6

TWO CASE STUDIES IN LEXICAL PRAGMATICS

Reinhard Blutner, Humboldt University, Berlin, Germany
Torgrim Solstad, University of Oslo, Oslo, Norway

1 INTRODUCTION

Lexical Pragmatics is a particular account of the division of labor between lexical semantics and pragmatics (e.g. Blutner 1998). It combines the idea of (radical) semantic underspecification in the lexicon with a theory of pragmatic strengthening (based on conversational implicatures). In the core of this approach is a precise treatment of Atlas & Levinson's (1981) Q- and I-principles and the formalization of the balance between informativeness and efficiency in natural language processing (Horn's 1984 division of pragmatic labor). In a roughly simplified formulation, the I-principle seeks to select the most coherent interpretation, and the Q-principle acts as a blocking mechanism which blocks all the outputs which can be grasped more economically by an alternative linguistic input. Recently, these mechanisms have been implemented within a bidirectional version of optimality theory (OT) which aims to integrate expressive and interpretive optimization (Blutner 1999).

The aim of this paper is to apply this framework to two different kinds of examples. First, we want to provide a concise treatment of the phenomenon of *negative strengthening* as arising in connection with gradable adjectives. Second, we want to resolve some puzzles of *dimensional designation* of spatial objects. In the first case, the optimality theoretic treatment can be seen as a reformulation and revivification of earlier approaches by Horn (1989) and Levinson (2000). In the second case the treatment is really new and crucially deviates from earlier approaches.

The paper is organized as follows. Section 2 introduces the *bidirectional optimality framework* and illustrates how the Gricean framework of conversational implicature can be reformulated by means of this technique. In Section 3 we give a concise introduction to the phenomenon of negative strengthening, and in the subsequent

section we account for the basic phenomena by using the bidirectional optimality framework. In section 5 some puzzles of *dimensional designation* of spatial objects are outlined. Finally, in section 6, the bidirectional optimality framework is applied to solve these puzzles, and in section 7 some preliminary conclusions are drawn.

2 BIDIRECTIONAL OT AND PREFERENCES FOR INTERPRETATION

Recently, de Hoop & de Swart (1998), Hendriks & de Hoop (to appear), and de Hoop (2000) have applied OT to sentence interpretation. They argue that there is a fundamental difference between the form of OT as used in phonology, morphology and syntax on the one hand and its form as used in semantics on the other hand. Whereas in the former case OT takes the point of view of the speaker (expressive perspective), in the latter case the point of view of the hearer is taken (interpretive perspective).

One obvious reason for this difference is that ambiguity, polysemy, and other forms of flexibility are much more obvious and manifested much broader in the area of interpretation than in the realm of syntax. The assumption that OT in sentence interpretation takes the point of view of the hearer is mainly motivated by this observation. Using the interpretive perspective, a mechanism for preferred interpretations is constituted that provides insights into different phenomena of interpretations, such as the determination of quantificational structure (Hendriks & de Hoop, to appear), nominal and temporal anaphorization (de Hoop & de Swart 1998), the interpretational effects of scrambling (de Hoop 2000), and the projection mechanism of presupposition (Zeevat 1999 a,b; Blutner 1999; Geurts, to appear).

However, Blutner (1999) argues that this design of OT is inappropriate and too weak in a number of cases. This is due to the fact that the abstract generative mechanism (called **Gen** in the OT literature) can pair *different* forms with one and the same interpretation. The existence of such alternative forms may raise blocking effects which strongly affect what is selected as the preferred interpretation. The phenomenon of blocking requires us to take into consideration what else the speaker could have said. As a consequence, we have to go from a one-dimensional, to a two-dimensional (bidirectional) search for optimality.

This bidirectional view was independently motivated by a reduction of Grice's maxims of conversation to two principles: the Q-principle and the I-principle (Atlas & Levinson 1981; Horn 1984, who writes R instead of I). The I/R-principle can be seen as the force of unification minimizing the Speaker's effort, and the Q-principle can be seen as the force of diversification minimizing the Auditor‘s effort. The Q-principle corresponds to the first part of Grice's quantity maxim (*make your contribution as informative as required*), while it can be argued that the countervailing I/R-principle collects the second part of the quantity maxim (*do not make your contribution more informative than is required*), the maxim of relation and possibly all the manner maxims.

In a slightly different formulation, the I/R-principle seeks to select the most coherent interpretation, and the Q-principle acts as a blocking mechanism which blocks all the outputs which can be grasped more economically by an alternative linguistic input (Blutner 1998). This formulation makes it quite clear that the Gricean framework can be conceived of as a bidirectional optimality framework which integrates expressive and interpretive optimality. Whereas the I/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. The important feature of this formulation within bidirectional OT is that although it compares alternative syntactic inputs to one another, it still helps to select the optimal meaning among the various possible interpretational outputs of the single actual syntactic input given, by acting as a blocking mechanism.

A strong version of bidirectional OT can be formulated as given in (1). Here, we relate pairs (*f*, *m*) of possible (syntactic) forms *f* and meanings (= semantic interpretations) *m*, by means of an ordering relation >, *being more efficient.*

(1) **Bidirectional OT** (*Strong Version*)

A form-meaning pair (*f*, *m*) is optimal iff it is realized by **Gen** and it satisfies both the Q- and the I-principle, where:

(Q) (*f*, *m*) satisfies the Q-principle iff there is no other pair (*f'*, *m*) realized by **Gen** such that (*f'*, *m*) > (*f*, *m*)

(I) (*f*, *m*) satisfies the I-principle iff there is no other pair (*f*, *m'*) realized by **Gen** such that (*f*, *m'*) > (*f*, *m*)

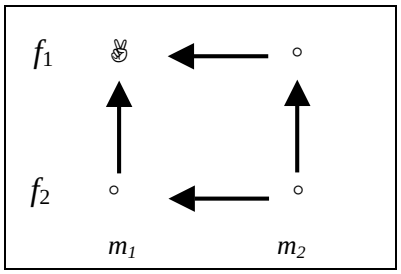
We will now give a very schematic example in order to illustrate some characteristics of the bidirectional OT. Assume that we have two forms f_1 and f_2 which are semantically equivalent. This means that **Gen** associates the same meanings with them, say m_1 and m_2. We stipulate that the form f_1 is less complex (marked) than the form f_2 and that the interpretation m_1 is less complex (marked) than the interpretation m_2 . From these differences of markedness with regard to the levels of syntactic representation / semantic interpretations, the following ordering relation between representation-meaning pairs can be derived:

(2) a. $(f_1 , m_1) > (f_2, m_1)$
b. $(f_1 , m_2) > (f_2, m_2)$
c. $(f_1 , m_1) > (f_1, m_2)$
d. $(f_2 , m_1) > (f_2, m_2)$

Using Dekker's & van Rooy's (1999) notation, the following bidirectional OT diagram can be construed, nicely representing the preferences between the pairs. More

importantly, such diagrams give an intuitive visualization for the optimal pairs of (strong) bidirectional OT: they are simply the hallows if we follow the arcs. (It should be noted that Dekker & van Rooy (1999) give bidirectional OT a game theoretic interpretation where the optimal pairs can be characterized as so-called Nash Equilibria). The optimal pairs are marked with the symbol ✌ in the diagram.

(3)



The scenario just installed 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 exist. According to Kiparsky (1982) *partial blocking* is realized in the case where the special (less productive) affix occurs in some restricted meaning and the general (more productive) affix picks up the remaining meaning (consider examples like *refrigerant - refrigerator, informant - informer, contestant - contester*). McCawley (1978) collects a number of further examples demonstrating the phenomenon of partial blocking outside the domain of derivational and inflectional processes. For example, he observes that the distribution of productive causatives (in English, Japanese, German, and other languages) is restricted by the existence of a corresponding lexical causative. Whereas lexical causatives (e.g. (4a)) 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, (4b) could have been used appropriately when Black Bart caused the sheriff's gun to backfire by stuffing it with cotton.

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

Typical cases of total and partial blocking are not only found in morphology, but in syntax and semantics as well (cf. Atlas & Levinson 1981, Horn 1984, Williams 1997). 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 labor"*.

There are two principal possibilities to avoid total blocking within the bidirectional OT framework. The first possibility is to make some stipulations concerning **Gen** in order to exclude equivalent semantic forms. The second is to weaken the notion of

(strong) optimality in a way that allows us to derive Horn's division of pragmatic labor by means of the evaluation procedure.

Blutner (1998,1999) argues that the second option is much more practicable and theoretically interesting. A *weak* version of two-dimensional OT was proposed, according to which the two dimensions of optimization are mutually related:

(5) **Bidirectional OT** (*Weak Version*)

A form-meaning pair (*f*, *m*) is *super-optimal* iff it is realized by **Gen** and it satisfies both the Q- and the I-principle, where:

(Q) (*f*, *m*) satisfies the Q-principle iff there is no other pair (*f'*, *m*) realized by **Gen** which satisfies the I-principle such that (*f'*, *m*) > (*f*, *m*)

(I) (*f*, *m*) satisfies the I-principle iff there is no other pair (*f*, *m'*) realized by **Gen** which satisfies the Q-principle such that (*f*, *m'*) > (*f*, *m*)

A more transparent formulation of *super-optimality* has been proposed by Jäger (2000):

(6) **Bidirectional OT** (*Weak Version, Jäger's variant*)

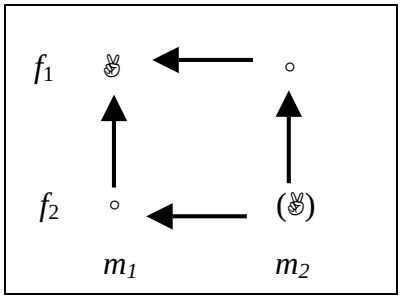
A form-meaning pair (*f*, *m*) is *super-optimal* iff it is realized by **Gen** and it satisfies the following two conditions:

(Q) there is no other *super-optimal* pair (*f'*, *m*) : (*f'*, *m*) > (*f*, *m*)
(I) there is no other *super-optimal* pair (*f*, *m'*) : (*f*, *m'*) > (*f*, *m*)

Under the assumption that > is transitive and well-founded, Jäger (2000) observes that both versions of weak bidirection coincide; that is a representation-meaning pair is super-optimal in the sense of definition (5) if and only if it is super-optimal in the sense of definition (6).

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. Consider again the situation illustrated in (3), but now apply the weak versions of bidirectional optimization (to make things more concrete we can take f_1 to be the lexical causative form (4a), f_2 the periphrastic form (4b), m_1 direct (stereotypic) causation and m_2 indirect causation).

(7)



We have seen that the *strong* version cannot explain why the marked form f_2 has an interpretation as well. The *weak* version, however, can explain this fact. Moreover, it explains that the marked form f_2 gets the atypical interpretation m_2. The form f_2 gets the interpretation m_2 because this form-meaning pair is super-optimal: (i) the alternative form f_1 doesn't get the atypical interpretation m_2, and (ii) we prefer to refer to the typical situation m_1 by using f_1 instead of f_2. In this way, the weak version accounts for the pattern called *"the division of pragmatic labor"*. It is not difficult to see that this pattern can be generalized to systems where more than two forms are associated by **Gen** with more than two interpretations. In the general case, we start with determining the optimal pairs. Then we drop the rows and columns corresponding to the optimal pair(s) and apply the same procedure for the reduced tableau.

The additional solutions are due to the flexibility and ability to learn which the weak formulation alluded to. The strong view is sufficient when it is enough to find *one* prominent solution. The weak view allows us to find out other solutions as well. In section 4 we will make use of this more general solution concept to explain the effects of negative strengthening, and in section 6 we will use it in order to explain the patterns of dimensional designation for spatial objects.

3 THE PHENOMENON OF NEGATIVE STRENGTHENING

Negation in natural language is a rich source of a variety of non-logical inferences (see Horn 1989). Standard examples are *scalar implicatures* (*Not all of the students came ≈> Some of them came*). Others are collected under the term *negative strengthening*. For an excellent discussion of the phenomenon of *negative strengthening* we refer to Horn (1989) and Levinson (2000).

One instance of the phenomenon of *negative strengthening* arises with gradable adjectives which typically occur as antonyms, such as {*good, bad*}, {*large, small*}, {*happy, unhappy*}. Semantically, the elements of antonym pairs are *contraries*, that is they are mutually inconsistent but do not exhaust the whole spectrum, permitting a non-empty *middle ground*.

What are the effects of negating gradable adjectives? For the sake of explicitness let us consider the gradable antonyms *happy* and *unhappy*, and assume three possible states of happiness – iconized by ☺, ☹, and 😐. Not unexpectedly, we

want to take *happy* as referring to the first state, *unhappy* as referring to the second state, and *neither happy nor unhappy* as referring to the third state.

Let's consider first the effect of negating positive adjectives, starting with a sentence like (8a). Obviously, the preferred interpretation of this sentence is (8c); this corresponds to a logical strengthening of the content of (8a) which is paraphrased in (8b). The discourse (8d) shows that the effect of strengthening (8c) is defeasible. This indicates that the inferential notion that underlies the phenomenon of strengthening ought to be non-monotonic.

(8) a. I'm not happy

b. It isn't the case that I'm happy (*Entailment*) ☹ 😐

c. I'm unhappy (*Implicature*) ☹

d. I'm not happy and not unhappy (*Defeasibility*)

Following Levinson (2000), the effect of negative strengthening for positive adjectives can be illustrated in the following way:

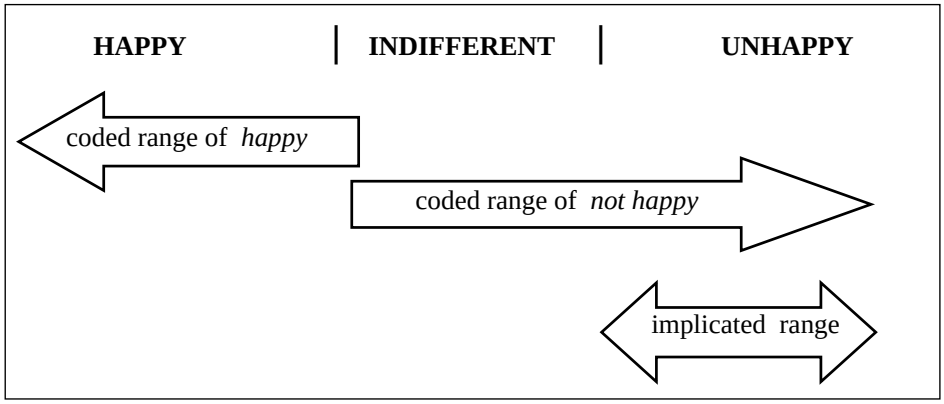


Figure 1: Negative strengthening as implicated contraries

It describes the effect of negative strengthening as implicating contraries from contradictions.

The illustrated shape of negative strengthening is restricted to the positive (unmarked) element of an antonym pair. When considering negative adjectives, deviations from this pattern may be found. The deviations are rather obvious for adjectives with affixal negation. This leads us to the well-known case of double negation (*Litotes*):

(9) a. I'm not unhappy

b. It isn't the case that I'm unhappy (*Entailment*) ☺ 😐

c. I'm neither happy nor unhappy (*Implicature*) 😐

d. I'm rather happy (but not quite as happy as using the expression *"happy"* would suggest) (*proper Implicature*)

e. I'm not unhappy, in fact I'm happy (*Defeasibility*)

Admitting only three states on the happiness scale allows only a rather rough approximation of the interpretational effects. The simplest approximation describes negative strengthening as a preference for the middle ground. This is what (9c) expresses. A more appropriate formulation of the effect is given in (9d). For the sake of precision, we would have to introduce intermediate states between ☺ and 😐 (on the scale of happiness). Other interpretational effects left out here, may be seen as cultural preferences, as for example in the case of understatement extensively used in British English, where *I am NOT happy* (negation phonologically marked) is often taken to implicate that I am unhappy. In the following diagram a fairly adequate illustration of the basic pattern is presented (as described in Horn 1989, Levinson 2000.)

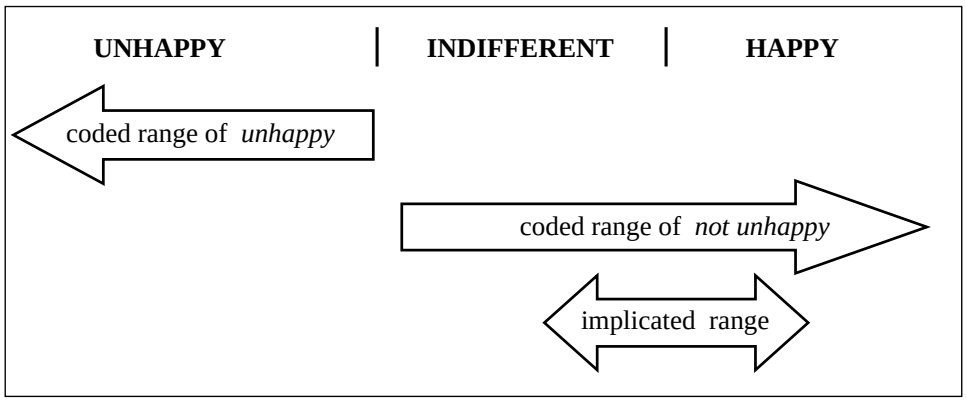


Figure 2: Litotes: when two negatives don't make a positive

As in the case discussed above, the effect of negative strengthening proves defeasible, a fact that requires the underlying inferential notion to be non-monotonic.

4 NEGATIVE STRENGTHENING AND BIDIRECTIONAL OT

In the previous section a concise description of the phenomenon of negative strengthening was given. This phenomenon will be brought into play in order to

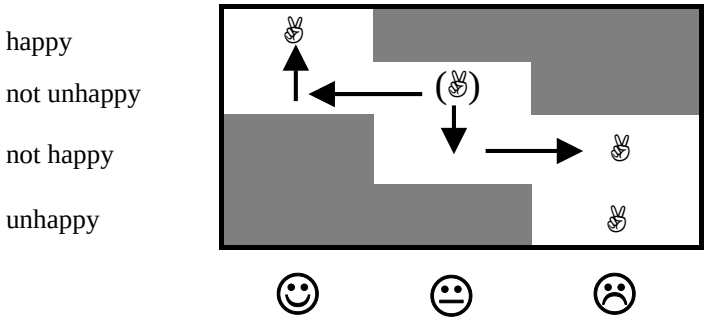
illustrate the general mechanism of pragmatic strengthening, which is formulated by using the method of bidirectional optimization.

In the analysis of Horn (1989) and Levinson (2000) there are some types of negative strengthening that are obviously attributable to the I/R-principle. A clear case is the negation of positive adjectives, which was described in connection with example (8). Here the I/R principle leads to a pragmatic strengthening effect excluding the middle ground and inferring the contrary.

The situation is not so clear in the case of adjectives with affixal negation such as in example (9). Whereas Horn (1984, 1989) seems to attribute the observed effect of negative strengthening to the interaction between Q and R, Levinson stipulates a third pragmatic principle, the M(anner)-principle: "what's said in a abnormal way, isn't normal; or marked message indicates marked situation." (Levinson 2000: 33). Obviously, this principle expresses the second half of Horn's *division of pragmatic labor*. In our opinion, Levinson (2000) tries to turn a plausible heuristic classification scheme based on the three principles Q, I, and M into a general theory by stipulating a ranking Q > M > I. Accepting the heuristic classification schema, we see problems for this theory, which is burdened with too many stipulations. Not unlike Horn's conception, we would rather like to see the M-principle as an *epiphenomenon* that results from the interaction of Zipf's two "economy principles" (Q and R in Horn's terminology).

Let us now have a look at how bidirectional OT accounts for the effects of negative strengthening. The bidirectional tableau (10) shows the competing candidate forms to the left. (Take the candidate entries as shortcuts for complete sentences; for example take *happy* as abbreviating *I'm happy*, etc.). The three columns designate the possible states of happiness considered in this simplified analysis. The gray areas in the tableau indicate which form-interpretation pairs are excluded by the compositional mode of truth-functional semantics, which is described by **Gen**. For example, *I'm not unhappy* is assumed to exclude the state iconized by ☹.

(10)



The preferences between the form-interpretation pairs are due to markedness constraints for forms and markedness constraints for interpretations, respectively.

With regard to the forms, we simply assume that the number of negation morphemes is the crucial indicator. The corresponding preferences are indicated by

the vertical arrows. (Note that *not happy* and *unhappy* aren't discriminated in terms of markedness – a rough simplification, of course.)

With regard to the states, we assume that they are decreasing in markedness towards both ends of the scale, assigning maximal markedness to the middle ground. Although this assumption seems not implausible from a psycholinguistic perspective, we cannot provide independent evidence for it at the moment. In the tableau (10), the corresponding preferences are indicated by the horizontal arrows.

Now it is quite easy to find the optimal solutions – indicated by ☞. One optimal solution pairs the sentence *I'm not happy* with the interpretation ☹. This solution corresponds to the effect of negative strengthening that is attributable to the I/R-principle. The other two optimal solutions are reflecting the truth condition of *I'm happy/unhappy*.

Most interesting, there is an additional super-optimal solution, indicated by (☞). It pairs the sentence *I'm not unhappy* with the interpretation 😐. This corresponds to the effect of negative strengthening in the case of *Litotes*, normally attributed to Levinson's (2000) M-principle or Horn's *division of pragmatic labor*. As already stressed, this solution comes out as a natural consequence of the weak form of bidirection, which can be seen as a formal way of describing the interactions between Q and I/R.

It's an interesting exercise to introduce more than three states of happiness and to verify that the proper shape of implicature as indicated in Figure 2 can be approximated. More importantly, in the context of *litotes* it seems necessary to account for the effect of gradient acceptability and continuous scales. Using a stochastic evaluation procedure, Boersma (1998) did pioneering work in this field, which should be exploited in the present case.

The other prominent class of examples that exhibit the effect of negative strengthening concerns the phenomenon of neg-raising, i.e. the tendency for negative main sentences with subordinate clauses to be read as negations of the subordinate clause (cf. Horn 1998, Levinson 2000). It seems desirable to analyze the phenomenon using the same technique as described before.

5 SOME PUZZLES OF DIMENSIONAL DESIGNATION

The term 'dimensional designation' refers to the contextual interpretation of a group of spatial adjectives such as *long, high, broad, deep, thick* and can be illustrated by the following example:

(11) a. The windowsill is 1 m long, 30 cm wide and 3 cm thick
b. The windowsill is 1 m wide, 30 cm deep and 3 cm thick

In (11a) the adjective *wide* refers to the secondary dimension whereas in (11b) it refers to the maximal (most salient) dimension. In order to explain the basic effects of dimensional designation we need the right combination of lexical stipulations and general principles of coherence, blocking and (perhaps) deblocking. In the following,

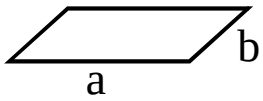
we want to illustrate how bidirectional OT solves this conceptual and methodological problem.

There is a thorough literature describing the linguistic facts of dimensional designation of spatial expressions (e.g. Bierwisch 1967, Lang 1989). It is not the aim of this section to extend this literature or to find out new observations that challenge the basic facts described there. As usual, the facts are described by using some (semi-) formal representational system. Certainly, there are good reasons for further improving the existing systems. However, that will not be of any concern in the present article. Let us concentrate on a small sector of the known observations by using standard representational systems. What is more important, we feel, is to provide a real explanation of these facts and observation. Our aim is to demonstrate that the framework of bidirectional optimization can be an appropriate tool for obtaining *explanatory adequacy*. Hopefully, this tool will help us to get a real understanding of the basic facts. A related point is a methodological one. It aims at the right relationship between lexical stipulations and general principles of economy. This is of great importance also for practical systems that should use lexical stipulations sparingly.

Suppose a physical object that our brain tries to encode. Suppose further that we can discriminate different dimensions (or axes) of spatial extent. It is the typical function of a spatial adjective to refer to a particular dimension of that object (in a particular contextual setting). The theoretical problem concerning the dimensional designation of spatial objects is to provide a mechanism that allows a realization of the mapping between dimensional adjectives on the one hand and the dimensions of physical objects referred to on the other hand. For simplicity, we will concentrate on two- and three-dimensional spatial objects where all axes are disintegrated (i.e. we don't considering objects like *tree*, *ball*, and *wheel* where two or more axes are integrated into one dimension. Furthermore, we are considering only a very restricted number of adjectives, namely the following: *long, high, wide, deep, thick.*

The facts we are considering aim at two different but interrelated phenomena: interpretational preferences and blocking. We start with the first aspect, preferences in interpretation. As an example, consider the following *question* in the context of a visually presented rectangle:

(12) How long and how wide is this rectangle?



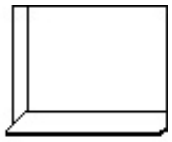
Obviously, there is a strong preference to refer to the maximal axis *a* with *long* (and to the secondary axis *b* with *wide*). However, as noticed in Lang (1989: 349), there are exceptions to the rule that *long* designates the maximal axis:

(13) a. The seed drill is wider than long
b. Our new double bed is 2 m long and 3 m wide
c. The velvet remnant is 1,3 m wide but only 0,5 m long.

Obviously, in cases where a non-maximal axis is designated, this axis is the most salient for other reasons than spatial extent (salient direction of movement / salient inherent orientation / prototypical designation). As a consequence, we should not characterize the adjective *long* as referring to a *maximal* dimension. Instead, we should take the lexical entry for *long* to be a candidate for radical *underspecification* and we should look for a mechanism of pragmatic strengthening that conforms to internal competition and aims to assign the most salient axis in the given context.

A more complex example which is appropriate to make a similar point was already presented at the beginning of this section and is repeated here for convenience:

(11) a. The windowsill is 1 m long, 30 cm wide and 3 cm thick
b. The windowsill is 1 m wide, 30 cm deep and 3 cm thick



The puzzling fact is that in contexts where we try to conceptualize an inherent observer (example b) the adjective *wide* designates the maximal dimension. However, in observer-free context (example a) there is a strong preference to designate the secondary (next to maximal) dimension. This makes clear that the adjective *wide* is another candidate for radical *underspecification.* The puzzling point (to be resolved in the next section) is how to manage the extreme context-dependency of this adjective.

Spatial adjectives provide an excellent area for studying the phenomenon of blocking. The examples are legion, and so we can restrict ourselves to a short list. In this list, the (a) examples block the corresponding (b) examples.

(14) a. The tower is 10 m high [vertical dimension]
b. ??The tower is 10 m long

(15) a. The pencil is 20 cm long [maximal dimension]
b. ?The pencil is 20 cm high (possible in particular contexts!)

(16) a. The tunnel is 2 km long [maximal dimension]
b. ?The tunnel is 2 km deep (possible in particular contexts!)

(17) a. The well is 10 m deep [vertical, observer direction]
b. ??The well is 10 m long / high

The examples (15b) and (16b) can be taken to illustrate the phenomenon of deblocking: in particular contexts, the anomalies may disappear. As is discussed elsewhere (e.g. Blutner 1998), the phenomenon of blocking / deblocking excludes a classical treatment of such examples as simple violations of definite conditions.

Another domain where the effects of interpretational preferences, blocking and deblocking come to the surface is the field of spatial prepositions (see Solstad 2000). However, reasons of space force us to drop this extension here.

6 OPTIMAL DIMENSIONAL DESIGNATION

It is a common assumption of all theories of form perception that the description of spatial relationships between the various parts of an object involves a frame of reference or mental coordinate system. According to seminal work by Jackendoff (1996) there are eight frames of reference to consider, four which are intrinsic (geometric frame, motion frame, canonical orientation, canonical encounter) and which determine the axis proper to an object, and four which are environmental (gravitational frame, geographical frame, contextual frame, observer frame). For the present discussion we can restrict ourselves to three frames of reference:

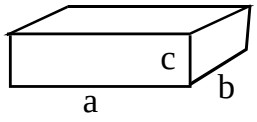
- ♦ geometric reference frame determining the axes proper to an object: **intrinsic**
- ♦ scene-based / **gravitational** frame
- ♦ **observer** frame

} **environmental**

Following Jackendoff (1996) we assume a domain-specific and informationally encapsulated module of *spatial representation* (SR). This module is distinct from *conceptual structure* (CS). The usual assumption is that CS encodes propositional representation and SR image schema or mental models (SR is geometric, but not imagistic). Spatial information is encoded in SR, rather than in CS.

For convenience, we adopt Lang's (1989) *object scheme* as a structural skeleton of SR. In short, an object scheme represents that part of SR which is relevant for specifying the spatial dimensions of the object. At first place, thus, an object scheme contains a specification of the involved coordinate system(s). So, the object scheme of the term *brick* may contain three perpendicular axes *a*, *b*, *c* which are ordered intrinsically (*a* is the axis with maximal extent, *b* the secondary and *c* the axis with minimal extent). Furthermore, the axis *c* is extrinsically specified as the vertical axis.

(18) < a b c> 'intrinsic'
Vert 'gravitational'



Similar object schemes apply for *tombstone, beam* (*girder*), *board, a*nd *windowsill*. As we have seen in example (11), the term *windowsill* (and analogously, the other terms) may trigger an additional object scheme which involves the observer coordinate:

(19) < a b c> 'intrinsic'
Vert 'gravitational'
Obs 'observer'

Following Jackendoff (1996), we can see the lexicon as the interface between the language module and the modules CS and SR. For spatial dimensional adjectives it is plausible to assume that an association with SR is most crucial. The question of how the association of an object scheme with some spatial object and thus with the lexeme or lexical material representing this object comes about, is an interesting one, but also obviously one that cannot concern us here. In the following we assume that dimensional adjectives are discriminated by two factors:

- the reference frame they trigger, e.g.
  - intrinsic: *long*, *wide*
  - gravitational: *high*
  - observer: *deep*, *wide*

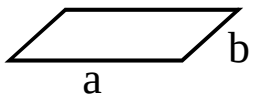
- specificity, e.g.
  {*long*, *deep*, *high*} > *wide*

The assumption that *wide* is the most unspecific adjective considered here derives from the fact that it is related to two different frames (intrinsic & observer) whereas the other adjectives refer to one frame only.

In subject-predicate expressions the reference frame triggered by the predicate must be present in the SR of the subject term. This is our basic assumption determining the **Gen**erator. Roughly spoken, it is a realization of all the potential pairings of a dimensional adjective with the designated axes of an object scheme given by the predicate term. The only condition is that the reference frame triggered by the adjective is compatible with the designated axis.

In the case of our earlier example (12) the generator leaves this correlation completely underspecified: Each of the two adjectives *long* and *wide* can be paired with either of the two axes *a* and *b*. This fact is due to the lexical entries of *long* and *wide* which only contain the information that both axes are to be intrinsic axes. The correct correlations *a-long* and *b-wide* are realized by the basic mechanism of bidirectional optimization (weak version). The discussion is completely analogous to that of the schematic example (7). The two axes are intrinsically ordered by salience: a > b, and the two adjectives are ordered by specificity: *long* > *wide*. From these two orderings of the inputs / outputs the ordering of the adjective-axes pairs can be derived. This is shown in (20):

(20) How long and how wide is this rectangle ?



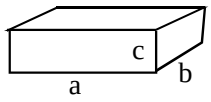
The framework of bidirectional optimization then produces the right correlation which can be seen as another reflection of the *division of pragmatic labor*: the salient axis correlates with the more specific adjective.

Earlier solutions (e.g. Bierwisch 1967, Lang 1989) crucially deviate from the present solution. They make use of features like MAX DIMENSION in case of *long* and SECONDARY DIMENSION in case of *wide*. The model theory of these features defines a kind of *internal competition* that simply stipulates the wanted result. The aim of the

present approach is to avoid such stipulations and to replace internal competition by an external one. Needless to say, external competition is defined by the overall framework of bidirectional optimization, and thus reflects claims that are motivated independently (see Wilson (1998) for a general discussion of the relationship between internal and external competition).

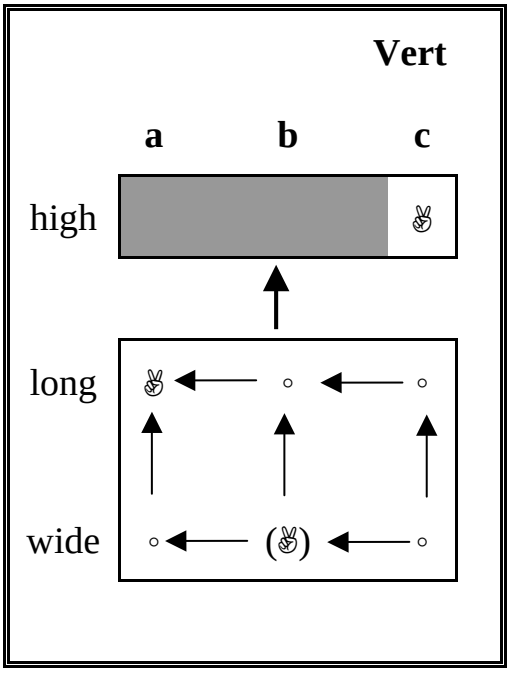
Let us finally present the analysis for a more complex example, (21).

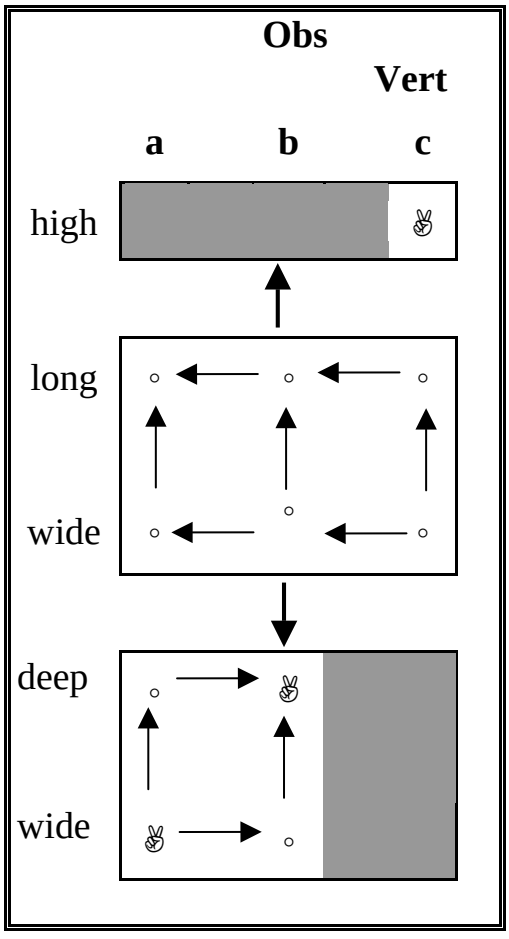
(21) a. The brick is 24 cm long, 15 cm wide, 8 cm high
b. The brick is 24 cm wide, 15 cm deep, 8 cm high



Similar to example (11), we assume that the module SR realizes two different object schemes for the term *brick*, one that doesn't involve the observer–represented in (18), and one that does – represented in (19). The tableau that corresponds to the first case is the one in (22a). It involves the intrinsic frame and the gravitational frame. The tableau that corresponds to the second case involves all three frames (intrinsic, gravitational, observer). It is (22b).

(22) a. b.





The ranking of the different sub-tableaus conforms to the following general assumption:

(23) If activated, the involved frames of reference are ranked as follows:
environmental > intrinsic

This assumption reflects the relative autonomy of the environmental frames relative to the intrinsic frame. It is an easy exercise to determine the super-optimal solutions in the tableaus (22a,b). In the first case, (21a), it comes out that the adjective *long* designates the maximal axis *a*, the adjective *wide* the secondary axis *b*, and *high* the vertical axis *c*. In the second case, (21b), it results that the observer-sensitive variant of *wide* designates the maximal axis *a*. The adjective *deep* designates the observer axis *b*, and high the vertical axis *c*. Notably, the use of the adjective *long* is **blocked** if an observer axis is involved. The treatment of example (11) is analogous. However, it involves a further dimension: *substance* (triggered by the adjective *thick*, cf. Lang 1989).

7 SUMMARY AND CONCLUSION

Investigating the interactions between the (mental) lexicon and pragmatics we have pointed out that situated meanings of many words and simple phrases are combinations of their lexical meanings proper and some superimposed conversational implicatures. The basic approach of lexical pragmatics combines the idea of (lexical) underspecification with a theory of pragmatic strengthening. The latter is formulated in terms of a bidirectional OT formalizing Grice‘s idea of conversational implicature. The mechanism of pragmatic strengthening crucially makes use of “non-representational” parameters that are described by preferential relations, such as information scales or salience orderings.

The main advantage of bidirectional OT is that it helps us to put in concrete terms what the *requisites* are for explaining the peculiarities of negative strengthening, dimensional designation and other potential phenomena that may be discussed. What are the relevant cognitive scales? How do we measure morpho-syntactic markedness? How do we measure the values of probabilistic parameters that control and organize conceptual knowledge (salience, cue validity)?

An important challenge for the present view is the work done in relevance theory (e.g. Sperber & Wilson 1986, Carston 1998, 2000, this volume). Although we prefer a variant of Atlas’, Levinson’s and Horn’s framework, that doesn’t mean that we are taking a stand against relevance theory. Rather, it seems desirable and possible to integrate most insights from relevance theory into the present view. As a kind of meta-framework, optimality theory can help to realize this integrative endeavor and to bring the two approaches closer to each other. Recently, van Rooy (2000a,b) made the first important steps in this direction.

The general conclusion that can be drawn from the present analysis is that weak bidirection can simplify the system of lexical stipulations rather radically. In the case of negative strengthening, the interpretational effects of negating graded adjectives. were treated by means of weak bidirection. This appropriately accounts for the differences between positive and. negative adjectives, thus avoiding unmotivated lexical stipulations. In the case of dimensional designation, on the other hand, bidirectional OT helps to eliminate internal competition and to replace it by external competition. Both analyses nicely illustrates Saussure's view that the semantics of natural language is partly determined by the inventory of lexical items.

REFERENCES

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

Bierwisch, M. (1967). Some semantic universals of German adjectivals. *Foundations of Language* **3**, 1-36.

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

Blutner, R. (1999). Some aspects of optimality in natural language interpretation. In: Papers on Optimality Theoretic Semantics (H. de Hoop & H. de Swart, eds.), Utrecht Institute of Linguistics OTS, Uil OTS Working Paper, pp. 1-21. Also available from http://www2.rz.hu-berlin.de/asg/blutner/pap.html

Boursma, P. (1998). *Functional Phonology*. Holland Academic Graphics, The Hague.

Carston, R. (1998). The semantics/pragmatics distinction: a view from relevance theory'. *UCL Working Papers in Linguistics*, **10**, 1-30.

Carston, R. (2000). Informativeness, relevances and scalar implicature'. Manuscript, University College London.

Dekker, P. and van Rooy, R. (1999). Optimality theory and game theory: some parallels. In: Papers on Optimality Theoretic Semantics (H. de Hoop & H. de Swart, eds.), Utrecht Institute of Linguistics OTS, Uil OTS Working Paper, pp. 22-45.

Geurts, B. (to appear). Buoyancy and strength. Paper available from http://www.kun.nl/phil/tfl/bart/

Hendriks, P. & Hoop, H. de (to appear). Optimal theoretic semantics. To appear in *Linguistics and Philosophy*.

Hoop, H. de (2000). Optimal scrambling and interpretation. In: *Interface Strategies* (H. Bennis, M. Everaert, and E. Reuland, eds.), pp. 153-168. KNAW, Amsterdam.

Hoop, H. de & Swart, H. de (1998). Temporal adjunct clauses in optimality theory. Manuscript, OTS Utrecht.

Horn, L.R.. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicatures. In: *Meaning, Form, and Use in Context* (D. Schiffrin, ed), pp. 11-42. Georgetown University Press, Washington.

Horn, L.R. (1989). *A Natural History of Negation*. University of Chicago Press, Chicago.

Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. In: *Language and Space* (P. Bloom, M.A. Peterson, L. Nadel and M.F. Garrett, eds.), MIT Press, Cambridge, Mass.

Jäger, G. (2000), Some notes on the formal properties of bidirectional optimality theory. In: *Studies in Optimality Theory* (R. Blutner and G. Jäger, eds.), pp. 41-63. Linguistics in Potsdam, Potsdam..

Kiparsky, P. (1982). Word-formation and the lexicon. In: *Proceedings of the 1982 Mid-America Linguistic Conference* (F. Ingeman, ed.).

Lang, E. (1989). The semantics of dimensional designation of spatial objects. In: *Dimensional adjectives* (M. Bierwisch and E. Lang eds.). Springer-Verlag, Berlin etc.

Levinson, S.C. (2000). *Presumptive Meaning.* The MIT Press, Cambridge, Mass.

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

Rooy, R. van (2000a). Comparing questions and answers: a bit of logic, a bit of language, and some bits of information. Manuscript ILLC, Amsterdam.

Rooy, R. van (2000b). Decision problems in pragmatics. Manuscript ILLC, Amsterdam.

Solstad, T. (2000). The optimal lexicon and methods in lexical semantics. Project outline.

Sperber, D. and Wilson, D. (1986). *Relevance: Communication and Cognition.* Blackwell, Oxford.

Williams, E. (1997). Blocking and anaphora. *Linguistic Inquiry*, **28**, 577-628.

Wilson, C. (1998). Bidirectional optimization and the theory of anaphora. Manuscript, Johns Hopkin University.

Zeevat, H. (1999a). Semantics in optimality theory. In: Papers on Optimality Theoretic Semantics (H. de Hoop & H. de Swart, eds.), Utrecht Institute of Linguistics OTS, Uil OTS Working Paper, pp. 76-87.

Zeevat, H. (1999b). Explaining presupposition triggers. Manuscript AC99, University of Amsterdam. Available from http://www.hum.uva.nl/computerlinguistiek/henk/