Optimality-Theoretic Pragmatics

Reinhard Blutner
ILLC, University of Amsterdam

Henk Zeevat
ILLC, University of Amsterdam

The article aims to give an overview about the application of Optimality Theory (OT) to the domain of pragmatics. In the introductory part we discuss different ways to view the division of labor between semantics and pragmatics. Rejecting the doctrine of literal meaning we conform to (i) semantic underdetermination and (ii) contextualism (the idea that the mechanism of pragmatic interpretation is crucial both for determining what the speaker says and what he means). Taking the assumptions (i) and (ii) as essential requisites for a natural theory of pragmatic interpretation, section 2 introduces the three main views conforming to these assumptions: Relevance theory, Levinson's theory of *presumptive meanings*, and the Neo-Gricean approach. In section 3 we explain the general paradigm of OT and the idea of bidirectional optimization. We show how the idea of optimal interpretation can be used to restructure the core ideas of these three different approaches. Further, we argue that bidirectional OT has the potential to account both for the synchronic and the diachronic perspective on pragmatic interpretation. Section 4 lists relevant examples of using the framework of bidirectional optimization in the domain of pragmatics. Section 5 provides some general conclusions. Modeling both for the synchronic and the diachronic perspective on pragmatics opens the way for a deeper understanding of the idea of naturalization and (cultural) embodiment in the context of natural language interpretation.

1   Introduction

Optimality Theory is an integrated approach to cognition that combines the advantages of symbolic, constraint-based models with the advantages of subsymbolic, neuron-style models of cognition (cf. Smolensky & Legendre, 2006). In the study of natural language, OT was successfully applied to the main linguistic disciplines phonology, morphology and syntax, and also to the

explanation of natural language acquisition and other performance traits. OT pragmatics is an application of the integrated approach to the domain of Gricean pragmatics. It has its origin in the attempt to explain certain phenomena of lexical pragmatics (Blutner, 1998) and is inspired by the optimal interpretation approach proposed by Hendriks & de Hoop (2001).

The view of seeing OT pragmatics within the scope of a naturalistic (explanatory) approach to cognition (as represented by the main proponents of OT) is not without problems. This has to do with the normative character that is attributed to the Gricean setting. Speakers, as Grice puts it, must

> *make their contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which (they) are engaged.* (Grice, 1975: 45)

It's obvious that this principle of cooperation is normative, and so are Grice's conversational maxims. If a person acts in a particular situation in a particular way we can ask *why* she did it the way she did; alternately, we can ask if it was *reasonable* what the person did, and if other options were possibly more reasonable in the given situation. Good Griceans are expected to ask the second type of questions whereas the first question is expected to be asked by cognitive scientists. While the normative and the naturalistic aspects of understanding human actions can be clearly separated from each other that does not mean they predict different action patterns in most cases. The idea of a rational world isn't so irrational to be excluded in ordinary affairs. Evolutionary game theory has presented us with many examples demonstrating that the reasonable is naturally arising (Axelrod, 1984). In other words, though there is a philosophical gap between Gricean pragmatics as a normative theory and OT as a scientific, explanatory theory of natural language there is not a deep empirical conflict between an interpretation oriented pragmatics and a speaker ethics. It seems the speaker better be cooperative or pretend to be cooperative if she wants to use language to bring about effects in hearers.

The naturalistic stance taken by OT pragmatics is one characteristic that brings it close to Relevance Theory (Sperber & Wilson, 1986/1995; Sperber & Wilson, 1995). Another point of agreement has to do with the way OT pragmatics views the division of labor between semantics and pragmatics. Taking the lead from Atlas (e.g. Atlas, 2005), both relevance theory (RT) and OT pragmatics reject the doctrine of literal meaning. And both approaches conform to the ideas of

(i)   semantic underdetermination

(ii)   contextualism (the suggestion that the mechanism of pragmatic interpretation is crucial both for determining what the speaker says and what he means).

In the broad view of OT, this framework can be seen as a general scheme that can be used for expressing many different and possibly diverging views. For instance, it is possible to give optimality-theoretic reconstructions of a speaker-oriented normative pragmatics like the one developed by Grice. It is also possible to reconstruct hearer-oriented naturalistic pragmatics as in RT (Hendriks & de Hoop, 2001; Zeevat, 2007b). These systems are important for online synchronic accounts of speaking and interpretation. But – perhaps most surprisingly – it is also possible to reconstruct the Neo-Gricean systems of Horn (1984), Atlas & Levinson (1981) and Levinson (2000). In contrast to RT where there is only one fundamental pragmatic principle (the presumption of optimal relevance), the Neo-Gricean systems have two opposing optimization principles, the Q- and the I-principle (Atlas and Levinson 1981, Horn 1984 who writes R instead of I) by two simultaneous optimization directions (the speaker and the hearer direction) and so obtain a bidirectional OT pragmatics. OT pragmatics in the narrower sense will start from this system and will show that Levinson's M-principle (iconicity) can be reduced to it. The system can also explain the emergence of mono-directional pragmatic systems that can account for online incremental interpretation in the style of RT. Given the divergences within the Neo-Gricean camp[1], it cannot be expected that a coherent theory like bidirectional OT-pragmatics can reconstruct all the views of all representatives of this camp.

The present chapter aims to give an overview of the application of OT to the domain of pragmatics. The assumptions (i) and (ii) are essential requisites for a natural theory of pragmatic interpretation. In section 2 we will introduce the three main views conforming to these assumptions: (a) RT, (b) Levinson's (2000) theory of *presumptive meanings*, and (c) the Neo-Gricean approach. In section 3 we explain the general paradigm of OT and the idea of bidirectional optimization. We show how the idea of optimal interpretation can be used to restructure the core ideas of these three different approaches. Further, we argue that bidirectional OT has the potential to account both for the synchronic and the diachronic perspective of pragmatic interpretation. Section 4 lists relevant examples of using the framework of bidirectional optimization in the domain of pragmatics. Section 5 provides some general conclusions. It argues that OT pragmatics has the potential to account both for the synchronic and the

---

[1] For instance, Horn (2005) points out that Levinson's (2000) view is very close in important respects to that of RT.

diachronic perspective in pragmatics. This bolsters the way for a deeper understanding of the idea of naturalization and (cultural) embodiment in the context of natural language interpretation.

2 The naturalization of pragmatics: three variations on Grice

The *naturalization of pragmatics* refers to a research program that aims to provide a cognitively realistic picture of utterance interpretation and production. Hence, the proponents of this program such as relevance theorists take the stance of seeing natural language interpretation as a cognitive phenomenon and thus considering the basic principles of communication as a consequence of the nature of human cognition. A prerequisite of this program deals with the levels of cognitive representations and the boundary between semantics and pragmatics. There is a strong tendency among current researchers to follow the tradition of radical pragmatics and to accept the following three claims:

1. There is a level of logical form or semantic representation. The representations of this level do not necessarily provide truth conditions. Rather, they underspecify truth-conditional content in a number of ways.

2. There is a mechanism of enriching underspecified representations; sometimes this mechanism is called development of logical form. The result of this development is propositional content. It expresses the utterance meaning of the expression under discussion.

3. There is a level of implicatures proper, understood as separate thoughts implied by the utterance. It is implicit propositional content that can be inferred from the explicit content mentioned in 2.

Obviously, the consensus is about rejecting the Gricean doctrine of literal meaning (logical form conforms to literal meaning), accepting the role of underspecification (logical forms are underspecified with regard to the expressed semantic content) and acknowledging that implicature is a graded category (some implicatures are closer to LF than others). Obviously, this view sharply contrasts with the paradigm of Generative Semantics – a view that tries to ground pragmatic phenomena by using particular syntactic stipulations.

Before we come to a discussion of three variations on Grice and the naturalization of conversational implicatures in utterance interpretation it is useful first to introduce the distinction between *global* and *local* approaches to conversational implicatures (cf. Chierchia, 2004). According to the global (Neo-Gricean) view one first computes the (plain) meaning of the sentences; then, taking into account the relevant alternatives, one strengthens that meaning by

adding in the implicature.' (Chierchia 2004: 42). This contrasts with the local view, which first introduces pragmatic assumptions locally and then projects them upwards in a strictly compositional way where certain filter conditions apply. Representatives of the global view are Atlas & Levinson (1981), Gazdar (1979), (Horn, 1984), Soames (1982), Krifka (1995), Blutner (1998), Sauerland (2004), and Sæbø (2004); the local view is taken by Chierchia (2004), Levinson (2000), and Relevance Theory (Sperber & Wilson, 1986/1995; e.g. Carston, 2002).

Usually, the globalists argue against the local view and the localists against the global view. We will argue, instead, that proper variants of both views are justified if a different status is assigned to the two views: global theories provide the standards of rational discourse and correspond to a diachronic, evolutionary scenario; local theories account for the shape of actual, online processing including the peculiarities of incremental interpretation. In this way, we will argue that seemingly conflicting approaches such as relevance theory and the neo-Gricean approach are much closer related than expected by its opponents. In section 3, once more OT will prove his power of unification in giving hints how to relate theses different frameworks in a systematic way.

RT assumes the representational/computational view of the mind, and, on this basis, gives a naturalization of pragmatics adopting Jerry Fodor's language of thought hypothesis (Fodor, 1975). The central thesis of RT is the Communicative Principle of Relevance, according to which utterances convey a presumption of their own optimal relevance. In other words, any given utterance can be presumed:

(i)   to be at least relevant enough to warrant the addressee's processing effort
(ii)   to be the most relevant one compatible with the speaker's current state of knowledge and her personal preferences and goals.

From these two assumptions relevance theorists derive the following general procedure that the cognitive system follows in comprehending an utterance (cf. Sperber, Cara, & Girotto, 1995: 95): (a) test possible interpretations in their order of accessibility, (b) stop once the expectation of (optimal) relevance is satisfied (i.e. a certain context-dependent threshold value of relevance is reached). The procedure makes sure that the wanted effect (a certain value of relevance) is reached with the minimal cognitive effort.

Levinson's (2000) theory of *presumptive meaning* is a chameleon that in a certain sense adapts general assumptions of RT and in another sense crucially conflicts with RT, for instance in assuming more than one basic principle (*maxim*) for formulating the interpretational mechanism. In short, these are the general assumptions:

(i) Differing from both RT and the standard neo-Gricean view, Levinson assumes *three* levels of meaning corresponding to sentence meaning, utterance-type meaning and utterance-token meaning

(ii) utterance-type meanings are in correspondence with Grice' generalized conversational implicatures. They are a matter of preferred interpretation calculated by a particular default mechanism. Basically, there are three such defaults or heuristics:
- Q-heuristic: *What isn't said is not the case*
- I-heuristic: *What is expressed simply is stereotypically exemplified*
- M-heuristic: *What's said in an abnormal way isn't normal*

(iii) In contrast to Grice' generalized conversational implicatures, which are calculated in a global manner, presumptive meanings are local, i.e. they arise at the point at which they are triggered (for instance, the word *some* triggers the default interpretation NOT ALL via the Q-heuristics). The feature of local pragmatics is essential to artificial intelligence pragmatics (e.g. Hobbs & Martin, 1987) and likewise to RT.

Presumptive meanings are very useful for understanding natural language interpretation, especially for explaining the predominantly incremental character of utterance comprehension.

Neo-Griceans (Atlas & Levinson, 1981; Horn, 1984; Blutner, 1998; e.g. Atlas, 2005; Horn, 2005) are assuming two countervailing optimization principles: the Q-principle and the R-principle.[2] The first is oriented to the interests of the hearer and looks for optimal interpretations; the second is oriented to the interests of the speaker and looks for expressive optimization. Here is a standard presentation of the two principles (cf. Horn, 1984, 1989, 2004, 2005):

The Q-Principle (Hearer-based):
  Make your contribution sufficient!
  Say as much as you can! (modulo R)
  (Grice's first quantity maxim and the first two manner maxims)

The R-Principle (Speaker-based):
  Make your contribution necessary!
  Say not more than you must! (modulo Q)

---

[2] In OT, these 'principles' correspond to different directions of optimization where the *content* of the optimization procedure is expressed by particular OT constraints. This will be pointed out in more detail in the following section.

(Grice's second quantity maxim, relation maxim and the second two manner maxims)

It is tempting to identify the Q-principle with Levinson's Q-heuristic and the R-principle with the I-heuristics. However, they are not identical though there is a correspondence between them. The difference has to do with the different status of *principles* in the global, neo-Gricean pragmatics on the one hand and *heuristics* (*defaults*) in Levinson's local pragmatics on the other hand. According to the neo-Gricean picture the principles constitute a kind of communication game – either between real speakers and hearers or between fictive speakers and hearers in the mind of a language user. In this game both principles are applied in a recursive way (corresponding to the modulo-clause in the formulation of the principles). In Levinson's theory, no such interaction between real or fictive Speakers/Hearers takes place. Instead, presumptive meanings are default interpretations and they are processed in a nearly automatic way. No 'mind reading' facilities or other mechanisms of controlled processing are required.[3] The difference will become quite clear in the following section when we give formalization in terms of bidirectional OT.

Sometimes it is stressed that there is a fundamental difference in perspective and goals between the neo-Gricean and the RT approaches to pragmatics. For instance, Horn (2005) claims the following:

> *Grice's goal of developing an account of speaker meaning (of which implicature constitutes a proper subpart) is distinct from Relevance theorists' goal of developing a cognitive psychological model of utterance interpretation, which does not address the question of how and why the speaker, given what she wants to convey, utters what she utters.* (194).

This seems to express the difference between the naturalistic stance and the normative stance mentioned in section 1. However, we agree with Carston (2005) that this statement is too strong as it stands since RT (as does Horn's theory) makes some predictions about **why** the speaker, given her communicative intention, utters what she utters. Further, the difference between the normative stance and the naturalistic stance should not be overestimated because in practice there is seldom a deep empirical conflict between the two

---

[3] However, presumptive meanings can demand a lot of effort as soon 'conflicts' arise and the corresponding assumption has to be cancelled. Conflict resolution can be very resource demanding. Hence, for the overall mechanism we have to take into account the peculiarities of controlled processing. Of course, this does not refer to any mind reading facilities.

stances (Spohn, 1993). A much more important question concerns the status of the theory with regard to synchrony versus diachrony. Both RT and Levinson's theory of *presumptive meaning* takes the synchronic view where neo-Griceans take both views (and, sometimes, confuse them). In the following section we will see how OT relates both views/perspectives.

3   The framework of OT

OT can be seen as a general framework that systematizes the use of optimization methods in linguistics.[4] One component of OT is a list of tendencies that hold for observable properties of a language. These tendencies take the form of violable constraints. Because the constraints usually express very general statements, they can be in conflict. Conflicts among constraints are resolved because the constraints differ in strength. Minimal violations of the constraints (taking their strength into account) define optimal conflict resolutions.

   Standardly, OT specifies the relation between an input and an output. This relation is mediated by two formal mechanisms, **GEN** and **EVAL**. GEN (for Generator) creates possible output candidates on the basis of a given input. EVAL (for Evaluator) uses the particular constraint ranking of the universal set of constraints **CON** to select the best candidate for a given input from among the candidate set produced by GEN. In phonology and syntax, the input to this process of optimization is an underlying linguistic representation. The output is the (surface) form as it is expressed. Hence, what is normally used in phonology and syntax is unidirectional optimization. Obviously, the point of view of the speaker is taken. This contrasts with OT semantics where the view of the hearer is taken (de Hoop & de Swart, 2000; Hendriks & de Hoop, 2001).

   Bidirectional optimization (Blutner, 1998, 2000) integrates the speaker and the hearer perspective into a simultaneous optimization procedure. In pragmatics, this bidirectional view is motivated by a reduction of Grice's maxims of conversation to two principles: the R-principle, which can be seen as the force of unification minimizing the Speaker's effort, and the Q-principle, which can be seen as the force of diversification minimizing the Auditor's effort. In a slightly different formulation, the R-principle seeks to select the most coherent interpretation[5] and the Q-principle acts as a blocking mechanism which blocks all the outputs which can be expressed more economically by an alternative linguistic input. This formulation makes it quite clear that the neo-

---

[4] A recent overview is given in Smolensky & Legendre (2006). For OT pragmatics the reader is referred to Blutner & Zeevat (2004) and Blutner, de Hoop & Hendriks (2005).

[5] What is meant by coherence has to be expressed by particular OT constraints, such as formulated by Zeevat (2007a, 2007b) for instance.

Gricean framework can be conceived of as a bidirectional optimality framework which integrates the speaker and the hearer perspective. Whereas the R-principle compares different possible interpretations for the same syntactic expression, the Q-principle compares different possible syntactic expressions that the speaker could have used to communicate the same meaning.

We will give a very schematic example in order to illustrate some characteristics of the bidirectional OT. Assume that we have two forms f_1 and f_2 which are semantically equivalent. This means that **GEN** associates the same interpretations with them, say m_1 and m_2. We stipulate that the form f_1 is less complex (less marked) than the form f_2 and that the interpretation m_1 is less complex (less marked) than the interpretation m_2. This is expressed by two markedness constraints: F for forms and M for interpretations – F prefers f_1 over f_2 and M prefers m_1 over m_2. This is indicated by the two leftmost constraints in table (1).

Table 1: Markedness and bias constraints in a 2-forms × 2-interpretations design

	F	M	F→M	*F→*M	F→*M	F*→M
$<f_1, m_1>$					*	
$<f_1, m_2>$		*	*			
$<f_2, m_1>$	*			*		
$<f_2, m_2>$	*	*				*

Besides the markedness constraints, four so-called linking constraints can be formulated. There are precisely four independent linking constraints in the present example. The linking constraint F→M says that simple (unmarked) forms express simple interpretations. Hence, this is a straightforward formalization of Levinson's (2000) **I**-heuristics as an OT constraint. The constraint *F→*M says that complex forms express complex interpretations, and this is an expression of Levinson's **M**-heuristics[6]. The two remaining bias constraints express the opposite restrictions. In the present case linking constraints can be seen as lexical stipulations that fix a form-interpretation relation in an instance-based way. With only two forms and two meanings, the
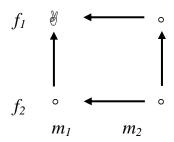
---

[6] Levinson's M-principle should not be confused with the markedness constraint M introduced in Table 1.

substance of the Q-heuristics is not really different from that of the M-constraint.[7]

In the so-called strong version of bidirectional OT, a form-interpretation pair $<f, m>$ is considered to be (strongly) optimal iff

- Interpretive Optimization: no other pair $<f, m'>$ can be generated that satisfies the constraints better than $<f, m>$ and

- Expressive Optimization: no other pair $<f', m>$ can be generated that satisfies the constraints better than $<f, m>$.

From the differences of markedness given by the constraints F and M the ordering relation between form-meaning pairs can be derived as shown in Figure 1. The preferences are indicated by arrows in a two-dimensional diagram. Such diagrams give an intuitive visualization for the optimal pairs of (strong) bidirectional OT: they are simply the meeting points of horizontal and vertical arrows.[8] The optimal pairs are marked with the symbol ✌ in the diagram.

Figure 1: Diagram to illustrate strong bidirection

The scenario just mentioned describes the case of *total blocking* where some forms (e.g., *furiosity, *fallacity) do not exist because others do (fury, fallacy). However, blocking is not always total but may be partial. This means that not all the interpretations of a form must be blocked if another form exists. McCawley (1978) collects a number of examples demonstrating the phenomenon of *partial blocking*. For example, he observes that the distribution of productive causatives (in English, Japanese, German, and other languages) is

---

[7] Harmonic alignment (Prince & Smolensky 1993, Aissen 2003) is precisely the fact that these two linking constraints hold.

[8] Dekker & van Rooy (2000), who introduced these diagrams, gave bidirectional OT a game theoretic interpretation where the optimal pairs can be characterized as so-called Nash Equilibria.
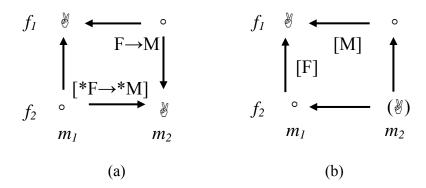
restricted by the existence of a corresponding lexical causative. Whereas lexical causatives (e.g. (1a)) tend to be restricted in their distribution to the stereotypical causative situation (direct, unmediated causation through physical action), productive (periphrastic) causatives tend to pick up more marked situations of mediated, indirect causation. For example, (1b) could have been used appropriately when Black Bart caused the sheriff's gun to backfire by stuffing it with cotton.

(1) a. Black Bart killed the sheriff
    b. Black Bart caused the sheriff to die

To make things concrete we can take f_1 to be the lexical causative form (1a), f_2 the periphrastic form (1b), m_1 direct (stereotypic) causation and m_2 indirect causation.

Typical cases of partial blocking are found in morphology, syntax and semantics. The general tendency of partial blocking seems to be that "unmarked forms tend to be used for unmarked situations and marked forms for marked situations" (Horn 1984: 26) – a tendency that Horn (1984: 22) terms *"the division of pragmatic labour"*.

There are two ways of avoiding total blocking within the bidirectional OT framework and to describe Horn's division of pragmatic labour. The first possibility makes use of linking constraints and fits the intended form-interpretation relation by stipulating the appropriate ranking of the constraints such that partial blocking comes out. Let's assume that the two bias-constraints $F{\rightarrow}M$ and $*F{\rightarrow}*M$ are higher ranked than the rest of the constraints. This can be depictured as in Figure 2a. Hence, strong bidirection can be taken as describing Horn's division of pragmatic labour when the appropriate linking constraints are dominating.

f_1 ☝ ← ◦
$F{\to}M$
$[{*}F{\to}{*}M]$
f_2 ◦ → ✌
m_1 m_2

f_1 ☝ ← ◦
$[M]$
$[F]$
f_2 ◦ ← (✌)
m_1 m_2

(a)   (b)

Figure 2: Two ways of describing Horn's division of pragmatic labour: (a) by assuming two dominant bias constraints; (b) by assuming markedness constraints and weak bidirection

The second possibility is to weaken the notion of (strong) optimality in a way that allows us to derive Horn's division of pragmatic labour by means of the *evaluation procedure* and without stipulating particular bias constraints. Blutner (2000) proposes a *weak* version of two-dimensional OT, according to which the two dimensions of optimization are mutually related:

Super-Optimality

A form-interpretation pair $<f, m>$ is called super-optimal iff
- Interpretive Optimization: no other super-optimal pair $<f, m'>$ can be generated that satisfies the constraints better than $<f, m>$;
- Expressive Optimization: no other super-optimal pair $<f', m>$ can be generated that satisfies the constraints better than $<f, m>$.

This formulation looks like a circular definition, but Jäger (2002) has shown that this is a sound *recursive* definition under very general conditions (well-foundedness of the ordering relation). The important difference between the weak and strong notions of optimality is that the weak one accepts super-optimal form-meaning pairs that would not be optimal according to the strong version. It typically allows marked expressions to have an optimal interpretation, although both the expression and the situations they describe have a more efficient counterpart.

Figure 2b shows that the *weak* version of bidirection can explain the effects of partial blocking without the stipulation of extra bias constraints; especially it can explain why the marked form f_2 gets the marked interpretation m_2. This is a

consequence of the *recursion* implemented in weak bidirection:[9] the pairs $\langle f_1, m_2 \rangle$ and $\langle f_2, m_1 \rangle$ are *not* super-optimal. Hence, they cannot block the pair $\langle f_2, m_2 \rangle$ and it comes out as a new super-optimal pair. In this way, the weak version accounts for Horn's pattern of *the division of pragmatic labour*.

The two parts of Figure 2 describe the same set of solution pairs but the calculation of the solutions is completely different in the two cases. In the first case unidirectional optimization (either hearer or speaker perspective) is sufficient to calculate the solution pairs. It is plausible to assume that this kind of OT systems can be used to construct cognitively realistic models of online, incremental interpretation (cf. Blutner 2006). The second case – using the recursion of weak bidirection (super-optimality) – has a completely different status. Because of its strictly non-local nature the proposed algorithm that calculate the super-optimal solutions do not even fit the simplest requirements of a psychologically realistic model of online, incremental interpretation (Zeevat, 2000; Beaver & Lee, 2004)[10]. The proper understanding of weak bidirection relates best to an off-line mechanism that is based on bidirectional learning (Blutner, Borra, Lentz, Uijlings, & Zevenhuijzen, 2002; Benz, 2003b; Van Rooy, 2004). In these approaches the solution concept of weak bidirection is considered as a principle describing the results of language change: super-optimal pairs emerge over time in language change. This relates to the view of Horn (1984) who considers the Q and the I principle as diametrically opposed forces in language change. This conforms to the good old idea that synchronic structure is significantly informed by diachronic forces.

For the sake of illustration let's go back to our example illustrated in (1). Let's assume a population of agents who realize speaker- and hearer strategies based exclusively on the markedness constraints F and M. In this population each content is expressed in the simplest way (f_1) and each expression is understood in the simplest way (m_1). Let's assume further that these agents communicate with each other. When agent x is in the speaker role and intends to express m_1, then expressive optimization yields f_1. Agent y is a hearer who receives f_1 and, according to interpretive optimization, he gets the interpretation m_1 – hence the hearer understands what the speaker intends: successful communication. Now assume the speaker wants to express m_2. With the same

---

[9] In the original formulation given in section 2, the recursion is indicated by the modulo-clause.

[10] There are several arguments why bidirectional OT cannot yield an online mechanism of linguistic competence. Beaver & Lee (2004) argue that if more rounds of optimization are allowed, the bidirectional OT-model severely overgenerates in the sense that in later rounds peculiar new form-meaning pairs will emerge as winners. Before the Beaver & Lee paper, Zeevat (2000) argued against the symmetric view of OT pragmatics starting from the famous rat/rad problem and its pragmatic counterparts.

logic of optimization he will produce f_1 and the agent y interprets it as m_1. In this case, obviously, the communication is not successful. Now assume some kind of *adaptation* either by iterated learning or by some mutations of the ranked constraint system (including the bias constraints). According to this adaptation mechanism the expected 'utility' (how well they understand each other in the statistical mean) can improve in time. In that way a system that is evolving in time can be described including its special attractor dynamics. In each case there is a stabilizing final state that corresponds to the system of Figure 2a where the two Levinsonian (2000) constraints **I** (= [F→M]) and **M** (= [F→M]) outrank the rest of the constraints. It is precisely this system that reflects Horn's division of pragmatic labour. The only condition we have to assume is that the marked contents are less frequently expressed than the unmarked contents.[11]

Hence, the important insight is that a system that is exclusively based on markedness constraints such as in Figure 2b is evolutionary related to a system based on highly ranked bias constraints such as in Figure 2a. We will use the term *fossilization* for describing the relevant transfer.[12]

Now we come back to the earlier goal of giving an OT reconstruction of the three variations on Grice (section 2). For reconstructing Levinson's (2000) presumptive meaning theory, unidirectional optimization is sufficient where a system of OT constraints has to be formulated conforming to his I, Q and M heuristics and Levinson's putative ranking Q > M > I. The unidirectional optimization procedure (interpretive optimization) is conform with a local approach to conversational implicatures, one which satisfies the requirements of incremental interpretation.

The neo-Gricean approach, on the other hand, is globalist in nature. Hence, the idea of (weak) bidirectional optimization fits best to this theory and can be used for a straightforward formalization. Unsurprisingly, this conception can be seen best from a diachronic perspective, at least so far we take a naturalistic stance towards Gricean pragmatics. As a model of actual language interpretation (or production) this approach does not make real sense and never was designed for this purpose.

Like Levinson's (2000) approach, RT conforms to the localist approach and can be formulated in terms of unidirectional optimization. Let's stipulate a constraint EFFECT for describing the wanted effect (a certain value of relevance) and a constraint EFFORT for describing the cognitive effort. Then the stipulation

---

[11] For more discussion of the role of frequencies in an evolutionary setting cf. Stalnaker (2006). The general conclusion is that the solution concept of weak bidirection can be seen as a rough first approximation to the more adequate solution concepts of evolutionary game theory that describe the results of language change.

[12] In a somewhat different context, Peter Cole (Cole, 1975) calls it "lexicalization of contextual meaning".

EFFECT > EFFORT makes sure that the wanted effect is reached with the minimal cognitive effort. Obviously, there are many questions left concerning the concrete content of the constraints EFFECT and EFFORT, the RT literature contains a number of specifications. These specifications typically have the character of linking constraints. It might be interesting to investigate recent OT models of pragmatics (see section 4) in the light of the general structure of RT – a task that goes beyond what can be done in the present paper.[13]

We have mentioned already that there is a relation between diachronic and synchronic systems, and we have introduced the term *fossilization* for describing the relevant transfer.[14] Taken the existence of this transfer, it can be demonstrated that the three discussed variations on Grice are much closer related than the occasional polemics let us expect.

In order to get an impression of OT pragmatics at work we shortly will discuss Zeevat's (2002, 2007a, 2007b) reconstruction of presupposition theory as formulated by Van der Sandt (1992) and Heim (1983). In both these theories there are two defaults or preferences. The first one prefers identifying the induced presupposition in the context of the utterance (resolution), the second one prefers the addition of the presupposition to the global context (Heim) or to the highest accessible context where that is possible (Van der Sandt). The reconstruction makes use of the following constraints that are used in finding an optimal interpretation: FAITH > CONSISTENCE > DO NOT ACCOMMODATE > STRENGTH. CONSISTENCE demands that there is no conflict of the current utterance with what is known already, FAITH asks for the presence of the presupposed information at an accessible position, DO NOT ACCOMMODATE forbids the addition of the presupposed information and STRENGTH forbids interpretations if there are informationally stronger ones available. The OT system improves in several ways on the theories that it tries to reconstruct. DO NOT ACCOMMODATE prefers partial resolutions to full accommodations and does not militate against copies of presuppositions. STRENGTH often gives a better

---

[13] Another important aspect concerns pragmatic acceptability. The RT account of pragmatic (un)acceptability is carefully worked out in connection with bridging phenomena (Wilson & Matsui, 1998). In RT, "unacceptability can result from (a) inadequate effects or (b) gratuitous effort" ((Wilson & Matsui, 1998: 19). That means there have to be thresholds for (a) effects and (b) effort, and when these thresholds are reached unacceptability results. This kind of argumentation is not possible within an OT approach because in OT the absolute strength of constraint violation is not of importance. What counts is the comparison with other expressions that lead to the same interpretation and the possibility of blocking an interpretation by a cheaper expression alternatives (e.g. Blutner, 1998). It's an open issue what are the empirical consequences of this view in case of bridging.

[14] Concerning the debate whether certain pragmatic inferences are truly conversational or whether they have become lexically encoded the reader is referred to Cole (1975) and Potts (to appear).

prediction of the accommodation site than van der Sandt (1992). Zeevat (2002) uses the reconstruction to explore particles like "too". These particles have an exceptional behaviour within accounts of presupposition: they do not allow (full) accommodation and are obligatory in the contexts in which they occur. The second phenomenon needs a max(F) constraint as in OT phonology: certain relations to the context need to be marked. But the other phenomenon seems to allow a bidirectional explanation. DO NOT ACCOMMODATE in a bidirectional interpretation prohibits candidate expressions that force accommodations, if there is a simple alternative that means the same and does not force the accommodation. For particles, the sentence without the particle always is an alternative of this kind.

4   Some applications of bidirectional OT in pragmatics

OT pragmatics has been used for describing a series of phenomena and observation in the domain of natural language interpretation. This section gives an overview of some of these applications without going into any technical or empirical details.

- *Centering theory*. (Beaver, 2004) is using bidirectional OT as framework for the reformulation of centering models of pronoun resolution.

- *Disambiguation*. Gärtner (2004b, 2004a) analyses Icelandic object-shift and differential marking of (in-)definite**s** in Tagalog addressing the issue of disambiguation and partial iconicity in natural language.

- *Differential object marking*: Aissen (2003), Nilsenova (2002), and Jäger & Zeevat (2002) discuss the relevant correlation between grammatical functions and semantic/pragmatic properties.

- *Binding theory*. Mattausch (2004a, 2004b) introduces the influential work of Levinson on the origin and typology of binding theory (summarized in Levinson, 2000) and reformulates the different historical stages assumed by Levinson in bidirectional optimality theory. Mattausch's work is of essential importance as one of the first in-depth studies showing the importance of the diachronic view for bidirectional OT. For early work on discourse anaphora in a bidirectional framework we refer to Buchwald, Schwartz, Seidl & Smolensky (2002).

- *Pragmatics for propositional attitudes*. Aloni (2001, 2005b) has argued that a number of seeming paradoxes emerging from logical analyses of attitudes and questions can be explained in terms of shifts in perspective over the universe of discourse. Shifts in perspective have a cost and, therefore, are generally avoided. However, under certain circumstances such shifts are

required to comply with general principles of rational conversation, which, for example, disallow vacuous or inconsistent interpretations. Aloni's work suggests a formulation of a perspective selection procedure in the framework of bidirectional OT.

- *Discourse particles and presupposition.* Zeevat (2002, 2004) treats discourse particles within an extended OT reconstruction of presupposition theory and concludes that more particles can be treated and the analysis becomes simpler if one starts from the fact that discourse particles are obligatory if the context of utterance and the current utterance stand in one of a number of special relations, like adversativity, additivity, contrast, etc. In another paper, Zeevat (2007a) provides a full solution to the projection problem for presuppositions. Jäger & Blutner (2000, 2003) suggest an bidirectional analysis of the different reading of German 'wieder' (again).

- *Complex implicatures.* Blutner (2007) gives an OT account of implicature projection and explains the relevance theoretic distinction between implicatures and explicatures in terms of a neo-Gricean framework.

- *Interpretation of stress and focus.* Several articles deal with a bidirectional perspective for stress on anaphoric pronouns and the interpretation of focus (Beaver, 2004; de Hoop, 2004; Hendriks, 2004; Aloni, Butler, & Hindsill, 2007)

- *Marking and Interpretation of negation.* Henriëtte de Swart (2004) provides a bi-directional OT approach to the syntax and pragmatics of negation and negative indefinites (see also, de Swart, in press).

- *Scalar implicatures and exhaustification.* Exhaustivity implicatures and also scalar implicatures depend on the issue under discussion which can be formalized using Groenendijk & Stokhof's theory of question and answers. Combining this framework with those of bidirectional OT, Aloni (2005b, 2005a) explains several puzzles in this area.

- *Permission sentences.* A series of other articles deals with the interpretations of permission sentences and the analysis of the particular conditions which constitute a so-called free choice interpretation (Sæbø, 2004; Aloni, 2005a, 2005b; Blutner, 2006).

- *Stage level/individual level contrast*: Maienborn (2004, 2005) argues against the popular view that the distinction between *stage level predicates and individual level predicats* rests on a fundamental cognitive division of the world that is reflected in the grammar. Instead, Maienborn proposes a pragmatic explanation of the distinction, and she gives, inter alias, a discourse-based account of Spanish *ser*/*estar*.

- *Aspectual interpretation of the Dutch past tenses*. Van Hout (2007) applied bidirectional reasoning about tense forms and their aspectual meanings.

- *Lexical pragmatics*: Lexical Pragmatics investigates the processes by which linguistically-specified ('literal') word meanings are modified in use. Well-studied examples include narrowing (e.g. drink used to mean 'alcoholic drink'), approximation (e.g. square used to mean 'squarish') and metaphorical extension (e.g. battleaxe used to mean 'frightening person'). Lexical Pragmatics can be formulated by using the formal instruments of OT-based pragmatics (Blutner, 1998; Blutner et al., 2005). Prototypical applications include the pragmatics of dimensional adjectives (Blutner & Solstad, 2000), the analysis of Dutch om/rond (Zwarts, 2006), the pragmatics of negated antonyms (Blutner, 2004; Krifka, to appear), the approximate interpretation of number words (Krifka, 2007), several examples of semantic change (Eckardt, 2002).

- *Language acquisition and learning*: There are several studies that test the role of weak bidirection in developing interpretation and production preferences in connection with indefinite NPs (deHoop & Kramer, 2005/2006) and pronominal anaphors (Hendriks & Spenader, 2005/2006; Hendriks, Rijn, & Valkenier, 2007; Mattausch & Gülzow, 2007). From a theoretical perspective, the problem of learning is investigated by Benz (2003a, 2003b).

5   Conclusions

The error in many formulations of pragmatic inferences is that synchrony and diachrony are confused. OT pragmatics accounts both for the synchronic perspective – by formulating a localist, incremental model based on unidirectional optimization using a emerging system of linking constraints – and the diachronic perspective – using the solution concept of weak bidirection which conforms to a strictly global view. The perspectives are connected by the idea of fossilization.

Many patterns in language have been proposed to be directly or indirectly influenced by the conflict between multiple influences on output form. Within phonology for example, the notion that conflict between minimization of articulatory effort and maximization of perceptual distinctiveness has an influence on grammatical patterns has held currency at least since Baudouin de Courtenay (1895). Contemporary work grounding phonological patterns in optimization of conflicting influences on output form include work done within Natural Phonology (Stampe, 1973), Grounded Phonology (Archangeli & Pulleyblank, 1994), and Optimality Theory (Prince & Smolensky, 1993/2004; Boersma, 1998) to name but a few. Weak bidirection describes the interaction of these forces in an approximate but simply to understand way. However, for fully understanding the bidirectional game that leads to the resolution of the conflict

between the opposite forces, evolutionary game theory provides a more adequate model (e.g. Van Rooy, 2004).

Relevance Theory and Levinson's theory of presumptive meaning account for the resolution of the conflict between effort minimization and effect maximization in different ways. In a certain sense, the crux of both approaches can be translated in OT pragmatics by making use of particular linking constraints. This translation makes the advantage of both approaches visible: both conform to the incremental, online character of natural language interpretation.[15]

We have argued that OT pragmatics has the potential to account both for the synchronic and the diachronic perspective in pragmatics, and for the way both are related to each other. We further have pointed out that the concepts of fossilization can help to understand the idea of naturalization and (cultural) embodiment in the context of natural language interpretation. However, there are also important open questions regarding the status of fossilization. In a by now classical paper Cole (1975) considered the following example of a true conversational implicature, where a girl called Pamela upon being asked (2) might reply (3):

(2) How are you doing in your new position at San Andreas Fault University?
(3) Well, I haven't been fired yet.

Although the logical content of (3) is roughly that of the proposition that Pamela has not yet lost her job, more than that is implicated, namely that Pamela is not doing well. In this example, the implicature is really novel. There is no construction involved whose frequent use could lead to the fossilization phenomenon (Cole's term is 'lexicalization'). Hence, this implicature is different from many other cases where a certain amount of fossilization is plausible. The important question is how to discriminate between offline implicatures that are not fossilized and their fossilized counterparts. Where is the boundary between aspects of interpretations that are truly conversational and aspects which have become lexically (or syntactically) encoded? We think the former aspect of interpretation can require some real mind reading capacities,

---

[15] In discussing processing characteristics, incrementality and automaticity of processing have to be discriminated. Whereas automaticity of processing implicates the incremental character of processing the opposite is not true: incrementality does not implicate automatic processing. RT explains the incremental character of processing and has good reasons for assuming controlled processing in order to account for the processing of conversational implicatures. That's different from Levinson's account which assumes automatic processing for generalized conversational implicatures. It seems that RT is better justified on empirical grounds (cf. Noveck & Sperber, 2005).

requires conscious reflections and proceeds offline. So far we can see, none of the discussed pragmatic theories has an interesting answer for this long-standing and intriguing question.

6   References

Aissen, J. (2003). Differential coding, partial blocking, and bidirectional OT. In P. Nowak & C. Yoquelet (Eds.), *Proceedings of the 29th Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.

Aloni, M. (2001). Pragmatics for Propositional Attitudes. In R. v. Rooy & M. Stokhof (Eds.), *Proceedings of the Thirteenth Amsterdam Colloquium*. Amsterdam: University of Amsterdam.

Aloni, M. (2005a). *Expressing ignorance or indifference. Modal implicatures in BiOT* Paper presented at the Batumi 2005 symposium Batumi.

Aloni, M. (2005b). A Formal Treatment of the Pragmatics of Questions and Attitudes. *Linguistics and Philosophy, 28*, 505-539.

Aloni, M., Butler, A., & Hindsill, D. (2007). Nuclear Accent in Bidirectional Optimality Theory. In M. Aloni & A. Butler & P. Dekker (Eds.), *Questions in Dynamic Semantics*: Elsevier.

Archangeli, D. B., & Pulleyblank, D. (1994). *Grounded Phonology*. Cambridge MA: MIT Press.

Atlas, J. D. (2005). *Logic, Meaning, and Conversation: Semantical Underdeterminacy, Implicature, and Their Interface*. Oxford: Oxford University Press.

Atlas, J. D., & Levinson, S. C. (1981). It-clefts, informativeness and logical form. In P. Cole (Ed.), *Radical Pragmatics* (pp. 1-61). New York: Academic Press.

Axelrod, R. (1984). *The evolution of co-operation*. London: Penguin.

Baudouin de Courtenay, J. (1895). *Versuch einer Theorie phonetischer Alternationen: Ein Capitel aus der Psychophonetik*. Strassburg/Crakow: Trubner.

Beaver, D. (2004). The optimization of discourse anaphora. *Linguistics and Philosophy, 27*, 3-56.

Beaver, D., & Lee, H. (2004). Input-output mismatches in OT. In Palgrave/Macmillan (Ed.), *Optimality Theory and Pragmatics*. Houndmills, Basingstoke, Hampshire.

Benz, A. (2003a). *Partial blocking and associative learning*. Unpublished manuscript, Berlin & Kolding. Available from http://www.anton-benz.de/Paper.html.

Benz, A. (2003b). Partial Blocking, associative learning, and the principle of weak optimality. In J. Spenader & A. Eriksson & Ö. Dahl (Eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory* (pp. 150-159). Stockholm.

Blutner, R. (1998). Lexical pragmatics. *Journal of Semantics, 15*, 115-162.

Blutner, R. (2000). Some aspects of optimality in natural language interpretation. *Journal of Semantics, 17*, 189-216.

Blutner, R. (2004). Pragmatics and the lexicon. In L. Horn & G. Ward (Eds.), *Handbook of pragmatics*. Oxford: Blackwell.

Blutner, R. (2006). Embedded implicatures and optimality theoretic pragmatics. In Torgim Solstad & A. Grønn & D. Haug (Eds.), *A Festschrift for Kjell Johan Sæbø: in partial fulfilment of the requirements for the celebration of his 50th birthday*. Oslo.

Blutner, R. (2007). Optimality Theoretic Pragmatics and the Explicature/Implicature Distinction. In N. Burton-Roberts (Ed.), *Pragmatics* (pp. 67-89). Houndmills, Basingstoke, Hampshire: Palgrave/MacMillan.

Blutner, R., Borra, E., Lentz, T., Uijlings, A., & Zevenhuijzen, R. (2002). Signalling games: Hoe evolutie optimale strategieen selecteert, *Handelingen van de 24ste Nederlands-Vlaamse Filosofiedag*. Amsterdam: Universiteit van Amsterdam.

Blutner, R., de Hoop, H., & Hendriks, P. (2005). *Optimal Communication*. Stanford: CSLI Publications.

Blutner, R., & Solstad, T. (2000). Dimensional designation: a case study in lexical pragmatics. In R. Blutner & G. Jäger (Eds.), *Studies in Optimality Theory* (pp. 30-40). Potsdam: University of Potsdam.

Blutner, R., & Zeevat, H. (Eds.). (2004). *Optimality Theory and Pragmatics*. Houndmills, Basingstoke, Hampshire: Palgrave/Macmillan.

Boersma, P. (1998). *Functional phonology*. The Hague: Holland Academic Graphics.

Buchwald, A., Schwartz, O., Seidl, A., & Smolensky, P. (2002). Optimality Theory: Discourse Anaphora in a Bidirectional framework citeseer.ist.psu.edu/buchwald02optimality.html.

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Carston, R. (2005). Relevance Theory, Grice and the neo-Griceans: a response to Laurence Horn's 'Current issues in neo-Gricean pragmatics'. *Intercultural Pragmatics, 2*, 303-319.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and Beyond* (pp. 39-103). Oxford: Oxford University Press.

Cole, P. (1975). The synchronic and diachronic status of conversational implicature. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics, Volume 3: Speech Acts* (pp. 257-288). San Diego, Cal.: Academic Press, Inc.

de Hoop, H. (2004). On the interpretation of stressed pronouns. In R. Blutner & H. Zeevat (Eds.), *Optimality Theory and Pragmatics*. Houndmills, Basingstoke, Hampshire: Palgrave/Macmillan.

de Hoop, H., & de Swart, H. (2000). Temporal adjunct clauses in optimality theory. *Rivista di Linguistica, 12*, 107-127.

de Swart, H. (2004). Marking and Interpretation of negation: A bi-directional OT approach. In R. Zanuttini & H. Campos & E. Herburger & P. Portner (Eds.), *Negation, Tense and Clausal Architecture: Cross-linguistic Investigations*: Georgetown University Press.

de Swart, H. (in press). *Expression and interpretation of negation* (Vol. 77). Dordrecht: Springer.

de Hoop, H., & Kramer, I. (2005/2006). Children's Optimal Interpretations of Indefinite Subjects and Objects. *Language Acquisition, 13*, 103-123.

Dekker, P., & van Rooy, R. (2000). Bi-directional optimality theory: An application of game theory. *Journal of Semantics, 17*, 217-242.

Eckardt, R. (2002). Semantic Change in Grammaticalization. In G. Katz & S. Reinhard & P. Reuter (Eds.), *Proceedings of the Sixth Annual Meeting of the Gesellschaft für Semantik, Sinn & Bedeutung VI*. Osnabrück: University of Osnabrück.

Fodor, J. (1975). *The Language of Thought*. New York: Thomas Crowell.

Gärtner, H.-M. (2004a). On Object-Shift in Icelandic and Partial Iconicity. *Lingua, 114*, 1235-1252.

Gärtner, H.-M. (2004b). On the OT-status of 'unambiguous encoding'. In R. Blutner & H. Zeevat (Eds.), *Optimality Theory and Pragmatics*. Houndmills, Basingstoke, Hampshire: Palgrave/Macmillan.

Gazdar, G. (1979). *Pragmatics*. New York: Academic Press.

Grice, P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics, 3: Speech Acts* (pp. 41-58). New York: Academic Press.

Heim, I. (1983). File change semantics and the familiarity theory of definiteness, *Meaning, Use and the Interpretation of Language* (pp. 164-190). Berlin: Walter de Gruyter.

Hendriks, P. (2004). Optimization in Focus Identification, *Optimality Theory and Pragmatics*. Houndmills, Basingstoke, Hampshire: Palgrave/Macmillan.

Hendriks, P., & de Hoop, H. (2001). Optimality theoretic semantics. *Linguistics and Philosophy, 24*, 1-32.

Hendriks, P., Rijn, H. v., & Valkenier, B. (2007). Learning to reason about speakers' alternatives in sentence comprehension: A computational account. *Lingua, 117*, 1879–1896.

Hendriks, P., & Spenader, J. (2005/2006). When production precedes comprehension: An optimization approach to the acquisition of pronouns. *Language Acquisition, 13*, 319-348.

Hobbs, J., & Martin, P. (1987). Local Pragmatics, *Proceedings, International Joint Conference on Artificial Intelligence* (pp. 520-523). Milan.

Horn, L. (1984). Towards a new taxonomy of pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form, and use in context: Linguistic applications* (pp. 11-42). Washington: Georgetown University Press.

Horn, L. (1989). *A natural history of negation*. Chicago: Chicago University Press.

Horn, L. (2004). Implicature. In L. Horn & G. Ward (Eds.), *Handbook of Pragmatics*. Oxford: Blackwell.

Horn, L. (2005). Current issues in neo-Gricean pragmatics. *Intercultural Pragmatics, 2*, 191-204.

Jäger, G. (2002). Some notes on the formal properties of bidirectional optimality theory. *Journal of Logic, Language and Information, 11*, 427-451.

Jäger, G., & Blutner, R. (2000). Against lexical decomposition in syntax. In A. Z. Wyner (Ed.), *Proceedings of the Fifteenth Annual Conference, IATL 7* (pp. 113-137). Haifa: University of Haifa.

Jäger, G., & Blutner, R. (2003). Competition and interpretation: The German adverb wieder ("again"). In C. Fabricius-Hansen & E. Lang & C. Maienborn (Eds.), *Handbook of Adjuncts* (pp. 393-416). Berlin: Mouton de Gruyter.

Jäger, G., & Zeevat, H. (2002). A reinterpretation of syntactic alignment. In D. d. Jongh & H. Zeevat & M. Nilsenova (Eds.), *Proceedings of the 3rd and 4th International Symposium on Language, Logic and Computation*. Amsterdam ILLC.

Krifka, M. (1995). The Semantics and Pragmatics of Polarity Items. *Linguistic Analysis, 25*, 209-257.

Krifka, M. (2007). Approximate interpretation of number words: A case for strategic communication. In G. Bouma & I. Krämer & J. Zwarts (Eds.), *Cognitive foundations of interpretation* (pp. 111-126). Amsterdam: Koninklijke Nederlandse Akademie van Wetenschapen.

Krifka, M. (to appear). Negated antonyms: Creating and filling the gap. In U. Sauerland & P. Stateva (Eds.), *NN*.

Levinson, S. (2000). *Presumptive meaning: The theory of generalized conversational implicature*. Cambridge, Mass.: MIT Press.

Maienborn, C. (2004). A pragmatic explanation of the stage level/individual level contrast in combination with locatives. In B. Agbayani & V. Samiian & B. Tucker (Eds.), *Proceedings of the Western Conference on Linguistics (WECOL), volume 15* (pp. 158 - 170). Fresno: CSU.

Maienborn, C. (2005). A discourse-based account of Spanish ser/estar. *Linguistics, 43*, 155-180.

Mattausch, J. (2004a). *On the Optimization & Grammaticalization of Anaphora*. Unpublished Ph.D. Thesis, Humboldt University, Berlin.

Mattausch, J. (2004b). Optimality Theoretic Pragmatics and Binding Phenomena. In R. Blutner & H. Zeevat (Eds.), *Optimality Theory and Pragmatics*. Houndmills, Basingstoke, Hampshire: Palgrave/Macmillan.

Mattausch, J., & Gülzow, I. (2007). A note on acquisition in frequency-based accounts of Binding Phenomena. In I. Gülzow & N. Gagarina (Eds.), *Frequency Effects in Language Acquisition: Defining the Limits of Frequency as an Explanatory Concept* (pp. 331-357). Berlin. New York: Mouton de Gruyter.

McCawley, J. D. (1978). Conversational implicature and the lexicon. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics* (pp. 245-259). New York: Academic Press.

Nilsenova, M. (2002). The Pragmatics of Differential Object Marking. In M. Nissim (Ed.), *Proceedings of the ESSLLI '02 Student Session*. Trento, Italy: University of Trento, revised version available from http://staff.science.uva.nl/~mnilseno/.

Noveck, I. A., & Sperber, D. (Eds.). (2005). *Experimental Pragmatics*. Houndmills, Basingstoke, Hampshire: Palgrave MacMillan.

Potts, C. (to appear). Into the conventional-implicature dimension. *Philosophy Compass*.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Rutgers University and University of Colorado at Boulder: Technical Report RuCCSTR-2, available as ROA 537-0802. Revised version published by Blackwell, 2004.

Sæbø, K. J. (2004). Optimal interpretations of permission sentences. In D. de Jongh & P. Dekker (Eds.), *Proceedings of the 5th Tbilisi Symposium on Language, Logic and Computation* (pp. 137-144). Amsterdam and Tbilisi.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy, 27*, 367-391.

Smolensky, P., & Legendre, G. (2006). *The Harmonic Mind: From neural computation to optimality-theoretic grammar*. Cambridge, Mass.: MIT Press.

Soames, S. (1982). How presuppositions are inherited: A solution to the projection problem. *Linguistic Inquiry, 13*, 483-545.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57*, 31-95.

Sperber, D., & Wilson, D. (1986/1995). *Relevance*. Oxford: Basil Blackwell.

Sperber, D., & Wilson, D. (1995). "Postface" to the second edition of Relevance, *Relevance*. Oxford: Blackwell.

Spohn, W. (1993). Wie kann die Theorie der Rationalität normativ und empirisch zugleich sein? In L. Eckensberger & U. Gähde (Eds.), *Ethik und Empirie. Zum Zusammenspiel von begrifflicher Analyse und erfahrungswissenschaftlicher Forschung in der Ethik* (pp. 151-196). Frankfurt a.M.: Suhrkamp.

Stalnaker, R. (2006). Saying and meaning, cheap talk and credibility. In A. Benz & G. Jäger & R. Van Rooij (Eds.), *Game Theory and Pragmatics* (pp. 82-101): Palgrave MacMillan.

Stampe, D. (1973). *A Dissertation on Natural Phonology*. Bloomington.: Distributed 1979 by Indiana University Linguistics Club.

van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics, 9*, 333-377.

van Hout, A. (2007). Optimal and non-optimal interpretations in the acquisition of Dutch past tenses. *Proceedings of GALANA, 2*, 159-170.

Van Rooy, R. (2004). Signalling games select Horn strategies. *Linguistics and Philosophy, 27*, 493-527.

Wilson, D., & Matsui, T. (1998). Recent approaches to bridging: Truth, coherence, relevance. *UCL Working Papers in Linguistics 10*

Zeevat, H. (2000). The asymmetry of optimality theoretic syntax and semantics. *Journal of Semantics, 17*, 243-262.

Zeevat, H. (2002). Explaining presupposition triggers. In K. van Deemter & R. Kibble (Eds.), *Information Sharing* (pp. 61-87). Stanford: CSLI Publications.

Zeevat, H. (2004). Presupposition Triggers, Context Markers, or Speech Act Markers. In R. Blutner & H. Zeevat (Eds.), *Optimality Theory and Pragmatics*. Houndmills, Basingstoke, Hampshire: Palgrave/Macmillan.

Zeevat, H. (2007a). A Full Solution to the Projection Problem for Presuppositions: University of Amsterdam.

Zeevat, H. (2007b). *Optimal Interpretation as an Alternative to Gricean Pragmatics* Unpublished manuscript, Universiteit van Amsterdam.

Zwarts, J. (2006). Om en rond: Een semantische vergelijking. *Nederlandse Taalkunde, 11*, 101-123.