BRILL

International Review of Pragmatics 4 (2012) 209–231

brill.com/irp

Rich Lexical Representations and Conflicting Features*

Lotte Hogeweg
Radboud University Nijmegen, the Netherlands
l.hogeweg@let.ru.nl

Abstract
This paper argues that interpretations are fine-grained and that, to come to a full understanding of meaning, it is important to find out more about how such detailed interpretations are derived. As a first step towards answering this question it is insightful to look at the interpretation of metaphors. Psycholinguistic experiments have shown that the interpretation of metaphors involves the suppression of irrelevant or incompatible features. These studies could be taken as an indication against the common view that word meanings are underspecified and enriched in a context. In contrast with this underspecification view, this paper suggests a view of the lexicon in which words come with very rich semantic representations. When two representations are combined, a conflict may arise when elements of the representations are incompatible. This paper argues that such a conflict is best analyzed in Optimality Theory. The optimization process of combining rich lexical representations is illustrated with an analysis of the adjective-noun combination *stone lion*.

Keywords
lexical semantics, underspecification vs. overspecification, Optimality Theory, metaphors

1. Introduction

Word meanings are flexible. The same word may have different interpretations dependent on the context. Consider sentence (1).

*) Most of this work developed during my visit at the Cognitive Science Department at Johns Hopkins University. This visit was funded by NWO by means of a Rubicon grant (446-09-016), which is gratefully acknowledged. I thank the members of the department for the fruitful discussions and feedback, in particular Géraldine Legendre, Paul Smolensky and Kyle Rawlins. In addition I thank Helen de Hoop and the Optimal Communication reading group. I also would like to thank Henk Zeevat for his support in developing this line of thought. Finally, I am grateful to the organizers and participants of the Theoretical Pragmatics workshop and to two anonymous reviewers and Uli Sauerland for their detailed and useful comments on an earlier version of this paper.

DOI: 10.1163/18773109-00040205

(1) My mouse makes a weird sound when I touch it

If (1) is uttered in a vet's practice, the hearer will probably interpret *mouse* as referring to a small animal but when uttered at a computer store, it gets interpreted as a complaint about the pointing device of the speaker's computer. This raises the question how words are stored in our internal lexicon. One option is that all possible meanings are listed. This may seem a plausible option for the above-mentioned readings of *mouse* in (1), which are easily distinguishable. However, words often have a flexible meaning without there being a clear cut distinction between the different senses. The adverb *fast* means something different in (2) than in (3).

(2) a fast car

(3) a fast meal

In (2) *fast* refers to the ability of the car to move fast. In example (3) on the other hand, *fast* refers to the time needed to eat the meal. It is unclear how much a meaning should differ for it to be mentioned separately in the list. Furthermore, words can be used in a novel and creative way. This aspect of language is extensively discussed in Clark (1983). Sentence (4) illustrates the class of what Clark calls *eponymous verbs (to do a Napoleon)*. Innovative use of words would not be possible if all possible uses of a word were enumerated, since such an innovation would not be in the list.

(4) The photographer asked me to do a Napoleon for the camera (Clark, 1983: 312)

The examples (1) to (4) show that it is not plausible that speakers of a language simply remember all possible meanings of a word (also see Pustejovsky, 1995 for an extensive argumentation against a sense enumeration approach).

A common theory about the lexicon entails that the semantics of words are underspecified and that more specific information is filled in by the context. However, it is seldom made explicit how such an enrichment process works precisely. I will discuss the underspecification approach briefly in Section 2. To come to a full understanding of meaning, I think is important to find out more about how a detailed interpretation is derived. As a first step towards answering this question I think it is insightful to look at the interpretation of metaphors. First of all, because metaphors are probably the most extreme example of the flexibility of word meanings. Furthermore, the interpretation of metaphors has been studied extensively in the psycholinguistic literature. I will go into some of the insights from those studies in section 3.

Based on the metaphor research I will suggest an alternative view on the lexicon in section 4. I will propose that the lexicon consist of very rich lexical representations and that features from this representation may be deleted to fit the

context. I will furthermore argue that the conflict between features that arises when words are combined can best be analyzed in Optimality Theory (Prince and Smolensky, 1993/2004). To illustrate this, I will model the interpretation of the adjective-noun combination *stone lion* based on rich lexical representations and a set of conflicting constraints in section 5. In section 6 I will say a bit more about the distinction between semantic and encyclopedic knowledge. I will end with a conclusion in section 7.

2. Underspecification and Contextual Enrichment

2.1. *Underspecification*

It is often argued that words have an underspecified semantics and that the precise interpretation is accounted for by conceptual knowledge or commonsense knowledge. By some, this underspecification is implemented by assuming free variables in a semantic representation. For example, Bierwisch (1983) accounts for the context dependence of gradable adjectives like *tall* with free variables that indicate a comparison class. Bierwisch and Schreuder (1992) argue for a two-level semantics with a conceptual level in addition to a semantic level. Bierwisch and Schreuder propose that the semantic form (SF) of a word is linked to a conceptual structure (CS). The interpretation in CS is determined by SF and an interpretation mapping *Int*.

The Generative lexicon by Pustejovsky (1995) is a special case of the underspecification approach. It assumes richer lexical representations than usual. By means of a restricted set of principles it can account for a large part of the flexibility of the meaning of words. In the Generative Lexicon, the semantics of a lexical item α is defined as a structure consisting of four elements:

$$\alpha = \langle A, E, Q, I \rangle$$

A is the argument structure, E is the specification of the event type, Q provides the binding of these two parameters in the qualia structure and I is an embedding transformation, placing α in a type lattice, determining what information is inheritable from the global lexical structure (p. 62). The Qualia Structure specifies four basic aspects of a word's meaning, based on the 4 modes of explanation by Aristotle: a constitutive, a formal, a telic and an agentive quale. By means of a well-defined set of mechanisms we can arrive at the right interpretation in a context. As an illustration, consider the mechanism of selective binding.

Selective binding

If α is of type $\langle a, a \rangle$, 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 $[[\alpha\beta]] = \beta \cap \alpha(q_\beta)$.

Pustejovsky argues that if we treat the adjective *fast* in for example *John is a fast typist* as an intersective adjective we get: λx[typist'(x) ∧ fast'(x)]. But this does not give us the right interpretation: "John is a typist who is fast at typing". In order to get this, the adjective is able to make available a selective interpretation of an event expression contained in the qualia of the head noun, in this case "typing". By means of the generative mechanism called selective binding, the denotation of the noun combined with the adjective {αβ} is the noun intersected with the adjective applied to the relevant quale of the noun. The result is that fast is interpreted relative to the activity a typist is typically engaged in: typing.

In discussing overspecification and underspecification of lexical meaning, the terms *narrowing* and *broadening*, as often used in the literature on lexical pragmatics (e.g. Wilson, 2003; Blutner, 2011), seem highly relevant. Narrowing is the phenomenon where a word is used to express a more specific sense leading to a restriction of the lexically specified denotation and broadening is the name for the phenomenon where a word is used to express a more general sense leading to a widening of the denotation of the linguistically specified denotation (Wilson, 2003). However, the notions broadening and narrowing are orthogonal to my argument for overspecification, which will be presented below. Narrowing and broadening pertain to the growth or shrinkage of the extension of a word while my proposal pertains to the loss of features of a mental representation. The result of the loss of features may lead to a wider or more restricted extension of the word.

2.2. *Detailed Representations*

The underspecification approaches I discussed briefly above do not (intend to) make explicit how we get from a underspecified lexical entry to a detailed representation. However, most authors agree that our interpretation of language is detailed and precise and includes what is mostly considered conceptual or encyclopedic knowledge. This is illustrated by the quotes by Bierwisch and Schreuder (1992) and Blutner (2004):

> ... the conceptual structure CS, in terms of which the **actual** interpretation of linguistic expressions is specified. (Bierwisch and Schreuder, 1992: 32, boldface mine)

> Of course it is not sufficient to postulate underspecified lexical representations and to indicate what the sets of semantically possible specifications of the variables are. In order to grasp natural language interpretation it is also required to provide a restrictive account explaining how the free variables are instantiated in the appropriate way. (Blutner, 2004: 10)

Furthermore, it has been shown by psycholinguistic experiments that sentence comprehension involves elaborate representations. Anderson and Ortony

(1975) performed a sentence recall test whereby cues were provided that would be effective for retrieval of a previously exposed sentence if the mental representation that resulted from this sentence was detailed but ineffective if no such details were represented. For each sentence they had two cues, one that was semantically relevant given the meaning of the sentence as a whole and one that was less semantically relevant. The semantically relevant cue would not be differently effective unless it related to the representation for the sentence constructed at the time of the initial encoding, Anderson and Ortony argue. An example of a test sentence set can be found in (5).

(5) A. Nurses are often beautiful

B. Nurses have to be licensed

C. Landscapes are often beautiful

D. Taverns have to be licensed

When the cue was *actress*, A was the target sentence, B was the subject control sentence, C was the predicate control sentence and D was the double control sentence. Anderson and Ortony found that when semantically relevant cues were provided, the test sentences were more often recalled than the control sentences. Anderson and Ortony conclude that their research "suggests that sentence comprehension and memory involve constructing particularized representations whose sense cannot be reliably predicted from the dictionary readings of the constituent words" (176).

As mentioned, most authors do not aim at making explicit how we arrive at a full and detailed interpretation of utterances. Bierwisch and Schreuder (1992) do give some requirements that a mapping from SF to CS must meet but they indicate that they cannot be explicit about the precise working of *Int* yet. The lexical representations that are assumed by the Generative Lexicon (Pustejovsky, 1992) are much more detailed than in the usual underspecification accounts but a lot of basic conceptual information, like color for example, is not represented (also see Blutner, 2004). One of the few studies I know of that does propose an explicit mechanism for meaning enrichment is Blutner (2004). Blutner argues for Radical Underspecifiaction, a view which he summarizes as follows (p. 18):

a. Every lexical unit determines an underspecified representation (i.e. a representation that may contain, for example, place holders and restrictions for individual and relational concepts).
b. The combinatorial system of language determines how lexical units are combined into larger units.
c. There is a system of type and sortal restrictions which determines whether structures of a certain degree of (under)specification are well-formed.

d. There is a mechanism of contextual enrichment (pragmatic strengthening based on contextual and encyclopedic knowledge).

To put it briefly, Blutner argues for radical underspecification of lexical knowledge and that the lacking information is filled in by means of abduction rules. For example, upon hearing *red apple*, abduction rules determine the "price" of interpreting *red* as referring to the color of the peel of the apple as opposed to the color of the stem or the pulp. The price is determined based on a Horn Clause Knowledge base containing clauses of the form $p_1, \ldots p_n \rightarrow q$, where the literals p_j in the antecedent are annotated with weights. I will come back to this study in the next section.

In sum, the most common view among formal semanticists who address the issue is that word meanings are underspecified and a detailed interpretation is ascribed to world knowledge. A interesting question to ask is then: How does such a detailed representation of meaning come about? I think an answer to this question is important to arrive at a full theory of meaning and interpretation. As a start in answering this question, it is insightful to look at research on metaphor interpretation. First of all because metaphors are probably the most extreme example of the flexibility of word meanings. Furthermore, the interpretation of metaphors has been studied extensively in the psycholinguistic literature. I will go into some of the insights from those studies in the next section.

3. The Interpretation of Metaphors

Psycholinguistic studies on metaphors provide evidence that semantic features that are incompatible with a metaphoric interpretation are initially activated and are subsequently suppressed in order to create a coherent representation. In this section I argue that these studies could be taken as an indication against an underspecification approach.

3.1. *Metaphors and Suppression*

As said, the most extreme examples of the flexibility of word meanings are probably metaphors. In example (6) the metaphor vehicle *mouse* is used to attribute a particular property to "my cousin", namely that she has a timid personality.

(6) My cousin is such a mouse

The interpretation of the vehicle of a metaphor diverges greatly from the usual interpretation of the word. To make sense of the metaphorical use of *mouse* in (6), a hearer has to interpret it with a meaning that greatly diverges from its typical interpretation. Nonetheless, at least some metaphors seem to be interpreted as

easily as literal language use (see Carston, 2010 for a discussion about different modes of metaphor processing).

A lot of psycholinguistic research has been conducted to test the interpretation of metaphors. This variety of experimental work led to several, sometimes conflicting frameworks.[1] Most views on metaphors, however, ascribe some role for the mechanism of suppression. One speaks of suppression when properties of a word that are initially activated, have to be suppressed if they are in conflict with the context the word appears in.[2] Rubio Fernández (2007) investigates the role of suppression in the interpretation of novel metaphors. Her lexical decision experiment was designed to answer the question whether metaphor interpretation involves enhancing metaphor-relevant properties of the vehicle while suppressing metaphor-irrelevant properties. Furthermore, Rubio Fernández investigates at which point in processing suppression reduces the activation of metaphor-irrelevant information. In the experiment, subjects were presented with contexts biased in favor of a metaphorical interpretation of a noun. For example, subjects heard the sentences: *Nobody wanted to run against John at school. John was a cheetah*. After hearing the metaphor, subjects had to decide whether a visually presented word was a existing word of English. Critical targets were either metaphor-inconsistent properties of the metaphor vehicle (e.g. "cat") or metaphor-relevant properties ("fast"). The time it took the subjects to respond to these targets were compared to the reaction time of an unrelated control. It is commonly assumed that faster reaction times with respect to a target items relative to an unrelated control is indicative of property activation. A set of 22 common nouns with predictable superordinates and distinctive properties were tested. Target words were presented after 0, 400 or 1000 milliseconds. Rubio Fernández found that both the metaphor-relevant and the metaphor-irrelevant properties were activated at 400 milliseconds. At 1000 milliseconds, however, the metaphor-irrelevant properties no longer showed shorter reaction times than the unrelated controls, while the metaphor-relevant properties still did. A previous study (Rubio Fernández et al., 2003) showed that superordinates are still activated at 1000 milliseconds in neutral context, which means that in this study the loss of activation was not due to a passive decay but they were actively suppressed.

---

1) A complete discussion of the literature on metaphors is beyond the scope of this paper but see Giora (2008) for an overview.

2) The precise nature and scope of the mechanism of suppression is under debate but most authors agree that semantic features that are initially activated need to be suppressed if they are disruptive in coming to a coherent interpretation of a word in context (see Giora, 2008 for a discussion).

In sum, Rubio Fernandez (2007) shows that relevant/compatible and irrelevant/incompatible properties of the metaphor vehicle facilitate lexical decision tasks up till 400 ms after the stimulus onset. However, after 1000 ms this effect is no longer present for incompatible vehicle properties while it is still there for compatible properties. Concretely, this means that upon hearing utterance (6), people would be faster at recognizing *grey* and *timid* after 400 ms compared to an unrelated word like *tree*. However, after 1000 ms this effect would disappear for *grey* while it would still be there for *timid*.

3.2. *Semantic Features and the Context Dependence of Meaning*

What does the study by Rubio Fernández tell us about lexical interpretation in general? To answer this question we first need to determine whether the features that are activated are part of "the meaning" of the expression or whether they are merely associations. This is a complex question that involves many assumptions about the nature of meaning and language. Although I realize that this issue requires much more research and elaboration, for now I assume that the features can indeed be seen as part of the meaning of the words. The fact that features are suppressed if they are incompatible with the context indicates that hearers make a selection of the activated features to create a coherent representation. The view of the lexicon I (tentatively) propose here is that a word is linked to a set of semantic features. This set includes possibly conflicting features. For example, a mouse can be grey or white or black or another color. They are all possible features of a mouse but they cannot be true at the same time (at least not for one particular part of the mouse). From the set of associated features a subset is selected each time the word is interpreted in a context. In this view, the meaning of words is fundamentally context-dependent. I realize that this view is not uncontroversial. Fodor and Pylyshyn (1988) for example argue that meanings must be context-independent. They criticize the view outlined by Smolensky (1991) that the representation of the meaning of for example *a cup with coffee* varies according to the context in which it appears. According to Smolensky, we can depict the representation of *a cup with coffee* as the combination of certain semantic features like "upright container", "hot liquid", "porcelain curved surface", "burnt odor" etc. Critics of this view on representations would argue that it cannot be right because the representations of *cup without coffee* and *coffee* should be subtractable from the representation of *cup with coffee*. Now, Smolensky argues that we *can* subtract the representation of *coffee* from the representation of *cup with coffee*, only this will be a representation of *coffee in a particular context*. There is not one representation for coffee, but a collection of representations knit together by family resemblance. The particular representation that will emerge in a given situation is therefore context dependent. Nonetheless, *coffee* is a constituent of

the representation of *cup with coffee*. However, this constituent relation is not part of the mechanism within the model. Fodor and McLaughlin (1991) argue that systematicity requires *context-independent constituents*. In Smolensky's solution, the representation of *coffee* you get by subtracting it from the representation of *cup with coffee* is not a representation of *coffee* when it stands alone. The representation of coffee that you get from *cup with coffee* does not give the necessary conditions for being *coffee* for a representation of coffee in a can with coffee would yield a different set of features. And, Fodor and McLaughlin argue, it is not a sufficient set of features either. So, Fodor and McLaughlin wonder, what does make a representation a coffee-representation? There is no single vector that counts as *the* coffee-representation and therefore there is no vector that is a component of all the representations, which in a classical system would have *coffee* as a classical constituent. Fodor and McLaughlin suggest that Smolensky confuses being "a representation of a cup with coffee with being a CUP WITH COFFEE representation":

> Espying some cup with coffee on a particular occasion, in a particular context, one might come to be in a mental state that represents it as having roughly the microfeatures that Smolensky lists. That mental state would then be a representation of a cup with coffee in this sense: there is a cup of coffee that it's a mental representation of. But it wouldn't of course, follow, that it's a CUP WITH COFFEE representation; and the mental representation of that cup of coffee might be quite different from the mental representation of the cup of coffee that you espied on some other occasion or in some other context. So, *which mental representation a cup of coffee gets is context dependent*, just as Smolensky says. But that doesn't give Smolensky what he needs to make representations themselves context dependent. (342)

Smolensky argues that the semantic "representation of a cup with coffee" can vary over contexts. Fodor and McLaughlin (1991) argue that the "representation of a cup with coffee" that for example arises upon seeing one may be context dependent but this is not the type of representation that is part of the combinatorial system of language and thought. This latter type of representation, a CUP WITH COFFEE representation, is context independent.

The view outlined in this paper on the distinction between the two types of representations is the following: intentions (to express something) and interpretations are of the type "representation of a cup with coffee". This representation consists of a set of features which may vary across contexts. Words *are* linked to an invariable set of features. However, these features are not directly accessible to language users but they may surface in semantic representations that form the intentions or interpretations. As for the relation between words and meanings, Fodor and Lepore (2002) say: "we assume, for the present discussion, that words express concepts, and that the content of a word is the content of the concept that it expresses" (43). I suggest that words do not contain concepts but that

concepts are the output to processes that take a word as their input or the input to processes that have a word as their output. The relevant question to ask then is not: what are the conditions for *being* COFFEE but what are the conditions for *calling* something coffee. The answer to this last question is that there are no necessary and sufficient conditions to label something as *coffee* but what matters is that the label *coffee* is better at expressing the intended meaning than the other available labels.

3.3. *Beyond Metaphors*

The second question that needs to be answered with respect to the value of the metaphor research for lexical interpretation in general, is whether the findings for the interpretation of metaphors can be generalized to non-metaphoric contexts that give rise to conflicting features. I suspect it can. After all, since we are able to interpret metaphors (and sometimes as quickly as literal language) the interpretive system has to encompass the mechanisms necessary for doing so and it seems implausible that these mechanisms are specialized for and only used in the interpretation of metaphors. But of course, this question is really an empirical matter. I am therefore currently designing an experiment to test the mechanism of suppression for two more common constructions: coerced nouns like *stone lion* and "simple" noun adjective combinations whereby the adjective is in conflict with a central property of the noun, such as *rotten banana* (which is not yellow).

Although the abovementioned experiments should confirm this, the experiment on metaphors already suggest that the underspecification view in combination with an enrichment process as described by Blutner (2004) seems to be on the wrong track. Under this view, features such as "cat" for the word *cheetah* are never activated and therefore need not be suppressed. In contrast, the metaphor experiment seems to point at overspecification rather than underspecification. In this paper I therefore entertain the hypothesis that words come with much semantic information, some of which may be deleted in a final interpretation.

A model of interpretation should not only be in line with what is known about psycholinguistic processes but should also be formally explicit. If we want to assume rich lexical entries, how can a process of interpretation that is based on such representations be formalized? The experiment on metaphors already shows that when words are combined, their semantic information can be in conflict and that this conflict has to be resolved somehow. I hope to show with my experiments that this is not particular to metaphors. What is important to realize, I think, is that this resolution is not random but there is systematicity behind it. In the example *my cousin is such a mouse*, for example, it is the features of *mouse* that would be suppressed and not the features of *my cousin*. So, even if you want

to call this information world knowledge or encyclopedic knowledge, the integration of the pieces of information associated with the several words seems to be governed by some kind of rules or constraints. To model this, we first of all need to assume very rich lexical representations for example like in the Generative Lexicon (Pustejovsky, 1995; also see McNally, 2006). However, the lexical entries in the Generative Lexicon do not include all the features which are shown to be activated in the metaphor experiments and which I hope to show are activated in the interpretation of coerced nouns and other adjective noun combinations. So we need even richer lexical representations. And we also need to allow for stored features to be absent from a final interpretation. When combining words, the features associated with those words may be incompatible. In that case there is a conflict in interpretation between on the one hand being faithful to the associated features and on the other hand the wish to avoid contradictions in interpretation. Optimality Theory is a linguistic theory which is very well suited for handling such conflicts. Just as in phonology, OT semantics takes an input and delivers an optimal output based on a set of ranked constraints. In the next section I will briefly introduce the basic characteristics of OT after which I will illustrate the OT-resolution of conflicts that arise between semantic features when words are combined in Section 5.

4. Optimality Theory

4.1. *The Architecture of Optimality Theory*

Optimality Theory (OT) forms an important part of the Integrated Connectionist/Symbolic Cognitive Architecture (ISC) (Smolensky and Legendre, 2006). ICS is a framework that integrates lower level connectionist representations and higher level symbolic representation. By doing so, symbolic theorizing has benefited from insights at the lower level of description. One of the most important insights was that networks can settle into a stable state through the interaction of conflicting forces (Soderstrom, Mathis and Smolensky, 2006). Optimality Theory is based on this principle. In OT, linguistic knowledge is described as a system of ranked constraints. The constraints are ordered according to a strict priority ranking and they are potentially conflicting. A constraint may be violated to satisfy higher ranked constraints. OT hypothesizes that every language shares the same set of constraints. The difference between languages is due to a different ranking of those constraints.

OT specifies the relation between an input and an output. GEN (for generator) generates the possible output-candidates on the base of a given input. EVAL (for evaluator) evaluates the different candidates. The output that best satisfies the set of ranked constraints emerges as the optimal output for the given input. There are

two types of constraints: faithfulness and markedness constraints. Faithfulness constraints order the output to be faithful to the input. Markedness constraints are solely concerned with the output. They indicate that an unmarked output is preferred over a marked output. To put it briefly, structures that are more complex are considered to be marked structures and structures that are less complex or more natural are considered to be unmarked. Faithfulness to the input may sometimes require marked structures. Therefore, faithfulness and markedness constraints are potentially conflicting (Prince and Smolensky, 1993/2004).

Phonology was the first area in linguistics to which Optimality Theory was applied. In phonology, constraints pertain to the relation between underlying form and surface form. Later the theory was also applied to syntax (e.g. Legendre, Grimshaw and Vikner, 2001) and semantics (Hendriks and de Hoop, 2001; de Hoop and de Swart, 2000). In OT semantics the input is an utterance and the output is an interpretation of that utterance.

4.2. *Optimality Theory, Words and Concepts*

OT has been applied to the interpretation of words or to model the choice of words in several studies. Zeevat (2002) analyses the interpretation of various discourse markers in OT. Fong (2003) shows that in Colloquial Singaporean English, the use of the adverb *already* is the result of the interaction of markedness and faithfulness constraints. Zwarts (2004) gives an OT-analysis of the interpretation of the preposition *(a)round*. Furthermore, in Hogeweg (2009a, b) the interpretation of the Dutch discourse particle *wel* is modeled in Optimality Theory. Common in these approaches is that a word is assumed to correspond to a fixed set of features. When a speaker wants to express a meaning, she compares the features in the input (the meaning she wants to express) to the bundles of features expressible by the lexicon of her language. Similarly, when a hearer interprets a word, she interprets the features that are stored for this word, provided that they are not in conflict with the (linguistic) context.

The central question in this paper is what aspects of a word's meaning are stored in our mental lexicon and how we get from this stored information to an interpretation in a context. An OT approach to this question entails that words do not have a one-to-one relation with an interpretation but the relation between words and meanings is the result of a process of optimization. A word is the output to the process that has a meaning as the input (production) or an input to the process that has a meaning as output (interpretation). In OT lexical semantics the input and output are both meanings. What is compared in the optimization of the interpretation of words is the fixed set of semantic features associated with the lexical item that forms the input and candidate interpretations for the word, which also consist of semantic features. In production, the input is a meaning

the speaker wants to express and the candidate outputs are the sets of features that are conflated into words in the lexicon of the speaker. Hence, when I use the word *concept* I refer to a set of semantic features (or, as I will argue in the next section, attribute value structures). This contrasts with for example the use of the term *concept* by Osherson and Smith (1981) who argue that concepts that underlie kind terms such as *animal*, *tree*, *tool* are represented as a set of objects with information about the degree of prototypicality and the degree to which the object is characteristic for the concept. Osherson and Smith (1981) and later Kamp and Partee (1995) discuss the problems that arise with this view once one starts looking at combinations of concepts. The present proposal is not about combining *concepts*, but about combining *words*. Prototypicality may play a role in interpretation when there is no more specific information available but this does not mean that a less prototypical concept is less well described by a word. This means that there is no reason why a combination of *words* should lead to a combination of *prototypes*. For example, the word *red* can express a range of different colors. If a speaker wants to describe a person as having red hair, the speaker herself will start with a representation of a particular, probably non-prototypical, shade of red and a particular representation of hair. By a process of optimization she will choose the optimal words to describe this representation: *red hair*. The hearer will choose the optimal interpretation from the range of colors expressible with *red*. There is no reason why the speaker or hearer should necessarily pick the prototypical concept from the set of concepts a word can express. Therefore, we do not run into the problem of compositionality discussed in for example Kamp and Partee (1995) and Osherson and Smith (1981).

In the next section I will illustrate how a conflict that arises when two lexical representations are combined, is solved by a small set of constraints in OT.

5. Combining Rich Lexical Representations: The Case for *Stone Lion*

Adjective-noun combinations like *stone lion* involve (metonymic) type coercion. The noun *lion* is coerced from referring to real life lions to certain representations of lions. This has always been a problematic example because here the adjective seems to change the denotation of the noun, which is the head. Partee (2010) already hinted at an OT solution for phrases like this.

5.1. *Attributes and Values*

To model the phrase *stone lion*, we first need a way of representing the rich lexical meanings. Previous lexical semantic analyses in OT, which mainly addressed function words like discourse particles or prepositions, represented the

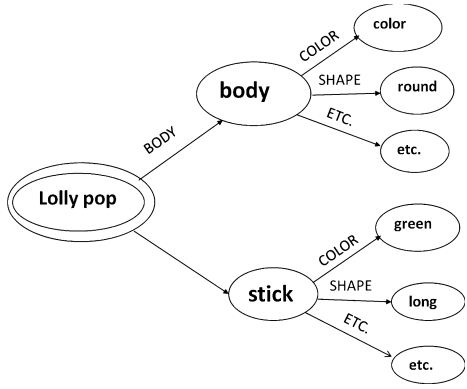

Figure 1. Lolly pop frame.

meaning of a word as a set of features (see section 4). If we translate this to content word like nouns, lexical knowledge could specify that for example an apple has a stem, pulp, a peel etc. It also specifies that apples, or to be more precise, the peel of an apple can be red or green. So, an apple could be represented as {pulp, stem, peel, red ∨ green}. However, this does not specify that it is the peel that is green or red and there is for example no way to indicate that the stem is most probably brown. In other words, there needs to be more structure in the set of features, in the form of recursive attribute value pairs. Barsalou (1992) argues that frames as recursive attribute-value structures provide the fundamental representation of human cognition. Rather than consisting of a feature set at a flat level of analysis, concepts are formed by attribute-value sets, with some characteristics (values) being instances of other characteristics (attributes). For example: "blue" and "green" are values of the attribute "color", "swim" and "fly" are values of "locomotion". Whereas the theory of frames as proposed by Barsalou was not formalized, nowadays formal theories of frames are available. Petersen (2007) for example, provides a formal theory of frames as connected directed rooted graphs with labeled nodes and arcs. The attributes in the frames assign unique values to concepts and are therefore functional relations. The values in frames can be atomic or composite (in which case the value is further specified by attributes of its own). Furthermore, they can be more or less specific. For example, the value of the attribute "color" can be "color" or "red". Frames come with appropriateness conditions restricting the possible values for an attribute. Appropriateness is determined based on a type signature and the condition that the possible values of an attribute are subtypes (e.g. "red", "blue", "indigo") of the type corresponding to the attribute ("color"). Concepts can also have more specific appropriate conditions. Since an apple is always round, for example, the appropriateness con-

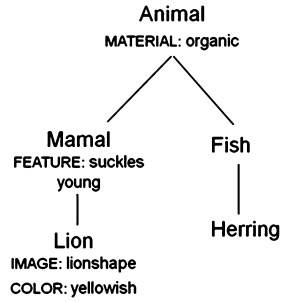

Figure 2. A simplified hierarchy for lion.

dition shape:shape for apple is tightened up to shape:round. The appropriate conditions in the type signature are ordered hierarchally. They are stored at the highest node and are inherited downwards. Figure 1 gives an example of a frame of a "lolly pop", based on Petersen (2007).

For the purpose of this article I will disregard most of the details of the formalization of frames and I will represent them as follows in (7):

(7) λx[body of (body, x) ∧ color of (green, body) ∧ shape of(round, body) ∧ taste of(apple, body) ∧ etc.]

Note that I do not include the head node ("lolly pop" in this case) because the head node is nothing but the combination of all its lower parts.

Say a lion is part of the (simplified) hierarchy in Figure (2). Like in the Generative Lexicon, it is indicated in the lexical representation how it relates the global lexical structure. I indicate this by giving the level from which the feature is inherited in subscript behind it, as in (8).

(8) $\lambda x[\text{animate}_2(x) \wedge \text{material of}_2(\text{organic}, x) \wedge \text{mammal}_1(x) \wedge \text{suckles young}_1(x) \wedge \text{color of}_0(\text{yellowish}, x) \wedge \text{image of}_0 (\text{lion image}, x) \text{ etc.}]$

For the adjective "stone" I assume the representation in (9):

(9) λx[stone(x)]

5.2. *The Constraints and Their Interaction*

Now that we have determined the two lexical representations, let me introduce the relevant constraints. The first constraint we need is a very general faithfulness constraints. This constraint is one of the key ingredients in OT (in all domains it is applied to: phonology, syntax, semantics and pragmatics) and it demands that

elements in the input must be present in the output. In our case the elements are semantic features.

> Faith: features in the input must be present in the output

The second important constraint is particular to OT semantics and was first introduced as Avoid Contradiction by Hendriks and de Hoop (2001) and later renamed as Fit by Zwarts (2005).

> Fit: Interpretations should not conflict with the (linguistic) context.

Fit is a markedness constraint that favors interpretations that do not conflict with the (linguistic) context over ones that do. If a possible interpretation does not fit the conversation or the context, according to this constraint it will not emerge as the optimal interpretation for the given utterance. In our case Fit is violated in three cases: 1: when there are two or more values for the same attribute; 2: when a value or attribute is of the wrong type as determined by a type hierarchy, for example the value "warm" for the attribute "color" or the attribute "color" for the type "smell", and 3: when a value of an attribute is in incongruence with an inherited value.

A third constraint is needed that makes sure that all words in a given utterance contribute to the overall interpretation. Such a constraint was proposed by Kamp and Partee (1995) as the Non-vacuity Principle (NVP): "In any given context, try to interpret any predicate so that both its positive and negative extension are non-empty" (161). This is reformulated as the following OT constraint:

> Non-vacuity Principle (NVP): all lexical entries in the input must at least contribute one feature to the output.[3]

I make one extra assumption, namely that the adjective only modifies the type denoted by the noun (properties at level 0). Adjectives cannot make changes inside the higher levels because information is only inherited downwards. If this were not the case, a statement about lions could lead to a reinterpretation of the whole class of mammals, for example. However, an adjective can cause, for example, a cancellation of whole node "animate" and all information that comes with it since this would not lead to inconsistencies in the representation of *lion*.

---

3) As Uli Sauerland (p.c.) suggested to me, the effect of this constraint could also be obtained by assuming a bidirectional Optimality Theoretic Framework. A bidirectional framework optimizes over forms and meaning simultaneously. In a situation where two structures lead to the same interpretation, the form-meaning pair containing the most economic form would come out as optimal. This would indeed lead to a situation where a form or interpretation would never be chosen if one of the words does not add to the meaning of the whole.

Tableau 1. Interpretation of *stone lion*.

	FIT	NVH	FAITH
λx [animate$_2$(x) ∧ material of$_2$(organic, x) ∧ mammal$_1$(x) ∧ suckles young$_1$(x) ∧ color of$_0$(yellowish, x) ∧ image of$_0$(lion image, x).] λx[stone(x)]			
λx [animate$_2$(x) ∧ material of$_2$ (**organic**, x) ∧ mammal$_1$(x) ∧ suckles young$_1$(x) ∧ color of$_0$(yellowish, x) ∧ image of$_0$(lion image, x) ∧ **stone**$_0$(x)]	*		
λx [animate$_2$(x) ∧ material of$_2$(organic, x) ∧ mammal$_1$(x) ∧ suckles young$_1$(x) ∧ color of$_0$(yellowish, x) ∧ image of$_0$(lion image, x)] (***stone* is not interpreted**)		*	
☞ λx [non-animate(x) ∧ color of$_0$(yellowish, x) ∧ image of$_0$(lion image, x) ∧ stone$_0$(x)]			****

This can be seen in Tableau 1. In the upper left box the input is given. In this case, the input consists of two representations, one of *lion* and one of *stone*. The two boxes below the input contain (three of) the possible outputs which are generated by GEN. Since the adjective only modifies the level at which the noun denotes (not its supertype), candidates in which the adjective overrides information at level one or two are not included. In the remainder of the upper row, the relevant constraints are given. The constraints are ordered according to their position in the hierarchy. An asterisk represents a violation of a constraint and the little hand points at the winner. Note that in OT, one violation of a higher ranked constraint is worse than any amount of violations of a lower ranked constraint.

If the adjective *stone* modifies *lion* with the properties "animate" and "mammal" intact, as in the first candidate, this leads to a violation of FIT because *lion* inherits the value "organic" for the attribute "material". Another option is to not to interpret *stone*, as in the second candidate, which leads to a violation of NVH. In the third candidate, the levels one and two are removed from the representation of *lion* due to which *stone* is not in conflict with the value "organic" anymore. This means that this particular representation expressed by the noun *lion*, is not placed in the hierarchy as a subtype of "animal" and "mammal", due to which it does not inherit the features specified for those types and due to which it is not interpreted as an instance of an animate entity. This candidate gives rise to the most violations of FAITH, but since this is the lowest ranked constraint it is still

the optimal candidate. The analysis I presented here is compatible with the analysis of privative adjectives and adjectives like *stone* in combinations such as *stone lion*, as subsective adjectives plus coercion by Partee (2010). The interaction of FIT and NVH causes the loss of a number of features in the representation of lion,[4] due to which the extension of *lion* shifts to include images or models of a lion.

In this section, I have shown how a conflict that arises when two rich lexical representations are combined is resolved by means of three constraints. As can be seen in Tableau 1, this leads to the preservation of some of the features while others have to be deleted from the final interpretation. For *stone lion* we get the desired interpretation of an artifact shaped like a lion. Admittedly, the details of this unification process need further elaboration. As a reviewer rightfully pointed out, a stone lion is not necessarily (and probably not even most likely) yellowish. A solution to this problem would involve a more fine grained representation of stone, involving information about the natural color of this material.

A question one might ask is: Where in the linguistic process does this optimization take place? Is it between a semantic level and a conceptual level? Or can the rich representations replace the semantic level altogether? In that case there would be a direct link between words (that is, their phonological representations) and conceptual structures.[5] I will go into this question in the next section.

6. Semantic and Encyclopedic Knowledge

Most formal semantic theories make a strict distinction between semantic knowledge and what is alternatively called world knowledge, common sense knowledge or encyclopedic knowledge. In the previous section, I illustrated how the combination of rich lexical structures leads to a coherent and detailed interpretation. It may seem very uneconomic to have such elaborate structures stored. However, we may safely assume that people have detailed conceptual informa-

---

4) Partee (2010) suggests that the shift in interpretation could be due to an OT-like interaction between the NVP and the *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" (p. 6). FIT in Tableau 1 has an effect similar to the HPP because it forces the representation of the adjective to attach to one of the appropriate slots in the representation of the noun.

5) Such a direct link is also argued for by Jackendoff in his Conceptual Semantics (1996, 2011).

tion about entities and events in the world and if it can be shown that this conceptual information behaves systematically and can explain phenomena for which otherwise extra semantic knowledge has to be assumed, it may eventually be the most economical system to have only a conceptual level.

Consider for example the difference between intersective and subsective adjectives. Intersective adjectives are adjectives that, when combined with a noun, lead to an intersection of the set that is the denotation of the adjective and the set that is the denotation of the noun. An adjective like *skilful* is not intersective for a skilful surgeon playing the piano is not necessarily a skilful pianist. *Skilful* is a subsective adjective because for every N it holds that "skilful N" ⊆ "N".[6] Intersectivity and subsectivity are usually considered to be a property of the adjective that is stored independently, for example as a meaning postulate (as in Partee 2010). However, if we allow for rich interpretations, as I have been arguing for, the difference between intersective and subsective adjectives may be the result of their meaning in combination with the meaning of the noun they are combined with rather than a semantic feature that is independently stored.

This idea entails an elaboration of the selective binding principle by Pustejovsky (1995), which was discussed previously in Section 2. Some examples of intersective adjectives given by Partee (2007) are *sick, carnivorous, blond, rectangular*, and *French*. Examples of subsective adjectives are *recent, good, perfect*. The subsective adjectives usually modify an activity in the typical examples they occur in, e.g. a good basketball player is good at playing basketball, a good surgeon is good at performing surgery. In Pustejovsky's analysis, the adjective *good* in for example *a good knife* modifies the telic role "cut" in the qualia structure of the noun. The meaning of the noun *knife* can be represented as follows λx[... Telic= λeλy[cut(e, x, y)]]. The adjective *good* modifies the event "cut" in the telic quale of the representation (Pustejovski, 1995: 129). Because the adjective *good* comes with a restriction that it is bound to an event, this leads to subsective interpretation when combined with a noun that has an event as part of its qualia structure. In the present proposal this effect can be ascribed to Fit, which only allows values to be bound to an appropriate attribute. The embedding of an adjective like *good*, x, in the representation of a noun A leads to a subsective interpretation because the event e_a that is part of the representation of A may be different from the event e_b in the representation of a noun B in a phrase xB, even though A and B have the same referent.

---

6) Adjectives like former and alleged are neither intersective nor subsective because a former senator is not a member of the set of senators. Nonsubsective adjectives may be plain nonsubsective (no entailments at all, such as *alleged*, *arguably* etc.) or privative (entailing a negation of the noun property, such as *past*, *would-be* etc.) (see for example Partee, 2007).

If we have even more elaborate lexical structures than in the Generative Lexicon, the combination of attributes of the noun and the restrictions of the attribute would cause a situation where the intersectivity or subsectivity is the result of the process of combining two lexical representations. Say we have an intersective adjective like *carnivorous* and a noun *pianoplayer*. The adjective *carnivorous* can only be attached to an attribute, say, "diet". The more elaborate representation of *piano player* will not have an attribute "diet" linked to the node "piano" nor will it be linked to the event node. Rather this attribute will be an attribute of the head node (the person) due to which the adjective can only be interpreted intersectively: if an adjective y modifies the complete concept expressed by a noun A and by a noun B, and if A and B have the same referent, then if yA is true, yB is also true.

If intersectivity or subsectivity is indeed the result of a combination of lexical representations, we would expect there to be adjectives that behave intersectively in some combinations and subsectively in others. Let us look at *French*, for example, which is mentioned as an example of an intersective adjective in Partee (2007). Indeed a French piano player is also a French athlete if *piano player* and *athlete* refer to the same person.[7] However, in *French restaurant*, *French* pertains to the type of food that is served rather than the nationality of the building or organization. It is therefore selectively bound an should be seen as a subsective: a French restaurant is not necessarily a French enterprise.

In this section I asked the question whether the rich lexical interpretations including what is usually considered encyclopedic or common sense knowledge exist in addition to or instead of a (underspecified) semantic level. Although this matter is far from settled, I think that if we can show that conceptual knowledge behaves systematically and is able to explain phenomena for which otherwise extra semantic knowledge has to be assumed, it may eventually be most economic to have only a conceptual level of representation.

7. Conclusion

In this paper I have tried to find an answer to an often neglected question: how do fine grained interpretations come about? Ideally, the answer to this question should be both formally explicit as well as in line with what is known about

7) As a reviewer pointed out to me, this may not even necessarily be the case. If a Frenchman (who also plays the piano) takes on the Belgian nationality because that makes it easier to qualify for the Olympics, perhaps one can speak of him as a Belgian athlete but a French piano player.

psycholinguistic processes. Research on the interpretation of metaphors suggests that upon hearing a word we initially activate a set of features, some of which may be deleted if they are incompatible with the context. This seems to suggest that an underspecification approach in combination with an enrichment process as proposed by Blutner (2004) cannot be right. I proposed that, instead, we start with a very rich representation from which features may be deleted. If two representations are combined this may lead to conflicting information. This conflict can be resolved in Optimality Theory by means of an interaction of faithfulness and markedness constraints. In Section 5, I gave an illustration of such an optimization procedure by providing an OT-analysis of the interpretation of *stone lion*.

A fine grained interpretation necessarily makes use of conceptual knowledge. If it turns out that this knowledge leads to compositional interpretation and no extra semantic knowledge has to be assumed, a direct link between word forms and conceptual information may eventually be the most economical organization of the lexicon.

It goes without saying that this paper does not come close to a complete answer to the question. Further research should confirm that the interpretation of other word combinations indeed involves the suppression of semantic features as well. Furthermore, more theoretical analyses should provide evidence that rich lexical representations can lead to a complete formal theory of interpretation.

References

Anderson, Richard C. & Andrew Ortony. 1975. On putting apples into bottles: A problem of polysemy. *Cognitive Psychology, 7*: 167–180.

Barsalou, Lawrence W. 1992. Frames, concepts, and conceptual fields. In E. Kittay & A. Lehrer (eds.), *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*. Hillsdale, NJ: Erlbaum.

Bierwisch, Manfred. 1983. Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In W. Motsch & R. Ruzicka (eds.), *Untersuchungen zur Semantik*. Berlin: Akademie Verlag.

Bierwisch, Manfred & Rob Schreuder. 1992. From concepts to lexical items. *Cognition* 42: 23–60.

Blutner, Reinhard. 2004. Pragmatics and the lexicon. In L.R. Horn & G. Ward (eds.) *Handbook of Pragmatics*, 488–514. Oxford: Blackwell.

Blutner, Reinhard. 1998. Lexical underspecification and pragmatics. In: P. Ludewig & B. Geurts (eds.), *Lexikalische Semantik aus Kognitiver Sicht*. Tübingen, 141–171. Gunter Narr Verlag.

Blutner, Reinhard. 2011. Some Perspectives on Lexical Pragmatics. In Archer, Dawn and Grundy, Peter (Eds.), *Pragmatics Reader*. London: Routledge.

Carston, Robyn. 2010. *XIII-Metaphor: Ad hoc concepts, literal meaning and mental images*. Proceedings of the Aristotelian Society, 110: 295–321.

Clark, Herbert H. 1983. Making sense of nonce sense. In GB Flores d'Arcais and R. Jarvella (Eds.), *The Process of Understanding Language*, 297–332. New York: Wiley. Reprinted in Clark, H.H. 1992. *Arenas of Language Use*, 305–340. Chicago: University of Chicago Press.

Fodor, Jerry & Zenon Pylyshyn. 1988. Connectionism and cognitive architecture: A critical analysis. *Cognition* 28: 3–71.

Fodor, Jerry & Ernest Lepore. 2002. *The compositionality papers*. Oxford, Clarendon Press.

Fodor, Jerry & Brian McLaughlin. 1991. Connectionism and the problem of systematicity: why smolensky's solution doesn't work. In Horgan, T. and Tienson, J. (eds.), *Connectionism and the Philosophy of Mind*, 331–355. Cambridge, MIT press/Bradford books

Fong, Vivienne. 2005. Unmarked *already*: Aspectual expressions in two varieties of English. In: H.J. Verkuyl, H. de Swart & A. van Hout (eds.), *Perspectives on Aspect*, 251–267. Dordrecht, Springer.

Giora, Rachel. 2008. Is metaphor unique? In: R. Gibbs, (ed.) *The Cambridge Handbook of Metaphor and Thought*, 143–160. New York: Cambridge University Press.

Glucksberg, Sam. 2003. The psycholinguistics of metaphor. *Trends in Cognitive Sciences* 7:92–96.

Hendriks, Petra. & Helen de Hoop. 2001. Optimality theoretic semantics. *Linguistics and Philosophy* 24: 1–32.

Hogeweg, Lotte. 2009a. The meaning and interpretation of the Dutch particle *wel*. *Journal of Pragmatics*, 41.3: 519–539.

Hogeweg, Lotte. 2009b. *Word in process. On the interpretation, acquisition and production of words*. PhD dissertation, Radboud University Nijmegen.

Hoop, Helen de & Henriette de Swart. 2000. Temporal adjunct clauses in Optimality Theory. *Rivista di Linguistica* 12.1: 107–127.

Jackendoff, Ray. 1996. Conceptual Semantics and Cognitive Semantics. *Cognitive Linguistics* 7: 93–129.

Jackendoff, Ray. 2011. Conceptual semantics. In K. von Heusinger & P. Portner (eds.) *Semantics: An International Handbook of Natural Language Meaning*, Vol. 1, 688–709. De Gruyter Mouton.

Kamp, Hans & Barbara Partee. 1995. Prototype theory and compositionality. *Cognition* 57: 129–191.

Legendre, Geraldine, Jane Grimshaw & Sten Vikner (eds.). 2001. *Optimality-theoretic Syntax*. MIT Press.

McNally, Louise. 2006. Lexical representation and modification within the noun phrase. *Recherches Linguistiques de Vincennes*, 34: 191–206.

Osherson, Daniel N. & Edward E. Smith 1981. On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9: 35–58.

Partee, Barbara. 2007. Compositionality and coercion in semantics: The dynamics of adjective meaning. In: G. Bouma et al. (eds.): *Cognitive Foundations of Interpretation*, 145–161. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Partee, Barbara 2010. Privative adjectives: subsective plus coercion. In R. Bäuerle, U. Reyle & T.E. Zimmermann (Eds.), *Presuppositions and Discourse. Essays offered to Hans Kamp*, 273–285. Amsterdam: Elsevier.

Petersen, Wiebke. 2007. Representation of Concepts as Frames. In *The Baltic International Yearbook of Cognition, Logic and Communication* 2: 151–170.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Technical Report, Rutgers University and University of Colorado at Boulder. Revised version published by Blackwell, 2004.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MIT Press.

Rubio Fernández, Paula. 2007. Suppression in metaphor interpretation: Differences between meaning selection and meaning construction. *Journal of Semantics* 24.4: 345–371.

Rubio-Fernández, Paula, Richard Breheny & Ming Wei Lee. 2003. Context-independent information in concepts: an investigation of the notion of core features. In F. Schmalhofer, R. Young, & G. Katz (Eds.), *Proceedings of the European Cognitive Science Conference (EuroCogSci) 2003*, 277–282. Mahwah, NJ: Lawrence Erlbaum.

Smolensky, Paul. 1991. Connectionism, constituency and the language of thought. In Loewer, M. & Rey, G. (eds.), *Meaning in Mind: Fodor and His Critics*, 201–227. Oxford, Blackwell.

Smolensky, Paul & Geraldine Legendre. 2006. *The Harmonic Mind: From Neural Computation To Optimality-Theoretic Grammar Vol. 1: Cognitive Architecture; vol. 2: Linguistic and Philosophical Implications*. Cambridge, MIT Press.

Soderstrom, Melanie, Donald Mathis & Paul Smolensky. 2006. Abstract genomic encoding of Univeral Grammar in Optimality Theory. In: P. Smolensky & G. Legendre (eds.) *The Harmonic Mind: From Neural Computation To Optimality-Theoretic Grammar. Vol 2: Linguistic and Philosophical Implications*, 403–471. Cambridge, MIT Press.

Wilson, Deirdre. 2003. Relevance and lexical pragmatics. *Italian Journal of Linguistics/Rivista di Linguistica*, 15.2: 273–291

Zeevat, Henk. 2002. Explaining presupposition triggers. In K. van Deemter & R. Kibble (eds.), *Information Sharing*: 61–87. Stanford, CSLI Publications.

Zwarts, Joost. 2004. Competition between word meanings: the polysemy of (a)round. In Meier, C. andWeisgerber, M. (eds.), *Proceedings of SuB8*, 349–360. Konstanz, University of Konstanz Linguistics Working Papers.