modifier. Type coercion takes place when two elements have conflicting semantic properties. The elements triggering type coercion in discourse differ in great respect; some are related syntactically, others can cause type coercion on a noun because they function as semantic markers in discourse. I will first discuss the syntactic related elements, than the semantic markers. This concerns the distance to the shifted N, I will close the section with the positions of the shifting N, in order to decide if coercion is an operation that necessarily works *forwards*, or if *backwards* coercion (i.e. 'coercion in retrospect') is possible as well.

The smallest distance over which type coercion can take place in this investigation, is between modifier and head, examples (1)-(5). The way a noun and an adjective are related syntactically and semantically, plays a role in the way a noun can shift from its original type to N_{ima}. This will be discussed extensively in later sections. For now, the most important fact is that in these instances of MTC, the MTC operator and the coerced N are tied together in a single phrase (NP). Another example of MTC within a syntactic relation is the case of the complex NP in (6), of the form N-PP, the PP containing a N that undergoes the type-shift, because the head N apparently has semantic properties that conflict with the N that occurs later in the PP construction. We can assume that the syntactic construction allows these semantic properties to percolate up to the upper node that dominates the shifted N. The last occurrence of MTC in a syntactical environment is when N is an argument of a V that has demands for its θ-roles that conflict with the semantic properties of the default type of N. This can be seen in (7) and (9). Within Adj-N, complex NP's and VP's, the shifting N is related to its trigger, and can be explained in terms of syntax and lexical knowledge. I leave the term lexical knowledge unspecified here, because what is lexical and what is syntax, is theory-dependent and will thus be discussed in Chapter 2.

What is even more interesting, are the examples in which the trigger for coercion is not

related syntactically to the shifting N. This is the case for (8)-(11). I will refer to these triggers as *discourse markers* in the discourse of N. These markers can be *location* (like *room* in (8)), but also certain subjects can influence the type of nouns that follow in a construction. Locations can be seen as discourse markers that have a strong influence on the story that is told. Pragmatic principles of text *cohesion* will probably be responsible for the type coercion that occurs in examples like (8) and (9). Once the story is for example located in a room, any object or event that is in that room must fit into that room. It follows that an elephant or a lion of N_{nat} are too big to fit in a normal room, causing a conflict of meaning that is resolved by operation MTC.[13] Note that in (9), the location marker does not precede the coerced N. I will come back to that later on. In the examples (10) and (11), we deal with certain subjects that seem to function as discourse markers. In (10) we saw that it is more likely for a child to have a collection of toy-cars and not real cars, whereas a man can have a collection of real cars. In these examples, the subject/AGENT of the sentence potentially causes a N with conflicting semantic properties to shift. It can be argued that this happens through the VP, if the θ-role of the first argument is fulfilled, it somehow predicts properties of the arguments of the VP. Leaving the syntactic issue aside, the difference in (10) between *the child* and *the man* can be caused by knowledge of the world. A principle that follows from that is that of *common sense knowledge*. The speaker assumes that the hearer has the same sort of vision of the world, in which a child does not own several real cars, but a man could. Note that common sense is a floating principle. In the highest classes of society it might just as well be that a child has two cars. Example (11) is similar, but differs on one important point. The subjects here are referred to with proper names that are assumed to be known. They are iconized persons, that seem to invoke sort of a setting. This setting could predict for example movies (series of images) or a world of puppets (all puppets). Intuitively, this is knowledge that goes beyond the knowledge that is required to interpret *car* in the context of *child* and *man* respectively. The latter are concepts referred to by nouns that require specific knowledge based on the lexical knowledge. In (11) the concepts are referred to by proper names that in the world have a certain status and concepts related to them. We can conclude that discourse markers can have the function of MTC operators and they will coerce nouns that fall under their scope when the default meaning of the noun conflicts with the demands or properties of the discourse marker.

For the examples (12)-(16), we saw that some of the triggers for MTC were somehow stronger than others. For example the PP-NP example resulted in $N_{art/nat} \rightarrow N_{ima}$ shifts in all cases. Nouns also shift into images when they are interpreted in the context of Bert and Ernie. However, the other examples seem to depend more on an interaction between the noun and the triggers. For a gun it is almost impossible to imagine a location in which a gun cannot be a real gun. When I worked out these examples I noticed that in the sentence where I differentiated between man and child as the subject, that the word *collection* triggered MTC as well, especially for N_{nat}. So that is an unexpected trigger, which will be discussed in detail later on. Also in these examples, I found that verbs can shift their arguments into the type that they demand in their argument structure. In fact, the examples that resulted from 'mixing up the data' in section 1.2, affirmed my expectations. What is more interesting, is to

[13]Note that a change of location introduces a new discourse marker, that could prevent for MTC to take place. If a location like room is embedded in a location that does not lead to a conflict with the default type of N, MTC will not take place: *I was working in the veterinarian hospital of the National Zoo at that time. One day I worked with this elephant that had been very ill. When I came into the room, I saw the elephant had fallen down.*

see what kind of interpretations are typical for MTC. Especially the extended examples of 1.2 open up the possibility to generalize the interpretations of MTC in discourse and how they differ from the interpretation of MTC that you can obtain by modifying the nouns explicitly with an adjective. An attempt to generalize the interpretations that are involved in MTC.

At the beginning of this section I addressed the issue of direction of MTC. Although my examples are all (straight)forward, we can assume that there is no restriction for MTC in terms of direction. First, this is supported by the English data in (24):

(24) a. *a lion stone

b. Jenny liked *the flowers* [on the dress]$_F$.

c. I saw *the elephant* had fallen down once I had entered [the room]$_F$.

d. It was *the truck* that I [tripped over]$_F$ when I walked in the garden.

e. *The dinosaur* was very popular in this episode of [Bert and Ernie]$_F$.

It is only (24a) that is problematic in English. English syntax demands adjectives to precede the nouns they modify. Evidently, MTC would work backwards in French for Adj-N constructions, since the French syntactic rule for adjectives is the exactly the opposite. A question that rises here is whether the possible distance between MTC operator and coerced N are equal for backward/forward coercion. Especially in cases where discourse markers function as MTC operators could offer some perspective on this issue: if you tell a story and specify LOC in the beginning, the hearer will assume that unless another LOC occurs, the story continues to take place at the same LOC. It might be the case that a whole series of events takes place between the specification of LOC and the moment a noun needs to coerce to N_{ima} because of the properties of LOC. It would be implausible for the same series of events to be told between the N and LOC when LOC is specified 'in retrospect'. Although on the borderline of pragmatics, this issue will certainly be revisited in chapter 3.

To conclude this section, I give an overview of the MTC operators that play a role in the data of this study, followed by a definition of MTC that I will use as a take off for this study:

MTC OPERATORS

`Adjectives`: semantic properties of Adj and N respectively can force type-shift on N in [A[N]]$_{NP}$. The coercion takes place at the phrasal level.

`Nouns`: semantic properties of N_1 and N_2 respectively can force type-shift on N_2 in [N[P[N]$_{PP}$]]$_{NP}$ and vice versa. The coercion takes place at the phrasal level.

`Verbs`: semantic requirements that V can impose on its arguments can force type-shift on N if the semantic properties of N itself do not answer to those requirements. The coercion takes place at the sentential level.

`Discourse Markers`: the semantic properties of TOPIC$_1$ and LOC$_1$ have scope in a context as long as they are not replaced by any other TOPIC$_n$ or LOC$_n$. When any N that falls under the scope of a discourse marker has incompatible semantic properties relative to the marker, this can result in a type-shift of N. The coercion takes place at the sentential level or beyond the sentential level.

METONYMIC TYPE COERCION

If in any linguistic context C there is an N such that the semantic properties of N are incompatible with those of an MTC operator that N is related to in C:

If N = $N_{nat/art}$,
$F_{\text{MTC}}(N_{nat/art}) \rightarrow N_{ima}$ and
If N = N_{ima},
$F_{\text{MTC}}(N_{ima}) \rightarrow (N_{ima})_{ima}$.

1.3.2 Different interpretations

In this section I will explore the different interpretations that occur with MTC. In other words, the type N_{ima} seems to have a set of possible interpretations. I will describe the interpretations generally and informally, since the sections about different frameworks will all offer their own formalizations of MTC. I will discuss the example in the order they have been described previously, and give a schematic overview at the end. Finally, I will sum the phenomena of MTC and reformulate the main questions of this part of the investigation, as a starting point for the following sections of this chapter.

In the Adj-N examples there are different interpretations of the nouns that have undergone type coercion. In example (1), *lion* is associated with statue or ornament, which I will refer to as STAT and ORNA respectively. I differentiate between those, because in other examples the context can provide information if a *stone lion* is an ornament (small) or a statue (big). In case of (2) the interpretation can only be described in terms of the modifier *fake*. I propose this as a separate possibility for interpretation of MTC: FAKE. In (3) I take the modifying *chocolate* to be an element of the category CANDY, since of all examples of food classes, candy is the sort of food that comes in different shapes (images) most often. For (4), I argue that *wannabe* belongs to FAKE. *Wannabe* means *fake*, but it is applied to concepts that can have intentions, i.e. humans. According to Franks (1995:497), the literal meaning of *fake* can be formulated as '*object the purpose of whose appearance is to make people think it is an N, but without the function of an N, where 'N' indicates the head N that is modified*'. Following this paraphrase, *wannabe* can be formulated as: '*human whose appearance is intended to make people believe he is an N, but who lacks basic features of N, such that people do regard him as a fake N.*' In fact, when someone refers to another person as being a *wannabe* N, he regards that person as a *fake* N, and thinks that person wants to be an N, but isn't. So, being a *wannabe* is being FAKE in the eyes of a second person. Note that both *fake* and *wannabe* have in their definition the predicate of *intended appearance*. The difference between the two is that fake is an objective property, whereas wannabe is subjective.[14]

The interpretation of *imaginary friend* in (5) is hard to define. I could place it under FAKE, but it does not follow from the definitions given above. A *fake friend* would be a person that intends to appear like a friend but truly has different intentions that do not belong to being a friend. An *imaginary friend* may be the most perfect friend that there is, but that is made of due to the lack of that perfect friend in reality. I place this interpretation under FANT, an interpretation of fantasy concepts. Nouns that refer to non-existent concepts often occur in fantasy-stories and sci-fi. In stories that belong to this category, it happens that nouns that in reality refer to real objects, are used to refer to an object that does not exist in reality. They could be idealized versions of that concept. Often the FANT concept shows similar properties to the concept that it is based on. [15] Of course, fantasy concepts also get

[14] See sections 1.1.1, 1.2.1 and 1.2.2 for more discussion on the difference between *fake* and *wannabe*.

[15] For example, take 'My little ponies', where ponies are not real ponies, they have blue hair and pink bodies and they can sometimes fly. In a story, they can be referred to by the noun *pony*, but when reading the story, the noun refers to a non-existent fantasy concept.

other names or show little resemblance to existing concepts, as we have seen in (23) with *unicorn*. We now have four interpretation categories: STAT/ORNA, FAKE, CAND and FANT. Let's see if these will suffice to describe the other examples.

In case of (6a), we need a new interpretation. This interpretation occurs with [NP[P-NP]$_{PP}$]$_{NP}$ constructions that have as their first noun an object that can have pictures printed on it, like dress, sweater, wallpaper, etc. The noun that falls under the PP that is attached to this N, will shift to the interpretation of a picture: PICT. This interpretation can be seen as a two-dimensional version of STAT/ORNA.[16] In case of (7a), we have a TOY interpretation. Toys will often be images of existing things, like cars, trucks, babies, bears. The interpretation of (8a) falls under ORNA, and is similar to *stone lion*, only location *room* seems to offer information that is sufficient to decide that N_{ima} *elephant* is an ornament rather than a statue. In (9a) duck is interpreted as a TOY, as is *car* in (10a). In (11a), the interpretation of dinosaurs is most similar to FAKE, because the intention of Spielberg's movies is to let the dinosaurs appear as full-sized real dinosaurs. In (11b), dinosaurs is interpreted as TOY. In sum, the occurrences of MTC outside constructions with modified NP's offered two new categories: PICT and TOY. We have are left with one example that falls somewhere between Toy and Pict. In (11) I introduced *Bert and Ernie* as a discourse marker. We have seen in section 1.2.1 that in the examples in which this discourse marker is used, some of the objects that fall under its scope change into puppets rather than toys. For example, when Bert and Ernie would have a show with guns, gun would at least been fake. However, when they have a show with punkers or dinosaurs, those would probably appear as puppets, and not as toys per se. This discourse marker is able to shift N_{nat} into N_{ima} with the interpretation of a *puppet*, or more generally, a *puppet*. Although this interpretation shares properties with both PICT, TOY and even STAT, I will regard it as part of the interpretation-class PICT, since this seems the most general interpretation. PUPP is the interpretation that is a 3-dimensional representation of a natural kind that is not STAT, and not necessarily TOY.

Note that in general, when enough information is provided by the context, interpretations might change. For example, a FAKE gun can become a TOY gun, if you provide more information. In other words, interpretations are formed in online-processing, and seem to be flexible. They need to be, otherwise backward type coercion wouldn't be allowed. I conclude this section with an attempt to generalize the interpretations in (informal) descriptions. Note that this is sloppy at the least, the interpretations of the data presented in this study are difficult to distinguish from each other.

MTC INTERPRETATIONS

`Toy`: coercion of N results in the interpretation where N (partially) loses its default function and in some cases its default appearance, and becomes an object for people to play with.

`Fake`: coercion of N results in the interpretation where N loses its functionality but not its appearance.

`Food`: coercion of N results in the interpretation where a kind of food has the shape of N (in most cases this will be candy).

`Stat/Orna`: coercion of N results in the interpretation where a 3-dimensional representation of the appearance of N has a decorative function.

`Pict`: coercion of N results in the interpretation of a 2-dimensional representation

[16] Also interpretations of sentences like 'My grandmother put me high on the wall of her hallway' fall under this category; paintings, photos, pictures, any two-dimensional representation of an N.

of N.
`Fant`: coercion of N results in the interpretation where N ceases to exist in reality.
`Pupp`: coercion of N results in the interpretation of a 3-dimensional representation of N (which is possibly decorative or toy but not necessarily).

1.3.3 Different questions

Now that we have more or less established a definition of Metonymic Type Coercion, illustrated by a variety of examples, it is time to step away from describing this phenomenon and start studying it from different points of view. This will be topic to the next chapter. Based on the data that we have seen, there are many questions that need to be answered. The following will be studied in the sections of chapter 2:

How and to what extend existing semantic frameworks account for Metonymic Type Coercion:

- Which examples of MTC the framework does account for,
- Which examples are impossible to solve within the framework,
- How the framework formalizes the relation between the MTC operator and the coerced N,
- To what degree we can extend the framework to account for examples that would be unaccounted for otherwise.

Chapter 2

MTC in perspective

> *'People say that* ***Orion*** *is called Orion because Orion was a hunter and the constellation looks like a hunter with a club and bow and arrow. (...) But this is really silly because it's just stars, and you could join up the dots in any way you wanted, and you could make it look like a lady with an umbrella who is waving, or the coffeemaker which Mrs. Shears has, which is from Italy, with a handle and steam coming out, or like a dinosaur.'*

Since the aim of this thesis is to look for a framework that accounts for the phenomenon that we call Metonymic Type Coercion, a first step will be to investigate existing semantic theories. This could take a lifetime when you consider the vast amount of semantic literature that has been produced over the last centuries, so my choice has been to use only the frameworks that discuss examples of this phenomenon (either explicitly or implicitly). It has been my choice to use theories that have very different views on natural language semantics. It will be clear that every theory described in this chapter has potential to resolve at least one of the examples of MTC (i.e. one of the constructions in which MTC can occur). It's my purpose that the different ways of resolving the MTC puzzle together form a sort of spectrum. This spectrum shows the rules and restrictions that should be included in a theoretic framework that accounts for all different cases of MTC.

From Cognitive Semantics in section 2.1 I move to Functional Semantics in 2.2. Each section will contain a summary of the theory in general, a discussion about the examples of MTC they account for in a direct way and an estimation of how the theory would account for the other examples. Section 2.3 will describe Generative Semantics, followed by 2.4 which concerns what I refer to as Formal Semantics. The final section will discuss the overall pro's and con's of the different frameworks.

2.1 Cognitive Semantics

In cognitive semantics, language is seen as part of cognition, and not primarily a separate (human) cognitive ability. So, a theory about language is a theory about cognition in a broader sense. Theories of meaning in particular in this framework are, then, not restricted to linguistic meaning alone. Not only do the theories offer a broad perspective on the context of linguistic expressions, they also intend to incorporate non-linguistic and non-'literal' meaning into a single framework. Whether this is possible or not will be no primary point to the discussion, but I will use the examples of MTC seen in this framework as an argument to limit semantics to meaning of language. In this section I will mainly focus on the general

view of Fauconnier, taken from Fauconnier (1994), Coulson&Fauconnier (1999) and Fauconnier&Turner (1996, 1998). In 2.1.1 I will give a short overview of the framework, 2.1.2 will represent the cases of MTC that are accounted for by Fauconnier. Finally, 2.1.3 will be a provisional prediction for other cases of MTC.

2.1.1 Senses and Spaces

In Fauconnier (1995) we find a challenging view on meaning. The construction of meaning is assumed to take place in what are called 'mental spaces'. Before I discuss their applications and operations, I will first give some general assumptions that underlie this framework. Fauconnier criticizes many viewpoints and methods in linguistic in 'common' linguistic research. In order to understand his theory, it is important to know against which common linguistic thoughts he's fighting.

First of, Fauconnier (1994:xviii) states that meaning is not part of language; language does not do 'cognitive building', nor does it carry meaning, it just provides minimal information to find the appropriate domains and principles in a given situation. The problem of most traditional approaches in linguistics is that meaning is analysed of isolated sentences, without (non-)linguistic context. The mistake lies in the fact that those observations of minimal context cannot be generalised to larger contexts. The rejection of the idea that sentences form propositions (as is assumed in most logical views) is subsumed by this. Finally, semantics is often approached truth-conditionally. Truth-conditions are problematic in a lot of cases of every day language-use, like metaphors, irony, utterances about beliefs and modality, because meaning is taken to be 'literal'.

In the framework of Mental Spaces, meaning is constructed in a discourse construction. Sentences are thus parts of discourse constructions, and are not viewed in isolation. Mental space configurations are not models of the ('real') world but they are mental models of discourse. According to Fauconnier(1994:xxiii) a discourse construction does not solely depend on linguistic information, it includes information from various 'grammatical devices':

> Grammatical devices in Mental Space Theory
> `Space builders`: typically express information regarding what new spaces are being set up.
> `Clues for accessibility`: what space is currently in focus, what is its connection to the base and how accessible is it. Typically expressed by grammatical tenses and moods.
> `Descriptions of new elements`: introduction in discourse, possibly with their counterparts.
> `Descriptions, anaphora` or `names`: identify existing elements in discourse.
> `Syntactic information`: typically sets up generic-level schemes and frames.
> `Lexical information`: connects mental space elements to frames and cognitive models from background knowledge; structures spaces internally by taking advantage of available prestructured background schemes. These can be altered or elaborated within the construction under way.
> `Presuppositional markings`:allow some of the structure to be instantly propagated through the space configuration.
> `Pragmatic and rethoric information`: typically signal implicit scales for reasoning and argumentation.

Although some of these 'devices' are very questionable, I will confine my criticism to this approach to the points that are problematic in solving the MTC puzzle. A clear advantage of Fauconnier's idea is that the theory will include all factors that might play a role in meaning construction. His criticism on current theories makes sense, in semantics it is no exception that meaning is analysed out of context of utterance, for the simple reason that a linguist does not always have access to the context of the utterance. Another advantage is that this system could account for semantic relations that lie beyond syntax; a notion like 'space builder', however vague, allows us to relate linguistic elements that influence one another in a bigger discourse. A last general remark would be that in Fauconnier (1994:xxxix), meaning is dynamic: '*Provisional categories are set up in appropriate spaces, temporary connections are established, new frames are created on line, meaning is negotiated.*' I will turn back to the idea of 'meaning negotiation' in Chapter 2.

Fauconnier (1994) does not explicitly treat the phenomena of MTC, this is done in Fauconnier&Turner (1996,1998) Coulson&Fauconnier (1999), with the notion of 'conceptual blending'. I will turn to that after a short overview of principles that account for other semantic puzzles, since these are the general principles that form the basis of later theories of blending. A central principle in Mental Spaces is the Identification Principle described in Fauconnier (1994:3):

> IDENTIFICATION (ID) PRINCIPLE
> If two objects (in the most general sense), a and b, are linked by a pragmatic function F ($b = F(\mathrm{a})$), a description of a, d_a, may be used to identify its counterpart b.

ID principle is a pragmatic function that can establish links between objects that are different in nature. This can be for psychological, cultural or 'locally pragmatic' reasons. Function F in ID subsumes a diversity of functions, like from person to body and from persons to their name. ID makes it possible to identify the 'target'; a description of a trigger (for example a proper name) can identify an object (the person) in a connected situation. A good example of application of ID is given by Fauconnier (1994:11-12) in an account for images: *In Len's painting, the girl with blue eyes has green eyes* and *Lisa saw herself in Len's picture.* In the model of Mental Spaces, it is possible to refer to an object in terms of its image. The cases in (18) are nonetheless different from the previously described examples of MTC. For instance, the relation between a person and a picture of that person can be explained in terms of an identity relation (by way of ID). Fauconnier's framework allows the target (the girl) to be identified in terms of the trigger's description, in other words: the 'real girl' is identified in terms of a description of the image. The representation of this sketched by Fauconnier are circles that represent mental spaces, with arrows interconnecting them that represent (pragmatic) functions. For details I refer to Fauconnier (1994). This representation is quite informal, and does not contain any semantic features, as described by Fauconnier&Turner (1996:113): 'Mental Spaces are small conceptual packets constructed as we think and talk, for purposes of local understanding and action.'. The aim of this framework is not to establish a theory of conceptual or lexical meaning, but to account for the construction of meaning based on knowledge, background frames and relevant pragmatic functions or connectors.

A few more issues concerning Mental Spaces need to be addressed before I explain its

application to MTC. Whereas other model-theoretic semantics base their notion of truth on a model, 'reality' in Fauconnier's sense is taken to be the speaker's reality and is therefore assumed to be a mental representation. Linguistic phenomena are then accounted for by links between mental representations. Spaces are set up by 'space builders', expressions that may establish a new space or refer back to one already introduced in the discourse. Mental spaces are structured sets with elements and the relations that hold between the elements. They are flexible; new elements can be added during meaning construction, and new relations can be established. Mental spaces are related by connectors (functions). For example, image connector F_i connects models to pictures and vice versa. Returning to the notion of space builders, Fauconnier (1994) provides a list of linguistic elements that typically function as space builders. Temporal adverbials are typical space builders for 'time space'. This allows a referent in another time than the time of utterance to identify its target at the time of utterance. The 'space space' works more or less the same. I will elaborate this some more, since this allows for locations to influence their 'occupants', which will be useful in accounting for the example of MTC in (8) and (9) as I will show later. Fauconnier (1994:30-31) gives the examples *In Moldavia, the president is a tyrant* and *In the other apartment, the chandelier looks fine.* It is argued that *in Moldavia* sets up a space, which forces the definite N *the president* to be interpreted as 'the president of Moldavia'. In (19a), *the other apartment* allows for an ambiguity, it could be that the chandelier of the other apartment looks fine or the chandelier looks fine when it is in the other apartment. Other spaces are the *domain* space, which accounts for (non-contradictory) interpretations that are derived based on the knowledge of a certain domain, like Fauconnier (1994:31) illustrates by '*In Canadian football, the 50-yard line is 55 yards away*'; the *hypothetical* space, accounting for linguistic forms as 'If p then q' and presuppositions. Finally there are *tenses* and *moods* that do not set up spaces, but give grammatical clues about spaces for the processed sentence.

All of this is a general overview of how the framework of Mental Spaces works and on what philosophy it is based. In this framework, although not explicitly in Fauconnier (1994), a lot of work has been done on MTC. This section was necessary to represent the basis from which the notion of 'conceptual blending' has been developed; the topic of the next section.

2.1.2 MTC in this framework

The examples *stone lion* and *fake gun* are treated explicitly in Coulson&Fauconnier (1999). The mechanism that accounts for the interpretation of these examples is called Conceptual Blending. This mechanism is described extensively in Fauconnier&Turner (1996,1998). This section will be a description of the analysis of these examples within Conceptual Blending; first I will summarize its fundamental properties.

Conceptual blending is seen as a general cognitive process, to which mental spaces function as inputs. Mental spaces, then, can be seen as conceptual packets that we use to think and talk.[1] The process of blending consists of two input spaces, say x and y, which are connected to what is called a 'generic space'. The generic space maps onto each of the inputs. In the representation x and y are circles that contain semantic properties. These properties are not represented in any kind of feature-structure, like they are for example in the work of Franks and Pustejovsky (see sections 2.2 and 2.4). They resemble most to the kind of structures we see in Discourse Representation Theory (DRT), however, the representations

[1] See section 2.1.1 for more explanation about this.

of mental spaces and operations over mental spaces remain unexplained in both Fauconnier (1994) and Fauconnier&Turner(1996,1998). The generic space is supposed to contain what the inputs have in common. The fourth space that is part of the process of conceptual blending, is the the blend itself. In the blend, x and y are 'mapped onto a single slope', as Fauconnier&Turner (1998:8) put it. The emergent blend partially inherits the structures from x and y respectively, but has an emergent structure of its own. That is, the blend space functions on its own. Dependent on the usage of a blend in everyday language-use, the blend can become conventional. I gather that these conventional uses of blends are part of the idiomatic expressions that a language contains. This basic system of conceptual blending is able to account for a lot of linguistic peculiarities, like metonymy, metaphors, etc.

Before I give details about the application to MTC, I will discuss some difficulties in this framework that occur even before we look at specific cases. Note that in the blend counterparts of x and y can be fused, however, not all of the elements of x and y need to be fused, as stated in Fauconnier&Turner (1996:115): '... any two counterparts may or may not be fused in the blend'. This means that if x and y for example both contain an element of 'speaking' (for the example of 'debate'), this element is projected into the generic space, and later fused in the blend. It remains unclear why elements that x and y have in common shouldn't be fused, and what this fusion entails in the first place. The articles seem to be built on these kind of 'weak propositions', they state that it is possible that some elements are not fused, and others are, without an attempt to generalize over this difference in behavior of the elements. Another weakness is the lack of generalization over concepts or elements that are suited for blending. In Fauconnier&Turner (1998:157) it is stated that blending is non-arbitrary: '... not just any discordant combination can be projected in the blend'. I am aware of the fact that I quote this out of context, but the article does not really provide clues for the limitations of the system, while it suggests they *do* exist. Other disadvantages of blending follow from the Mental Space theory in general, and will be discussed in section 2.1.3.

Coulson&Fauconnier (1999) is merely a reply to the framework of Franks and his proposals regarding MTC. I will abstract the framework from the critics Coulson and Fauconnier give to Sense Generation (see section 2.2). This article is focused on the 'classical' examples of MTC: *fake gun* and *stone lion*. The article compares the treatment of Franks with a treatment within the Mental Space framework. Since the example of fake gun is treated most extensively in this paper, I will start with that. In Coulson&Fauconnier (1999:117), blending is summarized to the following processes:

> Processes in Conceptual Blending
> `Composition`: attributing a relation from one space to an element or elements from the input spaces.
> `Completion`: pattern completion which occurs when structure projected inputs matches information in the long-term memory.
> `Elaboration`: involves performance and/or mental simulation of the event in the blend.

Conceptual blending is said to be rooted in the imaginative capacities of the speakers, suggesting that concepts are very flexible. The process of blending relies on the cognitive ability to make cross-space mappings, where it is possible that the properties of the elements that are linked differ. In case of *fake gun*, Coulson&Fauconnier (1999:151) argue that the adjective fake is a space builder, that '*prompts a mapping between an actual scenario in which the actor*

employs the fake gun, and a counter factual scenario in which his audience reacts as if it were a real gun'. The result of this view is that the interpretation and representation within blending and Mental Space Theory is highly dependent on the context in which the expression fake gun is used. However, it is assumed that a basic prototypical meaning is present, and that speakers/hearers use that information in the absence of particular knowledge in the domains. How these prototypes of concepts are to be seen, is not clear, the representation of the blend space (the 'Conceptual Integration Network') does not contain any specified semantic features, as I remarked previously. The construction of meaning is knowledge rich and flexible. In this respect, the 'definition' of *fake* cited above, is at the very least questionable. This definition already contains a certain context, namely *a robbery* (as is used in the article). In this particular context, a fake gun is intended to be perceived as a real gun. But as I argued in section 1.2.1, a fake gun can also be interpreted as TOY, which could be scarcely discernible from a real gun, but also be a yellow water pistol. The interpretation of *fake gun* according to this definition is not represented in a conceptual blend the way it is done in Fauconnier&Turner (1996, 1998). That is, the input spaces are not *fake* and *gun* respectively, but *Actor's Knowledge* vs. *Victim's Belief*. The problem of this representation is that this is not the meaning of *fake gun*, but the meaning of the situation. The question is if there is an individual who will grasp this meaning. In other words: in whose head are we looking when we represent the meaning of *fake gun* from both the Actor and the Victim? The victim thinks the gun is real, at least that is the intention, but how can the Actor know so? Another problem is that the model as drawn in Coulson&Fauconnier (1999:151), does not represent the processes of blending that are stated before. The core of blending is the 'derivation' of the generic space, which contains abstract properties that both the input spaces have in common. It seems impossible to determine what the Actor and Victim have in common, other than some sort of gun (which is exactly the problem). In sum, I claim that blending in this manner offers no solution for MTC interpretation in *fake gun*. If we adjust the model with input spaces *fake* and *gun*, blending becomes less of an option. After all, how could a generic space ever be derived for *fake* and *gun*? I will argue later on that it is for this reason that we have to approach MTC in a linguistic framework that takes into account both the contextual factors and the syntactic/semantic relations that are involved in the process of MTC.

Coulson&Fauconnier, *gun* in the frame of 'plays' is also elaborated. In this case, they argue that in a play the pointing of the index finger can be perceived as sort of a fake gun ('a finger gun'). I argue that this usage of the concept gun falls outside the boundaries of MTC. The 'finger gun' is a gesture, and I doubt that anyone would actually call this a 'finger gun'. This gesture does relate to the concept of *gun*, evidently, but this relation is symbolic, and lies in the field of semiotics rather than linguistics. In fact, the latter can be seen as a general critique on the framework of Mental Spaces. Its intention being to account for meaning in the broadest sense of the word, the theory becomes too sloppy to account for phenomena that are linguistic. In other words, in my opinion there should be a boundary between which meaning is linguistic (expressed by language and the context of the expression) and which meaning comes from a non-linguistic context. Of course, hereby I attack the core of Mental Spaces (and other examples of cognitive-semantic theories), and this can raise the question why I want to use this theory in the first place. My purposes to use this framework to some extend, will become clear later on in this study.

The case of *stone lion* is merely used as an argument against the usage of features in the semantic representations. Coulson&Fauconnier (1999:155-6) recognize the point that is made in Franks (1995) that stone in stone lion is a 'functional' privative (see section 2.2.2),

and argue that '*understanding functional privatives is not as much a problem of semantics as of understanding the representational conventions that allow us to map between spaces with different degrees of iconicity. Thus by positing general blending mechanisms constrained by conventions of representation, we make a special feature mechanism superfluous.*' Directly after this, they claim that blending provides a better account of the way *features* accommodate, and thereby refer to the work of Langacker. The problem is, that the representations used in the work within Mental Space theory has no clear way of specifying semantic features, at least this is not done in a consistent way. The case of *stone lion* is considered to have both a possibility for a privative and a non-privative meaning. In their view also the metaphorical and metonymic usages of *stone lion* are considered to be instances of privative meaning. Although this is arguable, I will restrict this thesis by limiting the scope to the interpretations that fall inside the scope of MTC, that is the type shift from N_{nat}/N_{art} to N_{ima}. One of the strong points of the Mental Space theory is that it can account for meanings of metaphors and metonymia, this is at the same time its weakness; for the same reasons I gave for meaning phenomena that are otherwise treated within semiotics and not semantics. Although Coulson&Fauconnier do not discuss stone lion as extensively as fake gun, it can be inferred how the Mental Space theory would account for the different meanings by blending. The representation would consist of two input spaces, for *stone* and *lion* respectively. Deciding which properties will appear in the generic space already offers us some problems that I encountered in the case of *fake gun*. I assume that Coulson&Fauconnier would not consider *stone* to function as a space builder like *fake*. After all, the adjectival use of *stone* does not necessarily entail a privative effect on the noun it modifies. Coulson&Fauconnier argue that *stone lion* can be interpreted in so many ways that just the nominal compound only offers minimal clues for its meaning. This is due to the 'contextuality of interpretation' (Clark, 1983), but seems a redundant point in this framework. I claim that *stone lion* will of course be interpreted differently depending on the context, but the modified nominal will have a basic meaning. Only if the context tells us otherwise, *stone lion* will be interpreted in this meaning. If we take that meaning within the framework of Mental Spaces, we can try to establish a blend for *stone lion*. Suppose *stone* and *lion* are respective input spaces. The conflicting properties of being a living natural kind (lion) and a non-living material (stone) cannot be projected into the generic space and will be unrepresented in the blend space. However, a lion will be made of a certain substance. The substance of stone will be specified as stone, the substance of the lion is difficult to specify, let's call it 'flesh'. In my view, it is difficult to see how the generic space will provide substance as stone for stone lion, and not 'flesh' as the blending theory says nothing about the priority of properties of one input space over the other.

So, in order to use the blending mechanism for *fake gun* and *stone lion*, we will need a better algorithm for the blending process. In my opinion, the relation between modifier and head must be regarded somewhere in the blending process. Next to that, in case of the modifier *fake*, I suggest that representing meaning from Actor vs. Victim is impossible. However, the thought that *fake* is a space builder (i.e. a sort of operator) is very plausible and will useful in solving the MTC puzzle. *Fake gun* is then not subjected to a process of blending, but is interpreted under the scope of space builder *fake*, i.e. the result would be a single space for *gun* that is set up by *fake*.[2] For *stone lion*, we need to find a better mechanism that leaves

[2]It would be challenging to work out a new way to represent the meaning of fake gun within Mental Space theory, but this is not the aim of this thesis. The aim is just to create a framework for MTC cases while adopting useful material of existing theories.

no possibility open that the substance of *stone* will not be inherited in the blend space.

2.1.3 Extension to other cases of MTC

The examples *stone lion* and *fake gun* are treated explicitly in Coulson&Fauconnier (1999), and the problems that come with the application of blending can be predicted for the other adjectival examples that are topic to this thesis as well. As I have noted before, the problem with blending is that the process does not give information about priority of properties from the input spaces that are projected in the blended space. In this section I will not dwell on this point, but give a basic idea of how the framework of Mental Spaces can be applied for other cases of MTC. Although Fauconnier (1994) does not refer to MTC explicitly, the principles of space-builders potentially offer us some tools to solve cases of MTC. As I have described in section 2.1.1, the ID principle allows us to refer to an object in terms of its image. This principle directly maps an image to its original, and is a useful basic assumption in a theory of meaning that we need in order to account for MTC. In chapter 1 I have argued that all cases of MTC contain a type-shift from N_{art}/N_{nat} to N_{ima}. The first are original objects, the latter their images. In this section I will try to find out if the examples in 1.1.2 can be accounted for by the model of Mental Spaces. We will see to what degree we can use the notion of Conceptual Blending (combined with the basic principles of Mental Spaces).[3]

Example (6) shows a case of MTC triggered in a PP-NP construction. The type-shift in this example could be explained by Mental Spaces using Conceptual Blending. However, we need to modify the mechanism of Conceptual Blending, since this is not a Adj-N construction. Since the framework of Mental Spaces and Conceptual Blending are barely restricted by rules of (generative) syntax, it is easy to integrate a different construction than N-N. We could say that the objects *rose* and *dress* in this particular construction blend into 'a dress that has images of roses on it'. *Rose* and *dress* form two input spaces, that project properties onto the generic space. Here we encounter the same difficulties as in the case of *fake gun* and *stone lion*, discussed in 2.1.2. But let's assume we blend the properties of rose and dress. *Dress* could, for example, project a property of being a piece of garment, and rose the property of being a natural kind flower. The knowledge of the world that is presumed to be present in the mind of the hearer, should include that dresses can have images of objects on them. The model of Mental Spaces is especially suited for including background knowledge into the account for meaning. This way, the information in the generic space would be based on input from not only the meanings of the words respectively, but also on the knowledge that dresses can have images of flowers on them. This would be an easy way out, since anything that cannot be explained by the information of the input spaces, could be explained by common knowledge. The question is, when does this knowledge come to play and how? As stated before, the model of Mental Spaces does not address questions of this type.

In example (7) we deal with the requirements that the verbs *to trip over* and *collide with* put on their arguments respectively. Although Fauconnier (1995) does not elaborate on them, verbs are seen in the traditional way: they establish the relation between different elements. In the framework of Mental Spaces, that means that these relations hold between space elements.[4] For example, *think* in 'he thinks...' is assumed to be a connector from reality to beliefs. This connector can clearly create mental spaces in this framework. However, the

[3]The problems of Conceptual Blending will not be elaborated in this section, but of course what was discussed in the previous section, also holds for the examples here.

[4]see Fauconnier 1995:17

problem is not words like *think*, *believe* and *say*. Where these verbs have propositions as their complements, the verbs in my examples connect objects to each other. It is clear that verbs that take propositions as complements, are suitable as space-builders in this theory, but it is unlikely that verbs that relate objects to each other can be explained in the same way. Verbs like *to trip over* and *to collide with* are not discussed in the theory of Mental Spaces. The theory doesn't make any assumptions about verbs and their arguments, let alone about a (formal) relationship between them. However, within this framework one might propose that for example the meaning of (7) is the result of Conceptual Blending. In this line of reasoning, we would have to treat verbs as concepts. In Conceptual Blending, concepts are assumed to be very flexible. The meaning that results from putting together two concepts is a blend in the most literal sense in the examples that are discussed in Fauconnier&Turner(1996,1998). Both elements that are blended lose some of their semantic properties. If we would assume that the meaning of *to trip over* in a context where *truck* is its argument, does result from such a blend, we would leave open the possibility of the verb losing semantic properties. Which is exactly the opposite of what is happening. As we have seen in the discussion in chapter 1, verbs are able to force their arguments into another type *because* they are not able to make concessions on their meaning. In other words, the requirements of a verb are absolute, and could be more easily explained by for example formal θ-roles. In the process of Conceptual Blending those requirements would lose their rigid force, which is exactly why we call it type *coercion*.

Mental Space theory offers more potential in case of (8), where we deal with a location that triggers MTC on an element that is part of that location. Fauconnier (1995:17) argues that locations can operate as space-builders. In his model, a space-builder builds the 'parent space'. Space-builders are 'expressions that may establish a new space or refer back to one already introduced in the discourse'. In this case, the location is *the room*. All events and object that follow, fall within this space, until a new space-builder is introduced in the discourse. Turning to the fallen elephant, based on Mental Spaces theory, we can only derive the N_{ima} interpretation if we assume that this is part of common knowledge. That is, we share the knowledge that a fallen elephant in a room will more likely be an ornament in the shape of an elephant than it will be a real elephant. It would be possible in the framework of to associate *room* with ornaments (ORNA), in order to derive the interpretation of *lion* in (8). However, we would still have to deal with the knowledge-problem I have discussed in the previous section. In Mental Spaces, there seems to be no categorization of knowledge. For the part of the theory that deals with what I would call linguistic meaning (semantics), there is no assumption of a separate sort of knowledge, like 'lexical knowledge' or 'conceptual knowledge'. This results in a theory for meaning that is not based on any features or other generalized forms of semantic properties. Suppose we try to solve it without the features and only by assuming 'rich and flexible knowledge' and 'pragmatic functions and connectors'(see 2.1.1). This could be successful for room; within this framework it functions as a space-builder and common/background knowledge can tell us what we consider as parts of a room (such as ornaments). But how do we account for an elephant becoming an ornament in (8)? Elephant falls under the space of room, but how does that make elephant an ornament? If we replace elephant by goldfish, there is no such interpretation. In other words, the idea of making locations and certain subjects space-builders (and thereby pragmatic markers) is useful, but useless when we cannot ascribe any general semantic properties to the linguistic elements that fall under the scope of these space-builders. So for MTC, we need more than just the notions of space-builders and the magical human knowledge to fill in the gaps.

Example (9) is problematic for the same reasons we saw before: here we have a location and a verb that influence the type of the noun *the duck*. We see here that verbs are stronger than locations. Mental Spaces would offer a way to deal with the location, but this case shows that we need to open up the possibility that a verb cancels the influence of the location. In Mental Space theory there is no ordering of the possible space-builders (or MTC triggers in this context), which is a tool we need in dealing with MTC.

In the two final examples, we see that the subjects of the sentences can alter the type of other elements in the sentence. Mental Spaces can handle these examples more or less in the same way it can handle the cases of MTC triggered by locations. In (10) we see that the jaguar is most likely to be interpreted as N_{ima} if the possessor is a child. However, as I stated before, this is not a strict case of MTC. That probably why Mental Space theory is able to account for this example. *The millionaire* sets up a mental space with the elements associated with being a millionaire, and the jaguar is real. *The child* sets up a mental space with elements that belong typically to being a child, in which case the jaguar is most likely a toy. But not necessarily, if the speaker and the hearer would both know that in case of this particular child the jaguar is actually real, this would be part of the shared knowledge that overrules 'common' knowledge. In (11) we see that Bert and Ernie and Steven Spielberg both coerce dinosaurs into N_{ima}. Also in this case, different spaces are set up that make it possible that dinosaurs are interpreted as replicas of dinosaurs (Spielberg) or puppets with the appearance of dinosaurs. All this is possible, because the Mental Space theory allows for common knowledge to create spaces that function like 'interpretational frames'[5], in which it is possible for instance that all participating animals in Sesame Street are automatically assumed to be puppets (N_{ima}).

In general, the theory of Mental Spaces and the related method of Conceptual Blending give us some useful insights about the extend of the theory that we need to develop for MTC. It is obvious that there are a lot of problems that we encounter in this framework, this merely has to do with the fact that the theory is intended to account for meaning in the broadest sense. In my approach I will use the insights of making location and certain subjects space-builders, by stating they are functors that can trigger MTC. Other factors can be explained in more formal terms, like verbs and the Adj-N constructions. This is the weakness of Mental Spaces, by assuming that all meaning can/must be accounted for in one single theory, the theory becomes to weak and knowledge-rich for cases in which the meaning can be explained in strictly linguistic terms. This is nicely illustrated by the examples in which MTC is triggered by a verb. This can easily be explained by lexical knowledge about the kind of arguments a verb selects, but since Mental Spaces is focused on meaning on a wide scale, the (relatively) simple linguistic explanation is overlooked. From this section it should be clear that I work with different frameworks like a opportunistic parasite: my aim is to see to what extend the theory can offer us tools to resolve the MTC puzzle. Although I will refrain from deep criticism, the main points of my disagreement will be explained.

2.2 Functional Semantics

This section will be focused on a framework that is offered by Franks (1995). This paper offers a framework that is generative for deriving multiple senses of lexical items. The theory is

[5]I refer to this as 'frames' because this term covers best the intended meaning. This has little to do with Frame Semantics introduced by Charles Fillmore.

based on a functional system that accounts for cognitive processing of semantic constructions. I will give a short summary of the basic idea that is the foundation of the system of Sense Generation in section 2.2.1, and move on to the ideas it offers about MTC in 2.2.2 In 2.2.3 I will attempt to extend this framework for my own purposes.

2.2.1 Features and Senses

As Franks (1995) is mainly focused on examples of MTC in Adj-N constructions, this general abstract will be short. The framework is based on the idea that meaning is generated from lexical items that have default content. The approach takes lexical concepts to be stable; the information that is part of a lexical concept is supposed to be located in the mental lexicon. To account for the flexibility and context-sensitivity of word meaning, Franks states that from every lexical concept, different senses can be generated. This is a process that is bottom-up, and generative in the general sense. In principle, an infinite set of senses could be generated from the lexical concept. If there where no restricting factors, every word could be associated with every possible meaning (any sense). To preclude this, Franks argues that pragmatic and communicative factors put constraints on the senses that are generated. These constraints work top-down. In order to prevent that first an infinite set of senses are generated, and then are restricted by the pragmatic constraints, Franks proposes that the bottom-up rules interact with the top-down constraints.

A lot of attention in this article is paid to the notion of concepts and classification. As this will be discussed extensively in chapters 3 and 4 of this thesis, I will restrict myself to the basics of Franks' view here. Lexical concepts are viewed in a 'quasi-classical' way. Conceptual contents are sets of single-valued attributes. A structure for conceptual content is described in terms of AVS: Attribute Value Structure. Although the attributes are single-valued, they are not translated in terms of necessary and sufficient conditions for category membership (like the classical view does). The generation process might alter the values in the AVS, when different senses for a lexical item are derived. In this way, context-sensitivity is accounted for without having to adopt a fuzzy-set theory for category membership, which here is regarded as binary due to the way AVS's are built.[6]

The last notion about concepts concerns one of the central claims of Franks. In an AVS, there are two kinds of properties that are considered to be separate: CENTRAL and DIAGNOSTIC features. Franks (1995:444) describes them as follows: '*Observable diagnostic properties are the basis for (non demonstratively) inferring the presence of (usually non observable) central properties, which allow the classification of an object as being of a given type.*' The distinction between these properties or semantic features plays an important role in how this framework accounts for concept combination in general and MTC in particular. The AVS forms a concept representation and forms the basis of later discussions about classification, since the classification of a concept makes use of its representation. A concept's AVS forms its *default* content, which in Franks' opinion is a stable but defeasible set of properties. Because this content is sensitive to the context in which it occurs, the AVS of the content can change during the generation of a sense. The generation is an operation over AVS's. Although Franks claims that this approach is motivated by the psychology of concepts, it does not say that the formal language of the conceptual contents is identical to the mental representations or the algorithms that can combine them. The framework of Sense Generation is a specification

[6]For further discussion about the way this is supposed to function theoretically, and criticisms on these viewpoints I refer to Chapter 3 and 4.

of the relation between different contents.

Before we turn to the application to MTC, I will first show how sense generation works, and how AVS's are construed in this framework. Franks (1995:456-7) proposes a sense generation system consisting of a triple, and an AVS of a quadruple:

> SENSE GENERATION SYSTEM
> ⟨AVS, R, START⟩, where AVS is a set of attribute-value pairs, R is a set of concept specification and combination rules and START is the starting constituent (the first lexical item processed).
>
> ATTRIBUTE VALUE STRUCTURE
> AVS = ⟨ATTR, VAL, LAB, δ⟩, where ATTR is a set of atoms called attributes, VAL is a set of atoms called values, and LAB a set of atoms called labels, where ATTR, VAL, and LAB are disjoint. δ is a partial function such that δ: LAB×ATTR→LAB∪VAL.

An AVS in this framework looks as follows:

$$\begin{bmatrix} \text{CENTRAL} & \begin{bmatrix} \text{A}: & \text{p} \\ \text{B}: & \text{q} \\ \text{C}: & \text{r} \\ \text{D}: & \text{s} \end{bmatrix} \\ \text{DIAGNOSTIC} & \begin{bmatrix} \text{K}: & \text{w} \\ \text{L}: & \text{x} \\ \text{M}: & \text{y} \\ \text{N}: & \text{z} \end{bmatrix} \end{bmatrix}$$

It should be noted here, that the values of attributes, for example, B has value q, are not specified according to a binary system. In the AVS's that will be discussed later, we shall see how 'values' are defined in this framework. When concepts are combined, the AVS's of the concepts are combined, in order to derive one sense. One sense cannot have incompatible features, which yields that the AVS of a sense that results from concept combination, can only be derived by certain operations. In the framework that is proposed by Franks, those are unification, priority union and type coercion. Franks states that any theory that tries to establish an account for concept combination, should be able to account for MTC, by means of a test case. How this works out, is to be discussed in the next section.

2.2.2 MTC in this framework

Main focus of this article are examples of MTC that are triggered by adjectival modification of N. The examples that are discussed are *stone lion*, *fake gun* and *apparent friend*. This summary will be restricted to the first two, since those are examples that are included in this thesis.

In Franks' view, adjectives like *fake*, *stone* and *apparent* have a privative meaning. FAKE is a proper negating privative, stone a functional negating privative and apparent a proper equivocating privative. Proper privatives will have the effect of a privative for any N they modify, whereas functional privatives are only privative when they conflict with the features

of the N they modify.[7] Negating privatives lead to MTC_R[8], by resulting of a failure of the inference from an entity's being of privative NP its being of head N type. In other words, the central features of the head N are negated. Equivocative privatives lead to MTC_U[9], the central features become unspecified after MTC_U. Franks (1995:451-4) describes the process of concept combination and classification in five principles, summarized below:

> Principles of Concept Combination
> P_1: `Compositionality`: The sense associated with an expression is a function of the lexical concepts of its parts and the way they are combined.
> P_2: `Default Content`: Unless there is a good reason, the content represented in a constituent's lexical concept will also be represented in any sense generated for that constituent or for a phrase of which the constituent forms a part.
> P_3: `Partial Information`: Unless there is a communicative or cognitive demand for additional information, the sense for an expression will comprise the minimum information necessary to support reference and communication.
> P_4: `Default Classification`: Unless there is information to the contrary, the classification judgement expressed by associating a referring expression with a referent indicates a default, general perspective comprising both central and diagnostic attributes.
> P_5: `Perspectival Classification`: Any divergences from default classification will involve a judgement under a perspective comprising the minimum information necessary to support reference and communication, unless there is additional supporting information available or inferable from the context.

Note that Franks' idea is that when interpreting a sentence, you do not use information other than the default content, combined by compositional rules, unless this leads to a conflict in interpretation or insufficient information for interpretation. Note that the notion of default content remains undefined here. Although Franks assumes that (default) lexical content can be divided in central and diagnostic features, it is not clear how many features are included in the default content. The consequences of this indistinction will be discussed in 2.2.3.

In order to account for the functional privative *stone* in *stone lion*, Franks assumes that in this case, the generation mechanism can use knowledge that is otherwise unaccessible. This is by way of what he calls 'implicit semantic attachment'. This process instantiates a noun that is understood, for example statue or ornament. This is possible because of the semantic relation 'head MADE OF modifier'; a lion made of stone is a statue or ornament. The latter process is called Noun Instantiation (NI). Respecting P_2, this is possible only if there is a demand for it. The context will provide further information to interpret *stone lion* as statue or ornament. For *fake gun*, possible NI's are *replica*, *toy* or *model*. NI's are further elaborated by adopting some of the proposals of Landman (1986). I will not complicate matters by dwelling on this analysis, but rephrase that part to the core points. One basic assumption is that senses have different roles in the discourse. The head N and (possibly a) NI have different roles and types.[10] When the type and role of the head N do not make sense in a certain context, a different classification takes place by assuming an NI that can adopt the type and

[7] Stone is 'functional' because in examples like *stone house*, it has no privative meaning.
[8] MTC with rebuttal
[9] MTC with undercutting
[10] The notion of type also remains undefined in the proposal.

role demanded by the context. In case of NI in a modifier-head relation with a privative, the N can be interpreted according to a privative classification (PC), opposed to the (default) affirmative classification (AC). AC and PC can be related to P_4 and P_5 respectively. Next to principles and categorisations that distinguish between default and non-default interpretation, Franks (1995:467-8) distinguishes between to types of perspective, type I and II:

> Types of Perspective
> `Type I`: based upon all of the central and some of or all of the diagnostic attributes of the categorising N type, under which the categorised referent is construed as being of the categorising N type in the general sense. This is implicit in AC and governed by P_4.
> `Type II`: restricted categorisation in which the criterion of indiscernibility does not include central features of the categorising N type. This is implicit in PC and governed by P_5.

The notion of perspectival relativity makes it possible to adopt a binary (i.e. non-fuzzy) classification system for concepts, which will be discussed in chapter 3 and 4. We have seen a general picture of how the framework is founded, I will demonstrate shortly how it works for fake gun and stone lion.

Starting with *fake gun*, *fake* is assumed to be an MTC_R operator; it coerces head N AVS, and ascribes the output of coercion to NP sense. Processing of interpretation is approached in a functional way by representing different information states. I restrict myself to Franks' information state 1, and will describe later information states. This is what information state 1 of *fake gun* looks like:

$$\begin{bmatrix} \textsc{Central} & \begin{bmatrix} \neg\text{fires} : & \text{bullets} \\ \neg\text{made of} : & \text{metal} \\ \neg\text{mechanism} : & \text{explosion} \\ \neg\text{function}: & \text{kill} \end{bmatrix} \\ \textsc{Diagnostic} & \begin{bmatrix} \text{trigger} : & + \\ \text{size} : & \text{portable} \\ \text{barrel} : & + \\ \text{handle} : & + \\ \text{colour} : & \text{grey} \end{bmatrix} \end{bmatrix}$$

In this information state, the NI is not specified. The context will provide the information for specific NI, by demanding for a certain role. Until that information is provided, the NI can be any NI that is associated with modifier *fake*. In this view, the lexical concept of the NI can be regarded as relational concepts. Once the context has provided the necessary information for NI, the central features of the AVS of fake gun are extended by inserting the diagnostic features of *gun*, and the aim of deception (central feature of NI *replica*). The diagnostic features are extended by inserting the *negation* of the central features of *gun*. The latter seems peculiar, since one of Franks' central claims is that there is a distinction between central and diagnostic features. To insert the diagnostic features of gun as central features for fake gun seems plausible, because the function of a fake gun is to appear as a real gun. It then seems unlikely that the negation of central features of a gun are diagnostic

(i.e. observable). I will come back to you on this later.

For the example of *stone lion*, the process is different. Since stone is a functional privative, it is not inherently a MTC_R operator. The operation of information state 1 is as follows:

$$\begin{bmatrix} \textsc{Central} & \begin{bmatrix} \text{organic} : & - \\ \text{animate} : & - \\ \text{solid} : & + \end{bmatrix} \\ \textsc{Diagnostic} & \begin{bmatrix} \text{hard} : & + \\ \text{texture} : & \text{rough} \\ \text{weight} : & \text{heavy} \\ \text{colour} : & \text{grey} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textsc{Central} & \begin{bmatrix} \neg\text{organic} : & + \\ \neg\text{animate} : & + \\ \neg\text{genus} : & \text{lion} \\ \neg\text{biological essence} : & \text{lion} \end{bmatrix} \\ \textsc{Diagnostic} & \begin{bmatrix} \text{legs} : & 4 \\ \text{tail} : & + \\ \text{texture} : & \text{soft} \\ \text{colour} : & \text{tawny} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textsc{Central} & \begin{bmatrix} \text{organic} : & - \\ \text{animate} : & - \\ \text{solid} : & + \\ \neg\text{genus} : & \text{lion} \\ \neg\text{biological essence} : & \text{lion} \end{bmatrix} \\ \textsc{Diagnostic} & \begin{bmatrix} \text{hard} : & + \\ \text{texture} : & \text{rough} \\ \text{weight} : & \text{heavy} \\ \text{colour} : & \text{grey} \\ \text{legs} : & 4 \\ \text{tail} : & + \end{bmatrix} \end{bmatrix}$$

The first matrix is the AVS of stone, the second represents the MTC_R operation for lion. In this operation, the central attributes of lion are negated. That is not enough though, next to MTC, we need Priority Union to complete the AVS for stone lion, which is visible in the last matrix. This is so, because there are conflicts between the diagnostic features of stone and lion. According to Franks (1995:480), '*the resulting sense describes a stone object with many diagnostic properties of a lion. Its nonspecificity renders it compatible with specific central and diagnostic features of statue, ornament or rock-formation.*' In other words, we need more contextual information to specify the appropriate NI. This is all part of information state 2, in

which the NI is selected. Based on the NI, the correct perspective (type II) and classification (PC) can be inferred.

2.2.3 Extension to other cases of MTC

So, where does that leave us? We have seen that Franks (1995) is able to give reconstructions of sense generation for examples (1)-(3). There will be ways in this system to account for (4) and (5) as well, since it appears that features can be assigned quite informally and *ad hoc*, in a way. So, the framework of sense generation might well work for MTC that is triggered in a modifier-head relation. It will also work for N-N combinations that are not part of the present paper, because these are examples of type coercion within a single clause. For cases like (6), it is imaginable that sense generation would offer a solution as well, after all, *a dress with roses*, is an example of modification as well, only by an attached PP, instead of Adj. The only difference here is that the shift takes place on the N that is contained in the PP, and not on the head N. More importantly, the notion of privative adjectives vacillates in this context; we are not dealing with a privative adjective, only what in Franks' view would be regarded as a privative meaning (of a construction). Let's assume that in this case there is a functional trigger for MTC_R. I will not take this to be any kind of privative, since the term privative, if used for this kind of meaning effect, is only applied to adjectives. If we assume that a the MTC operation that takes place in (6) is similar to the one in *stone lion*, there is a problem for Franks. Since he assumes that next to a MTC_R operation, we also need priority union. Using priority union, it would be the modifier that takes priority over the head, and features of the modifier would be inherited by the head. In this case the operations MTC_R and priority union have effect on the modifier and not on the head. It could be argued that this only holds for modifiers that precede the head, but this could be refuted by taking examples from Dutch like (18) and (19):

(25) *de ge/bebloemde jurk*
the flowered dress
'the dress with flowers'

(26) *de *ge/bebloemde tuin*
the flowered garden
'the garden with flowers'

The adjective *bebloemde* is, however, a very complex form. It would be an adjectival use of a verbalised N. I use it as an argument because in case of (19), the garden is covered by N_{nat} flowers, and in case of (18), dress is covered with N_{ima} flowers.[11] So, it is the head that coerces the property of the modifier in this case, resulting in the same meaning when modified by an attached PP. To sum, the definition priority union should be either abandoned completely or revised, in order to account for (6) and (18). Another line of reasoning would be to look at the argument that Franks has to use priority union in this case in the first place. In the AVS of stone lion, we saw that this was needed to account for inheritance of diagnostic features of modifier *stone* in *stone lion*. At first sight, the division in central and diagnostic features seems appealing. However, when taking a closer look, we see that there is no sharp-cut line between them. For example, let's take a look at the diagnostic features of stone. Franks takes

[11] Since Franks does not offer any definition of the notion type, I will use types as defined previously.

stone to be grey. A majority of stones is not grey: soapstone, of which many African and Asian crafts are made, can have a tawny colour. Franks assumes stone to be grey by default, which could be questionable. Another problematic diagnostic feature is weight. A stone lion is not necessarily heavy, moreover, it can be a hundred times as light as a N_{nat} lion. The inference from *heavy* plus *heavy* equals *heavy* does not seem to hold in practice; which is a weakness in this system. Franks resolves this by assuming that once it is the case that the NI ornament is selected by contextual information, priority union will change this feature to possibly 'not heavy'. This means that if the context does *not* provide further information, a stone lion is supposed to be heavy, which seems highly implausible. A last remark is that in the last AVS of *stone lion*, the central features of 'genus' and 'biological essence' are negated only with respect to the value of lion. I think that these central features should be negated in their whole existence, and be valued binary with a '-', since it is the case that no stone object could have the property of any genus or biological essence.

In general, I would say that distinguishing between central and diagnostic properties for the default content of a concept has its advantages in treating privatives. I think the problem with *stone lion* (and later on *the dress with roses*), is that the alternation of diagnosting features is part of the AVS that is used to represent the information states. If we take into account the type II perspective that is the result of the construction, we see that diagnostic properties are no part of that. These properties will be altered, but in my view, this could take place when there is enough contextual information to support a certain NI. If that is not the case, the diagnostic features of *stone lion* will be left to imagination, simply because they are not accessible for the hearer. Turning to (6), we'd have to assume that dress is a functional MTC_R operator. The N that is modifies dress, undergoes type coercion, in which the central features, such as [alive: +] and [natural kind: +] are negated. Diagnostic features can be altered as more information is released in context. So, Franks' theory can account for nouns that trigger type coercion of nouns that modify them in a PP adjunction. But only if we assume that nouns can be MTC_R operators as well, and altering the diagnostic features is left to contextual information.

It will be more difficult in this framework to account for examples that have MTC outside the semantic relations of Mod-Head and Head-Mod that are present in (1)-(6). Franks' framework accounts for concept combinations, which could be interpreted as complex concepts, but not for conceptual demands of verbs. This also holds for MTC that is triggered by a certain location and demands of topics. The basic framework could potentially be extended to those, but that would mean that most of the central claims of Franks will be abandoned. Verbs, locations and the 'special' topics that can cause MTC require AVS's that are different from nouns, and are not divisible in central and diagnostic features. However, there is a vast amount of literature in LFG and HPSG that provides similar AVS's that might offer the possibility of linguistic elements occurring as functional MTC operators. Another difficulty can be seen when you look back at the possibilities is the notion of privative meaning. In this approach, privatives are taken to be adjectives. We have seen that the same meaning effects can occur in constructions that do not involve adjectives. Franks could probably account for these by constructing an MI (modifying instantiation) next to an NI. This MI could serve as functional MTC operator, and coerce the N instead of the trigger. It would be difficult to work out the formal relation between the trigger and a MI, and the construction of AVS's would be rather complicated. Because it is quite far-fetched to resolve the problem this way, I will for now assume that Franks is able to deal with examples (1)-(6), but not for the others.

To conclude this section, let us look at the advantages and disadvantages of this approach.

The strength of this framework lies in its ability to account for context-sensitivity in concept interpretation, and its relative simple notion of conceptual content. The quasi-classical notion of concepts allows us to adopt default content as basis next to non-default senses that can be generated. This is an appealing view with respect to the phenomenon of MTC in the present investigation. Also the notion of perspectival relativity and two types of interpretation could be useful in resolving the MTC puzzle. In other words, the theoretic assumptions that are made in this article, could offer a set of tools for us. It is the application in the functional AVS's that is bothersome to me. I have noted earlier that a lot of things remain undefined in this framework, such as the notion of 'type', 'default' and 'value'. The latter can either be binary (+ or -), or descriptions. This makes the AVS's unclear and difficult to apply. Furthermore, not having accounted for default content, the features that appear in Franks' AVS's seem *ad hoc*. For example, it is problematic to use relative features such as size, weight and colour, as we have seen in the representation of stone lion. If those relative notions present difficulties as described, the implication would be that diagnostic features (and some of the central features as well), cannot be taken to be part of the AVS. Finally, the distinction between the central and diagnostic features is not clear, and due to the problems it faces, we could question if there really is one.

2.3 Generative Semantics

This section will give an impression of the work of Pustejovsky (1995).[12] His focus is on lexical semantics in a generative framework. Pustejovsky (1998:103) describes his goal best himself: '*What we hope to achieve is a model of meaning in language that captures the means by which words can assume a potentially infinite number of senses in context, while limiting the number of senses actually stored in the lexicon. The mechanism responsible for this polymorphic behavior of language is a set of generative devices connecting the different levels of lexical semantics, providing for the compositional interpretation of words in context.*' The first paragraph will describe the general ideas presented in Pustejovsky (1998). In 2.3.2 I will explain how this framework accounts for type coercion in general and to what extend the framework resolves the puzzle of MTC. Finally, in 2.3.3 I will explain the limits of this framework with regard to the examples of this thesis.

2.3.1 Lexical meanings

Before I start, I need to emphasize that the work of Pustejovsky (1998) is too voluminous to capture in one short paragraph. The abstract here merely functions as an introduction to the next paragraph, so subjects that fall within the theory of the Generative Lexicon but outside the scope of this thesis, will be left aside. The core of Pustejovsky (1998:5) is to address the problems faced by lexical semantics:

> Goals of Generative Lexicon (GL)
> Explaining the polymorphic nature of language,
> Characterizing the semanticality of natural language utterances,
> Capturing the creative use of words in novel contexts,
> Developing a richer, co-compositional semantic representation.

[12] Note that my references to this work are based on the paperback edition of 1998.

The theory shows a contrast with the framework of Fauconnier (1994), in the sense that Pustejovsky makes a distinction between common sense knowledge and lexical structure. This is merely methodological but interesting since it captures my need for the separation of linguistic meaning and other forms of meaning that could influence the interpretation of an utterance). Nevertheless, the semantics of natural language should be related to the 'nonlinguistic conceptual organizing principles'. This corresponds to the aim of this thesis, since we will look at the consequences for conceptual theory in chapter 3. Pustejovsky (1998:7) argues that '*the representation of lexical semantics, therefore, should be seen as just one of many levels in a richer characterization of contextual structure.*' In other words, the focus of his work is lexical semantics, but this can only be established as part of a richer (generative) theory about meaning.

Pustejovsky starts with describing two kinds of ambiguity; *contrastive ambiguity* and *complementary polysemy*. The first entails that one lexical item carries two unrelated meanings. In case of the latter, the lexical item has senses that are systematically related. It will become clear that the senses in the examples of MTC in this thesis, can be explained in terms of complementary polysemy. This type of polysemy can be divided into two subtypes: *category preserving* and *category changing*. We are dealing with the first one, also referred to as *logical polysemy*. A good illustration of this is the noun *lamb*. You can eat lamb for dinner (N_{mass}) or see a lamb walking in a field (N_{count}). Pustejovsky explains that this can be formalised in an SEL (Sense Enumeration Lexicon) as follows:

$$\begin{bmatrix} \textbf{lamb} & \\ \text{SENSE}_1 = & \begin{bmatrix} \text{CAT} = \textbf{mass\_noun} \\ \text{GENUS} = \textbf{meat} \end{bmatrix} \\ \text{SENSE}_2 = & \begin{bmatrix} \text{CAT} = \textbf{count\_noun} \\ \text{GENUS} = \textbf{animal} \end{bmatrix} \end{bmatrix}$$

Pustejovsky (1998) uses this kind of mechanism to describe the lexical structures, although it is not SEL in the strict sense. SEL is too primitive to account for the creative use of words in new contexts, and it only accounts for data in a post hoc fashion, so it's not possible to make predictions of any kind in this lexicon. Another important problem in using SEL to describe a lexicon, is that the senses are fixed. For the interpretation of a natural language utterance it is not always clear how to select the correct sense in a given context. In reality, the different senses of a word can overlap, while the senses are distinct in SEL. The approach of SEL does not offer a possibility to abstract over meanings, that is, there is no core meaning for a lexical item on which all other senses are based.[13]

The Generative Lexicon (GL) is based on the following levels of representation (Pustejovsky (1998:61):

> LEVELS OF REPRESENTATION IN GL
> `Argument Structure`: Specification of number and type of logical arguments, and how they are realized syntactically.
> `Event Structure`: Definition of the event type of a lexical item and a phrase. Sorts include STATE, PROCESS, and TRANSITION, and events may have subeventual

[13]It might seem redundant to describe the failures of SEL here, the intention of this will be clear in later chapters, where we will establish a suitable framework for conceptual meaning.

structure.
`Qualia Structure`: Modes of explanation, composed of FORMAL, CONSTITUTIVE, TELIC and AGENTIVE roles.
`Lexical Inheritance Structure`: Identification of how a lexical structure is related to other structures in the type lattice, and its contribution to the global organization of a lexicon.

Pustejovsky (1998:61) assumes three generative devices that connect these four levels. These devices make it possible to interpret a word in its context:

CONNECTING GENERATIVE DEVICES
`Type Coercion`: where a lexical item or phrase is coerced to a semantic interpretation by a governing item in the phrase, without change of its syntactic type.
`Selective Binding`: where a lexical item or phrase operates specifically on the substructure of a phrase, without changing the overall type in the composition.
`Co-composition`: where multiple elements within a phrase behave as functors, generating new non-lexically senses for the words in the composition.

I will give a short impression of how these levels and devices work in GL, the more specific functioning for the examples of this thesis will be described in section 1.4.2. The semantic structure of a lexical item α is defined as $\alpha = \langle A, E, Q, I \rangle$, where A stands for argument structure and E represents the event type. These two are related by Q, the qualia structure which provides the 'binding' of A and E. I represents the lexical inheritance structure. This places α in a type lattice, and determines to what extend information is inherited from the global lexical structure.

To describe the argument structure, Pustejovsky (1998:63) introduces four types of arguments: *true arguments* (syntactically realized), *default arguments* (participate in logical expressions in qualia, not necessarily syntactically realized), *shadow arguments* (parameters that are semantically incorporated into the lexical item) and *true adjuncts* (parameters which modify the logical expression, but are part of the situational interpretation). Especially the true arguments and shadow arguments will be important in the attempt of resolving cases of MTC in GL, as I will describe in the next section.

The event structure used in GL is similar to the event variable of verbs in event-based semantics (for example Davidson (1967)). However, Pustejovsky (1998:68) argues that this atomic view does not represent the event structures fully. He proposes a way to represent subeventual structures as well. These are PROCESS, STATE and TRANSITION. The relation between an event and its proper subevents are represented in an 'extended event structure' $\langle E, \leq, <, \circ, \subseteq, * \rangle$. This tuple represents the possible relations between the subevents. Since this is only important for the global idea of howGL works, I won't discuss the event structure in detail. A last important thing to note is that not only are the subevents represented in GL, they are also put under restrictions. That is, the relations between the subevents described above, form a restriction on the event structure.

For the purpose of this thesis, the qualia structure[14] (next to the argument structure) is most important. Moreover, the assumption of the existence of such a structure is what makes GL unique in semantic theories. The qualia structure represents the relational force of a lexical

[14]The term 'qualia' in this framework has nothing to do with the term 'qualia' in cognitive sciences, where qualia are used to describe concepts that we can only know by experience, like colours, smells and tastes.

item. In other words, the qualia of lexical items in a context account for the operations that these items undergo if they are put together (in a linguistic context). The qualia structure of a lexical item explains what a word means. Pustejovsky (1998:76,84) proposes four essential aspects that are described by qualia structures:

> QUALIA STRUCTURES IN GL
> `Constitutive`: the relation between an object and its constituent parts (i.e. material, weight, parts and component elements);
> `Formal`: that which distinguishes it within a larger domain (i.e. orientation, magnitude, shape, dimensionality, colour and position);
> `Telic`: its purpose and function (i.e. purpose that an AGENT has in performing an act and built-in function or aim which specifies certain activities);
> `Agentive`: factors involved in its origin or 'bringing it about' (i.e. creator, artifact, natural kind and causal chain).

Pustejovsky notes two characteristics of qualia structures. The first is that every category expresses a qualia structure, the second is that not all lexical items carry a value for each qualia role. Qualia can be seen as a set of properties or events that are associated with the lexical item. The qualia structure is not just a list, in GL it functions as starting point for semantic operations, such as type coercion, as we will see in section 2.3.2. All this is possible because the values of qualia are defined as types and relational structures. For example, Pustejovsky (1998:79) illustrates qualia structures by describing the adjective *tall*:

$$\begin{bmatrix} \textbf{tall} & \\ \text{EVENTSTR} = & \left[E_1 = e_1\textbf{:state} \right] \\ \text{QUALIA} = & \left[\text{FORMAL} = \textbf{tall}(e_1,\textbf{x}) \right] \end{bmatrix}$$

The true function of these structures, and their role in GL is explained (Pustejovsky (1998:82)) in the representation of the verb *build*. This representation forms the starting point of the discussion in the next section:

$$\begin{bmatrix} \textbf{build} & \\ \text{EVENTSTR} = & \begin{bmatrix} E_1 = e_1\textbf{:process} \\ E_2 = e_2\textbf{:state} \\ \text{RESTR} = <_\alpha \\ \text{HEAD} = e_1 \end{bmatrix} \\ \text{ARGSTR} = & \begin{bmatrix} \text{ARG1} = \boxed{1} & \begin{bmatrix} \textbf{animate\_ind} \\ \text{FORMAL} = \textbf{physobj} \end{bmatrix} \\ \text{ARG2} = \boxed{2} & \begin{bmatrix} \textbf{artifact} \\ \text{CONST} = \boxed{3} \\ \text{FORMAL} = \textbf{physobj} \end{bmatrix} \\ \text{D-ARG3} = \boxed{3} & \begin{bmatrix} \textbf{material} \\ \text{FORMAL} = \textbf{mass} \end{bmatrix} \end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix} \textbf{create-lcp} \\ \text{FORMAL} = \textbf{exist}(e_2, \boxed{2}) \\ \text{AGENTIVE} = \textbf{build\_act}(e_1, \boxed{1}, \boxed{3}) \end{bmatrix} \end{bmatrix}$$

We can see that build has three arguments. The default argument is a logically distinct argument, namely that the argument that represents the material the artifact is made of. In short, this representation says everything there seems to be said about the meaning of *to build*. We can see here that only the formal and agentive qualia are described. The qualia structure here represents the relation between the arguments that the verb selects. Both events are represented (process and state), with the arguments specified and described above. This example of representation in GL forms the basis for semantic operations, which will be topic to the section below.

2.3.2 MTC in this framework

My first impression of GL was that it could be especially suitable in accounting for cases of MTC that are triggered by a verb. Pustejovsky (1998:111-112) proposes that the verb has no different semantic types: '*the verb's type remains constant, where the internal argument is uniformly typed as proposition; if the syntactic form appearing in the complement position matches this type, then the resulting structure is well-formed. If, however, the appropriate type is not present in the complement position, it is coerced by the verb to match the type required by the typing restrictions on the verb.*' In my criticism on Mental Spaces in section 2.1, I already noted that at least for the cases in which the verb is the MTC operator, we need a framework that describes the demands of verbs as being rigid (or, as we will see later, inviolable). It seems that Pustejovsky agrees with me on this point. Before I will discuss the possibilities of resolving the issue of MTC in this framework, I should note that the notion of *type* is different in GL. Since GL is mainly organized around the coercive force of verbs, the types in this framework are conform the type ladder proposed by Partee&Rooth (1983). It is clear though, that the definition of type coercion according to Pustejovsky (1998:111) can work for the types that I assume (N_{nat}, N_{art}, N_{ima}) as well:

FUNCTION APPLICATION WITH COERCION (FAC)
If α is of type `c` and β is of type `<a,b>`, then,

(i.) if type `c` = `a`, than $\beta(\alpha)$ is of type `b`.

(ii.) if there is a $\sigma \in \sum_{\alpha}$ such that $\sigma(\alpha)$ results in an expression of type `a`, then $\beta(\sigma(\alpha))$ is of type `b`.

(iii.) otherwise a type error is produced.

In the previous section, I gave a representation of the two types of *lamb*. As explained in the preliminaries of this thesis, the types of nouns I distinguish here, are similar to the distinction mass-count. In the examples of MTC the verb coerces arguments into the type that they require, because otherwise there is a type error. Is it as simple as that? I will show to which degree GL accounts for examples of MTC. Intuitively, GL should be able to account for the examples where the type coercion is triggered by a verb (example (7)).

First we will take a closer look at the generative operations that are the core of GL. Pustejovsky(1998:113-122) distinguishes between two kinds of type coercion: *subtype coercion* and *true complement coercion*. The first is said to be the simplest kind of type coercion, where a function selects for τ_1, but the realized form is τ_2, and τ_2 is a subtype of τ_1 ($\tau_2 \leq \tau_2$). The simplicity of this can be seen in the example that illustrates it well: *Mary drives a Honda to work*. In this case the internal type selected by drive is *vehicle* (τ_1), but the realized word is *Honda* (τ_2). *Honda* is a brand of car, and therefore a subtype of the required vehicle. Pustejovsky (1998:115) gives an important characteristic that distinguishes subtype coercion from true complement coercion: '*Subtype coercion follows the inferences available in a single type lattice, while true complement coercion requires reference to multiple type lattices, making use of information through the qualia.*' True complement coercion involves the strict shifting of one type to another specified type. The most striking property of true complement coercion is that this is only licensed by lexical governance, as Pustejovsky (1998:115) An example will make this clear: *John began a book*. The verb *to begin* requires an event-denoting expression. This captures the intuition that *to begin a book* means *to begin writing/reading a book* (events). The noun shifts to an event in the interpretation because the verb requires for this. In other words, the true complement of to begin is an event, not a physical object. The coercion is made possible by the qualia structure of *book*. The idea seems to be that a verb can look into this structure in order to find the type it requires. Later on in this section I will discuss if MTC belongs to one of these operations or if we will need another generative operation for MTC.

Other operations are co-composition and selective binding. Co-composition means that a structure allows more than one function application. An example of this is *John baked the potato* as opposed to *John baked the cake*. It is assumed that here, that the complement co-specifies the verb. That is, potato makes that *bake* means only **state_change**, while cake makes that *bake* means to **create**. The argument selects one of the possible senses of the verb, and this is done by function application with qualia unification. The issue of selective binding is more important for my purposes, as it could be applied to the adjectival cases of MTC. This will be elaborated in section 2.3.3.

So, how can we place MTC within the generative operations proposed in Pustejovsky's Generative Lexicon? A first consideration would be that MTC is an instance of subtype coercion. This would mean that an α of type N_{ima} can be inferred from a β of type N_{art}/N_{nat}. I argue that this is not the case. Later on, in chapter 4 I will reconsider the

possibilities of placing the image type within a set. In Pustejovsky (1995) the subtypes are for example `Honda` $\leq$ `car` $\leq$ `vehicle`. In my examples we refer to an image in terms of its original, which makes it difficult to place the different types in a single type lattice. If the image of a concept form a subset to the set of that concept will not be addressed here. An other argument would be that in terms of functionality, a Honda relates in a different way to vehicle than for example a toy truck does to a truck. A Honda is a sort of vehicle, while a toy truck is, functionally seen, not. The main difference between the examples of subtype coercion and MTC is that in the first, the qualia structures of both are equal. As a result, the coercion is quite simple, instead of *vehicle*, *Honda* is realized, which has the same requirements as *vehicle* for the verb *to drive*. This obviously is not the case for *truck* in the vehicle sense and *truck* in the toy sense. We need an operation for this that alters the qualia structure of truck in terms of all its roles, and the question is if GL is able to provide for an operation like this.

The operation of true complement coercion could come closer to what we need for MTC in GL. In the examples of true complement coercion, the NP that is an argument for the VP is coerced into an event type. Since I work with different types here, the definition of true complement coercion is not applicable directly. However, an adjusted version of this type of coercion would be able to describe MTC. The example of *John began the book*, allows two readings: *John began reading the book* and *John began writing the book*. The argument *the book* thus describes an event and not an object, because the verb to begin requires for an event. In example (7) the verb *to trip over* requires (informally described) for an object that is no higher than uptill a human knee.[15] We can assume that, when assuming GL as point of view, this will be represented in the qualia structure of the verb:

$$\left[\begin{array}{ll} \textbf{to trip over} & \\ \text{EVENTSTR} = & \left[\begin{array}{l} E_1 = e_1\textbf{:process} \\ E_2 = e_2\textbf{:state} \\ \text{RESTR} = <_\alpha \\ \text{HEAD} = e_1 \end{array}\right] \\ \text{ARGSTR} = & \left[\begin{array}{ll} \text{ARG1} = \boxed{1} & \left[\begin{array}{l} \textbf{animate\_ind} \\ \text{FORMAL} = \textbf{physobj} \\ \text{CONST} = \textbf{legs} \end{array}\right] \\ \text{ARG2} = \boxed{2} & \left[\begin{array}{l} \textbf{physobj} \\ \text{FORMAL} = \textbf{count} \\ \text{FORMAL} = \textbf{height} \leq \textbf{25\% of } \boxed{1} \end{array}\right] \end{array}\right] \\ \text{QUALIA} = & \left[\begin{array}{l} \textbf{state\_change\_lcp} \\ \text{AGENTIVE} = \textbf{trip over\_act}(e_1, \boxed{1}, \boxed{2}) \end{array}\right] \end{array}\right]$$

I am aware that this figure is 'sloppy' in some respects. This is due to the fact that the CONSTITUTIVE qualia are not treated extensively in Pustejovsky (1998). The most important

[15] I am aware that this is relative to the height of a person, but let's for simplicity's sake assume that the subject of this sentence is an average adult.

thing is that this verb requires from the second argument to be no higher than the knee of the first argument. It is easy to see that truck in its default meaning is no candidate for the second argument position:

$$\begin{bmatrix} \textbf{truck} & \\ \text{ARGSTR} = & \begin{bmatrix} \text{ARG1} = \textbf{x:artifact\_vehicle} \\ \text{D-ARG1} = \textbf{y:physobj} \end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix} \text{FORMAL} = \textbf{x} \\ \text{FORMAL} = \textbf{height: } \geq\textbf{270cm} \\ \text{TELIC} = \textbf{drive(e,x,y)} \end{bmatrix} \end{bmatrix}$$

How do we account for the coercion from N_{art} to N_{ima} in this generative system? I will argue that the only way to resolve this, is to assume that every noun of type N_{art} or N_{nat} has a possibility to have two different qualia structures. One belongs to the default type (τ_1), the other belongs to the type that can occur after coercion (τ_1). All together this means a new kind of type coercion (MTC), which operates when the qualia requested by the verb conflict with the qualia of τ_1. The following representation is merely a sketch of how this would look 'officially' in GL.[16] Note that in this matrix, I have 'filled in' the actual sentence, and not the basic meanings of the lexical items part of that sentence. I will discuss this later on in this section, as one of the problems of GL.

$$\begin{bmatrix} \textbf{I trip over a truck} & \\ \text{EVENTSTR} = & \begin{bmatrix} E_1 = e_1\textbf{:process} \\ E_2 = e_2\textbf{:state} \\ \text{RESTR} = <_\alpha \\ \text{HEAD} = e_1 \end{bmatrix} \\ \text{ARGSTR} = & \begin{bmatrix} \text{ARG1} = \boxed{1} & \begin{bmatrix} \textbf{x:human} \\ \text{FORMAL} = \textbf{physobj} \\ \text{CONST} = \textbf{legs} \end{bmatrix} \\ \text{ARG2} = \boxed{2} & \begin{bmatrix} \textbf{y:artifact\_vehicle} \\ \text{FORMAL} = \textbf{count} \\ \text{FORMAL} = N_{ima}\ \textbf{of y} \\ \text{FORMAL} = \textbf{height} \leq\textbf{25\% of } \boxed{1} \\ \text{TELIC} = \textbf{toy} \end{bmatrix} \end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix} \textbf{state\_change\_lcp} \\ \text{AGENTIVE} = \textbf{trip over\_act}(e_1, \boxed{1}, \boxed{2}) \end{bmatrix} \end{bmatrix}$$

We have seen from the matrices above that the verb *to trip over* has a demand concerning the height of the object, relatively to the length of the subject. This is successful, because this

[16] The aim of this thesis is not to resolve the puzzle of MTC within the existing frameworks, but to show the limits of the possible solution in any given framework. Therefore it's beyond my scope to represent the matrices extensively.

way we are able to adjust truck to any subject that trips over it. Suppose the subject is an ant, the truck would be no higher than the knee of an ant (an estimated height of 0.125mm). In this extreme case, we would still assume that the truck is an image of a real truck, however very, *very* tiny. It also prevents the truck from shifting to an image of a truck in case the subject is giant. In a story about a giant walking in a city and tripping over trucks, we would interpret the trucks in their default manner of N_{art}, because 25% of the height of a giant would be assumed to be enough to trip over them. So far, this is the best result in the attempts of solving MTC. But of course the problems of this approach are lurking somewhere.

There are three problems with the GL approach for this example of MTC. First, there is no explanation for the TELIC and FORMAL qualia being changed in ARG2. Second, it is not possible to show the generative *process* in this way. The third problem is the *interpretation* of *truck* that is now described in the TELIC quale as TOY. The first is a 'theory-internal' problem. Part of the goal of Pustejovsky's Generative Lexicon is to avoid the problem of polysemy by allowing lexical entries to have more than one sense. These different senses can be seen in the qualia structure of the entry, and, when needed, generative operations take place in order to adjust the lexical entry to demands of the linguistic context. In the case of (7) represented in the last matrix, we see that both the TELIC qualia and the FORMAL qualia of *truck* have undergone some dramatic changes relative to the representation of *truck* itself. The TELIC quale changes from **drive(e,x,y)** to toy (or, put in an event structure: **play with(e,x,y)**) This would mean that there has to be more than one lexical entry for truck in examples like these. This would ruin the goal of the whole theory, which primary goal was to resolve polysemy by having one lexical entry for a word and generate senses from there. More or less the same holds for the FORMAL qualia of truck. Somewhere, although not visible in this matrix, there must have been a demand for changing the type of the truck. I am aware of the fact that this typing system is the fruit of my own mind and not Pustejovsky's, but suppose he would use these types. Then we would have a specification in the matrix of *truck*; one of the FORMAL qualia would be assigned N_{art} or this would be assigned as part of the existing FORMAL quale **count**. Even if this was the case, there would still be no explanation for the type-shift, nor for the interpretation of that shift. The verb does not require the type to be N_{ima}, it only requires the height its object to be at most 25% of the height of its subject. Even if we would specify the heights of respectively ARG1 and ARG2, there is no way that we could do this in one single matrix the way we have now, which relates to the second problem.

The lack of visible *process* in this representation is quite a big obstacle. As I indicated above, we would need at least two matrices for (7): one for the collision of meanings, in which it is visible why coercion is needed, and one for the result. Aside from this, I already deviated from the matrices in Pustejovsky (1998) by representing an actual sentence. In his work, Pustejovsky merely describes the general properties of lexical items. This is a problem, because this way, the matrices don't represent the process of coercion. You can either describe the default meaning in this kind of matrix, or the result after coercion. My aim is to find a framework that represents the process in between as well. That way, it would be more clear in which linguistic context a noun coerces and why.

Related to both of these problems is the fact that from this representation, we cannot deduce why the change of type goes hand in hand with the change of the TELIC quale. Nor do we know from this why it is the TOY interpretation that comes into mind first when processing example (7). Again this is a problem for the Generative Lexicon, since it would be impossible to specify the TELIC quale with both **drive(e,x,y)** and **play with(e,x,y)**. However, this does not mean that I am looking for a framework that allows more than one entry per lexical

item. Even from the conceptual point of view I would not say that the two senses of truck are unrelated. It's just impossible to relate them in the Generative Lexicon.

Overall, I doubt that we can rely solely on the qualia structures of the different lexical items for the derivation of the meanings that result from MTC, nor the process that *is* MTC. The type N_{ima} can be contained in the qualia structure of a noun, and we can suggest to use different matrices to represent different stages of meaning generation, but I wonder how we could ever account for the other examples of MTC if already the example that contains the one of the most obvious syntactic relationships, is so problematic. Let's first take a look at the possibilities for adjective constructions in GL.

2.3.3 Extension to other cases of MTC

Pustejovsky (1998:128) argues that '*the adjective is able to make available a selective interpretation of an event expression contained in the qualia for the head noun.*' This is illustrated by the examples *fast typist* and *fast driver*. *Fast* in the first means hits per minute, in the second it means mile per hour. The meaning is thus determined by the semantics of the head. In my examples it is the head that undergoes a type shift because of the modifying adjective. Later in this section, I will discuss if adjectival MTC should be resolved by another version of selective binding:

> SELECTIVE BINDING:
> If α is of type `<a,a>`, β is of type `b`, and the qualia structure of β, QS_β, has quale, q of type `a`, then $\alpha\beta$ is of type `b`, where $\|\alpha\beta\| = \beta \cap \alpha(q_\beta)$.

As described before, the process of selective binding accounts for polysemous adjectives like *fast*. Examples (1)-(5), however, do not show polysemy of the adjective. It are the heads in the construction that seem to be polysemous by coercion. I cited the definition of selective binding above. What this says informally, is that the meaning of Adj-N construction with adjectives that are selective binders, is based on only one quale of the noun. In the case of *fast typist*, *fast* only applies to the TELIC quale of typist **type(e,x,y)**. In this representation, we see the use of the intersective symbol, which is confusing in a sense, since we are looking at the class of adjectives that are otherwise called *subsective*. What is meant here, is that the intersection holds only for one quale of the modified N; this operation of selective binding holds for subsective adjectives.

So how can we make a qualia-based account of adjectival meaning that is suitable for the examples of adjectival MTC? First, there is an intuitive correspondence between for example good knife and stone lion. In the first, *good* only applies to the TELIC quale; in the latter, *stone* applies to the CONST quale of *lion*, but not the other qualia. The difference is that selective binding is not a process of coercion (Pustejovsky (1998:129-130)), while the adjectives in my examples *coerce* the nouns they modify. This coercion can be described as the result of incompatible qualia. *Stone* has **stone** as CONST quale, lion has '**flesh of lion**' as CONST quale. The resulting meaning still has properties of the head noun, i.e. a stone lion still has features by which it can be recognized as a (sort of) lion.[17]

In order to account for MTC in GL, we will need more than selective binding, which is not more than a foundation for adjectival meaning generation. The following can be considered an addition to selective binding:

[17] Whether a stone lion *is* in fact a sort of lion or not, is part of a conceptual issue, that sadly enough cannot be addressed in this study. However, a short note will be made on this in chapter 4.

METONYMIC BINDING
If α is of type `<a,a>`, β is of type `b`, and $Ima(\beta)$ is of type `b`, and the qualia structure of β, QS_β, has quale, q of type `a`, such that q_β is semantically incompatible with α, then $\alpha\beta$ is of type `b`, where $\|\alpha\beta\| = Ima(\beta) \cap \alpha(q_\beta)$.

As noted before, selective binding is an extension of what Kamp&Partee (1995:137) refer to as *subsective* adjectives. Pustejovsky adds the idea of underlying qualia-structures, but in fact the idea is the same. In this definition of MTC, I have added 'type' *Ima* as I have used in this study. This notion of type is used at a different level, as described in the preliminaries to this study. The types of Pustejovsky, I assume, are abbreviations of the types in the tradition of Montague `<<s,<<s,e>,t>>,<<s,e>,t>>/<<s,<e,t>>,<e,t>>` (= `<a,a>`) and `<s,<e,t>>/<e,t>` (= `b`). This description of MTC as Metonymic Binding in GL can be informally paraphrased as follows: if there is a quale in N of Adj-N, that is incompatible with the qualia in Adj, this quale takes the value of Adj, which results in a type shift to N_{ima} of N (which is of the same semantic type). This definition accounts for the cases of subsective adjectives which would be unaccounted for in selective binding because of the conflict between qualia of Adj and N respectively. Note that the quale of Adj 'wins' this conflict; the conflicting quale of N adopts the value of Adj for that quale and N is interpreted as the image of N. It would be interesting now to see if all the examples of Adj-N constructions with MTC can be dealt with using Metonymic Binding (MB). In case of (1), the CONST quale of *lion* is incompatible with the CONST quale of *stone*. The result of MB is that this quale is changed, while the remaining qualia of lion stay the way they are. *A stone lion* is now a representation of a lion made of stone; consistent with the intuitive interpretations I formulated in section 1.1.1. For (2), the incompatibility that causes coercion is not the CONST quale (although the CONST quale of *fake gun* might be different in certain interpretations). *Fake* 'denies' the functionality of *gun*, and because of this, the TELIC quale **shoot(e,x,y)** of *gun* is somehow incompatible with *fake*. Note that this is less obvious than in (1). In (1) it is clearly visible in the QS's of *stone* and *lion* that they are incompatible. This does not hold for *fake*. *Fake* does not mean 'no capacity to function in the default manner' necessarily, it might actually operate on a CONST quale of a N as well. *Fake fur* is an example of this. *Fake* seems to apply to the CONST quale and not to the TELIC quale, i.e. *fake fur* has the same function as real fur, but is made out of for example polyester. The relative successes of Metonymic Binding for (1) and (2) can be generalized over (3)-(6), *chocolate* will function like *stone*, *wannabe* and *imaginary* more or less like *fake*. In general, I would say that this definition works for MTC. The case of *fake* remains difficult though. For now, I propose that MB only applies to adjectives such as *stone* if the CONST quale of the head is incompatible with the adjective, and adjectives in the class of *fake* ('privatives') will always behave according to MB. Within GL it could be argued that *fake* has the lexical property of denying at least one of the qualia of the head N it modifies. To which quale this applies might depend on the noun.

I assume that the other examples of MTC will be very tough to account for in GL. This is not as wild an assumption as it seems. First, we have seen how difficult verbal MTC was for GL, due to the lack of process in the representation and the fact that the image interpretation was unaccounted for. The other examples will be even harder, because of the restriction Pustejovsky (1998:111) proposes for type coercion: '*... which undergoes a type shifting operation by virtue of lexical governance of the verb.*' Type coercion is regarded as a generative mechanism, but is restricted to verbs in GL. In the summer of 2001 I attended a guest lecture of Annemarie Mineur, who found a way to integrate Pustejovsky's qualia

structures in Kamp&Reyle's (1993) Discourse Representation Theory (DRT). A full account of her lecture can be seen in Bos *et al.* (1995). One of the examples in Mineur's lecture was: *I looked into the room. The ceiling was very high.* She argued that in a DRS, it is possible for room and ceiling to be *bridged*.[18] This study is too restricted to give a full picture of this, but in a nutshell: it can be argued that *room* has *ceiling* as CONST quale. This (and also the example) is based on Clark (1977), who makes a distinction between *necessary*, *probable* and *inducible* parts of an object. A ceiling is a necessary part of a room, and therefore belongs to the CONST quale of *room*. Mineur assumed that only necessary parts can be considered as CONST qualia for an object. In the resulting DRS, ceiling would occur as CONST quale in the structure of *room* and could thus be bridged to *ceiling* in the discourse. It was this example that made me think of elephants as ornaments in rooms in the first place, which in the end led to this study. Example (8) cannot be addressed the same way. *The elephant* is at most an inducible part of a room. Actually, this example was discussed at the lecture, and there was no solution for it in this way. As long as *the elephant* and *room* cannot be bridged, there is no GL/DRT account for examples of this type. And even if we would be able to bridge *the elephant* to room, we would again meet the problem of the interpretation of *the elephant* as N_{ima} here. This holds for examples (9)-(11) as well, although there is no *bridging* involved here, these examples would be difficult for a process in discourse. My objection against bridging in this way, is that this would lead to an enormous amount of lexical information that is necessary to resolve this phenomenon. *If* we would regard the elephant of type N_{ima} is somehow part of the lexical information (as would be the case in a GL solution), this will lead to generation of many 'unsolicited' senses. This claim is supported in Asher&Lascarides (1998:89), who we will meet again in section 2.4.3.

A last possibility for MTC resolution in GL would be the (6) and its variations in section 1.2.2. However, as I have stated before, operations of coercion are based on verbs in GL. For the adjectives there was an indirect way to make an adjustment for MTC, but for the 'picture' examples in these PP-examples, I would have to create a complete new GL function, which is not within the scope of this study. When I attended the GL conference in 2003 in Geneve, I talked to James Pustejovsky himself about examples of MTC that are caused by discourse markers like *room*. He agreed that there would be no way in GL to account for this, however he added that GL didn't need to account for examples like this, since discourse markers as room and certain topics are part of pragmatics and not (lexical) semantics. I agree to disagree; I will not understand how examples like *stone lion* and *fake gun* are semanticists favourite puzzles, but when the same meaning effect occurs at discourse level, this is suddenly *pragmatics*, and not semantics.

In sum, Pustejovsky's Generative Lexicon has a powerful way of explaining the polysemic nature of lexical meaning. It will be useful to use his ideas about verbs and type coercion in dealing with MTC, however, his way of representation doesn't bring us far. The need of more than one TELIC quale in many of the examples creates polysemy at the level of lexical entries in GL. If we would 'adjust' GL to this, lexical items should have unrelated TELIC qualia that are regarded as different senses of the same entry. As I noted before, I am not against this idea myself, however, in GL this would undermine the core of the generative system. I am aware of the fact that Pustejovsky's framework is sort of 'generative' in itself, and that it's possible to adjust the system in more than one way in order to account for MTC. However, I doubt that we will be able to do so *without* affecting the foundation of GL. It might be

[18] For more information about bridging, see Clark (1977) and Matsui (2000).

possible to come up with other ways of representation, and certainly my matrices can be worked out more properly. The biggest problem is the restriction on type coercion to lexical governance of verbs, and hereby passing similar semantic effects that take place on other levels so easily at pragmatics to deal with. And it is clear why, GL deals with the generative processes that can be built on *basic* assumptions about the meanings of lexical items. We have seen that this is hard when the interpretation of a lexical item is on the border of the concept it represents by default.

2.4 Formal Semantics

In this section I will explore the possibilities for a solution of the MTC puzzle within formal semantics. In the field of semantics, a lot of theories can be covered by the term *formal semantics*, for example the work of Montague, Davidson and Dowty. For the purposes of this study I have restricted the possible 'representatives' of these theories to Kamp (1975), Kamp&Partee (1995) and Partee (2001). It will be clear from this section that in the focus in these works are mainly on the semantics of *adjectives*, and are thereby less likely to apply to other cases of MTC. This does, however, not mean that formal theories of semantics would not provide possibilities for these other cases. I will show that the assumptions about presuppositions can be useful in the extension of the articles that are central to this section. In section 2.4.1 I will summarize the general ideas about the semantics of adjectives of Kamp (1975) and Kamp&Partee (1995). In section 2.4.2 I will discuss the treatment of *stone lion* in Kamp&Partee (1995), followed by a discussion about *privatives* and *coercion* as proposed in Partee (2001). Section 2.4.3 will give an impression about how one could deal with MTC in cases other than Adj-N constructions. For this purpose, I will shortly address the relevant assumptions on presupposition in DRT made by Geurts&Van der Sandt (1997), (2004) and the principles of SDRT/DICE in Asher&Lascarides (1995), (1998) and Lascarides *et al.* (1992).

2.4.1 Adjectives as functions

To understand the line of reasoning that led Barbara Partee to write her interesting note on privative adjectives, I have chosen to summarize (and discuss when necessary) Kamp (1975) and Kamp&Partee (1995). The work of Kamp (1975) is a rather compact article that describes two theories of adjectives. It is mainly an attempt to capture all kinds of adjectives *and* their comparative/superlative forms under one theory. I will refrain from describing the types of logics that are discussed in detail, since this is beyond the scope of the present study.[19]

Kamp (1975:124) starts with describing the assumptions on adjectival meaning by Montague (1970):

> MONTAGOVIAN DOCTRINE
> (a) Each possible interpretation (for the language in question) is based upon (i) a non-empty set W of possible worlds (or possible situations, or possible contexts) and (ii) a set U of individuals.
> (b) A property relative to such an interpretation is a function which assigns to

[19]The reason I give a very short and global idea of the proposed supervaluation theory is that I intend to come back at this in chapter 4.

each $w \in W$ a subset of U (intuitively the collection of those individuals which satisfy the property in that particular world (or context) w).
(c) The meaning of a noun phrase in such an interpretation is always a property. Thus the meanings of adjectives in an interpretation of this kind will be functions from properties to properties.

The last statement is essential; adjectives are not regarded as *extensional*, they are now functions from intensions to intensions. Kamp (1975:124-5) then distinguishes four kinds of adjectives: *predicative*, *privative*, *affirmative* and *extensional*. It is evident that all predicative adjectives are extensional, but the reverse does not necessarily hold: 'relative' adjectives like *tall*, *heavy* and *small* are extensional but not predicative (as argued in Kamp (1975:126) and illustrated in Kamp&Partee (1995:142)). So, all adjectives can and should be regarded as functions *from properties to properties*, although some adjectives have a property that makes them behave like predicates. The main advantage of this is that this theory doesn't treat adjectives as predicates. This is best illustrated by sentences as *every alleged thief is a thief* and *every small elephant is small* (examples cited from Kamp (1975:123)). If adjectives were treated as predicates, these would be logically true, which is problematic. In fact Kamp (1975:124) argues that only a small amount of adjectives can be treated as predicates (hence the *predicative* adjectives): '*those whose extensions are not affected by the nouns with which they are combined*'.

For the account of the normal usage of adjectives, this is theory is sufficient. The problem is that this theory fails to account for comparatives and superlatives: Kamp (1975:127) points out that for comparatives we need the meaning of an adjective to be '*such that the comparative can be understood as a semantic transformation of that meaning into the right binary relation.*' This is problematic for the basic assumption that adjectives are functions from intensions to intensions. The idea of comparative meaning is that adjectives do not apply to things in one single way; their meaning can be applied in different degrees and is thereby, often, *vague*. The aim of Kamp (1975:128) is to '*develop a semantical framework in which the idea of a predicate being true of an entity to a certain degree can be made coherent and precise.*' The proposal of Kamp accounts for vagueness using the theory of *supervaluations*, which later is elaborated in Kamp&Partee (1995). For this purpose it seems plausible to assume a multi-valued logic, which answers to the intuition of 'degrees of truth' by having a set of truth values with a cardinality of ≥ 2. The problem of using this multi-valued semantics is that according to Kamp (1975:131) '*the truth value of a complex formula -say $\phi \wedge \psi$- should depend not just on the truth values of the components -i.e. ϕ and ψ- but also on certain aspects of these formulae which contribute to their truth values but cannot be unambiguously recaptured from them*'.

In order to solve this, it is suggested that we need the set of truth-values to be ordered according to the structure of the propositional calculus, and not linearly. Kamp (1975) argues that there are two theories that qualify for this: *probability* theory and *supervaluation* theory. In the first, the set of truth-values is associated with a real number in the interval [0:1]. This number then represents the probability of a proposition instead of a truth-value. In order to account for the probability of conjuncts, the option is to regard the sets as possible worlds (and thus using a type of intensional logic in addition to probability theory).[20] Supervaluation theory works with two basic principles: *partial interpretations* and *comple-*

[20] Kamp (1975:132-3) argues that without a form of intensional logic, the probability of the conjunction $\phi \wedge \psi$ is the intersection of the probability of ϕ and ψ.

tions. Simply put, the theory is two-valued but allows for sentences to be neither false nor true (see Kamp 1975:133 for a core description of SVT). This theory is a suitable approach for vagueness, since the partial interpretations can have various completions. The quantity of completions that makes the sentence true, are then a the degree to which the sentence is true in the original interpretation. The various completions 'resolve' the elements that are vague (not assigned with a truth-value) in the partial interpretation. To make this theory work, Kamp proposes a relatively simple first order predicate logic for classical and partial models. In standard supervaluation theory, partial models are regarded in conjunction with *all* their completions. Kamp (1975:136) suggests a model that considers only a subset of the completions, to which a probability function is applied. The final model is a 'graded context-dependent model' (Kamp 1975:142-3).

Kamp concludes that this model accounts for comparatives and the classes of adjectives that led to this hypothesis in the first place: the adjectives that are extensional but not predicative (*tall*, *heavy*, *small*, etc.). However, this model is highly problematic for the extensional adjectives that *are* predicative, such as *four-legged* and *blue*, and for 'privative' adjectives like *fake*, which are both part of this study.

Kamp&Partee (1995) is basically a reaction to the assumptions made by Oherson&Smith (1981), who argue that a prototypical approach for conceptual theory is impossible because of its incompatibility with fuzzy-set theory. Their main line of reasoning is that fuzzy-set theory is in adequate rather than prototype theory, and the latter fits well in the theory of supervaluation. Kamp&Partee (1995:140) honour the tradition of Montagovian semanticists 'to generalize to the hardest case'. This means that we will still regard adjectives as function from intensions to intensions. The classification of adjectives in Kamp (1970) is modified to an extend:

> Adjectives in Kamp&Partee (1995)[21]
> `Intersective`: when $\|AdjN\| = \|Adj\| \cap \|N\|$ holds for any N. For each intersective adjective meaning ADJ' it holds that $\exists P_{\langle e,t\rangle}\ \forall x_e$[ADJ'(Q)(x)$\leftrightarrow P(x)\&^{\vee}$Q(x)].
> `Subsective`: when $\|AdjN\| \subseteq \|N\|$ holds for any N. For each subsective adjective meaning ADJ' it holds that $\forall Q_{\langle s,\langle e,t\rangle\rangle}\ \forall x_e$[ADJ'(Q)(x)$\rightarrow^{\vee}$Q(x)].
> `Non-Subsective`: when $\|AdjN\| \neq \|Adj\| \cap \|N\|$ and $\|AdjN\| \not\subseteq \|N\|$ holds for any N. For these adjectives there is no meaning postulate, they can be described as *non-committal*.
> `Privative`: when $\|AdjN\| \cap \|N\| = \emptyset$ holds for any N. For each privative adjective meaning ADJ' it holds that $\forall Q_{\langle s,\langle e,t\rangle\rangle}\ \forall x_e$[ADJ'(Q)(x)$\rightarrow \neg[^{\vee}$Q(x)]].

It is argued that the main difference between a subsective and an intersective adjective is that subsective adjectives are sensitive to information of the noun that is part of its intension. However, intersectives as a class are also not quite straightforward: examples like *tall*, *big*, *heavy* and *small* are considered intersective but *vague* and *context-dependent*.

The main focus of this article is the defense and extension to the supervaluation theory that I summarized earlier. Kamp&Partee add a *c*-function which is split up in c^e and c^p. When applied to a concept, like *teenager*, the first is the degree of membership in the extension of a concept and the latter the degree of prototypicality. The most relevant aspect

[21] The additional meaning postulates are cited from Partee (2001:3)

of SVT that is described in Kamp&Partee (1995) is that apparent contradictions can be re-interpreted as non-contradictory. The possibility of re-interpretation in case of an expression that would otherwise be 'semantically impossible' might be of use in accounting for the apparent impossibility of for example *fake gun* and *stone lion*. This is related to the assumption in supervaluation theory that concepts can be *vague* and *context-dependent*. As noted before, Kamp (1975) already assumed that *somehow* the context picks out the particular model out of the set of partial models. Kamp&Partee (1995:161) describe effects that the context has in certain linguistic constructions:

> CONTEXT EFFECTS
> `Parallel structure effect (PSE)`: In a conjoined structure, each conjunct is interpreted in a parallel way relative to their common context.
> `Head primacy principle (HPP)`: In a modifier-head structure, the head is interpreted relative to the context of the whole constituent, and the modifier is interpreted relative to the local context created from the former context by the interpretation of the head.
> `Non-vacuity principle (NVP)`: In any given context, try to interpret any predicate so that both its positive and negative extension are non-empty.

HPP is a principle that relates to the possibility of 're-interpretation', which is later referred to as *recalibration*. This principle is explained by Kamp&Partee (1995:160) that in an Adj-N construction, the adjective is such that it is '*able to make distinctions* within *the class of possible referents for the head noun*'. After this recalibration of the modifier in constructions like these, it is possible to simply interpret the construction as a conjunction. That is, assumed that the adjective would normally be intersective.[22]

This seems to work sufficiently for Adj-N constructions in which the N is 'sharp' (i.e. where its extension has clear boundaries). However, Kamp&Partee (1995:166) admit to themselves that they '*do not have very clear intuitions about what recalibration should be like in situations in which N fails to be sharp*'. It will be interesting to see if this modified version of supervaluation theory and the principles of interpretation proposed in Kamp&Partee (1995) offer a suitable solution for at least the adjectival cases of MTC, which will be discussed next.

2.4.2 MTC in this framework

This section will discuss the treatment of adjectival MTC in the formal framework of Kamp and Partee. The example *stone lion* is explicitly discussed in Kamp&Partee (1995), and in addition to that, Partee (2001) discusses the example *fake gun*. As I have argued before, this literature is specifically focused on the semantics of adjectives, so this section will only

[22] In order to adjust supervaluation theory to this recalibration mechanism, Kamp&Partee illustrate this with the example of *striped apple*. It is proposed that c is a function that assigns a graded extension to the primitive predicates, such that every object in $U_{\mathscr{M}}$ is assigned a number in the interval [0,1]. It is assumed that the core model $\langle \mathscr{M}, c \rangle$ determines a set of 'supermodels' $\langle \mathscr{M}, \mathscr{S}, \mu \rangle$. In the modification of supervaluation theory, the core model is seen as a 'presupermodel'. Assumed than we have a sharp extension of $\|N\|_{\mathscr{M}}$ in a Adj-N construction, we can consider the 'best case' of Adj-N within N as being definitely in the positive extension of N, and the worst in its negative extension. This way, we have a recalibrated c-function c', which leads to a presupermodel $\langle \mathscr{M}, c' \rangle$. The set of supermodels that are determined, are now based on the recalibrated Adj in the *context* of N. Applied to the case of *striped apple*, this means that the meaning of *striped* is first recalibrated in the extension of *apple*. This creates a presupermodel and supermodel that can simply be analysed as a conjunction.

discuss the adjectival examples of MTC (see section 1.1.1). It will be made clear that we can associate *chocolate teapot* with *stone lion* and *imaginary friend*, *wannabe punker* with *fake gun*. Let's start by looking at the possibilities for *stone lion*.

The first thing that is considered in Kamp&Partee (1995:162) is '*is a stone lion a lion?*'. Their answer is that although scientifically, the answer would be *no*, but argue that any kind of material can replace *stone* in this construction and any sortal noun can replace *lion*. In other words, this is not an idiomatic expression with a fixed meaning. The explanation that is that NVP seems to override HPP (see section 2.4.1): the head N is reconstructed in order to include the adjective into its positive extension. More than that, Kamp&Partee (1995:163) argue that we need to change the meaning of *lion* before we can interpret the meaning of the whole construction. This need for change of the meaning of N is problematic for any compositional account of meaning: *stone lion* means *a lion made of stone*. Within compositional theories it would be unlikely to regard this meaning as part of the lexicon.[23] When this example is analysed in terms of supervaluation, Kamp&Partee (1995:167) note: '*What seems to drive the reinterpretation in a case like* stone lion *is the fact that on the original or primary meaning of the head noun* lion, *there is no way to construe the modifier as having a positive extension that overlaps the positive extension of the head noun at all*'. The problem is that we need to account for two changes in meaning if we follow Kamp&Partee. They assume that not only do we need to recalibrate the adjective, the noun undergoes a 'shift of interpretation' (what I refer to as MTC, a shift of meaning type). This is quite a problem for HPP, which would recommend us to interpret the head noun relative to the external context (outside of the NP) and recalibrate the adjective relative to the context that results from this. It seems that in some respect, the reverse is going on in this example.

Kamp&Partee (1995:167) say literally that it cannot be expected from a compositional theory of meaning to '*tell us what is the most salient available secondary meaning will be if the primary meaning fails*'. So, compositional theories of meaning are 'excused' from the task to account for the shift of *lion* (in this case) and for the interpretation of the construction. According to Kamp&Partee this is part of a complete theory of understanding, together with metaphors and other typical ways we use words to express non-literal meaning. Or is it? Later, in the same section (Kamp&Partee(1995:167)), it is argued that *stone lion* can be simply interpreted by a simple predicate conjunction '*once the meaning of the head noun has been shifted*'. This seems odd to me, since the shift of *lion* is directly related to the meaning of *stone*. It would indeed be very difficult to adjust supervaluation theory in order to account for cases as stone lion. First of all, HPP doesn't allow the noun to be re-interpreted relative to the meaning of the adjective. To allow this, would lead to problems in this compositional framework. It would be an option to consider this as a secondary rule, rather than a secondary meaning alone, perhaps. It will be part of the discussion in the next chapter to re-analyse the semantic rules in Mod-Head/Adj-N constructions. To close the case of *stone lion*, it is clear from Kamp&Partee (1995) that they have no real solution within the theory of supervaluations for this example, nor do they intend to find a way to resolve this phenomenon. That would rule out a solution for chocolate teapot as well. However, the 'interpretation'

[23]Note that the meaning type N_{ima} is not a type that is used in any of the semantic theories. Although Kamp&Partee (1995) suggest that the case of *stone lion* is by no means a single occurrence of this kind of meaning, they make no suggestions about systematically integration of this type of meaning. This is kind of an interesting fact, given that the theory under discussion in this article is prototype theory, one which especially allows for 'deviation' of the prototypical meaning. These issues will be further addressed in section 2.5 and chapter 4.

shift that takes place according to Kamp&Partee is from *animal* to *artefact*, I wonder if they would consider *teapot* to undergo a similar shift in *chocolate teapot*, since both can be seen as artifacts as long as there is no interpretation (sub)type specified for images. Let's for now assume that the problem with chocolate will not be solved by using the theory as proposed in Kamp&Partee(1995).

In her short note[24], Partee (2001) admits that although Kamp&Partee (1995) made useful adjustments in the theory of adjectives by allowing them to 'coerce' in context (I assume this refers to 'recalibrate'), the case of *stone lion* remains unresolved. Her work is mainly focused on the group of adjectives that are standardly called *privative*. Based on an analysis of 'NP-splitting' data in Polish, Partee (2001:1) proposes an a new account for adjectives, in which privative adjectives are '*reanalyzed as subsective adjectives*'. The new account includes the possibility for the head N in Adj-N constructions to coerce; that can be motivated by certain adjectives that represent presuppositional meanings. Partee (2001) in a nutshell seems a very promising development in formal accounts for (adjectival) semantics. It will be worthwhile looking at this proposal more closely.

As in Kamp&Partee (1995), Partee (2001) describes privative adjectives as part of the class of the non-subsectives. Plain non-subsectives are adjectives that do not yield any entailments and have no meaning postulates, such as *potential*, *alleged*, *arguable*, *likely* and *questionable*. The notion of privativity as described in the previous section, is debatable. Partee (2001:4) argues that if privative meaning would indeed exist, the combination of the sentences *a fake gun is not a gun* and *is that gun real or fake?* wouldn't be problematic. Partee argues that there is no such thing as a privative adjective. This is based on both examples with the usage of fake and data from Polish. The Polish data lead to the conclusion that 'privative' adjectives behave similar to intersective and subsective adjectives when it comes to NP-split. Only non-subsective, non-committal adjectives, that cannot occur predicatively, cannot be split in Polish. For Partee this phenomenon supports the view that adjectives like *fake* and *imaginary* should be analysed as *subsectives* and not as privatives (that are non-subsective). In the end, this leads to the assumption that the only distinction in adjectives is *subsective* vs *non-subsective and non-committal*. Looking at the interpretation of *I don't care whether that fur is fake or real* seems to be sufficient: Partee (2001:5) proposes that in this sentence '*we actually expand the denotation of 'fur' to include both fake and real fur*'. When a noun is modified by a privative, it coerces the noun such that fake N is included in the interpretation of N. Therefore, adjectives like *fake*, *imaginary* and *fictitious* are subsectives with the special property that allows them to coerce the nouns they modify. Partee (2001:5-6) adds that '*without the coerced expansion of the denotation of the noun, not only would* fake *be privative, but the adjective* real *would always be redundant*'. This coercion would be impossible if the earlier established HPP would be seen as an absolute principle. In her new proposal, Partee suggests that HPP may be violated when it's impossible to obey NVP without shifting of the head N.

Partee (2001) notes that the possibility of coercion in Adj-N constructions is an extremely productive one, since constitutive material modifiers (like stone) often and easily lead to coercion with the interpretation of representation of N (which comes close to my meaning type N_{ima}). This is where the formal approach for MTC becomes attractive, at least for resolving the adjectival cases. Especially the shifts that lead to representational interpretations are supported by the prototype theory of concepts that was part of the discussion in Kamp&Partee

[24]Barbara Partee wrote this note for the occasion of the 60th birthday of Hans Kamp.

(1995); which would also mean that the extensions of concepts (here nouns) are adjustable to their context. Partee concludes that adjectives are still to be seen as functions from intensions to intensions, however constrained to be subsective. Looking back at the phenomena of Polish NP-split, the only real distinction that we would need would be for the plain non-subsective 'non-committal' adjectives (*alleged*, *potential*, etc.).

We now seem to have a better proposal for the analysis for MTC in the adjectival cases. However, where Partee (2001) starts out by describing the coercion in terms of *presupposition accommodation*, this is not further elaborated for the adjectives. I assume that she means that the adjectives that lead to this interpretation effect carry a sort of presupposition.[25] For example in *stone lion*, *stone* as a modifier presupposes the head to be inanimate. When the head noun has is not inanimate by default, coercion (expansion such that the denotation of the noun includes this meaning) takes place. This can be seen as a form of presupposition accommodation. I will contemplate some more on how this could form a basis for the resolution of MTC later in section 2.4.3, and in chapter 3.

Although we seem to be getting nearer to a suitable account for MTC, of course we should keep in mind that this short note is a proposal rather than an elaborated solution for the issue of *fake gun* and *stone lion*. For instance, it seems useful to study the expansion of the denotations of N: to see where the limits are set, or if they are bound to a certain interpretation. In Partee (2001:6) it is mentioned that the shifts that can take place when a noun is modified by a constitutive material modifier, '*it is evidently so easy to shift nouns from their literal meaning to a meaning 'representation/model of...' that we hardly notice the shift*'. I would like to add that also for the old 'privative' adjectives modifying N, it holds that the result is an interpretation of a representation of N. This way, the shift is generalized, and we could assume that the limit of the expansion of in these cases is the representation of N.[26] This goes along with the question how lexical information is seen in this framework: if we would assume that the interpretation is somehow the result of *presupposition accommodation*, it would follow that the extended meaning of N is not present in the lexicon, but is added somewhere along with the process of accommodation. However, if the shift to the representation-interpretation is so easy, I could imagine that the meaning is somehow included in the lexical item.

Overall, the proposal that was developed from Kamp (1975) to Partee (2001) give a satisfying account for the way MTC meanings in Adj-N constructions are derived. Especially the interacting principles HPP and NVP will be very useful in the development of a complete account of MTC. As will be clear from the next section, we need to think more about the notion of *presupposition accommodation* in MTC.

2.4.3 Extension to other cases of MTC

This section will be focused on the question if the formal framework as presented so far, will have potential to be extended to account for the other cases of MTC. As I noted in the introduction of 2.4, the papers I discussed deal with adjectives only. I will not start over by

[25] Note that this can also be seen in the second statement that Kamp (1975:124) describes for the Montagovian Doctrine.

[26] Note that there could be other expansions of nouns, for instance when they are interpreted in a metaphorical manner or other non-literal ways. Since these interpretations are not caused by lexical items in the context of N, I will propose that we take the representation of N not as a non-literal meaning but as a secondary meaning of N. This would explain why we 'hardly notice the shift' as well. More on secondary and non-literal meanings will be discussed in chapter 3.2.2.

summarizing articles on verbs and other linguistic elements that we have seen are capable of being an MTC-functor. Instead, I will build some ideas from parts of the articles I have used so far, and use the principles of presupposition from Geurts&Sandt (1997) and (2004) to create more 'foundation' for the extension.

In all the articles discussed so far in this section, it has been clear that the context of a linguistic element can play a role in its interpretation. The most relevant example discussed in Kamp&Partee (1995:142) is '*My 2-year-old son built a really tall snowman yesterday*' versus '*The D.U. fraternity brothers built a really tall snowman last weekend*'. It is argued that the intersective *tall* is recalibrated in the context created by the head N, which in turn is interpreted relative to the context of the whole constituent (HPP). This works well, *snowman* is interpreted relative to the subject (its creator), so *tall* for a 2-year-old can be *small* for a fraternity member and still result from simple predicate conjunction in both sentences. If we assume that the part of HPP that explains the interpretation of N can be isolated, we would conclude that a NP (modified or not) is always interpreted relative to its context. This could be a very fortunate principle for other cases of MTC. The question is, though, how many adjustments we need for this sub principle of HPP in order to relate NP to its (linguistic) context. For instance, it is not clear from neither Kamp&Partee (1995) nor Partee (2001) *how* the contextual factors that play a role in the interpretation of context-dependent adjectives, are *recognized*. In other words, we would need a more formalized definition of context (C) to explain the interpretation shifts in my examples.

Before we focus on that, there is one example in particular that correctly predicted by the proposals of Kamp and Partee. That is the case of (6): in this example the head N of the construction is modified by a PP. The N contained in the PP undergoes the interpretation shift to a representation interpretation. Although shifts on nouns that function as modifiers of other nouns are not explicitly mentioned in the articles discussed above, we can assume that HPP allows for shifts on the modifier if this is necessary to fit into the extension of the head (recalibration). The question here is if we should assume that NVP is the responsible factor here, since the 'normal' interpretation would lead to an empty positive extension of the predicate. I assume that NVP allows for the shift that is needed in order to satisfy HPP, so the two principles operate together in this case.

The other cases of MTC are more complicated to account for, the examples where TOPIC, LOC and VERB function as MTC operators for N (see section 1.1.2). As I mentioned in the previous section, Partee (2001:1-2) suggests that the coercion in cases as *stone lion* and *fake gun* is a form of presupposition accommodation. It is suggested to treat '*the constraints on possible adjective meanings as presuppositions that must be satisfied by any use of an adjective*'. Within the domain of formal semantics/pragmatics, I assume we could treat other instances of MTC as presupposition accommodation as well. First, we need to interpret what is meant by Partee (2001:1).[27] I take the views in Geurts&Van der Sandt (1997) and (2004) as standard views in theories of presupposition. Their approach is to represent presuppositions in DRS, for more details on DRT I refer to Kamp&Reyle (1993). Geurts&Van der Sandt (2004:7) explain that presuppositions '*want to be bound*'; their account for presuppositions is called binding theory (in terms of behavior of presuppositions can be compared to pronouns). In a nutshell, this binding theory can be summarized to two core points: 1. if there is a suitable antecedent for the presupposition, the presupposition is bound to that antecedent, 2.

[27] Note that it would take a complete elaboration of presupposition theory to work this out, I have chosen to just give my impression of what is proposed in Partee (2001).

if there is no suitable antecedent, the preposition is accommodated. Accommodation preferably takes place in the main DRS, if that is not possible, '*the binding theory predicts that there is a preference for accommodating it further down in the chain of DRSs linking the presupposition's home DRS with the main DRS*' (Geurts&Van der Sandt (2004:9)). Moreover, accommodation is *constrained*, as described in Geurts&Van der Sandt (1997:1), '*accommodation must yield a coherent interpretation*'. I won't discuss the DRS's and the location of the accommodation for the cases of MTC, let's just see how far we get if we regard MTC as a form of presupposition accommodation. For this purpose, the assumptions that any presupposition that is accommodated, apparently has not had the opportunity to be bound and that accommodation should benefit discourse coherence, should be sufficient.

I suppose the presupposition projected by an adjective A is that there is an individual such that the property A holds for that individual.[28] In my examples of MTC in Adj-N constructions, it is obvious that we have no complete sentences, so it will be hard to define a real presupposition based on these exact data. Suppose we add *There is...* to examples (1)-(5). I assume that Partee (2001) means that an adjective presupposes a noun that it modifies. So the presupposition of *there is a stone object* can be represented as [x: object(x), y:made of stone(y), x = y].[29] In this representation the presupposed part is underlined, as in Geurts&Van der Sandt (2004). Apparently, the presupposition of the adjective, can only be bound to a *semantically compatible antecedent*. So, if we replace *object* by *house*, we get [x: house(x), y:made of stone(y), x = y]. The default meaning of *lion* is semantically incompatible with stone. In order for *stone lion* to be interpreted in a coherent way, Partee (2001) proposes coercion, what I call coercion to the meaning type N_{ima} (representation of lion). This coercion can be seen as a form of accommodation, the default semantic properties of *lion* are incompatible to be bound to *stone*, so *lion* coerces to *representation of lion* that *is* compatible with *stone*. How this should be worked out in proper DRSs is beyond the scope of this study.

This account of MTC in terms of presupposition accommodation can be applied to the examples where verbs are MTC functors as well. We have seen that verbs have particularly strong demands for their arguments. These demands, in syntactic theory θ-roles, can be seen as presuppositions. For example the verb *to begin* presupposes an event, as we have seen in section 2.3.2, where *book* is interpreted as an event. In example (7), the verb to trip over presupposes the height of its object to be of at most 25% of the height of its subject (see section 2.3.2 again). This presupposition is bound whenever the argument that is realized, fulfills this requirement. In case of (7a), this is not the case. In order for the presupposition to be accommodated, the realized N *truck* shifts to the interpretation of *representation of truck*. Although the details of the representation in DRSs will need more work, this seems very promising.

A next step would be to see if this presuppositional account will hold for the other cases as well, like examples (8)-(11), where either the TOPIC or Location in the discourse are MTC functors. If we follow Geurts&Van der Sandt (2004:10), I can be rather short on this part. They say that in *When I came home last night, I noticed immediately that the guillotine had been fiddled with*, '*the problem with this is not so much that 'the guillotine' is not specific enough but rather that it is* remarkable *that in our day and age a private person should own*

[28] Note that I use *individual* here, and not property. In DRT it is common to assign properties to individuals, which does not contradict the views on adjectives by Kamp and Partee.

[29] I use *made of stone* here in order to avoid confusion between nominal and adjectival use of *stone*.

one- and presupposed information is expected to be unremarkable'. Note that this examples is similar to example (8) in this study. Geurts&Van der Sandt refer to Asher&Lascarides (1998) for the resolution of these cases, which should then lie, again, in *bridging*. We can see bridging as an inference that relates objects or events in a discourse, that are not explicitly defined to be related that way. If this inference does not take place, the text (discourse) would be incoherent. In other words, bridging serves the purpose of making discourses *maximally coherent*. Asher&Lascarides (1998) claim that bridging can also occur in absence of presupposition triggers, and it should be seen as a by product of discourse interpretation. In the representation, they use SDRT, which is an extended form of DRT that incorporates rhetorical relations. These are especially important when it comes to solving the cases of MTC that are left. The DRSs are 'glued' together using the formal pragmatic theory DICE (Discourse in Commonsense Entailment), such that computing the rhetorical relations is possible. It would be far beyond the scope and aim of this thesis to describe this theory in detail, but it is interesting to see what kind of relations can now be formalised. Just for the sake of being complete, DICE works with a discourse *update function*, the functioning and explanation of which can be found in Asher&Lascarides (1998:96-97), Asher&Lascarides (1995:80) and Lascarides et al. (1992). In this framework, it is common to use schemata based on pragmatic maxims, along with their constraints on those schemata. Here I will only consider the *informal* rules for bridging that are proposed in Asher&Lascarides (1998):

> INFORMAL RULES FOR BRIDGING IN SDRT[30]
> `Bridges are Plausible`: a bridge must be built from consequences of the semantic content of the constituents.
> `DS Determines Bridging`: when the rhetorical relation used to connect the constituents offers a way to resolve a bridge, we use that way.
> `Maximize Discourse Coherence`: resolve the bridge to obtain maximal discourse coherence.

Before Asher&Lascarides start to explain their account of bridging, they first explain the shortcomings of the theory of bridging proposed by Bos *et al.* (1995), also described in section 2.3.3 of this study. The problem here is that all inferences that we need to bridge information, are based on lexical information. It is true that in GL allows other domains of knowledge to become part of lexical information. However, as I stated before, if you would have to explain example (8) of this study by bridging as proposed in Bos *et al.* (1995), you would have to allow for *elephant* to be some quale of *room*. If we would have to include this kind of knowledge in the lexicon, there would be a lot of redundant information in there. Asher&Lascarides (1998) state that there are many forms of knowledge that are used in the process of bridging. The set of rules that I described above are based on the assumption that bridging is necessary to obtain discourse *coherence*.[31] We see this in the statements listed above: the bridge is resolved to make a discourse coherent, and if the discourse offers a way to connect the constituents, we use that way (a rhetorical relation). It could even be the case that there is more than

[30]The rules are quite complicated for bridging, in order to avoid an even longer elaboration about the meaning of the formal denotations, I have chosen for the informal description of them in a summarized form. See Asher&Lascarides (1998:96) for *Bridges are Plausible* and *DS Determines Bridging* and *ibid.* p. 97 for *Maximize Discourse Coherence*.

[31]Asher&Lascarides (1998:89) illustrate that with the example *I've just arrived. The camel is outside and needs water*. The usage of rhetorical relations and coherence constraints makes it possible to use both lexical knowledge *and* world knowledge.

one rhetorical relation that could connect the constituents, `Maximize Discourse Coherence` tells us to use the relation that results in the 'best coherence'. `Bridges are Plausible` is probably the constraint that will leave us with a problem in resolving (8). This constraint is based on the assumption that the constituents that are bridged can be related within the taxonomy of Clark (1977). We see that this allows for *camel* to be bridged to *arrive* in the example of Asher&Lascarides (1998), *arrive* is a motion verb (*lexical* knowledge), *camel* a means of transport (*world* knowledge); the two forms of knowledge are used to meet the coherence constraint.

In (8) we have another issue to deal with, namely that *elephant* is most likely interpreted as a (small) representation of an elephant. My assumption is that in this case, Asher&Lascarides would use LEXICAL IMPOTENCE (Asher&Lascarides 1995:96-97). This principle basically says that a constituent should interpreted such that it doesn't create an incoherent discourse interpretation. In other words, it ensures that whenever there is a *conflict* between the discourse information and the a sentential rule for a specific interpretation of a given lexical item, the discourse information wins in order to avoid discourse incoherence. If we bridge *elephant* to *room*, we would end up with an incoherent discourse if we interpret *elephant* in the default way. LEXICAL IMPOTENCE could be seen as a way to let the discourse coerce this constituent to a sense that leads to more coherence, it this example a representation of an elephant that can be bridged to room as it serves the purpose of ornament for room. The latter is a difficult point, if we take the taxonomy proposed by Clark (1977) into account, which is fundamental for most literature on bridging: as I pointed out in 2.3.3, elephant and room cannot be related the way that is required when one wants to bridge. We are left with two issues that cannot be fully dealt with in this proposal for bridging: there is no clarity about the 'whereabouts' of the coercion in (8), and there is no clarity about the (rhetorical) relation that should support this bridge. This doesn't mean that this approach fails to account for MTC in discourse, it only means that more phenomena should still be captured by its rules.

A last suggestion made in Asher&Lascarides (1995) that may be of (future) use, concerns the way they see discourse interpretation in *narratives*. This is expressed in for example the DICE law SPATIAL CONSTRAINT ON NARRATIVE (Asher&Lascarides 1995:93), which captures the way expectations about locations in a narrative are exploited. It states that unless the narration contains information about changed location, location of x at these spatio-temporal points are the same. In a sense, this would mean that we are forced to interpret *elephant* as part of the *room* in (8), since the location in the narration is not changed. The adaption of elephant is then motivated by LEXICAL IMPOTENCE. Still, we would have to elaborate these laws some more to account for *meaning coercion* in these cases. Examples (10) and (11) are still somewhat unaccounted for, intuitively *child* in (10) and *Bert and Ernie, Steven Spielberg* in (11) set up certain expectations in the discourse. These are *expectations* that the hearer has due to the topic of the discourse, based on world-knowledge.[32] I expect that also these phenomena could be translated into DICE laws, but they aren't as far as I know.

On the whole, the frameworks discussed in this section are very promising. The advantage of the way formal semanticists tend to work, is that the rules that are formulated are very specific in regard to the phenomenon they account for, yet they are general because of the way they are formulated. We have seen that the extended version of DRT as proposed in Asher&Lascarides (1995), (1998) has great potential because of its focus on the discourse

[32]It is assumed in the work of Asher&Lascarides that once an expectation is triggered in the mind of the hearer, it is not easy to cancel or undo that.

effects on meaning. As I have argued above, within this framework there are still some questions left unanswered. Partee (2001) suggests the coercion being a form of presupposition accommodation for the cases of adjectival MTC. In the extension of the framework, this does only apply to the cases of MTC that are triggered in a syntactic relation. In a complete account we will have to figure out whether this distinction between syntactic triggers and discourse triggers is a necessary one or if there is a way to avoid this. This will be further elaborated extensively in chapter 3. A last important advantage of the DRT approach is that it justifies the intuition that lexical meaning or default meaning of in this case N, is somehow violable when it leads to incoherence or 'semantic unwellformedness' in its context. This is also clear from the principles stated in the work of Kamp&Partee (1995); NVP and HPP. Knowledge of the world is included in the process of interpretation, however, we need to ask ourselves to which domain of knowledge the process of MTC can be ascribed. In Asher&Lascarides (1995), (1998) it is assumed that lexical knowledge is quite restricted. So, the interpretation that follows from coercion, and the reason for this coercion should *both* come from knowledge of the world. This way, they avoid a lexicon with unnecessary and unwanted information. We can see here already that the conclusion of this section is more or less an introduction to a discussion of the results within existing frameworks and the questions we need to address.

2.5 Discussion

So far, I have addressed a few directions within semantics that have potential to resolve the MTC puzzle. Before we make an attempt of creating a complete account for MTC, I will shortly discuss the relevance of the theories discussed so far.

In 2.1 I described Mental Space Theory (MST) as proposed by Fauconnier (1994). I have argued that although its conception of meaning is too broad in my view, MST has some useful insights regarding the interaction of background knowledge and meaning processing. The theory is really focused on meaning in context, and is thereby (to a certain degree) comparable with the underlying thought of DRT. I have argued that this approach is problematic for my purposes because of its lack of any kind of features in the specification of (lexical) meaning. This is consistent with the assumption of Fauconnier that there is no such thing as a separate *linguistic* meaning, i.e. interpretation is by all means the result of a general *cognitive* process. It wouldn't be impossible to describe the process of MTC in terms of mental spaces. My reluctance to use this framework is merely due to my presumption that a lot of linguistic hinges that could be part of MTC would be ignored. This, in turn, depends on my own perspective on what a semantic theory should involve and what not. For this, I refer back to the preliminaries to this study. A useful 'tool' that we should keep in mind though, is the way Fauconnier considers `Space builders`. As we have seen in the last section, we need to account for discourse MTC functors in general (i.e. not restricted to explicit *locations*, like I discussed in section 2.4.3 for Asher&Lascarides (1995)). In this respect the 'sloppiness' (i.e. being too broad) that I hold against the framework of mental spaces, can be an advantage. In the next chapter, I will explore how `Space builders` can be involved in the process of MTC.

The framework of Sense Generation (SG), as described in section 2.2, is directly focused on the Adj-N examples like presented in this study. Also for this framework, I have described my reluctance in using it as an account for MTC. We have seen that the distinction of *diagnostic* versus *central* features as proposed in Franks (1995), is not clear enough to account for *stone lion* and *fake gun* in a similar way. Where DRT and MST are mainly focused on contextual

factors that may influence the interpretation of a lexical item, SG is organized around the lexical properties and how a context may lead to a different interpretation that the default content they have. This is clear from the principles that Franks formulates for SG (cited on p. 38 in this study) that the context variables that could play a role in shifts in meaning, are not specified. It is simply stated that unless there is a reason *not* to, the concept can and should be interpreted according to its default content. There are some similarities between the approach of Franks and the one Pustejovsky takes in GL. Both choose to represent meaning in matrices (AVS in Franks (1995)). The difference is that Franks does try to sketch different information states that should be part of meaning processing. Although I have argued in the section about Pustejovsky's GL that it lacks representation of process, I am not satisfied with the information states proposed by Franks as well. This is merely due to the nature of these kind of matrices. They are quite useful in the representation of (default) lexical meaning, and even for generation of complex concepts, as long as they can be explained in compositional terms. In other words, what I have also seen in the matrices in GL, it is really difficult to see what actually happens when the meaning of a lexical item collides with the semantics of items in the context (syntactically related or not). Contrary to the assumptions made in Partee (2001), Franks (1995) assumes not only that privatives exist, but also that that affirmative adjectives can occur as *functional* privatives (as opposed to *real* privatives like *fake*). This assumption is useless for my purposes, since MTC can also be triggered by elements outside the NP. It seems more plausible to regard adjectives ('privative' or not) as part of the class of potential MTC operators or functors. I tend to prefer the assumption made by Partee concerning the non-existence of absolute privative meaning in natural language, which will be discussed later in this section. Franks (1995) comes up with a couple of assumptions that I think will be useful in the full account for MTC. Especially the basic meaning for a lexical item that is flexible for change caused by context (`Default Content` and `Default Classification`) will be elaborated some more. Next to that, it seems plausible to follow Franks idea about `Noun Instantiation` to a certain extent. That is, its basic assumption appeals to the intuition that a N_{ima} interpretation that is the result from for example *stone lion* is in principle just a 'representation' interpretation. What kind of representation it is, has to follow from the (possibly non-linguistic) context.

The theory of GL as outlined in section 2.3 is mainly focused on the lexical aspect of shifts in meaning. The basic assumption is that different senses can be generated from the information that is stored in the lexicon for the lexical item in question. We have seen that the type coercions treated by Pustejovsky (1995) are coercions of semantic (Montagovian) types. The kind of type coercion we are dealing with in this study is a coercion of what I call *meaning* type. However, the way of structuring the information in the lexicon into `Qualia Structures` could be very useful. In the work of Asher&Lascarides (1995:89) that is discussed in 2.4 this way of structuring is used to disambiguate lexical items that can have both a *mass* and a *count* sense in context, using DRT. In Asher&Lascarides (1997) it is claimed though, that GL is not suitable in the account of bridging phenomena. This is in line with my own claims when I explored the possibilities of extending GL to resolve discourse-triggered cases of MTC. What especially concerned me in studying this approach, was that the possible resolutions for MTC that seemed to work, could only work if we would allow more than one entry per lexical entry. Which contradicts the aim of GL in the first place. This might be avoided if we analyse N_{ima} the way mass/count are analysed in Asher&Lascarides (1995). As for the adjectival examples, I have formulated the GL definition `Metonymic Binding`. Where Pustejovksy's `Selective Binding` accounts for subsective adjectives binding only to a certain quale of the

noun, `Metonymic Binding` should have potential to account for coercion of certain qualia. But it remained difficult to account for the coerced interpretation of the head N. We have seen that GL is especially suited to structure and explain the demands of verbs. However, when we want to construct the meaning of a verb that is a MTC operator in relation to one of its arguments, we face some serious problems. In my opinion, these difficulties are due to the way of representation as well as to the overall impression that GL is not a suitable framework to 'visualize' the *process* of interpretation. In other words, GL is a framework that is mainly focused on the organization of the lexicon. In an account of MTC it will be necessary to make assumptions on what is lexical information (or knowledge) and what are other forms of information that can play a role in interpretation. The qualia-based approach to the lexicon could be of use here. My idea would be to use GL as lexicon *underlying* processes such as MTC, but not to account for those processes. It could be of use to distinguish between lexical knowledge and other forms of knowledge that are used in interpretation.

The power of the formal theories presented in section 2.4 lies merely in their precise yet general descriptions. We have seen that the examples of *stone lion* and *fake gun* are quite problematic in the first (Montagovian) account of adjectives described in Kamp (1975). From there on there has been a 'development of thought' where the influences of context where included more and more (Kamp&Partee (1995)), to the point where Partee (2001) proposes a new analysis of privatives, i.e. that there is no adjective that is strictly privative in natural language. This sheds a new light on the issue of MTC, except that it just changes the idea of what Franks (1995) calls real privatives. Privatives like *fake* and *imaginary* are now re-analyzed as subsective adjectives that somehow lead to coercion of the head they modify. It should be noted that this still doesn't offer a solution for substance-adjectives like in stone lion (Franks's functional privatives). The rather unsatisfying explanation for this is that compositional semantics cannot be expected to explain this type-shift. In my search for the MTC solution it will be necessary to find a way to solve this, a temporary suggestion would be to assume that these kind of adjectives are intersective and can possibly lead to coercion of the head in case of conflicting properties. The suggested principles `Head Primacy Principle` and `Non Vacuity Principle` play an essential role in this explanation. This way the so-called functional privatives could be accounted for as well. Another thing that would still need some work is the interpretation as representation after coercion, which also lacks in GL, and is slightly addressed in MST and SG. This can be connected to the assumptions made in Kamp&Partee (1995) about Prototype Theory (PT) and to the suggestion made by Partee (2001) to see the coercion in the revised privatives-as-subsectives as a form of presupposition accommodation. Seen in this light, it is possible to analyse the examples in which the coerced N and the MTC operator are syntactically related, as presupposition accommodations. Making steps in this direction, the DRT proposals made by Asher&Lascarides (1997) for `Bridging`, the case of *the elephant in the room* (example (8) in this study) remains difficult to explain. However, this and also other cases of MTC could be resolved when we use principles from Asher&Lascarides (1995) `Spatial Constraint on Narrative` and `Lexical Impotence`. The latter can be related to `NVP` in Kamp&Partee (1995). On the whole, I would regard the theories presented for formal semantics as the most suitable way to resolve the MTC puzzle. However, we still have no real account for the interpretation that takes place; the shift in meaning from object to the representation of that object. The frameworks of DRT/DICE are especially suited to relate different parts of discourse, which is essential in a full account of MTC, however, also here we lack an account for the shift in meaning that is the result from MTC.

To end this chapter of a staggering amount of laws, principles, rules and ideas, it should be clear that there are very useful tools for the solution of MTC that are already present in different semantic theories. In discussing the different theories in this chapter, I have argued that for this kind of meaning effect, it's really essential we adopt a framework that represents the process of interpretation such that the shifts of meaning are somehow visible as well as accounted for. This is merely due to the fact that MTC occurs at different levels of natural language expressions, from phrase to discourse. The only framework that comes close to this 'request' so far is DRT, however, this is still a framework that is primarily focused on the process of discourse. Over the last decades, Optimality Theory (OT) has conquered the field of linguistics. Where it started out being a suitable theory for phonology, nowadays its range includes morphology, syntax and semantics as well. The next chapter will be an attempt to formulate constraints in OT to account for MTC. The main reason to chose this framework is because it's not a theory about language by itself. In a way, OT could be regarded as a new way to look at rules and principles. They are formulated as constraints, which are violable. This makes the concept of 'rule' more flexible, and allows for them to be violated when circumstances demand for that. As a consequence, we can adopt the rules, principles and laws and reformulate them into constraints. The advantage is that these rules are now more flexible, i.e. violable and within OT, they are ranked. This ranking is in accordance with the relative strength of the constraints. This will all become more clear in the first section of the next chapter. For now, let's just list the requirements the ot approach will have to meet in order to account for MTC:

A FULL ACCOUNT FOR MTC IN OPTIMALITY THEORY SHOULD:

- allow for meaning shifts of the type $N_{def} \rightarrow N_{ima}$,
- be based on clear assumptions about lexical information and meaning types,
- allow for other kinds of information to be included in the process of interpretation,
- account for different MTC operators and their nature,
- allow for N to be coerced by syntactically related elements as well as elements in the discourse of N.

Chapter 3

An Optimality Theoretic proposal for MTC

> *'And I said that some people who worked in an office came out of their house in the morning and saw that the sun was shining and it made them feel happy, or they saw that it was raining and it made them feel sad, but the only difference was the weather and if they worked in an office the weather didn't have anything to do with whether they had a good day or a bad day.'*

This chapter will be organized around the requirements for the resolution of MTC that haven't been met by the theories discussed so far. There are many theoretic reasons why an OT approach can be preferred over any other approach, but I think I should admit here that above all, I like the challenge of trying something relatively new here. Next to that, the choice for an exploration within OT is based on the fact that the phenomenon of MTC requires a theory that constitutes *soft* rules for the interpretation of nouns in context.

This first section is merely meant to explain the architecture of OT that allows for rules to function and behave in this way, when solving issues of natural language semantics. This section is the theoretical foundation for sections 3.2 and 3.3 where I will formulate OT constraints for MTC. I close this chapter with a discussion in 3.4.

3.1 Optimality Theory and Semantics

This section will discuss the main characteristics of OT, in order to see the way it works (for semantics in particular). This will be based on Prince&Smolensky (1997) (the article that caused the tidings in linguistics to change), Zeevat (2000) and Hendriks&De Hoop (2001). Section 3.1.1 will summarize the basic principles of the OT approach, i.e. constraints and the way output candidates are generated from input and evaluated. In section 3.1.2 this will be narrowed down to the current research in the field of semantics, illustrated by some examples. In 3.1.3 I will start working towards the issue of MTC by looking at the work of Zwarts (2003).

3.1.1 A general impression of OT

Optimality Theory, first proposed by Prince&Smolensky (1997), steps away from the idea that natural language is organized by *rules*. Optimality Theory finds it source in Neural Networks and Harmony Theory for Linguistics. Its first linguistic application was in the field

of phonology, nowadays a lot issues in syntax and semantics are approached within in an OT framework. The basic idea of Optimality Theory is that languages consist of a universal set of violable constraints. These constraints are highly conflictive and differ in strength, i.e. a constraint can be violated in order to satisfy a higher-ranked constraint. The rankings of constraints are different across different languages, which accounts for language-variation. It has been argued that languages can be typologized based on the ranking of the constraints. In this respect, OT is not unlike the basic assumptions that underlie Principles and Parameters as proposed in the 50's by Noam Chomsky. The difference is, as argued before, 'strict' rules with their many exceptions are now translated in more 'laid-back' constraints. This perspective makes sense considering the many enigmatic phenomena and erratic patterns we come across analysing natural language.

A device called *generator* GEN generates different output-candidates from a given input. A device called EVAL (evaluator) 'assesses' the candidates by looking at the constraints they violate. Some constraints are fatal if violated, candidates that violate those constraints are ruled out instantly. The *optimal* output candidate is the candidate that violates the least higher-ranked constraints relative to the other candidates. In OT, the ranking of constraints has a function of domination: the satisfaction of a weaker (lower-ranked) constraint can not make up for what is lost if a stronger (higher-ranked) constraint is violated.

Note that in principle, OT states that the number of possible output candidates is infinite.[1] In practice, though, the possible outputs that are discussed in linguistic research, are candidates that are close to the input. Input *f*, output-candidates and constraints are represented in OT-*tableaux*, like represented below:

	input *f*	CONSTRAINT 1	CONSTRAINT 2	CONSTRAINT 3
	candidate 1	*!		
	candidate 2		*	
☞	candidate 3			*

The optimal candidates for an output are indicated by ☞, violations are indicated by *, fatal violations by *!. It's possible for a candidate to violate a constraint more than once. Two violations of one constraint weigh heavier than one violation of that constraint. In this abstract example, candidate 3 is the optimal output for input *f*. It is possible for two outputs to be optimal, this is called a *tie*.

The constraints form a grammar can be divided in two sorts: MARKEDNESS constraints and FAITHFULNESS constraints. I will refrain from a detailed discussion about which forms in language ought to be marked. I will just shortly discuss the both classes of constraints, as they shall be of importance in later parts of this study. It is generally assumed that an element in a linguistic structure is *marked* if there exists an alternative that is less complex. Marked forms are restricted because they arise from violations of the markedness constraints. In other words, grammar has constraints that restrict marked forms. Some languages do not have marked forms at all, those that do have them, have them for special purposes. In general, marked forms give rise to marked meanings. In literature, MARKEDNESS constraints are also known as STRUCTURE constraints, if STRUCTURE constraints are violated, a marked

[1] this is by no means an irrelevant or unproblematic property; in section 3.1.2 this will be addressed more detailed.

form arises.

FAITHFULNESS constraints are focused at the relation between the input and the output candidates. According to Prince&Smolensky (1997:1605) '*...*FAITHFULNESS *constraints tie the success of an output candidate to the shape of the corresponding input; each* FAITHFULNESS *constraint asserts that an input and its output should be identical in a certain respect.*' FAITHFULNESS constraints compete with STRUCTURE constraints, the first favoring reduplication of the input, the latter favoring simple structures. I will discuss in the section below how these constraints are to be seen in the semantic and pragmatic domain. The constraints that linguists use in OT are usually not new. OT has the very convenient property of being a flexible theoretic framework, that is, rules that are formulated in other theories are easily reformulated into OT constraints.

3.1.2 Meaning, Interpretation and Optimality

In this section I will give an impression of how OT is used in current semantic and pragmatic research by taking a closer look at Hendriks&De Hoop (2001) and Zeevat (2000). Next to that, the constraints that are proposed in these articles could already provide some us with some fundamental OT constraints we can use for the MTC issue.

Hendriks&De Hoop (2001) is a good example of OT used in the domain of semantics. The aim of this paper is to show semantic issues that are unresolved in a compositional framework, can be dealt with in OT. The phenomena that are discussed in particular are *determiner quantification* and *elliptical comparatives*. I will not discuss these phenomena in detail, however the reason Hendriks&De Hoop (2001:1) give for their being problematic for compositional semantics is essential: '*... in order to yield the intended interpretation for quantified, but incomplete or anaphoric expressions, a compositional interpretation based on syntactic structure alone is not always possible.*' The leading hypothesis of their study is that the interpretation of a syntactically well-formed structure is free; the *Free Interpretation Hypothesis* (Hendriks&De Hoop 2001:13). This freedom is then restricted by violable and conflictive (OT) constraints, in order to avoid an infinite number of possible interpretations for every given input in OT.

Although focused on a different topic, the proposal of Hendriks&De Hoop (2001) consists of a few useful insights. For example, the DOAP constraint tells us to *anaphorize* text as much as possible. This constraint consists of a set of constraints, one of which should be satisfied in order for DOAP to be satisfied (see Hendriks&De Hoop 2001:22-23). This constraint captures the behaviour of presuppositions described in Geurts&Van der Sandt (2004) as well; if a presupposition can be bound to an element in the discourse, it should be bound. In Zeevat (2000) this constraint is formulated as its negative: *ACCOMMODATION (or AVOID ACCOMMODATION). The DOAP constraints are claimed to be conflicting with two relevant (pragmatic) constraints: EMPTINESS and AVOID CONTRADICTION. The latter occurs as a constraint in most of the literature on OT semantics/pragmatics, in fact I think it's implicitly assumed in most theories of interpretation. I will show in later sections how AVOID CONTRADICTION is a principle that underlies the constraints that I will propose for the account of MTC. I assume say that AVOID CONTRADICTION outranks the DOAP constraints. In other words, my intuition is that DOAP can be violated in order to obtain a *coherent* interpretation. We have seen in section 2.4.3 how the need for coherence in discourse leads to other phenomena, such as BRIDGING. It will be part of this study to see how discourse coherence is obtained by constraints on interpretation, and how these constraints predict coercions when

necessary.

On the whole, Hendriks&De Hoop offer us a good idea of how OT can work for semantic phenomena that compositionality fails to account for. It can already be seen here, that the fields of semantics and pragmatics sort of melt together in OT. This is encouraging; I have argued in chapter 2 that it's often the case that a semantic theory will use pragmatics as a bin for things that are beyond the scope of compositionality. It is my opinion that at least the influences of *linguistic* context on meaning should be accounted for within a proposed semantic theory. Constraints like AVOID CONTRADICTION and the consequences for accommodation of DOAP will be explored in the OT approach for MTC.

The constraints that are discussed in Zeevat (2000) are of a more general semantic/pragmatic nature. It is for this reason that I will cite them explicitly here; later in this study it will be clear that they It will be clear from the next section that Zeevat (2000) does not propose these constraints himself, he merely argues that it's difficult to make true *semantic* OT constraints. However, the semantic constraints that are discussed are quite general and probably useful. The constraints are described by Zeevat(2000:249-53), I reconstruct them below:

> *ACCOMMODATION
> If there is a presupposition, nothing should be added if the context (or one of the local contexts) already has the material for an antecedent.

*ACCOMMODATION prohibits presuppositions accommodating their antecedents. In other words, it can only be violated if the context has no candidate antecedent for the presupposition that needs to be bound. It could be stated that this is subsumed by *INVENT:

> *INVENT
> Don't overinterpret by adding some extra feature(s) to the sentence.

This constraint simply forbids the hearer to add extra material to what is said without a proper reason. Evidently, we will need more than this by defining and constraining the 'proper reason' for adding extra material. Next to that, it is questionable what constitutes 'overinterpretation', in the pragmatic sense of the word. Does it mean you can extend the meaning in order to make sense out of the utterance by filling in gaps, or does it also include for example shifting types (features) of elements that are already present in the utterance, as in MTC? This constraint will be revisited in later sections. Another semantic OT constraint according to Zeevat (2001) is:

> STRENGTH
> If possible, choose a reading that is informationally stronger over an informationally weaker reading.

To my understanding, this constraint means that the hearer should always choose the candidate reading that expresses the strongest truth conditions. This constraint is mainly used for OT approaches for the interpretation of presupposition triggers and reciprocals. I will use STRENGTH as a constraint for lexical meaning, which will be elaborated in 3.2. and explained in 3.1.3. Overall, we shall see that these constraints are interpretational principles rather than the constraints that interact in the specific interpretations of natural language expressions.

CONSISTENCY
If possible, interpret the sentence such that there is no conflict with its context.

According to Zeevat, this constraint comes into play in case of ambiguity, when there is more than one way for resolution of anaphora, and in ruling out accommodations. Since it's not *always* the aim of the speaker to be consistent, the constraint is considered to be violable. We have seen that MTC is mainly due to conflicts between different elements in context, I assume CONSISTENCY or a form of this constraint will be necessary in the OT resolution.[2]

FAITH-INT
Interpret all the speaker has said.

Zeevat assumes that this last constraint comes into play when CONSISTENCY is violated. So, in case the hearer cannot select a consistent reading of a sentence, the interpretation he has will be violating this constraint but satisfy Faith-Int (because he interpreted the conflictive reading that resulted from that sentence).

The ranking of this set of semantic/pragmatic OT constraints as proposed by Zeevat (2000:251) is FAITH-INT > CONSISTENCY, ANCHOR[3] > *INVENT, *ACCOMMODATION > STRENGTH. The main conclusion of Zeevat is that there are few *pure* semantic/pragmatic OT constraints that cannot be formulated or considered as syntactic constraints. This could mean that *meaning* (semantics) cannot be analysed apart from *form* (syntax). In the present semantic OT studies, this has led to the idea that should be analysed in terms of *pairs* of meanings and forms, rather than with one of them as the input and the other as the output. This is especially elaborated by Reinhard Blutner, in for example Blutner (1999). For here, it suffices to summarize his proposal of bidirectional OT and state why I don't have to rack my brains over its application in this study.

The core of Blutner (1999) is that an OT approach for syntax on the one side and interpretation on the other, leaves some relevant issues out in the open. He refers to the application of OT in the work of Hendriks&De Hoop (2001) as *unidirectional* OT. Unidirectional ot separates syntax (expression) and interpretation. A unidirectional form of OT has an input, an output and the selected optimal candidate for the input. For syntax the input would then consist of a meaning that needs to be expressed and the output would be a range of candidates forms. For interpretation it's the other way around; the input is a form, the output is a set of candidate meanings. In other words, unidirectional OT for syntax has the speaker as point of view, OT for interpretation the hearer. Blutner claims that at the process of syntax is related to the one of interpretation, and argues both should be integrated in a form of OT that he calls *bidirectional* OT. Interestingly, one of the arguments that would favour a bidirectional approach, is indirectly related to MTC: 'conceptual grinding', the phenomenon of a count noun having a mass noun reading. There are cases in which the mass reading has a separate term, the mass reading from the count noun is then *blocked*. However, in context it's possible to cancel blocking. This, as Blutner (1999:4) argues, can most simply be explained in 'a *bidirectional* OT *that takes into account the production perspective*', but would be impossible to even *point out* in an unidirectional approach.

Bidirectional OT as proposed by Blutner (1999:7) is based on work of Horn (1984) and Levinson (1987) on the Gricean Maxims. The *Q-principle* blocks the outputs that have

[2]Note the similarity of this constraint in comparison to LEXICAL IMPOTENCE Asher&Lascarides (1995).
[3]See Zeevat (2001:250) for the definition and explanation of ANCHOR

alternatives that are more economic, the *I-principle* is responsible for the selection of the most coherent interpretation. The bidirectional approach integrates the Gricean maxims on conversation. In its *strong* version, the Q-principle is absolute, in the sense that it blocks interpretations that can be more economically be derived from other forms. The *weak* version relates the Q-principle to the I-principle *within* the separate principles that need to be satisfied in the super-optimal candidate. This is based on the assumption we have seen in section 3.1.1, that marked forms are associated with marked meanings, (*Division of Pragmatic Labour*, Horn 1984:22). This is only possible if we adopt the weak version of bidirectional OT, where blocking can be partial because of the 'co-dependent' formulation of Q-principle and I-principle (see Blutner 1999:11). Considering the reasons for developing the weaker version, it is no surprise that this is the version is commonly used when applying bidirectional OT. In Blutner (1999) it is applied to presupposition projection in quite a satisfactory manner, by a relatively simply set of constraints that we have seen before: AVOID ACCOMMODATION which here 'counts' the number of discourse markers used in the accommodation and BE STRONG (STRENGTH).

For this study, I will stick to a unidirectional approach in OT. The main reason is that the examples that are the core of this study, are not examples that are conflictive at the level of pairs of forms and meanings. In fact, I think it's safe to say that the phenomenon of MTC is merely an interpretational phenomenon. We can have an endless theoretic debate here, and discuss the pros and cons of OT for semantics. I refer to Zeevat (2000), discussed previously, for a more existential discussion about the bidirectional approach. He addresses the issue of ineffability, next to the supposed lack of semantic constraints. Finally he proposes a form of OT where the set of interpretation candidates is restricted by the result of production. It is interesting to see that this 'architecture' is both opposed as similar to the model of Wilson (2001) where interpretability is a constraint on production, as pointed out in Beaver&Lee (2003). The latter is a quite pessimistic but insightful study on OT semantics and the *asymmetry* between input and output that is unaccounted for in the unidirectional ('naive') approach and the bidirectional approach in the sense of Blutner (1999).

This is just to say that we probably still have a long road to travel before we will have a common understanding of the architecture of OT in the semantic/pragmatic domain. In my application of OT in the account for MTC I will try to obey the research constraint SIMPLICITY. It is obvious that at some points I have already violated it, in an attempt to satisfy COMPLETENESS.

3.1.3 Requirements for an OT approach

This section is a combination of assumptions that we need to make about interpretation and lexicon and a hypothesis about the constraints that could be part of the OT solution for MTC that will be developed in the sections to come. Based on what is described in the previous chapters and what we have seen about possible OT constraints for MTC sofar, I think there are a couple of issues that we need to address before we even start to think about creating OT constraints. I will discuss the particular *interpretation* that is associated with MTC and the different forms of knowledge that are supposedly involved in the process of MTC. This is preceded by a discussion about Prototype Theory (see also section 2.4) which I claim to provide a notion of concepts that is flexible enough for my purposes. Finally, I will outline the principles of the work of Joost Zwarts on OT on the different interpretations of *(a)round* (Zwarts 2003), and use his work as a model for my proposal of resolving MTC within an

optimality approach.

We have seen that the interpretation of the different cases of MTC can be generalized to a *representation* or *image of* interpretation. We have addressed many examples of the phenomenon, and overall I have argued that the interpretation after coercions of this sort result in N_{ima}. I have described this as a meaning type, but did not discuss what happens exactly when a noun is coerced *metonymically*. In the fields of arts studies, *metonymy* is defined as a figure of speech, in which is a word is replaced by another word based on some relation between the two of them. These relations can be divergent, i.e. *pars pro toto*, *totum pro parte*, *eponymic* and *metaleptic*. In my opinion, the use of the word *metonymic* in the description of MTC means that a part of the concept that is referred to, is coerced. That is, one of the properties the concept normally has, is changed to another property in the interpretation of MTC.

I have claimed that this phenomenon leads to a change in 'meaning type' (hence, type coercion). As noted previously, a meaning type in this study is in fact an *ontological* type, hence the need for at least an assumption about a conceptual framework in the resolution of MTC. The interpretation resulting from MTC is more of a *representation of the object* that is referred to, than the object itself. In this sense, it could be regarded as metonymic as well, as long as we regard a representation of an object as a part of the whole concept of that object. This is kind of a difficult issue, that could lead to a whole different discussion on conceptual theories. We will soon see that the conflict that causes MTC, takes place between what will be referred to as STRENGTH constraints and FIT constraints. In order to formulate what it means to be *Be Strong* (STRENGTH) to a concepts description, it will be necessary to first decide what is *inside* the borders of that concept.

We have already seen that Kamp&Partee (1995) and Partee (2001) touch this issue in the discussing a the theory of *prototypes*. The theory of prototypes is in my opinion best explained in Rosch (1999)[4]. In short, Prototype Theory can be described as a theory of categorization based on principles of psychology (rather than for instance linguistics per se). The most important principles that are fundamental to this theory, and in my view make it especially suited in the perspective of OT, are COGNITIVE ECONOMY and PERCEIVED WORLD STRUCTURE. Rosch (1999:190) explains the first as '*to categorize a stimulus means to consider it, for purposes of that categorization, not only equivalent to other stimuli in the same category but also different from stimuli not in that category*'. The latter means that '*the material objects of the world are perceived to possess high correlational structure*' (Rosch 1999:190). The two of these principles allow for maximum information to be achieved with the least cognitive effort, which is basically what most linguistic theories would strive for. The value of this research is merely the fact that it is based on data that are acquired from experiments. The investigations that have been done, the structure of concrete objects were defined as having a set of aspects of form and function. For abstraction of the basic level the following things were analysed in terms of convergence: attributes in common, motor movements in common, object similarity in shape and identifiability of averaged shapes (see Rosch 1999:193). To avoid this section from becoming overly rich of information, I will skip the parts about the actual experiment and move to the most important results that led to the assumption of prototypes in the first place.

The data of the investigation in Rosch (1999) lead to the conclusion that prototypes of

[4]The original publication was in 1978, the complete reference is described in Bibliography. The specific references in this study are to the re-publication in 1999.

categories are related to psychological processes (under which language understanding). But what *is* a prototype of a category? Rosch (1999) explains that most categories don't have clear cut boundaries. Her findings are basically that there is people perceive typicality differences within a category. It follows that '*in terms of the basic principles of category formation, the formation of category prototypes should, like basic levels of abstraction, be determinate and be closely related to the initial formation of categories*' (Rosch 1999:197). A *prototype* for a category is then the member of that category that has the most attributes in common with the other members of that category and the least attributes in common with the contrasting categories. In my opinion, these are useful findings and assumptions for this study in two ways. First of all, in terms of MTC it could mean that a representation of an object can still be seen as a member of the category of that object.

A theory based on categories with prototypes allows for members to be 'scaled'; which means that some members of a category are better examples of that category than others. For *mtc* this could mean that when we interpret *stone lion* or *fake gun*, those are 'degrading' in typicality but still part of the category. In other words, *a lion* interpreted in the default sense would be a better example of the category of lions than *a stone lion* without denying the latter its membership to the category as a whole. The second advantage of a prototype based theory of concepts relates to the way STRENGTH constraints for lexical items work, as described by Zwarts (2003), where the different interpretations of *(a)round* are addressed, which I will discuss next. To close this part of the discussion, I quote Barbara Partee (Partee 2001:6, footnote 5) to support the view that images of objects should not be seen as belonging to a different category than that of the object itself: '*In fact, in the literature of prototype theory, one can observe that many of the reported experiments on judgments of prototypicality are carried out with pictures of objects rather than actual objects, but all of the language of the experiments and of the discussion of the experiments refers to the corresponding objects, not to pictures of objects. And normally we don't even notice; starting with our children and their picture books, we say things like, 'Where's the doggy? There's the doggy!'. Presumably no normal parent would say 'Where's the picture of the doggy? There's the picture of the doggy!*'.

The work of Zwarts (2003) explains a competition between constraints that require a word to be interpreted as closely as possible to its most elementary meaning, and constraints that come from the linguistic elements surrounding that word. Although Zwarts does not address the theory of prototypes explicitly, it's safe to say that this theory would fit with a prototypical theory on concepts. Zwarts (2003) explains that the different possible interpretations of *(a)round* are the result of competition between a STRENGTH and a FIT constraint. Note that the constraint STRENGTH is similar to the description of Zeevat (2001) and what Blutner (1999) calls BE STRONG, but here applied to the meaning of a lexical item, rather than the reading of a complete sentence. The work of Zwarts (2003) functions as a mere *model* for this study, although there is one fundamental difference, which in my opinion makes this study more complicated. This has to do with the example *(a)round* in Zwarts (2003) versus my examples of MTC; *round* is a mathematical concept, of which the prototype can be described with one strict definition. The result is that the constraints that account for the polysemy of this concept, are also very strict. This is all very beautiful, it makes his study very precise and somehow 'hopeful'. On the meta-level, Zwarts (2003) can be seen as the prototype or even the ideal for an optimality approach of the polysemous behaviour of nouns. In reality, there are really few concepts that we use in natural language that behave like *round*. By far most concepts do not have strict definitions, categories don't have clearly marked boundaries

and to make things worse, words can be used to refer to something else all together. Still, the procedure of constructing constraints in this study, will be the similar. Like Zwarts, I will start out by formulating the constraints that are part of STRENGTH. I will work on a more general level, since the set of constraints should ideally apply to all examples of meaning shifts on N that we have seen in this study. Zwarts (2003) focuses on simple constructions (for example: the meaning of around in around the door and around the block), which again is less complicated: the FIT constraints are narrowed down to one type of construction. In his work, the OT tableaux are relatively simple, the outputs can be described as resulting from a single STRENGTH constraint that depending on the words that surround *(a)round*, is violated a number of times. We will see that I have to adopt a slightly different approach for this. To account for the different kinds of interpretations, STRENGTH will consist of different, ranked 'subconstraints' that can be violated depending on the constraints that play a role in the context (in my case the properties of the MTC trigger). This will discussed in the next section.

3.2 In defense of the strongest interpretation

This section will be focused on the formulation of the constraints that together form a STRENGTH constraint in the sense of Zwarts (2003). In order to formulate constraints that are abstract enough to account for all shifts in meaning that result from the examples of MTC, I will adopt a prototypical view on concepts. It will become clear that in addition to that, I need to rely on a structured idea of the (mental) lexicon as well. This is because of the 'non-uniformity' of the meaning-shifts. For instance, in *stone lion*, the *constitution* of *lion* is changed from 'lion flesh' to stone. This can be seen as a change of *attributes*, whereas in the case of *fake gun*, the *functionality* of *gun* is taken away rather than any of the attributes (*a fake gun* can have the exact appearance of a real gun). This means that it will come in quite handy to work with a structured lexicon rather than a vague and undefined notion of lexical meaning. This will all be explained in the respective sections, starting with a description of default interpretation in section 3.2.1,

3.2.1 Default Interpretation

Translating perceptions, concepts and thoughts into words would be impossible if we wouldn't have some principles about word meaning. Those principles are needed to avoid the situation that we can interpret attach any meaning to any word. This is why a lot of semantic theories are based on the assumption that our minds have a structured lexicon. One of those theories about the mental lexicon is Pustejovsky (1995), which I have discussed in section 2.3. However, at the same time, the data of this study show that the boundaries of word meaning are flexible. It seems that it's more important to assign *a* meaning, than to conclude that an utterance is meaningless because one of the elements cannot be interpreted according to the information in the lexicon. In order to 'protect' the most common meaning of a word in language use, we need at least one STRENGTH constraint.

In the analysis of MTC we have seen that the interpretation of N is affected by its context. I have argued that instead of $N_{art/nat}$, MTC forces the noun into N_{ima}. It's reasonable to take the meaning types $N_{art/nat}$ as default for a noun that refers to an entity (see preliminaries for the type lattice). From now on, the two types will be referred to as N_{def}; the default meaning type of a given noun. The following STRENGTH constraint can be formulated:

DEFAULT INTERPRETATION I
Interpret a given N as default N_{def}.

Note that this is similar to the principle `Default Content` of Franks (1995) discussed in section 2.2.2. I have left away his 'hedge', since in this OT approach other constraints will provide the 'good reason' to violate this constraint. I have argued back then that this seemed to be a sensible principle, however, it will only be useful when we know the constraints that interact with it.

The question is, if this constraint would suffice for the interpretation of N. Especially when DEFAULT INTERPRETATION would be violated to satisfy another constraint, it would mean that N has an interpretation different than its default meaning. To differentiate between the interpretation that follows from satisfaction of this constraint and the one that would follow from its violation, we need a basic description of what it means to 'interpret N according to its default meaning type'. I have earlier argued in favour of a prototypical view on concepts. For default interpretation we could assume that any concept C is interpreted as its *prototype*:

DEFAULT INTERPRETATION II
Interpret a given C according to its prototype P (C_p)

In this way, the interpretation of the concept that N_{def} refers to, is its prototype. Suppose there is a conversation about an elephant. If it's possible to satisfy this Faithfulness constraint, that means *elephant* is interpreted as being a prototypical, N_{nat} elephant. The use of prototypicality as a part of interpretational constraints, is especially appealing for the case of mtc, which we will see in the next section. Unfortunately, this does not mean it's unproblematic when formulating a constraint that will be explanatory for the interpretation that we are looking for. Let's first finish the construction of the complete constraint that prohibits nouns to be interpreted in any way other than their prototype. I think it's better to see the meaning type as a semantic part of the constraint, whereas the prototype is a conceptual part. This is merely a personal assumption. We have seen that Fauconnier regards meaning processing (interpretation) and building as cognitive tasks rather than a part of language by itself. I will assume here that prototypes are part of both the mental lexicon and world knowledge. I have not addressed the mental lexicon in terms of *linguistic* versus *cognitive* information that it (at least in my opinion) both contains. I take the default meaning type as being a part of linguistic information, whereas the conceptual information that is part of the lexicon, is a part of cognition. To be more clear: I assume that the meaning type that belongs to N is acquired as a part of grammar, whereas the information that is acquired in the formation of a concept is acquired over time, and can be changed at all points in life. The meaning types are less flexible in terms of violation than the prototypes of the concept of N, which is what I want and what will be difficult to translate in a single constraint for STRENGTH for the interpretation of N. We will return to this is both sections that follow. I will formulate the constraint as follows:

DEFAULT INTERPRETATION III
Assume that a given N is of meaning type N_{def}, and interpret N_{def} as C_p of N.

We will soon see more of the necessity of this 'division' in this constraint. As I have noted before, this constraint is similar to a principle in Franks (1995). The classical view he adopts divides the default content of a concept into diagnostic and central features. I have claimed

this gives rise to many problems, some of which related to the solution of MTC being too restricted this way, others that are intrinsic to the classical view itself (see section 2.2) Another way of describing the lexicon is GL, in Pustejovsky (1995), who divides the information in the lexicon in terms of qualia structures. We need to have a more specific description of what a prototypical meaning of N looks like, since we have seen in the analysis of the examples that MTC functors do not work on *all* the properties or features of N. This makes sense of course, MTC could not exist the way it does if it would deny everything that is associated with the coerced N. If this were the case, not only would the word *real* be redundant as explained in Partee (2001), in the cases of MTC in other constructions than modified NP's, it would be strange that we can use a N to refer to something that is entirely *not* N.

My proposal is to include a default QUALIA STRUCTURE (QS_{def}) for C_p in this STRENGTH constraint. This way, the prototype that is the default interpretation of N_{def} constitutes *structured* information. It will be clear when we discuss the conflicts and negotiations between STRENGTH and FIT constraints that we need to see *where* this conflict arises. It is clear that meanings can collide- but if we are still able to interpret the expression anyway, we must look closely at the nature of the collision. For QS_{def} I will assume the different qualia as described in section 2.3: CONSTITUTIVE (CONST), FORMAL, TELIC and AGENTIVE. This is implicit in the formulation of the constraints to come though, since it would complicate things too much here and make the constraints data-specific rather than general. Taking this into account, a final version of DEFAULT INTERPRETATION for N is:

> DEFAULT INTERPRETATION (DEFINT)
> Assume that a given N is of meaning type N_{def}, such that $N_{def} = C_p$ of C_N, where C_p has QS_{def}.

I will construe similar constraints in the next section, for the interpretation that are not default. To avoid vagueness on the constraints, let me restate DEFINT in normal words: when one encounters a noun in an expression, assume that this noun is of default meaning type, and should be interpreted as the prototype of the concept that this noun normally represents. This prototype is associated with the default qualia structure. We will see in the next section, that this constraint is the strongest part in the defense of default meaning. It is the best, but not the only way to meet the demands of STRENGTH. As we will see, STRENGTH and the preference for the strongest interpretation of a concept is *relative* in actual language use.

3.2.2 Secundary versus figurative interpretation

Evidently, natural language meaning allows us to interpret words in other senses than the default one. We can assume that the default interpretation is the most common interpretation, however other interpretations are not necessarily 'rare' or uncommon. In fact, it seems that the IMAGE INTERPRETATION (elsewhere also referred to as 'representation interpretation') that is the result of MTC is done without much effort on the hearer's side. This leads me to the assumption that the interpretation of N in relation to its C_p is constrained but flexible. In other words; it's natural to understand an object in terms of its image. I see images as part of the concept itself, although we'd have to assume that one concept can have more than one meaning type. In the strict sense of GL this would be problematic, and for OT we'd need to have clear assumptions on which kind of meaning types are mutually exclusive and which aren't. By way of sketching a temporary constraint, I assume the following is needed for the interpretation resulting from MTC:

IMAGE INTERPRETATION I
Interpret N_{ima} such that $\|N_{ima}\| \cap \|N_{def}\| \neq \emptyset$ and $N_{ima} \in \{C_N\}$.

In other words, the interpretation of N_{ima} shares properties with N_{def} and although it is not interpreted as the prototype, it must not be interpreted outside the concept of N. Like I stated in the previous section, MTC works on different *parts* of the default interpretation. It either denies them or changes them such that the meaning type of N shifts to N_{ima}. We have seen that it's quite easy to use the qualia structures proposed by Pustejovsky (1995) to obtain a more precise view on the type shift. In the cases of MTC caused by a material adjective leads to a change in the CONST quale of N. A discourse marker `Loc` changes the default size contained in the FORMAL quale of C_p if the location cannot contain the default C_p.

Sofar, I have claimed that the interpretation of N_{ima} should fall within the borders of the concept that N represents. How is this to be seen if we take this interpretation as a violation of the STRENGTH constraint that I formulated in the previous section? It follows from my line of reasoning that IMAGE INTERPRETATION of N does not entail a strong violation of the Faithfulness constraint for N. My proposal is to use IMAGE INTERPRETATION as a secundary possibility for N to be interpreted as part of C_N. Suppose there is a reason why a given N cannot be interpreted as being N_{def}. In OT, this reason will have the appearance of a higher ranked constraint that will be violated if DEFINT is satisfied. Note that I use the term *secundary* interpretation. I think that there is a difference between a non-literal/figurative interpretation and the interpretation that we encounter after MTC. We could describe this in the constraint FIGURATIVE INTERPRETATION for the interpretation of meaning type N_{fig}, a type that I think we can use for most common nouns when they are used metaphorically, symbolically or metonymically (in other ways than MTC). The requirements for FIGURATIVE INTERPRETATION could be:

FIGURATIVE INTERPRETATION I
Interpret N_{fig} such that $\|N_{fig}\| \cap \|N_{def}\| \neq \emptyset$ and $N_{fig} \notin \{C_N\}$.

Formulated like this, this is in line with the intuition that any figurative interpretation is a partial interpretation of the semantic properties or qualia that are associated with N_{def}, but it's not in the category C_N. A lot of concepts can be learned by exposure to their images, but I doubt you can learn about a concept by figurative uses in utterances. I am aware that put like this, I do no justice to the interpretation of metaphors. This discussion is just to make clear that image interpretation is not the same as figurative interpretation, nor is it to be regarded as an instance of it.

Now that I have established the temporary constraint for IMAGE INTERPRETATION, and the difference with FIGURATIVE INTERPRETATION, we are left with some issues regarding MTC and qualia changes that are involved in the shift to N_{ima}. Like I argued before, MTC works on a quale in a way that N_{def} needs to be seen as N_{ima}. However, the question is if there are qualia that you can change without this result. The answer is yes, which is illustrated if we oppose *lion* with both *stone lion* and *baby lion*. *Lion* in *baby lion* undergoes a change in *Formal* quale, since the size of a baby lion is obviously not the same as the size of C_p of *lion*. I will regard cases like *baby lion* (and *dead duck*, see the discussion about example (9))as satisfaction of CLOSE INTERPRETATION:

CLOSE INTERPRETATION I
Interpret N_{def} such that $N_{def} \neq C_p$ and $N_{def} \in \{C_N\}$.

Note that the meaning type is unchanged in CLOSE INTERPRETATION, it merely says that N_{def} is not to be interpreted as the prototype of the concept, but it shares enough properties to be an element of the concept of the default type. Of course, this constraint could in turn be divided in different degrees of closeness to C_p of C_N. Suppose we analyse *three-legged dog* versus *baby dog*, I'd suppose that the latter would be interpreted closer to C_p. This is debatable of course, next to the fact that it's beyond the purpose of this thesis. It's an interesting thing to keep in mind though, when working with a theory of prototypes for lexical meaning.

We need to look closer at the examples and the way they shift, i.e. which qualia are changed such that the result is N_{ima}. If we look at the examples that involve a shift from N_{nat} to N_{ima}, we see that in all cases the default interpretation is impossible because of the simple reason that N cannot be interpreted as a living natural kind type, which is its default type. This can be due to changes in different qualia, even more than one when we consider example (8) where *the elephant* is not only reduced in size, but also in material (although not explicitly). We could answer the question by turning it around and assume that N_{ima} and IMAGE INTERPRETATION are preferred over other options, which would be FIGURATIVE INTERPRETATION or no interpretation, assuming that CLOSE INTERPRETATION is ruled out.

To make the OT approach for MTC as exact as possible, it could be a good idea to include the qualia of the lexical items that are involved. This is the only way to see the interaction of different meanings that lead to MTC. However, this will be quite complicated to establish if we would for example formulate constraints for each different qualia structure. This would entail different constraints for the different examples, which is not my aim. To resolve this, and keep things general to a certain extent, I will formulate the different interpretational constraints for N which include the notion of qualia. Let's first look at the interpretation closest to DEFINT:

> CLOSE INTERPRETATION (CLOSEINT)
> If $QS_N \neq QS_{def}$ such that N is of type N_{def}, interpret N such that $N \neq C_p$ and $N \in \{C_N\}$.

This constraint will take care that for example the combination baby elephant is not interpreted as being of another type, in fact I assume that it will account for all examples of nouns in combination where one or more qualia in QS_{def} are not met (such that DEFINT is violated) but the meaning type is still the same. The violation of DEFINT will be counted as a violation of the Faithfulness constraint on N. So what to do if the changes in QS_N are such that N cannot be of type N_{def}? I propose this is the time the following constraint may be satisfied:

> IMAGE INTERPRETATION (IMAINT)
> If $QS_N \neq QS_{def}$ such that $N \neq N_{def}$, assume that $N = N_{ima}$ and interpret N such that $\|N_{ima}\| \cap \|N_{def}\| \neq \emptyset$ and $N_{ima} \in \{C_N\}$.

Note that there is no specification here of which qualia must be changed in order for the type change to occur. Although it would seem like the more accurate thing to do, I think that to account for mtc as a process, we need to keep the constraint quite general. Another thing that will be no part of the ot solution when we use this constraint, will be the sort of interpretation that is assigned to the N_{ima}. We have seen that this type can have various interpretations, from the plain Fake to Toy. Once it has been concluded that the speaker

refers to N in terms of its image, it seems that the interpretation of N_{ima} is free to a certain extent. That is, a fake gun might be a toy gun or a replica. An elephant in a room may be a poster or an ornament. This is where Franks (1995) suggested Noun Instantiation, as I explained in section 2.2. I think the problem of making a constraint including this NI or a process similar to NI, is that it would yield a certain interpretation, while it is possible that the context simply doesn't provide the clue for this instantiation. This might mean that it is not necessary to narrow the possible interpretations for images down to one.

The last constraint that be satisfied in FAITHFULNESS of N, is where N is interpreted figuratively:

> FIGURATIVE INTERPRETATION (FIGINT)
> If $QS_N \neq QS_{def}$ such that N $\neq N_{def}$ and N $\neq N_{ima}$, assume that N $= N_{fig}$ and interpret N such that $\|N_{fig}\| \cap \|N_{def}\| \neq \emptyset$ and $N_{fig} \notin \{C_N\}$.

To limit the possible interpretations of N to use in this OT approach, I will assume that this constraint functions as the last resort for the interpretation of N. This constraint means that it's not possible to interpret N in terms of N_{def} (both DEFINT and CLOSEINT). Next to that, this N cannot be interpreted as an element of the set C_N, which rules out IMAINT. If circumstances are such that N cannot meet the requirements of FIGINT, FAITHFULNESS is violated. It could even be that it's violated in a fatal manner, suppose there is really no way to interpret a noun in a given context, it would result in *no* optimal candidate for the complete input (N and its context).

We now have a basic set of lexical constraints that I assume will function properly as a set of FAITHFULNESS constraints in the OT approach of MTC. Note that if this will work, we cover one of the most important principles of noun interpretation proposed by Kamp&Partee (1995) and Partee (2001): NON-VACUITY PRINCIPLE. Next to that, this set of constraints that together form a STRENGTH constraint, answers in my opinion to the principles of Grice. That is, at least we know that if we follow these constraints, we have done anything in our power as a hearer to have looked for a way to interpret a given N in its context. Before I will move to the constraints that lead to the different ways of STRENGTH violation, I will discuss this set of constraints in terms of relative *ranking*, as well as ranking compared to the set of FIT constraints.

3.2.3 Degrees of Faithfulness: STRENGTH versus FIT

The way the constraints are formulated sofar, suggests that FAITHFULNESS could be violated in different degrees. Zwarts (2003) defines a set of constraints accounting for the polysemy of (a)round. In his view, the two competing constraints are FIT and STRENGTH. I think that both terms cover the sort of constraints that we use in OT semantics better than the usual MARKEDNESS and FAITHFULNESS constraints. STRENGTH applies better to an approach that deals with different senses of a word in context when assuming a prototypical view on concepts. FIT is in fact a 'new' sort of constraint in the theory of Optimality. FIT is a constraint that is ranked over STRENGTH (Zwarts 2003:9), and can be seen as a set of constraints that form the *context* of the lexical item that is considered by STRENGTH. We will see in sections 4.3 and 4.4 that the constraints of FIT can come from different levels of the linguistic structure of an utterance. The basic idea of ranking FIT over STRENGTH, is that STRENGTH can be violated n order to satisfy FIT. In case of noun interpretation, this means we can step away from the

DefInt because elements in the context would be violated if DefInt would be satisfied.

From the way the lexical constraints are formulated, we can already see that there is a ranking within the set of Strength. One could say that a given lexical item, in this case N, prefers to be interpreted as strong as possible. I discussed the constraint Strength as formulated by Zeevat (2000), and questioned the notion of 'informationally richer readings'. In the perspective of Zwarts (2003) and my own examples, I'd say that for the interpretation of nouns, a stronger reading should be seen as the one that is closest to the prototype of the concept that N refers to. I doubt that this should be seen in terms of informational richness of some kind. Instead, I will reformulate Strength as Strong Interpretation N:

> Strong Interpretation N (StrongIntN)
> Interpret N such that there is no stronger interpretation of N available.

This constraint will then be satisfied by the constraints as proposed in the following ranking, based on the interaction with Fit:

> Constraints StrongIntN
> DefInt > CloseInt > ImaInt > FigInt

In the next section, I will formulate the constraints that are part of the set that interacts with StrongIntN. A last remark concerns the degrees of Strength and the proposed ranking. It follows from the 'gradeability' of StrongIntN that the only way to really violate StrongIntN is when all of the constraints are violated. I will assume that this would be a fatal violation because according to these constraints, there is no interpretation possible that StrongIntN would allow. This does not mean that these constraints are supposed to be a complete account for the interpretation of nouns, the constraints that are formulated here are basically formulated for the purposes of this study.

3.3 MTC: a result of meaning negotiation

This section concerns the Fit constraints that play a role in mtc interpretation. As the title suggests, I will formulate these constraints as part of a negotiation. In other words, while construing the constraints that are part of Fit, I will show the interaction of the particular constraint with Strength by applying both of them to mtc. Before that, let me define the general Fit Interpretation which interacts with StrongIntN:

> Fit Interpretation (FitInt)
> Interpret a given utterance such that there are no conflicts between its elements.

Where StrongIntN is construed to 'fight' for the strongest interpretation of N, FitInt makes sure N is not conflictive with the elements of the utterance in which N is contained.

I will discuss the Fit constraints that play a role in mtc according to the type of construction they are part of. As can be inferred from the discussion in chapter 1, I will start with a section on *Modifier-Head constructions*, which is followed by *Verb constructions* in section 3.3.2. Finally, 3.3.3 deals with the discourse examples, which I here call the *Narrative constructions*. To avoid confusion, when I refer to example (n), I mean the example of (n) containing mtc and not the complete set of examples in (n). If not, I will explicitly say so in the text.

3.3.1 MTC in Modifier-Head constructions

We have encountered two types of modified constructions, which each need a different constraint. The examples of Mod-Head constructions in the form of AdjN are (1)-(5) of this study, the example of Mod-Head in the form of a PP modifying N is (6). Remember that mtc takes place on the head in (1)-(5), and on the modifying N in (6). To start with the adjectival examples, I will formulate a constraint based on the analysis of *privative/subsective* adjectives in Kamp&Partee(1995) and Partee (2001) and the suggestion of SELECTIVE BINDING for adjectives by Pustejovsky (1995).

I will stick to the argument of Partee (2001:1) that it is better to analyse 'privative' adjectives as *subsective* adjectives, as described in Kamp&Partee (1995:140). However, privative adjectives have the property of coercing the head they modify. It remains unexplained how this should work, which is where Pustejovsky (1995) comes in. As I argued in section 2.3.3, the notion of SELECTIVE BINDING could in fact be regarded as a qualia-based description of subsective adjectives. I made an attempt to reformulate this into Metonymic Binding, which wouldn't work within the boundaries of GL as a whole. However, both approaches have elements that we need. We an account of mtc in which the complete meaning of the head N is not denied. Hence, we have to consider the adjectives that are considered privative in both Coulson&Fauconnier (1999) and Franks (1995), as subsective. Whereas Partee (2001) remains vague on the properties that are changed in N when combined with such an adjective, Pustejovsky's qualia-based lexicon could work here again. Consider the premature constraint INTERPRET ADJECTIVES I:

> INTERPRET ADJECTIVES I
> Interpret [Adj[N]]$_{NP}$ where it holds that $\|AdjN\| \subseteq \|N\|$ according to QS_{Adj}.

The problem of the constraint formulated this way, is threefold. First of all, it includes a result of the interpretation of AdjN which we don't know, since the status of Adj is unknown. The second is more or less related to this, not all cases of mtc involve adjectives that are privative in the sense of *fake*. Other cases are caused by what Franks (1995) refers to as *functional* privatives, like *stone*. In section 2.4 we have seen that there are in fact many problems in classifying adjectives. The conclusion of Partee (2001) based on Polish data is that the only true distinction is between subsective and non-subsective adjectives (see section 2.4.2). Cases of MTC with material adjectives are not part of the proposal of Partee (2001), although she does mention them. The problem is that *stone* and *chocolate* are intersectives, so a constraint of adjectival interpretation based on the assumption that the adjective is subsective by default is not useful in the account for mtc. This is why I think it's better to formulate a general constraint to avoid the issue of classification:

> INTERPRET ADJECTIVES II
> Interpret [Adj[N]]$_{NP}$ according to QS_{Adj}.

This constraint tell us that whenever we encounter a construction of the form [Adj[N]]$_{NP}$, we should interpret the qualia of the adjective. This should be the right, if we hadn't still a third issue left. Obviously, QS_{Adj} only partially accounts for the interpretation of the construction. In fact, an adjective can be seen as the *meaning* of a quale, rather than a qualia structure. Seen in this way, a constraint could be formed based on SELECTIVE BINDING after all:

INTERPRET ADJECTIVES (INTADJ)
If in $[\text{Adj}[\text{N}]]_{NP}$, QS_N has quale q, interpret the construction such that $\|AdjN\| = \text{N} \cap \text{Adj}(\text{q}_N)$.

The only thing left now is to see if this works. Take for example *blond girl*. The noun *girl* has a FORMAL quale that can be specified by the adjective. Note that whereas *blond* is an intersective adjective for Kamp&Partee (1995), this in fact is of little matter for Pustejovsky: *blond* applies to a FORMAL quale of girl, its intersective properties are another issue. The same holds for the interpretation of *stone lion*, in fact. The only thing that happens, and which couldn't be established within GL, is that there is a violation of STRONGINTN.

The way this works can be seen if we consider the following OT tableau:

	input = $\mathbf{U}_{(1-5)}$	FITINT	STRONGINTN			
		INTADJ	DEFINT	CLOSEINT	IMAINT	FIGINT
	Int $(1\text{-}5)_i$	*!				
	Int $(1\text{-}5)_{ii}$	*!	*			
☞	Int $(1\text{-}5)_{iii}$		*	*		

For clarity's sake, the input is here marked as consisting of examples (1)-(6), the outputs are not written out fully because that would not exactly allow for the tableau to both fit on this page *and* be readable. I think we should see the selection procedure of the optimal interpretation of the input as sort of a trading game. There are two parties: FIT and STRENGTH, that have to evaluate and judge certain outputs according to their principles. Suppose the input is (1) *stone lion*. STRENGTH notices: 'Hey, there is a noun, here we need an interpretation that satisfies DEFINT', because STRENGTH wants to secure that the strongest possible interpretation will be selected as the optimal interpretation. FIT has a broader perspective on things and compares the properties of the noun with the properties of its context, which in this case in the adjective *stone*. It's the task of FIT to convince STRENGTH that his candidate is not optimal due to conflicts in its context: this is the point at which it will be noticed that there is a collision between the respective QS's that are implicit for stone and lion in DEFINT. This information is contained in the lexicon, which, I must assume, is accessible during EVAL. The lexicon would at this point reveal that stone has a CONST quale that is incompatible with the CONST quale that is specified in the QS of *lion*. DEFINT results in a (fatal) violation of INTADJ, which is higher ranked. But not all is lost for STRENGTH. His next proposal is CLOSEINT, but again FIT tells him INTADJ is violated, since CLOSEINT for *lion* would still be incompatible with *stone* due to their respective CONST qualia. This is how we come to IMAINT, it's next to the next best thing, but still an option where STRENGTH is not completely violated.

In fact, this short illustration will hold for all the other examples of Adj-N constructions with mtc that I described in this study. Because the issue of different kinds of meaning shifts is resolved in by adopting Pustejovsky's qualia-based lexicon, I will just give short explanations of the other examples of this construction. In case of (2) *fake gun*, ADJINT will collide with DEFINT because of the incompatibility with the TELIC quale, which specifies a certain function for *gun*. This will be resolved only if we interpret N in a way where we don't need that function, and the best candidate for that is IMAINT. The case of (3) *chocolate teapot* is similar to (1), in the sense that the collision is caused by incompatible CONST qualia of

chocolate and *teapot* respectively. For (4) *wannabe punker* and (5) *imaginary friend* the OT solution is in fact enough, although it should be noted that the incompatibility is difficult to point out in terms of qualia. This will be an issue to be studied further in theories about the lexicon, I assume that the problem is in fact not a failure of the OT proposal.

Things are somewhat easier for the other type of construction. We have already seen in section 2.4.3 that the HEAD PRIMACY PRINCIPLE as proposed by Kamp&Partee (1995) could account for MTC on the N that is part of the modifying PP for N in (6). However, I need to be consistent in the use of qualia-based meanings. This means that I will use HPP as an underlying assumption while taking into account the more specific information we need based on QUALIA STRUCTURES:

> INTERPRET MODIFIERS (INTMOD)
> In a Mod-Head construction, where N_1 modifies N_2 by some relation R that holds between them, interpret N_1, according to R, such that N_1 is a quale q, and it holds that $q \in QS_{N_2}$.

Informally, this means that if you encounter a modifying N, this N is interpreted relative to the QS of the head N in the construction. Translated like this, it looks more like HPP again. It depends on the QS of the realized modifying N if STRONGINTN will be violated. Note that the function of HPP for the interpretation of AdjN constructions is taken over by INTADJ here, as I assume all interpretations of AdjN can be covered by this constraint. Turning back to INTMOD, let's look at an OT tableau for (6):

input = $\mathbf{U}_{(6)}$	FITINT	STRONGINTN			
	INTMOD	DEFINT	CLOSEINT	IMAINT	FIGINT
Int $(6)_i$	*!				
Int $(6)_{ii}$	*!	*			
☞ Int $(6)_{iii}$		*	*		

Here, $(6)_i$ is the output candidate where the input has the interpretation of a dress with flowers of meaning type N_{nat} on it: it satisfies STRENGTH constraint DEFINT, which, in turn, is a violation of the higher ranked FIT constraint INTMOD. The same holds for CLOSEINT: the interpretation may not be the prototype of the category denoted by N, it is still associated with the same meaning type, which dimension (FORMAL quale) is incompatible with the dimension of *dress*. For the modifying N to become compatible, and satisfy INTMOD, the N must be able to be interpreted as part of the QS of the head N, which here means it should be 2-dimensional (hence an image). It should be noted that this is closely related to the kind of relation that holds between head and modifier. There is a difference between *a dress of flowers* and *a dress with flowers*. The first would be interpreted without coercion, because preposition *of* here means *made of*. Somehow, the coercion seems to be absent for the same reason it is present in *stone lion*. I admit that the constraint is quite ad hoc in this respect, since I just vaguely mentioned 'according to R'. I will leave it like this, but we need more study and data on this matter, in order to improve INTMOD.

In a nutshell, after this negotiation, the first candidate satisfying INTMOD, without completely violating STRONGINT, is the interpretation where *flowers* are 2-dimensional images.

3.3.2 MTC in Verb constructions

As is clear from many linguistic studies, the verb is one of the most powerful elements in an utterance. The work of Pustejovsky (1995) deals with many cases in which a realized argument undergoes type coercion because the verb specifies for a type different than the type of the realized argument, like I illustrated in section 2.3 with *John began the book*, in which *the book* coerces to an event type. The examples of MTC show that coercion can also take place at the level of qualia, i.e. a change in the QS of the realized argument possibly (but not necessarily) leads to a type in meaning shift. This happens for example when a meaning of a verb requires one of the arguments to have maximum size, like in example (7). Suppose we would form what we see in for example the GL matrices in 2.3, into an OT constraint. That could result in:

> OBEY VERBS I
> If N is realized as argument A_n of a given V with AS_V, where AS_V requires a quale q in QS_{A_n}, interpret N such that $q \in QS_N$.

It is assumed that V not only has a qualia structure QS_V of its own, but it also has an argument structure AS_V. The latter contains information about the arguments *and* the QS's of the arguments. Formulated this way, the constraint is specific for N, which a given verb does not necessarily select as (one of) its arguments. In fact, since it seems the requirements of a verb *must* be met, the formulation of the verbal constraint can be simply as follows:

> OBEY VERBS (OBEYV)
> All requirements of $QS_{arg} \in AS_V$ of V must be met.

In words, this would be described as: if there is a verb V, all the qualia structures that are part of the argument structure of V must be met. Let us look at the representation of example (7) in the following OT tableau:

input = $\mathbf{U}_{(7)}$	FITINT	STRONGINTN			
	OBEYV	DEFINT	CLOSEINT	IMAINT	FIGINT
Int $(7)_i$	*!				
Int $(7)_{ii}$	*!	*			
☞ Int $(7)_{iii}$		*	*		

This tableau represents the interpretation process of (7). Outputs $(7)_i$ and $(7)_{ii}$ both yield interpretations where at least the TELIC and the FORMAL quale for 'magnitude' correspond to those of the prototypical truck. In case of $(7)_i$, the interpretation of truck *is* the prototypical interpretation. For $(7)_{ii}$, the interpretation is considered not the prototype truck, but still an object that can be defined as a truck in the functional sense. These two possible interpretations both result in a fatal violation of FITINT constraint OBEYV. DEFINT and CLOSEINT for *truck* would both yield a truck of a (proto)typical size, say about 220 cm. For the verb *to tumble over*, AS_V requires an ARG2. For this argument there is a set of qualia that is already determined by to tumble over. One of those qualia is obvious in this example: this verbs requires ARG2 to be no higher than 25% of the height of ARG1. Leaving aside the specific heights, no human ARG1 will be four times as high as a regular truck. In other words, if

truck is interpreted respecting its FORMAL quale for height, this means a fatal violation of OBEYV.

For IMAINT, it holds that N can be interpreted as a representation of the concept that is denoted by N. For (7), that means that a formal quale like the specification of height can be changed. In this example we are forced to interpret *truck* as being no higher than 25% of the length of an average human being (or of the length of the specific speaker *I* refers to, if we possess that knowledge).

3.3.3 MTC in Narrative constructions

This last section is focused on the manifestations of MTC in discourse. Since 'discourse' is a term that is applied to many things, I will refer to it as *narrative constructions* here. This way it is clear that although we may be outside the boundaries of syntactic structures, also set of utterances (say, a story) can be explained in terms of structure.

In section 2.1 I have discussed the idea of *space builders*, as proposed in Fauconnier (1995). I have stated that the approach of MST is too broad to be precise. Still, I'd say that part of the formulation of constraints for narrative constructions can be based on MST ideas. Suppose we regard locations as markers for a certain expression. This marker can than be regarded as a space builder (literally), it sets up a space in which the rest of the expression should be interpreted. In a sense, LOC has scope: it determines (part of) the *setting* of a story.[5] I assume that LOC holds until the narration reveals another marker for location. If this other marker has to be an explicit location is beyond this study, let's for here assume that a given LOC_1 holds until a LOC_n presents itself in the story. This is related to SPATIAL CONSTRAINT ON NARRATIVES described by Asher&Lascarides (see section 2.4.3); we should take into account that once a story is set somewhere, as a hearer we have expectations. To create consistency, we will assume there is no change of location until it is stated explicitly. A constraint that preserves the scope of LOC, should then request the hearer to interpret the elements of the story that fall under the scope of LOC, as being part of LOC:

> INTERPRET LOCATION I
> If an expression E is marked for a given location LOC_1, interpret all other elements within LOC_1 as long as LOC_1 holds for E.

Evidently, this is too vague to function as OT-constraint. First of, what would it mean for an element to be interpreted *within* a given LOC? At least we will need to specify the demands of LOC. In language, locations are commonly denoted by proper names (like *Holland*) or nouns. The latter are usually understood as locations due to prepositions or verbs, for example *tree* in *the cat was stuck in the tree*, and *street* in *I crossed the street*. In this study the focus is just on locations derived from nouns, I assume they behave like nouns and have a qualia structure QS_{Loc}. We can now assume that in case of a LOC marker in E, the following constraint plays a role in the interpretation of E:

> INTERPRET LOCATION II
> If an expression E is marked for a given location LOC_1, interpret all other elements in LOC_1 as long as LOC_1 holds for E, such that QS_{Loc} is respected.

[5] Or a Fauconniesk 'space', and in some cases even a Fillmorian 'frame'.

In this version, at least the nature of the role of LOC is revealed. Still, I didn't do anything about how 'the other elements' fit into this story. I think the most straightforward way to do this, is to see this part of the interpretation as a given relation R. R could, for example, be the BRIDGING relationship we saw in 2.4.3. In the constraint I formulate for this study, I will leave R unspecified. In case R *is* a BRIDGING relationship, like in (8), it could be suggested we use the laws of BRIDGING (as proposed by Asher&Lascarides 1998, see section 2.4.3 of this study) as OT constraints. For now, the final version INTLOC looks as follows:

> INTERPRET LOCATION (INTLOC)
> If an expression E is marked for a given location LOC_1, such that a relation R holds between LOC_1 and a given element in E, interpret R such that QS_{Loc} is respected.

INTLOC should now account for the examples where elements undergo coercion due to a LOC in discourse. Since the effect of certain topics on the interpretation of a discourse is similar, I will first discuss INTTOP, and then start the application of INTLOC and INTTOP on the examples of MTC in OT tableaux. We have seen in some of the examples, that topics possibly influence the meaning of elements in their context. Like Loc, a Top marker can be either a proper name or a common noun. In the examples I have shown that certain proper nouns are associated with a set of properties, which determine the interpretation of the entire discourse that topic holds for. I'd say that the properties that are associated with proper nouns are no part of lexical knowledge like those of common nouns. This could mean that INTLOC is not complete yet. My proposal is that although this knowledge may not be *lexical*, we can still think of it in terms of the Pustejovskian QUALIA STRUCTURES. I will discuss the issue of different kinds of knowledge, their structure and accessibility some more in the evaluation of this chapter. If we assume that it doesn't matter what kind of knowledge provides the information contained in the QS of a given lexical item, and that the effect of TOP is similar to LOC, INTTOP can be described as follows:

> INTERPRET TOPIC (INTTOP)
> If an expression E is marked for a given topic TOP_1, such that a relation R holds between Top_1 and a given element in E, interpret R such that QS_{Top} is respected.

Note that for INTTOP, I maintained the idea of a *relationship* between TOP and elements in its discourse. To illustrate this: these are relationships like *possession*, *part of* and can, like LOC, be part of a semantic BRIDGING relationship.[6]

Let's see how INTLOC and INTTOP interact with STRONGINTN constraints in the cases of MTC. In example (8) we see that the QS of the location *room* collides with the QS of the element that is contained in that LOC: *the elephant*. Intuitively, *room* has a FORMAL quale that specifies a size such that *elephant* cannot be contained in it. This is due to the specification of the *Formal* quale of the latter, which, as I assume specifies for a bigger size. The tableau below represents the consequence of this:

[6] For example, in *The car was yellow. The left door was lighter than the rest*, *part of* is established by BRIDGING. In this respect, we might even consider to regard locations as a type of topic.

	input = U$_{(8)}$	FitInt	StrongIntN			
		IntLoc	DefInt	CloseInt	ImaInt	FigInt
	Int (8)$_i$	*				
	Int (8)$_{ii}$	*	*			
☞	Int (8)$_{iii}$		*	*		

Since in (8) we are dealing with a noun that is of type N_{nat} by default, we can now see more of the role of CloseInt. Candidate interpretation (i) of example (8) yields a violation of IntLoc because of the specifications of sizes in the respective Formal qualia. This does not necessarily hold for the interpretation in which CloseInt is satisfied. That is, one could argue that non-matured elephant would fit in a room without much squeezing. And since StrongIntN asks us to interpret a given N in the strongest meaning that does not violate the Fit constraint at stake, this would lead to the expectation that Int (8)$_{ii}$ is the optimal interpretation for (8). This is not what this tableau tells us. I think in this case, the possible CloseInt that is available, is still not the best interpretation because there is an interpretation available that is better for Fit. As I mentioned before, the information of room and elephant in this example is established by means of a Bridging relationship. I have summarized the core points of this in sections 2.3.3 and 2.4.3. It comes to mind here that the prerequisites for Bridging as proposed by Clark (1977) could be an explanation for the choice of ImaInt over CloseInt. After all, a small wooden ornament in the shape of an elephant is quite a common object to be found in a room. More importantly, it is *more* likely to have a wooden replica of an elephant (N_{ima}) in ones room than to keep a baby elephant (N_{nat}) in it. I suppose that if the identity of R in IntLoc/IntTop would be studied more closely, we might end up formulating different constraints for different relationships.

Other than in the OT tableaux presented sofar, I have marked the violations of IntLoc as non-fatal violations. The thought behind this is twofold. First of all, the examples of MTC where Loc or Top (as discourse markers) are MTC functors are less convincing than the other examples. Intuitively, it is *strange* to interpret *elephant* according to DefInt in (8), whereas DefInt for (7) is merely impossible. This can be seen in the perspective of the more general communication constraints we have seen in section 3.1: I'd imagine that (7)$_i$ and (7)$_{ii}$ are a violations of the pragmatic constraint Avoid Contradiction as described by Hendriks&De Hoop (2001), and (8)$_i$ and (8)$_{ii}$ as violations of the Gricean 'constraint' Avoid Obscurity of Expression (see Grice 1975:159). In a nutshell, it seems that the different Fit constraints have different *forces*, and should be ranked. This will become clear in the discussion of example (9). Before going there, let's look at examples (10) and (11) when represented in OT:

	input = U$_{(10-11)}$	FitInt	StrongIntN			
		IntTop	DefInt	CloseInt	ImaInt	FigInt
	Int (10-11)$_i$	*				
	Int (10-11)$_{ii}$	*	*			
☞	Int (10-11)$_{iii}$		*	*		

Note that this tableau represents the interpretation process of (10a) and (11a/b). None of these constructions contain a Bridging relation, but there are different types of R established.

In (10) TOP is related to N by the POSSESSION relation I described earlier. R in (10) tells us that an object belongs to another object. TOP is then the object that possesses another object. In (10a) TOP requires for a secundary interpretation of the selected 'possessee'. This is not exactly because of *lexical* knowledge, however, if we assume that an entire concept can be seen as a complete set of qualia that are stored at different levels or departments of knowledge[7], we could imagine some quale of *child* colliding with some quale of *jaguar*. It is most likely that a FORMAL quale of *child* specifies something like 'under the age of 12' or 'not having reached maturity'. My guess is that when this N occurs in a POSSESSION relation with jaguar, the relevant (Formal) quale would be 'not having reached legal age'. This is then incompatible (in the sense of likelihood rather than semantic sense per se) with the TELIC quale of *jaguar* that says it's a vehicle you need a drivers license for and a possible FORMAL quale that describes *jaguar* as one of the most expensive vehicles a person can possess. Both these factors make an interpretation that satisfies DEFINT unlikely. Also a CLOSEINT would be odd; something close to a jaguar would first of all still be a car that functions as a car (because the TELIC quale has to be preserved). And it would probably also be an expensive car; maybe an old or damaged jaguar. The topic determines that the possession relation should be optimal: if there is an interpretation of jaguar available that is less problematic for TOP, that would be preferred over DEFINT. And that's how I think the interpretation of (10) is brought about.

The case of (11a/b) is a bit different. This this is due to both the nature of the topics that trigger MTC on *dinosaur* in (11), and the concept of *dinosaur* itself. The topics *Steven Spielberg* and *Bert and Ernie* are both proper nouns. In previous parts of this study I have referred to this kind of topics as being 'iconized' in a sense. In section 2.1.3, I suggested that Fauconnier's notion of SPACE BUILDERS could be of use for this interpretation effect. The knowledge that we use here is certainly no lexical knowledge. In fact, this example is the most clear example of the need of access to world knowledge in the process of interpretation. In a nutshell, the proper name *Steven Spielberg* is highly associated with things like 'making movies' and in particular 'being famous for movie about dinosaurs'. It could be argued that there is no such thing as a QS for a proper name. However, I think in this approach we can see how fruitful the concept of QS's can be. I would propose for this case (and for similar effects in interpretation of natural language expressions) that also this kind of encyclopedic knowledge is part of the 'interpretation-database' we use. If we decide to work with an ontology that divides things into different categories, and assign different QS's to the objects that belong to those categories, we might just as well assume a structured form of encyclopedic knowledge. For this case, it would mean that there is a *correspondence* between the QS of TOP *Steven Spielberg* (director of Jurassic Park, The Dinosaur Movie) and the realization of one of the parts of this QS in this utterance. So, in a sense, MTC is not the result of conflicting specifications in respective QS's, but due to a strong *preference* based on those QS's. We could consider to constrain this in a separate way, I will turn back to this possibility in section 3.4.2. The effect of placing (animated) objects in this context, is that the preferred interpretation is not DEFINT. I think the explanation of (11a) holds for (11b) as well: FITINT does not forbid the interpretation of DEFINT, but when STRONGINTN proposes the possibility of interpreting *dinosaur* as N_{ima} (either a replica or a puppet/muppet), it is by all means a better candidate because of corresponding QS's.

The only example that I haven't discussed yet, is (9). This is merely because (9) is the

[7] I will come back to this in the next section and in chapter 4.

'evidence' this study has for the existence of a *ranking* of the constraints as proposed in this chapter. In (9a) there is a (preferred) interpretation put forward by LOC, which leads to MTC on object N *duck*:

	input = U$_{(9a)}$	FITINT		STRONGINTN			
		OBEYV	INTLOC	DEFINT	CLOSEINT	IMAINT	FIGINT
	Int (9a)$_i$		*				
	Int (9a)$_{ii}$		*	*			
☞	Int (9a)$_{iii}$			*	*		

This is now accounted for by FITINT, and the explanation I just gave for corresponding QS's. *Bath* is associated with some objects, one of which is *duck* of meaning type N_{ima}. If FITINT negotiates enough about this, STRONGINT might give in, and give away the DEFINT to satisfy INTLOC. However, as we shall see now, LOC can be overruled by another FITINT constraint, namely OBEYV:

	input = U$_{(9c)}$	FITINT		STRONGINTN			
		OBEYV	INTLOC	DEFINT	CLOSEINT	IMAINT	FIGINT
☞	Int (9c)$_i$		*				
	Int (9c)$_{ii}$		*	*			
	Int (9c)$_{iii}$	*!		*	*		

Here we see what happens if we change the verb in (9) into a V that requires a N_{nat} as its argument: the demand of V expressed in OBEYV rules the preference of LOC in INTLOC out. This is why the optimal candidate for the interpretation of (9c) is the one where ObeyV is satisfied. In this example, STRONGINT is satisfied because OBEYV demands for a natural kind interpretation. This is still visible: the optimal output in fact violates a constraint that is higher ranked than StrongInt.

In this section more has become clear on the nature and the ranking of the constraints that I have proposed for mtc. In fact, whenever the negotiation between FITINT and STRONGINTN was discussed and represented in tableaux, we were facing new issues of OT and different dimensions of MTC. Old and new information will be evaluated in the section below.

3.4 Evaluation of the Optimality approach

On the whole, I think this approach is quite satisfactory; OT hasn't disappointed me when applied to a complex phenomenon like MTC. Another thing, in a way even more satisfying, is that I am left with more questions and problems than when I started this study. Call me crazy, but what would be the fun if this would be THE END? Before I will conclude this study, I will evaluate the OT approach I adopted in this chapter. I will focus on three topics that surfaced during the discussion in the previous sections. In 3.4.1 I will address the issue of knowledge and lexicon, relating to the Pustejovskian direction I took in my proposal. In 3.4.2 I will discuss the precise nature and ranking of the OT constraints I formulated. Finally, the last section will be a short note on the representation in the OT tableaux.

3.4.1 Lexical and cognitive foundation

The constraints I proposed in the optimality approach for MTC, rely heavily on assumptions about lexicon and general cognition. I have chosen to use part of the framework of Pustejovsky (1995) as a foundation to build my OT constraints on. However, a lexicon alone is not sufficient to carry the weight of what follows from the constraints. The lexicon as proposed in Pustejovsky (1995) ought to be taken literally; it contains the *lexical* information we use when we process utterances. We have seen especially in the discussions about discourse phenomena, that we need to distinct between the different kinds of knowledge in terms of *accessibility* and *usage* in interpretation of language.

Based on the data and OT constraints I developed for those data, I assume there are (at least) three kinds of knowledge that we use during interpretation: *lexical* knowledge, *world* knowledge and *background* knowledge. As I already suggested in the previous section, I think that it could be very useful to assume that these 'kinds of knowledges' are structured in similar ways. That would mean that if we adopt the Qualia-based approach of Pustejovksy (1995) for lexical knowledge, it would be very convenient to assume the other levels can be approached the same way. Next to that, the OT account for MTC shows that the different constraints on interpretation stem from different kinds of knowledge. In a sense, the use of the respective knowledge-kinds is even *ranked*.

To start out with what can be called *lexical knowledge*, I think the least that should be part of this is *grammatical* information of lexical items. That is, word type, semantic type and, if necessary, argument structure. For example, lexical knowledge of *to begin* should at least contain the information that the word belongs to the class of verbs, has the semantic type `<e,t>` and selects an event as its second argument. This is more or less similar to the information that is represented in the lexicon described in Pustejovsky (1995), (see section 2.3 of this study). Other sorts of information that could be regarded as being lexical, can be borderlining between world knowledge and lexical knowledge. I will refer to this as the *conceptual* part of lexical knowledge. It contains the basic meaning of a lexical item that is not expressed in its 'grammar', so to say. This is where things are starting to be more complicated. The lexical information of a word regarding its function and behaviour in a sentence is *finite*. In my opinion, this information is part of the *syntactico-semantic* level, like a bridge between the level of semantics and syntax. The same sort of position can be ascribed to conceptual knowledge, that could be part of the information in the lexicon, as knowledge of the world.

For purposes of simplicity and consistency with the OT constraints, I assume that lexical knowledge includes at least the *prototypical* meaning of the concept that a given lexical item represents (C_p). The meaning type associated with this meaning is N_{def}, the default meaning type of a lexical item. In short, the lexical part of the knowledge we use, is more than the qualia structure of a given lexical item:

> LEXICAL KNOWLEDGE
> `Grammatical properties`: wordclass, argumentstructure, eventstructure and default meaning type;
> `Default Meaning Type`: for class N either N_{art} or N_{nat}, and other choices based on the type lattice (AGENTIVE quale);
> `Prototypical Meaning`: attributes and function of the prototype of C_N. (CONSTITUTIVE, TELIC and FORMAL qualia)

This basic idea of lexical knowledge is mainly based on Pustejovsky (1995) combined with the assumption that we use prototypes (Rosch 1978) in addition to defaults. The division of properties into attributive and functional is based on the work of Franks (1995), as discussed in section 2.3.

I take *world knowledge* to have two important characteristics in the interpretation of utterances. First of all, world knowledge is linked to the formation on *prototypes* in lexical knowledge. This is partially related to the way concepts are learned, not only during language acquisition stages but also after language maturation. The number of concepts will be growing over time, and the prototypes of a given concept may change once more knowledge on the prototype has been acquired. This knowledge is based on what you perceive in the world about certain concepts. Take for instance the concept of an entity that I have used in this study: *elephant* (see example (8), section 1.1.2). Suppose there is a situation of a European child learning the word *elephant* in the way that Partee (2001:6, as quoted in section 3.2.1) describes. Most children in the world will have their first encounter with 'an elephant' by learning to assign the correct word to a picture of an elephant. In most cases, I think the picture will be no more than a basic drawing of an elephant (as opposed to a photo). I think this picture forms the initial prototype, until the child/person is exposed to more information about an elephant, going from photo's to maybe seeing an elephant in the zoo. In other words, this would mean that the prototypical meaning of elephant can be changed over time based on knowledge derived from perceiving the world.

Knowledge of the world can also play a role for certain interpretations, like the ones that are, according to Fillmore (1985), due to certain *frames* we are familiar with in life. These frames can for example be settings in a restaurant (where you can refer to persons as '*number 9*', meaning the *persons sitting at table number 9*). This kind of knowledge is also used when talking about someone who is famous, known things from television, et cetera. What happens is, amongst other things, that this world knowledge (that the speaker estimates the hearer(s) to have) allows the speaker to leave away certain words, or use a word and expect them to be interpreted correctly but in a non-default way without having to explicitly use a non-default word. In case of the examples of this study, this could be true for (10) and (11) where certain knowledge of the world can make the default interpretation *less likely* at the least (see section 1.1.2). In short, I see world knowledge as follows:

> World Knowledge
> `Extension of concepts and prototypes`: learning new things about a concept can extend and change the features that are considered part of the prototype of that particular concept.
> `Provision of 'settings'`: this information can be so relevant that it has scope over elements that in the utterance, are associated with that information.

These two sorts of world knowledge seem to be quite distinct. I have chosen to group these forms together under world knowledge because they are both kinds of 'knowledges' that are acquired at any stage in life without being completely independent from lexical knowledge. However, `Extension of concepts and prototypes` underlies the Strength constraint, whereas `Provision of settings` is part of the Fit constraints.

Last but not least, I take *background knowledge* as being the knowledge about the non-linguistic context in which the utterance takes place. This context can influence the interpretation of the utterance. As addressed in 1.1.2, again turning to the case of the elephant;

suppose this is an utterance made by someone in India who has a house in the middle of a wilderness where it's common for elephants to roam. In this case, the MTC interpretation that is the result of room can be avoided. We can then interpret (8) as the subject seeing the elephant lying on the ground from his window. This is only possible when an 'elephant-containing' location is the background of the specified LOC *room*. However, this should either be a marked location in a text (with for example the geographical name of the area in India as a discourse marker for LOC), or the hearers are supposed to know the circumstances of the speaker's home.

These layers of knowledge are not separated by clear-cut boundaries, maybe that is why they seem to be interactive and have access to each other's information if needed. For many reasons, a qualia-based approach to the lexicon has been very useful. It will be interesting to work some more on the idea that in fact all kinds of knowledge are structured in a similar way. Of course, not all parts of QS's that are described in the lexicon need to be represented in the other forms of knowledge. I think the discussion of the effect of INTLOC and the intuition that LOC in a discourse can create a setting with a certain scope could support this idea. We have seen that those parts can cause mtc by stating a mere preference for a certain interpretation. This preference is due to certain expectations that come with either INTTOP or INTLOC. Central to this is the intuition that the preference is realized in some cases (for example N_{ima} duck in (9a)) based on correspondences in respective QS's. The details of a study like this are beyond the scope of this work, and belong to the work of a linguist or philosopher working in the field of language and cognition.

I promised a 'ranking' of knowledge. This ranking is related to the *cognitive* aspect of the interpretation of MTC, and, in my opinion, could reveal something on the way different levels of knowledge have access to each other's database, so to say.

RANKING OF KNOWLEDGE DATABASES
`Lexical Knowledge > Background Knowledge > World Knowledge`

In my opinion, this answers to the intuition that when we process the meaning of an utterance, we will always try to interpret it first according to the demands stored in the lexicon. It is only in case of semantic incompatibility, collision between the meanings or when the utterance just doesn't make sense based on that knowledge, we will start digging around in the database to find information that offers us the optimal interpretation. Note that this ranking has only to do with the processing of the examples, and in fact doesn't provide anything about the outcome. In that sense, it's not an optimality ranking, but just an idea about the order in which the mind might work.

3.4.2 Relative ranking and levels of constraints

While discussing possible constraints for MTC, I already hinted at a ranking that holds for them. In this section I will dwell on this some more. In order to do so, I will discuss the difference between FITINT on the one hand and STRONGINTN on the other, and place them in the perspective of more general conversational principles.

In section 3.2.3 I have suggested a ranking for the constraint that promotes the strongest interpretation for N, STRONGINTN. A section later, I have argued that FIT constraints INTLOC and INTTOP do not have the same force in context. This means that they can be violated to satisfy a higher ranked FIT constraint, like OBEYV. There is a big difference between the

type of ranking of FitInt and StrongIntN respectively. The ranking of the Fit constraints is similar to the OT constraints that I presented earlier in this chapter. A constraint can be violated if its satisfaction yields a violation of a higher ranked constraint. The different OT tableaux for (9) in the previous section are a good illustration for this. The set of FitInt constraints is *competitive*. This does not hold for the constraints of StrongIntN, which I will discuss later. I think it's best to divide Fit constraints into two groups. The first can be seen as constraints that play a role in the interpretation of a given *syntactic* construction. These are in this study IntAdj, IntMod and ObeyV. The other group consists of the constraints that are part of the interpretation of the discourse, which I have called *narrative* constructions. In this study, we have seen the role of IntLoc and IntTop. I would imagine that things like *time* (not *tense*) or *narrator* are part of this group as well. For now, I will regard these groups as *governing* constraints. For the constraints playing a role in certain constructions, I propose the following governing constraint:

> Interpret Syntax I
> If an utterance U contains or consists of a syntactic construction Con_{syn} interpret U according to Con_{syn}.

This constraint should then take care that syntactic restrictions on interpretation are respected. We have seen that a violation of this results in a fatal error, marked by !* in the OT tableaux. I think the governing constraint should be formulated like ObeyV, since you cannot violate the demands on interpretation coming from syntax. This constraint should be violated under any circumstance:

> Obey Syntax (ObeySyn)
> Obey all syntactic restrictions on interpretation.

This means what it says, if there is a constraint in interpretation that finds its source in the syntactic construction of the utterance (and thereby belongs to ObeySyn), the optimal candidate interpretation of that utterance *cannot* violate that constraint. Note that the nature and behaviour of ObeySynt is similar to Syntactic Structure as described in Hendriks&De Hoop (2001).[8]

ObeySyn competes with a constraint that governs the narrative constraints. As we have seen in the tableaux of example (9) in the last section, the narrative constraints in this study can be violated to satisfy ObeySyn constraints if they play a role in the interpretation of the same utterance. Before I will explain this difference in force, let me propose a governing constraint for narrative constructions (narratives):

> Interpret Narratives I
> If an utterance U contains a narrative construction Con_{nar} interpret U according to Con_{nar}.

Since the constraints that are part of this governing constraint are explicit here, I can simplify this like I did in construing ObeyV, while taking into account the 'non-fatality' of violation of this constraint:

[8] I am not sure about the relation between the different (interpretational) constraints of ObeySynt. That is, I have no clear intuitions about examples like *The wooden dog was running in the garden*, or *When you enter the zoo, a stone lion will grawl at you.* At this point, I don't have enough material to work with. I think we need more than possible instances of MTC to see a ranking of these constraints, but it would be interesting to study this some more in the future.

Follow Narratives (FollowNar)
Follow all narrative restrictions on interpretation.

Now, we have governing constraints that account for examples where different FitInt constraints occur at the same time. In my opinion (and to my relief), I don't have to create new ideas and come up with innovative frameworks to explain this different nature. I have argued in the beginning of this study that it would be practical to have an understanding of interpretation where pragmatics isn't used as a bin for things we can't explain within semantics. At this point of this study I can put some pieces of the puzzle together, finally.

The constraints in ObeySynt are fatal when violated because of their *presuppositional* force. Partee (2001:1-2, see also section 2.4.3) suggested for adjectives that trigger mtc on the head N that the interpretation of an Adj-N construction with conflicting semantic properties is the result of presupposition accommodation. My claim is that this is true for all the constraints of ObeySynt. A verb presupposes certain semantic properties to be true for the realized argument. If this realized argument doesn't have these properties, it will be assigned these properties no matter what. If that means replacing one or more qualia in QS_N, concessions and negotiations take place with StrongIntN. This holds for adjectives as well; as explained in the early work of Montague, and adjective tells us that some property is true for an object . Accommodation in this sense means for StrongIntN providing the strongest interpretation that can bind the presupposition resulting from the given syntactic construction. If not, a higher leveled and fatal constraint, is violated. I'd describe this constraint as:

Bind Presuppositions

This can be seen as the ot variant of Binding (see Van der Sandt&Geurts 2004:7, and section 2.4.3 of this study).

The constraints of FollowNar that we have seen here, do not project a presupposition onto the discourse. I'd describe the behaviour of these constraints more as 'stating a preference'. This *preference* must be fulfilled as long as this does not result in violation of a higher ranked FitInt constraint. The constraints IntLoc and IntTop include a relation R. I have argued that in this study, the only clear R has been a bridging relation. Assuming that there are more types of R that can hold between discourse markers and other linguistic elements, I propose that we at least see R as some sort of *coherence* relation. In section 3.3.3 I suggested that mtc in some of the examples occurred because of corresponding semantic properties of mtc trigger and coerced N. As suggested in the work of Asher&Lascarides, relationships like Bridging serve the purpose of discourse *coherence*. We could then see a violation of FollowNar where this does not result in satisfaction of a higher ranked constraint, as a violation of:

Maximize Coherence

We have seen this constraint as a principle for Bridging in the dice/drt framework discussed in Asher&Lascarides (1998). I take this direction to show that ot has a convenient property of being applicable at different levels of linguistic theory. For this study, we see that FitInt can be seen in the perspective of presupposition binding and discourse coherence. In a way, this means that we should be able to ascribe different phenomena in natural language to the same 'higher-leveled' constraint.

We are left with the constraint that has to put up with the erratic demands of FitInt: StrongIntN. While the constraints of FitInt are competitive and thereby *autonomous*, the 'constraints' of StrongIntN are more like a set of pawns that are used and highly controlled by StrongIntN. That is to say, the different interpretations that StrongIntN allows are not competitive in nature, StrongIntN is in fact satisfied as long as the selected optimal candidate yields the strongest interpretation of N that does not violate a Fit constraint. In this study, I have only focused on the interpretation of nouns, and I have spoke of them in terms of *concepts*. Of course, concepts are not necessarily expressed by using a noun. In order to place StrongIntN in the same perspective as I just did for FitInt, I assume that StrongIntN is part of a general interpretational constraint that preserves *conceptual strength*:

Maximize Conceptual Strength

This constraint is then competitive with the general Fit constraint. It is clear that I chose to use more specific constraints than Zwarts (2003) did in his approach. Although we both had to account for violations of Strength in context (Fit), the data of this study were divergent and could not be approached in the same mathematical way. On top of that, I dealt with different kinds of contexts, which led to assumption of different member-constraints of Fit. The same holds for the ranking I propose for the constraints that describe mtc, on the highest and simplest level, we can say that:

Ranking Interpretational Constraints
`Fit > Maximize Conceptual Strength`

The Fit constraint splits into more levels than the Strength constraints, which is not surprising, given the autonomous status of Strength constraints and the fact that Fit concerns a context, and Strength only one linguistic element. So, within Fit (FitInt), we have seen the following competition:

Ranking Fit Constraints
`Bind Presuppositions > Maximize Coherence`

Since it is impossible to rank the constraints that fall under ObeySyn, the final ranking that ii propose for ot of mtc, is as follows:

Ranking MTC Constraints
`ObeySyn > FollowNar ▷ StrongIntN`

Here, I use ▷ to indicate that the interranked ObeySyn and FollowNar do not simply outrank StrongIntN, but that the selection of the optimal candidate interpretation- and thereby the phenomenon mtc- is the result of *negotiation* between FitInt constraints and StrongIntN rather than competition.

3.4.3 Dynamics in interpretation and linguistic representation

The final topic that is up for discussion in this chapter, is the representation in OT. I will use this section to discuss the advantages of OT tableaux as opposed to the modes of representation in the frameworks we got acquainted with in chapter 2, as well as the particular mode of representation in tableaux that I chose for this study. After that, I will relate the representation of OT to the dynamic nature of natural language interpretation, with a special focus on the interpretation of MTC.

In chapter 2, I have pointed out for more than one theory, that the representation of interpretation is an important part of what I see as their weakness. This was especially true for the proposals of Franks (1995, see section 2.2) and Pustejovsky (1995, see section 2.3). In both cases, interpretation is represented in static matrices which, especially in Franks (1995), represent *information states*. It is true that different information states at least represent *one* of the axes of the dynamics of interpretation. In fact, one could say that 'functionalist' matrices represent different points on a vector that moves from information state 0, where there is just an utterance to information state 1, where the utterance has been assigned an interpretation. I have argued in both 2.2 and 2.3 that if we'd choose a SG/GL representation, we'd remain clueless about the where and especially the *why* of the interpretation of MTC. This is mainly due to the fact that we move from one state to the other, and the point at which the mind decides to throw in a N_{ima} interpretation can be any point. It should be noted that I didn't address the representation that is used in Fauconnier's MST or DRT explicitly in this study. It holds for the general MST processes that are represented in Fauconnier (1994), as well as the representation of conceptual blending (see Fauconnier&Turner 1996), that they are static and example-specific. As a consequence, it will be complicated to use this mode of representation of general processes in natural language interpretation. Representations within the framework of DRT are in fact quite practical, the boxes give the impression you look into a sort of mental device, and because of their strong foundation in logics, generalizations are not problematic. However, they are still static compared to the tableaux in OT. This is merely due to the theoretic difference between DRT and OT. In the first it is assumed that language interpretation is guided by *rules*, whereas OT works with violable constraints. I think in the end, it is this new and flexible view on rules in OT, which allows for a representation that visualizes part of the dynamics of interpretations. By using the simple yet adequate OT tableaux, it is possible to see the actual meaning negotiation taking place.[9]

In this study, I have added a new dimension to the unidirectional OT tableaux. As I mentioned in section 3.4.2, the constraints can be formulated, and data can be analysed from different perspectives. In this study, we have seen a couple of members of a family of FitInt constraints. Next to that, I have argued that there is a scaled set of StrongIntN constraints. These are different, since they are not competitive constraints amongst themselves. In the work of Zwarts (2003), there is only one Strength constraint, which gets violated a number of times depending on the output interpretation. Although this is Optimality-theoretically correct, the process is a bit opaque. Which is no problem for Zwarts, since he chose data involving only one word with different interpretations, and a word that could easily be explained in terms of mathematical rules. That is why I added the different constraints involved in StrongIntN. It allows both generalization over the different examples of MTC and a more detailed view on the behaviour of StrongIntN in the negotiation with FitInt constraints.

[9] I am aware of the fact that this appraisal for OT tableaux is biased in this study; if you assume that the process of interpretation is linear, which I evidently don't, a representation in matrices will be sufficient.

The latter is relevant in the discussion about concepts (see section 3.1.3); I think that by including the different levels of interpretation of N this way, it is visible how StrongIntN wants the strongest possible interpretation given FitInt, which is the reason why in some cases, coercion takes place. The nature of concepts is flexible, and I think that if we adopt a prototypical view on concepts- as I have done in this proposal- most utterances will have interpretations in which DefIntN is violated, but CloseIntN is allowed by FitInt.

In a nutshell, the representation and approach in this study is based on a flexible notion of interpreting and understanding concepts in the context they are used in.

Chapter 4

Conclusion and Afterthought

> *'I like this fact. It is something you can work out in your own mind just by looking at the sky above your head at night and thinking without having to ask anyone.'*

Writing a final chapter means having to agree to a boundary of research. What ends here, may form a start for future research, because as usual there are more questions than answers and the solution of one problem usually means a set of new challenges somewhere else. This chapter consists of two parts, the first being a conclusion of this study, the second being a set of ideas that came into mind when working with this material.

4.1 Concluding this study

Metonymic Type Coercion is an effect in the process of natural language interpretation where the context can force a concept to alter its ontological type. I have discussed many examples of this phenomenon. The interpretation that is generally associated with this phenomenon is that a certain object has to be interpreted in terms of its image, hence the ontological type can shift from N_{def} to N_{ima}. Although explaining MTC in these terms is new, the phenomenon has been topic to studies in very different parts of semantic theory. While discussing the work around MTC from the perspectives of cognitive, functional, generative and formal semantics, it became clear that different semantic theories tend to focus on different instances of this phenomenon and have very different thoughts on lexicon, knowledge, rules and interpretational processes. I concluded that resolving the MTC puzzle within one of the current semantic theories will be very difficult, and that in order to come to a proper account for this phenomenon, we should aim for a framework that is flexible enough to account for all the examples of MTC and, more importantly, including an explanation for the different MTC functors in natural language.

The central part of this study is my proposal for an *optimality* approach for MTC. The OT framework has the convenient characteristic of being a view on how (linguistic) rules work rather than a set of rules in itself. Because of this characteristic, I was able to reconstruct rules, laws and principles as proposed by my predecessors into violable OT constraints.

For this purpose, Pustejovsky's GL serves as a solid foundation for the constraints, mainly because of the assumption that lexical items are structured in terms of QS's. I have proposed that we should extend this model of qualia structures to the level of both conceptual and world knowledge. This is the key in the resolution of MTC triggered by different linguistic elements: if we use more than *lexical* knowledge in the interpretation

process, and interpretation can be affected by other forms of knowledge, it is most practical to approach the structure of that knowledge in a similar way. Qualia Structures have become especially important in the construction of constraints that conserve conceptual integrity, which I have called StrongIntN). The idea is that the context in which a given noun occurs, can be seen as a set of constraints that are higher ranked. The lower ranked StrongIntN constraint will have to start negotiating with FitInt when QS's of a given N collide with QS's of an element in the context of N. The proposed subconstraints for FitInt represent the different MTC functors that we have seen in the examples of this study.

The combination of QS's with different a prototypical view on conceptual meaning allowed me to constrain the interpretational effect of MTC, namely that the coerced noun shifts to an *image-of* interpretation of the noun. I scaled the different possibilities for noun interpretation in a minimal way- from a default/prototypical to a non-literal interpretation. It seems plausible that StrongIntN is only violated when a given N cannot possibly be assigned any of the interpretations that are subsumed by StrongIntN.

The application of an OT methodology and perspective on MTC has been fruitful: not only does the violability of constraints answer to the intuition that conceptual meaning is weaker than the force of context, the representation in simple OT tableaux allows us to actually *see* the negotiation between FitInt and StrongIntN taking place. Above that, the constraints that in my proposal account for MTC, can be placed in the perspective of general interpretational and conversational constraints that have been formulated in other optimality-driven studies in the field of semantics and pragmatics.

4.2 Suggestions for further research

After more than a year of thinking almost non-stop about the topics addressed in this study, even dreaming frequently about elephants in my room or stone lions that wouldn't stop walking loudly over my roof, it is no surprise to me that I can think of quite some possibilities for extended research. I have narrowed it down to two directions: further theoretical research that concerns the consequences of this OT approach of meaning in the philosophy of language and theories about semantic compositionality and an idea for an experimental psycholinguistic study on processing MTC and other instances of apparent incompatible combinations in natural language expressions.

4.2.1 Optimal interpretations and conceptual theory

It was originally part of my research plan to investigate the conceptual side of MTC and similar effects in interpretation. Since this would have led to a thesis twice this size, I thought it wiser to just focus on OT and make this into a suggestion. As is clear from the proposal in this theory, I had to make some assumptions about a fitting conceptual framework. I chose the prototypical view because it is in line with the work of Zwarts (2003), but haven't defended this choice extensively here.

I'd like the idea of taking this study as a starting point for a more theoretic investigation in the field of conceptual theory. This 'field' doesn't have clear boundaries, theories stretch from Aristotelian philosophy on sorts, kinds and categories to the psychological investigations as mentioned in relation to the work of Rosch (1978). It can be inferred from this study that concepts cannot be explained in terms of strict definitions or sets of semantic values.

The data show that even non-lexical knowledge can alter the interpretation of a certain concept, if it is placed in a completely different setting or dimension. If we assume that an optimality approach works well in the account of cases where the context forces a concept to be interpreted in a non-default way, which consequences does this have for our view on concepts? And which conceptual framework fits to the idea of a context-based approach of conceptual semantics?

It would be interesting to divide a study that is concerned with questions like these, into different levels. After all, assumptions on conceptual theories coincide with assumptions on categories and set theories. These seem to be inseparable, but distinct. We have seen in the data of this study, that concepts and the boundaries of categories tend to be *vague*. Vagueness in natural language meaning has led to many debates amongst both linguists and philosophers. For example, the work of Kamp&Partee (1995), discussed in section 2.4 of this study, discusses the prototypical view on concepts in relation to fuzzy set theory. In fact, Kamp&Partee (1995) make some interesting suggestions of adopting a supervaluation theory as set theory, which would not be incompatible with PT (if we would be convinced prototypes would be the way to go).

A study should then involve current theories about concepts and sets, and view them in the perspective on conceptual flexibility that we encountered in this study. I suggest working some more on the possibilities of supervaluation and optimality, to see if there we could integrate flexible theories on sets and concepts into OT. Literature that comes to mind here are the interesting volumes of Laurence&Margolis (1999) and Keefe&Smith (1996) on VAGUENESS, and the important paper of Blutner et al. (2003) on COMPOSITIONALITY.

It has been pointed out a couple of times in this study, that if adopt a flexible notion of (conceptual) meaning, which we can't avoid given the vast amount of data, a compositional view on meaning and interpretation is challenged. So, in a complete 'conceptual account' of the results of the present study, a discussion about the *status quo* of compositionality seems inevitable.

4.2.2 Processing semantic collisions: an experimental foundation for OTS

What has been no part of this study at all, is some kind of experiment. I was the only subject in 'testing' the interpretations of the data, so that could be a starting point for an experimental study on processing meaning incompatibilities.

For starters, a test could be designed to test the examples of this study. This can be a test in which subjects are presented a sentence that supposedly has a coerced interpretation, and ask the subjects what their first understanding was. This won't be an easy test to design. In research methods of first language acquisition there is a test called 'picture representation task'. Maybe this could be a design to test interpretations, although this would be kind of *circular*, images and image-interpretations also being part of the problem. Another possibility is to give the subjects a set of possible, paraphrased interpretations of the data. This is a multiple choice test, which is not my favourite way of testing anything due to chance and people not knowing but filling in something for the sake of answering.

Suppose we would have developed a test and it would be safe to assume the results are valid, we could create a bigger set of examples (of the same kind of course) and perform a completely different test on another group of subjects. The first test and test group were, after all, a pilot to see if what I regard semantic incompatibilities is judged the same way by other speakers of the same language.

The actual test will be expensive but very interesting. It would be of interest to see if interpreting degrees of StrongInt violations requires more of our mind. If theorists who favour the optimality approach are correct, the language organ contains a device Eval. Eval consists of a set of ot constraints that are activated depending on the input that enters Gen. I'd think that an output candidate violating one or more of the relevant constraints would take *more* time to process than an output violating less or no constraints. An Eye-Tracking test would be a possibility to see if it takes more time to read a sentence like (7a) than it takes for (7b). Or any other pair of similar sentences where the first contains related elements that are semantically incompatible, and the second doesn't. My hypothesis for an experiment like this, would be that it takes longer to process sentences with semantic collisions.

Bibliography

Asher, N. and A. Lascarides (1995): 'Lexical Disambiguation in a Discourse Context'. In: *Journal of Semantics* **12.1**, pp.69-108.

Asher, N. and A. Lascarides (1998): 'Bridging' In: *Journal of Semantics* **15.1**, pp.83-113.

Beaver, D.I. and H. Lee (2003): 'Input-Output Mismatches in OT'. In: R. Blutner and H. Zeevat (Eds.), *Optimality Theory and Pragmatics*, Palgrave, Macmillan.

Blutner, R. (1999): 'Some aspects of Optimality in Natural Language Interpretation'. In: H. de Hoop and H. de Swart (Eds.) *Papers on Optimality Theoretic Semantics.* Utrecht Institute of Linguistics, OTS, pp.1-21.

Blutner, R., P. Hendriks and H. de Hoop (2003): 'A New Hypothesis on Compositionality'. In: P.P. Slezak (Ed.) *Proceedings of the Joint International Conference on Cognitive Science*, Sydney, Australia, pp.53-57.

Bos, J., A-M. Mineur and P. Buitelaar (1995): 'Bridging as coercive accommodation'. In: *Proceedings of the workshop on Computational Logic for Natural Language Processing*, Edinburgh.

Clark, H.H. (1977): 'Bridging.' In: P.N. Johnson-Laird&P.C. Wason (Eds.) *Thinking. Readings in Cognitive Science*, Cambridge University Press, pp.411-420.

Coulson, S. and G. Fauconnier (1999): 'Fake Guns and Stone Lions: Conceptual Blending and Privative Adjectives.' In: B. Fox, D. Jurafsky & L. Michaelis (Eds.) *Cognition and Function in Language*, CSLI, pp.??-??

Fauconnier, G. (1994): *Mental Spaces.* Cambridge University Press, Cambridge.

Fauconnier, G. and Turner, M. (1996): 'Blending as a Central Process of Grammar.' In: A.E. Goldberg (Ed.) *Conceptual Structure, Discourse, and Language*, Cambridge University Press, pp.113-130.

Fauconnier, G. and Turner, M. (1998): Conceptual Integration Networks. *Cognitive Science* **22**(2), pp.133-187.

Fillmore, C.J. (1985): 'Frames and the semantics of understanding.' *Quaderni di Semantica* **VI**(2), pp.222-254.

Foucault, M. (1983): *This is not a pipe.* University of California Press, Berkeley.

Franks, B. (1995): 'Sense Generation: A "Quasi-Classical" Approach to Concepts and Concept Combination.' *Cognitive Science* **19**, pp.441-505.

Geurts, B. and R. van der Sandt (1997): 'Presuppositions and backgrounds'. In: *Proceedings of the 11th Amsterdam Colloquium*, University of Amsterdam.

Geurts, B. and R. van der Sandt (2004): 'Interpreting Focus'. In: Krifka, M. (Ed.) *Theoretical Linguistics, Special Issue: Interpreting Focus*, Vol. **30**(1), pp.1-44.

Grice, H.P. (1975): 'Logic and Conversation'. In: P. Cole and J. Morgan (Eds.) *Syntax and Semantics*, reprinted in E. Margolis and S. Laurence (Eds.) (1999), *Concepts: Core readings*, MIT Press, Cambridge Mass.

Haddon, M. (2003): *The curious incident of the dog in the night-time.* Vintage Books, New York.

Hendriks, P. and H. De Hoop (2001): 'Optimality Theoretic Semantics.' *Linguistics and Philosophy* **24**, pp.1-32.

Kamp, H. (1975): 'Two theories about adjectives.' In: E.L. Keenan (Ed.), *Formal semantics for natural languages.* Cambridge University Press, pp.123-155.

Kamp, H. and B. Partee (1995): 'Prototype theory and compositionality.' *Cognition* **57**, pp.129-191.

Kamp, H. and U. Reyle (1993): *From Discourse to Logic.* Kluwer Academic Publishers, Dordrecht.

Keefe, R. and P. Smith (1996): *Vagueness: a reader.* MIT Press, Cambridge Mass.

Lascarides, A., N. Asher and J. Oberlander (1992): 'Inferring Discourse Relations in Context'. In: *Proceedings of the* 30^{th} *Annual Meeting of the Association of Computational Linguistics*, pp.1-8.

Laurence, E. and S. Margolis (1999): *Concepts: Core readings.* MIT Press, Cambridge Mass.

Matsui, T. (1995): *Bridging and relevance.* John Benjamins Publishing, Amsterdam.

Montague, R. (1970): 'English as a formal language.' In: *Linguaggi nella Societá e nella Technica. Edizione di Communitá*, Milan.

Oherson, D. and E.E. Smith (1981): 'On the Adequacy of Prototype Theory as a Theory of Concepts. In: *Cognition* **9**, reprinted in E. Margolis and S. Laurence (Eds.) (1999), *Concepts: Core readings*, MIT Press, Cambridge Mass.

Partee, B. (2001): 'Privative Adjectives: Subsective plus Coercion'. To appear in: T.E. Zimmerman (Ed.) *Studies in Presupposition.*

Partee, B. (2004): *Compositionality and Coercion in Semantics. Case Study: The Dynamics of Adjectival Meaning.* Lecture at the KNAW Masterclass 'Cognitive Foundations of Interpretation', Amsterdam.

Prince, A. and P. Smolensky (1997): 'Optimality: From Neural Networks to Universal Grammar'. In: *Science*, **275**, pp.1604-1610.

Prince, A. and P. Smolensky (2004): *Optimality Theory: Constraint Interaction in Generative Grammar.* Blackwell Publishing, Oxford UK.

Pustejovsky, J. (1995): *The Generative Lexicon.* MIT Press, Cambridge Mass.

Pustejovsky, J. (2001): 'Type Construction and the Logic of Concepts.' In: P. Bouillon and F. Busa (Eds.), *The Syntax of Word Meaning*, Cambridge University Press.

Rosch, E. (1978): 'Principles of Categorization.' In: E. Rosch and B. Lloyd (Eds.), *Cognition and Categorization*, reprinted in E. Margolis and S. Laurence (Eds.) (1999), *Concepts: Core readings*, MIT Press, Cambridge Mass.

Williamson, T. (1996): 'Vagueness and ignorance.' In: R. Keefe and P. Smith (Eds.), *Vagueness: a reader*, MIT Press, Cambridge.

Zeevat, H. (2000): 'The asymmetry of Optimality Theoretic Syntax and Semantics'. In: *Journal of Semantics*, **17**, pp.243-262.

Zwarts, J. (2003): 'Competition between word meanings: The polysemy of (a)round'. Paper presented at *Sinn und Bedeutung VIII*, Frankfurt

Samenvatting in het Nederlands

In dit werk staat het verschijnsel 'Metonymic Type Coercion' (MTC) centraal. Voorbeelden van MTC in het Nederlands zijn onder andere '*Ik wil een jurk met bloemen aan vandaag*' en '*Jan maakte een harde klap toen hij over de vrachtwagen struikelde*'. In het eerste geval wordt *bloemen* geïnterpreteerd als 'afbeeldingen van bloemen', en in het tweede geval vindt iets vergelijkbaars plaats. Hier wordt *vrachtwagen* geïnterpreteerd als een object wat niet hoger komt dan tot aan iemand's knie. De meest voor de hand liggende manier om deze zin dan zinnig te interpreteren is om vrachtwagen te zien als een speelgoedvrachtwagen. De belangrijkste overeenkomst tussen deze voorbeelden is dat er een verschuiving plaats vindt in de interpretatie. Normaal gesproken zou *bloem* tot de categorie *natuurlijke soort* gerekend worden, en *vrachtwagen* tot artefact. Beiden worden echter geïnterpreteerd als *afbeelding* van hun oorspronkelijke betekenis. Ik noem dit een coercie van ontologisch type: blijkbaar is het mogelijk voor een linguïstisch element om een bepaalde interpretatie van een ander element af te dwingen. In dit onderzoek richt ik me op nomina die zo'n ontologische verandering ondergaan wanneer ze in een bepaalde linguïstische context geïnterpreteerd worden. De coercie die centraal staat, wordt gezien als een functie F:

> METONYMIC TYPE COERCION I
> $F_{\text{MTC}}(N_{nat/art}) \rightarrow N_{ima}$

Dit verschijnsel in de interpretatie van natuurlijke taal roept vragen op, van zuiver taalwetenschappelijke tot meer taalfilosofische. Deze studie is met name gericht op het zoeken van een passende semantische theorie voor MTC. Mijn veronderstelling is dat de bestaande semantische theorieën veel te bieden hebben, maar een te specifieke en daardoor te nauwe visie hebben op betekenis.

Het is niet zinnig om te zoeken naar een verklaring wanneer je slechts een paar voorbeelden hebt van het te verklaren mysterie. Wat *is* MTC nu in de talige praktijk? Uit nadere bestudering van data, wordt duidelijk dat de omgevingen waarin MTC zich voordoet, niet uniform zijn. In deze studie is de kleinste afstand waarover coercie kan plaatsvinden, het niveau van de NP bestaande uit een constructie van Adj-N. Goede voorbeelden hiervan zijn *stenen leeuw*, en ook het meer abstracte *denkbeeldige vriend*. Het wordt duidelijk dat wat adjectieven kunnen, een eigenschap is van modifiers in het algemeen: *een jurk met bloemen* is daar een voorbeeld van. Nog een stapje verder, maar nogsteeds binnen de grenzen van syntactisch gerelateerde onderdelen van een zin, zien we dat ook werkwoorden in staat zijn tot het plegen van coercie. In zulke gevallen (zoals het struikelen over vrachtwagens) kan een werkwoord een bepaald ontologisch type afdwingen bij één van zijn argumenten. Tenslotte zijn ook elementen die traditioneel tot de discourse gerekend worden, potentiële gevaren voor het oorspronkelijk ontologisch type van een nomen. Zo kan de bepaling van een locatie maken

dat een object in die locatie opeens een stuk kleiner moet zijn, wat voor sommige objecten betekent dat de oorspronkelijke betekenis niet geïnterpreteerd kan worden. Een voorbeeld hiervan is '*Ik kwam de kamer binnen. De olifant was gevallen.*', of '*Bert en Ernie spelen vandaag met dinosaurussen*', waarin de associatie met Sesamstraat maakt dat *dinosaurus* als een muppet wordt voorgesteld.

De theorie die een verklaring en zelfs voorspelling van dit verschijnsel kan bieden, zal dit idealiter doen voor het complete scala aan voorbeelden. De eerste stap is om te kijken wat de mogelijkheden zijn om dit binnen de bestaande semantische leer te doen. Binnen het domein van semantiek hebben zich in de loop van de tijd veel verschillende scholen ontwikkeld. Een wat excentrieke visie op taal- voor linguïsten die gewend zijn aan de Chomskyaanse traditie- is te vinden in het werk van Fauconnier (1994). Zijn idee van Mental Spaces (MST) zou op zichzelf genoeg ruimte bieden voor MTC, zeker voor de voorbeelden waarin locatie markers of bepaalde onderwerpen MTC tot gevolg hebben. Naar mijn idee heeft een benadering als deze twee grote nadelen: er is amper aandacht voor het syntactisch onderdeel van een betekenisrelatie en er is geen duidelijk idee over wat voor informatie er lexicaal is. Dat laatste is voor mij erg bezwaarlijk, MST kent geen gestructureerd idee van lexicale kennis. Dat is moeilijk in de benadering van MTC, waar we te maken hebben met botsende kenmerken (features) van concepten. Dit geldt in zekere zin ook voor de theorie van Sense Generation (SG), die voorgesteld is door Franks (1995). Dit onderzoek is voornamelijk gericht op de adjectivale gevallen van MTC. Franks gaat er vanuit dat je vanuit een quasi-klassieke benadering van betekenis meerdere interpretatie mogelijkheden kunt genereren. Essentieel zijn hierin het onderscheid tussen *centrale* en *diagnostische* eigenschappen van een object en de aanname dat interpretatie een ontwikkeling van *informatie-toestanden* is. Dat laatste is wat moeilijk te bevatten met het oog op de matrices die zo'n toestand representeren. Deze zijn in mijn mening te statisch en kunnen de dynamiek van interpretatie die nodig is in de benadering van MTC niet aangeven. Daarnaast is de theorie van Franks voor wat betreft het lexicon vrij onhandig- objecten worden omschreven met talige kenmerken bij wijze van features, wat een generalisatie moeilijk maakt. De problemen van Franks (1995) zijn ten dele vergelijkbaar met die in de benadering van Pustejovsky (1995). In diens Generative Lexicon (GL) wordt uitgegaan van een sterk gestructureerd lexicon. Dit vormt, net als in SG, de basis voor het genereren van verschillende betekenissen. De informatie in het lexicon is georganiseerd in qualia structuren, die uniform toe te passen zijn op elk onderdeel van het lexicon binnen de lexicale categorie. Dat maakt het werken met GL in mijn ogen een stuk gemakkelijker dan werken met SG. Echter, ook Pustejovsky maakt gebruik van matrices voor de representatie van betekenissen en interpretatieprocessen. Dit heeft tot gevolg dat de betekenisverschuiving van MTC onzichtbaar (en daarmee ongrijpbaar) is in de GL benadering. Een laatste richting van de semantiek die in dit onderzoek belicht wordt is de formele semantiek. Hieronder valt het gedachtegoed van onder andere Kamp&Partee (1995) en een theorie als DRT. In de ontwikkeling van de formele semantiek is er een duidelijke uitbreiding te zien van het domein. Waar Kamp (1975) een Montegoviaans idee uitwerkt voor adjectieven, worden betekeniseffecten in een meer algemene zin geformaliseerd in een later werk als Kamp&Reyle (1993). Wat vooral van belang is in deze ontwikkeling is het idee over de handhaving van *compositionaliteit* van betekenis. Toegepast op MTC, zowel door Kamp&Partee (1995) en Partee (2001) worden zinnige suggesties gedaan over principes van adjectivale betekenis. Zo suggereert Partee (2001) dat MTC in Adj-N constructies een soort van presuppositie accommodatie is. In het meer discourse-georiënteerd werk van Asher&Lascarides (1995, 1998) zouden andere gevallen van MTC verklaard kunnen worden door principes over *presuppositie*, *bridging* en *narratie*.

Dit alles geeft het idee dat het aantal verschillende constructies waarin MTC plaats kan vinden, evenredig is aan het aantal verklaringen. Desalniettemin valt er voor de theorieën die aan bod komen, allemaal wel iets te zeggen.

De kern van dit onderzoek is de uitwerking van bestaande voorstellen die van toepassing kunnen zijn in de benadering van MTC, in een theorie die flexibel met regels en principes van taal omgaat: Optimality Theory (OT). In mijn voorstel is MTC het resultaat van conflicterende, schendbare *constraints*. Het idee is dat er in het proces van interpretatie een competitie is tussen twee factoren: het zo sterk mogelijk interpreteren van elk concept dat deel uitmaakt van het interpretandum en het behalen van een zo coherent mogelijke interpretatie van het geheel. Deze twee factoren worden vertaald naar constraints StrongInt en FitInt respectievelijk. Deze bestaan in mijn voorstel uit verschillende, geordende deelconstraints. StrongInt verdedigt als het ware de oorspronkelijke betekenis van een gegeven concept. Voor *olifant* zou dat bijvoorbeeld betekenen dat StrongInt ervoor vecht dat de interpretatie volgens het ontologisch type N_{nat} is en geïnterpreteerd wordt als *prototype* van zijn categorie. Er is meer nodig dan dat. Om echt te zien wat er gebeurt in het proces van MTC hebben we een gestructureerd idee van lexicale kennis nodig. In mijn voorstel zijn de constraints die onder StrongInt vallen, deels gebaseerd op qualiastructuren zoals deze beschreven worden in Pustejovsky (1995). Hierdoor is het mogelijk om een nauwkeurig beeld te geven van MTC. De qualiastructuren zijn ook aanwezig in de constraints die er voor zorgen dat een interpretatie een coherent geheel vormt. De constraints die deel uitmaken van FitInt stellen bijvoorbeeld dat je werkwoorden altijd moet 'gehoorzamen'. Als er een conflict is tussen de qualia die deel uitmaken van het werkwoord van het interpretandum en die van één van de argumenten van dat werkwoord, begint er een onderhandeling tussen FitInt en StrongInt. Deze onderhandeling gaat net zo lang door tot dat FitInt niet meer geschonden wordt. Schending van FitInt is *fataal* en de constraints van FitInt zijn hoger geplaatst dan die van StrongInt. De optimale interpretatie van een gegeven input is dus de interpretatie met de minste schendingen, waarbij schendingen van 'hogere' constraints meer gewicht hebben dan 'lagere'.

Op deze manier is MTC zowel te verklaren als te voorspellen: het is het resultaat van conflicterende eigenschappen in de *qualiastructuren* van elementen die respectievelijk tot FitInt en StrongInt horen in het interpretandum. StrongInt is bereid om in de onderhandeling steeds een stukje afbreuk te doen van de sterkste interpretatie, om FitInt tegemoet te komen. Er is nu slechts één relevant onderscheid tussen de verschillende voorbeelden van MTC die de revue gepasseerd zijn, de gevallen die binnen een syntactische relatie plaatsvinden en de gevallen die tot de discourse behoren. De eerste kunnen, zoals Partee (2001) voor adjectieven voorstelt, als accommodaties van presupposities gezien worden, terwijl coercie door discourse elementen meer te maken heeft met de voorkeur voor een zo coherent mogelijke interpretatie.