

Editor's Introduction: Pragmatics in Optimality Theory

Reinhard Blutner and Henk Zeevat

Based on the tenets of the so-called 'radical pragmatics' school (see, for instance, Cole, 1981), this book takes a particular view with regard to the relationship between content and linguistically encoded meaning. The traditional view embodied in the work of Montague and Kaplan (e.g., Kaplan, 1979; Montague, 1970) sees content being fully determined by linguistic meaning relative to a contextual index. In contrast, the radical view takes it that, although linguistic meaning is clearly important to content, it does not determine it, as pragmatic principles also play a role. The central issue of this book is how to give a principled account of the determination of content. Seeing linguistic meanings as underdetermining the content (proposition) expressed, there must be a pragmatic mechanism of completion which can be best represented as an optimization procedure. It is demonstrated that the general framework of Optimality Theory (OT) makes it possible to formulate the desired explanatory principles.

The first section of this general introduction outlines the basic framework of OT as applied to phonology, syntax and morphology. The second section takes a historical perspective and shows that the idea of optimization was present in the pragmatic enterprise right from the beginning. Further, it explains the main advantages of the general framework of OT when applied to the field of pragmatics, and it puts the whole idea into concrete terms by demonstrating how Horn's (1984) theory of conversational implicature can be implemented within a bidirectional optimality theory. In section 3, we raise several basic questions underlying the whole volume and discuss them from a theoretical and empirical perspective. This part

gives a overview of the different topics treated in the book, and it explains in which respects the single contributions aim to satisfy our cooperative goal: to give a new impulse to the tradition of radical pragmatics. Section 4, finally, outlines basic open question of future research.

1 Optimality Theory

OT was initiated by Prince & Smolensky (1993) as a new phonological framework that deals with the interaction of violable constraints. In recent years, OT was the subject of lively interest also outside phonology. Students of morphology, syntax and natural language interpretation became sensitive to the opportunities and challenges of the new framework. The reasons for this growing interest in OT are empirical and conceptual. First, it turned out that a series of empirical generalizations and observed phenomena can be expressed very naturally within this framework; this holds especially for phonology where in-depth analyses of many languages have provided a much better insight into cross-linguistic tendencies than we had before the invention of OT. Second, and perhaps much more important in linking scientists into a new research paradigm, there are the conceptual reasons, which are many in the present case: (i) the aim to decrease the gap between competence and performance, (ii) interest in an architecture that is closer to neural networks than to the standard symbolist architecture, (iii) the aim to overcome the gap between probabilistic models of language and speech and the standard symbolic models, (iv) the problem of learning hidden structure and the logical problem of language acquisition, (v) the aim to integrate the synchronic with the diachronic view of language.

OT respects the generative legacy in two important methodological aspects: the strong emphasis on formal precision in grammatical analysis and the goal of

restricting the descriptive power of linguistic theory. Seeing themselves within the Generative tradition, many representatives of OT adopt the fundamental distinction between Universal Grammar (UG) and a language-specific part of Grammar. UG describes the innate knowledge of language that is shared by all normal humans, and aims both to describe the universal properties of language and the range of variation possible among languages. The language-specific part of grammar typically consists of the lexicon and a system reflecting the specific structural properties of the particular language. Within the generative tradition, the concrete theoretical realization of this distinction has changed over the years. In the principles and parameters model, for example, UG is conceptualized as a system of (inviolable) principles, which are parameterized to demarcate the space of possible forms (see, for instance, Chomsky, 1981). The fixation of these parameters (triggered by language specific data) determines the grammar of the particular language. OT realizes an essentially different view of this distinction.

At this point we must emphasize that optimality theory is rooted, at least in part, in connectionism, a paradigm that makes use of neurobiological assumptions – in an extremely simplified way. As a consequence, OT does not assume a strict distinction between representation and processing. More than ten years ago, there was a lively debate in cognitive linguistics concerning the true architecture of cognition – the debate between connectionists and symbolists. The proponents of a symbolic architecture, among them Fodor and Pylyshyn (e.g., Fodor & Pylyshyn, 1988), had the clever idea to take the arguments for connectionism as showing that symbolic architecture is *implemented* in a certain kind of connectionist network. This idea corresponds to the strategy of maintaining classical architecture and reducing connectionism to an implementation issue. The development of OT demonstrates that

the opposite strategy is more exciting: augmenting and modifying symbolist architecture by integrating insights from connectionism.

Let's take a closer look now at the background and the nature of OT. Like other models of grammars, OT sees a grammar as specifying a function that assigns to each input (underlying representation of some kind) a structural description or output. For example, in Grimshaw's and Samek-Lodovici's theory of the distribution of clausal subjects (e.g., Grimshaw & Samek-Lodovici, 1995), an input is a lexical head with a mapping of its argument structure into other lexical heads, plus a tense specification. The input also specifies which arguments are foci and which arguments are coreferent with the topic. An example is

(1) <sing(x), x=topic, x=he; T=pres perf>

It represents the predicate *sing*, with a pronominal argument that is the current discourse topic. A possible output is an X-bar structure realizing an extended projection of the lexical head. Examples are

- (2) a. [_{IP} has [sung]] a clause with no subject
b. [_{IP} he_i has [t_i sung]] a clause with subject *he*, co-indexed with
a trace in SpecVP
c. [_{IP} has [t_i sung] he_i] *he* right-adjoined to VP, co-indexed with
a trace in SpecVP

The general idea of standard versions of generative syntax is to define the acceptable (grammatical) input-output pairs via a system of rules and transformations. In order to restrict the descriptive power of linguistic theory, constraints are added. All of

these constraints have been viewed as inviolable within the relevant domain. The idea of inviolable constraints has itself proved to be problematic and this has led to the “parameterization” of certain constraints, with one parametric setting for one language and another parametric setting for another language.

In OT the “generative part” of the Grammar is reduced to a universal function *Gen* that, given any input *I*, generates the set *Gen(I)* of candidate structural descriptions for *I*. The central idea of OT is to give up the inviolability of constraints and to consider a set *Con* of violable constraints. Furthermore, a strict ranking relation \gg is defined on *Con*. This relation makes it possible to evaluate the candidate structural descriptions in terms of the totality of the violations they commit, as determined by the ranking of the constraints. If one constraint C_1 outranks certain constraints C_2, \dots, C_i , written $C_1 \gg \{C_2, \dots, C_i\}$, then *one* violation of C_1 counts more than as arbitrarily many violations of C_2, \dots, C_i . The evaluation component selects the optimal (least offending, most harmonic) candidate(s) from the set *Gen(I)*. The grammar favors the competitor that best satisfies the constraints. Only an optimal output is taken as an appropriate (grammatical) output; all suboptimal outputs are taken as ungrammatical. This idea makes the grammaticality of a linguistic object dependent on the existence of a competitor that better satisfies the constraints.

Constraints are of two different kinds: *markedness constraints* that affect outputs only and *faithfulness constraints* that relate to the similarity between input and output. The main representatives of the faithfulness family are (i) PARSE prohibiting *underparsing* (“underlying input material is parsed into output structure”) and (ii) FULL prohibiting *overparsing* (“the elements of the output must be linked with correspondents in the underlying input”). In OT-syntax the latter constraint is also called FULL-INT(ERPRETATION): the elements of the output must be interpreted.

Markedness constraints are inherently connected with the domain under discussion. By way of example, we consider the following two constraints in the case of our OT syntax (distribution of clausal subjects):

- (3) a. SUBJ “the highest A-specifier in an extended projection must be filled”¹
 b. DROP-TOPIC “arguments coreferent with the topic are structurally unrealized”

To complete this short introduction to OT, let’s consider a typical OT tableau relating to the input-output pairings (1) & (2). In the present example, the following constraint hierarchy is assumed:

- (4) FULL-INT >> DROP-TOPIC >> PARSE >> SUBJ

As can be seen from tableau (5), this ranking yields an "Italian" behavior in which topicalized subjects are suppressed; this is exemplified by the optimal parse (5a).

(5)

<sing(x), x=topic, x=he; T=pres perf>	FULL-INT	DROP-TOPIC	PARSE	SUBJ
☞ (a) [IP has [sung]]			*	*
(b) [IP he _i has [t _i sung]]		*		
(c) [IP has [t _i sung] he _i]]		*		*

This behavior would change to "English" if we would chose the following hierarchy:

- (6) PARSE >> SUBJ >> FULL-INT >> DROP-TOPIC

where PARSE and SUBJ outrank DROP-TOPIC. In this case, (5b) would arise as the optimal candidate.

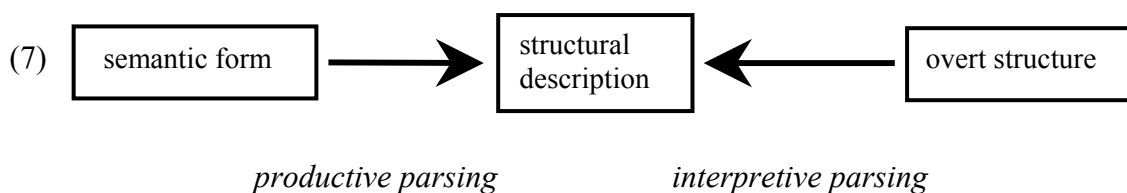
The architecture of OT suggests a simple realization of the fundamental distinction between UG on the one hand and the language specific part of Grammar on the other hand: UG consists of *Gen* (the generator) and *Con* (the set of constraints); the language-particular aspect of Grammar is determined by the particular ranking of the constraints. This proposal bolsters the way for defining a factorial typology:

Typology by reranking: Systematic crosslinguistic variation is due entirely to variation in language-specific total rankings of the universal constraints in *Con*. Analysis of the optimal forms arising from all possible total rankings of *Con* gives the typology of possible human languages. UG may impose restrictions on the possible rankings of *Con*. (Tesar & Smolensky, 2000, p. 27)

As already shown in (Prince & Smolensky, 1993), analysis of all rankings of the constraints considered in the basic CV syllable theory reveals a typology that explains Jacobson's (1962) typological generalizations. In the case of OT-syntax, Grimshaw & Samek-Lodovici (1995) were the first who performed an analysis involving all rankings of the above constraints and derived a typology of subject distribution on this way.

Typology by reranking is the most famous but not the only pleasant consequence from the general architecture of OT. Another consequence is the idea of *robust interpretive parsing*, which is substantial for many purposes, such as psycholinguistic applications of OT in describing online language production, comprehension, and natural language acquisition.

Although the term *parsing* is used more commonly in the context of language comprehension, in the OT literature it is treated as the general issue of assigning structure to input, an issue relevant to both comprehension and production. To be sure, the canonical perspective of an OT grammar is related to production – taking the input as an underlying form, and the output structural description as including the surface form. This type of parsing is called “productive parsing”, and it is schematically represented in diagram (7) – the term “overt structure” is used instead of “surface structure”:



In the context of language comprehension, another mapping comes into play. It maps a given overt form to an optimal structural description *SD* whose overt portion matches the given form. The process of computing the optimal *SD* for an overt form is called *interpretive parsing*.

It is a common observation that competent speakers can often construct an interpretation for utterances they simultaneously judge to be ungrammatical. Whereas it is notoriously difficult to account for this kind of “robustness” of natural language interpretation within rule- or principle-based models of language, the interpretation of ungrammatical sentences is much simpler when using an OT architecture. *Robust interpretive parsing* is the idea of parsing an overt structure with a grammar even when that structure is not grammatical according to that grammar. It is important to recognize that the presence of interpretable but ungrammatical sentences immediately corresponds to mismatches between productive and

interpretive parsing. Consider an interpretive parse that starts with some overt structure *OS* and assigns an optimal structural description *SD*. Paired with *SD* is a certain semantic form *SF*. The grammaticality of *SD* (and its overt structure, *OS*) depends on whether the outcome of productive parsing leads us back to *SD*, when starting with *SF*. In case it does, then *SD* is grammatical; otherwise it is ungrammatical.

As a simple illustration we reconsider the earlier example from OT syntax. Let's take the constraint hierarchy (4) that accounts for "Italian" syntactic behavior. In Italian, sentences such as *he has sung* are unacceptable if the pronoun refers to a discourse topic. Using the hierarchy (4), this is demonstrated in tableau (5), where the sentence *he has sung* comes out as suboptimal. Despite its unacceptability, the sentence is parsed into a structural description, namely [_{IP} he_i has [t_i sung]]. An important point in all examples of this kind is that both in productive parsing and in interpretive parsing the same constraint hierarchies are used. The difference arises solely from the different candidate sets that are relevant for the different perspectives of optimization.

The idea of robust interpretive parsing is crucial for the mechanism of language learning in OT when it is combined with another idea – the idea of constraint demotion (cf. Tesar & Smolensky, 2000). The latter idea conforms to a mechanism that reranks the constraints in a particular way, such that one prearranged candidate becomes the winner over the rest of the candidates (cf. **Vogel**²). The combination of both ideas gives the following picture of children's language acquisitions. Becoming confronted with some overt datum, the child tries to understand this datum (on the basis of her current grammar). She performs interpretive parsing, resulting in a structural description that includes an underlying structure. Next, the child turns to the production perspective: she starts with the

underlying form and performs productive parsing. If the results of productive and interpretive parsing are different, then this information is used to correct the grammar. The child applies constraint demotion taking the interpretive parse to be the winner (correct analysis) and the productive parse to be the loser. The child has succeeded in learning the target grammar if interpretive and productive parsing always give the same structural descriptions. Note that an overt form will allow the learner to improve his grammar just in case the current grammar (incorrectly) declares it to be ungrammatical.

There is an important consequence of this view of learning. The OT learning algorithm establishes an interesting type of equilibrium: *what we produce we are able to understand adequately and what we understand we are able to produce adequately*. This equilibrium corresponds to strong conception of bidirectional optimization: a logical combination of optimal comprehension and optimal generation (cf. Blutner, 2000; Zeevat, 2000; **Beaver & Lee**). Hence, bidirectional optimality can be seen as a kind of synchronic law describing the results of language learning. It should be mentioned that Tesar & Smolensky's (2000) mechanism for learning hidden structures is one aspect of language learning only. Acquiring conventions that link structured forms and conceptual contents (via lexical entries and idiom chunks) is another aspect. Most interestingly, empirical investigations have shown that also in this case the general pattern of bidirectionality or symmetry seems to apply.³ In the present volume, **Jäger** explores this possibility of bidirectional learning within an evolutionary setting.

Before we apply OT to the domain of pragmatics we must clarify what the general conditions are that every OT system has to satisfy. The following three conditions are the core of OT. They are a necessary basis for the family of procedures that performs grammar learning in OT (Tesar & Smolensky, 2000).

- (A) Universal Grammar is assumed to be determined by a generative part *Gen* and a system of violable constraints *Con* ($UG = Gen + Con$). The language-specific part of Grammar relates to a particular ranking of the constraints in *Con*. Only this part of the Grammar is learnable. Language learning simply reduces to inferring the ranking of the constraints in *Con*. This excludes both the possibility that the constraints themselves are learned (in part at least) or that aspects of the generator are learnable. On the other side it excludes the possibility that the set of the possible rankings is constrained on a universal basis.
- (B) The force of strict domination \gg : A relation of the form $C \gg C'$ does not merely mean that the cost of violating *C* is higher than that of violating *C'*; rather, it means that no number of *C'* violations is worth a single *C* violation. The force of strict domination excludes cumulative effects where many violations of lower ranked constraints may overpower higher ranked constraints.
- (C) The OT grammar of the language that has to be learned is based on a *total* ranking of all the constraints: $C_1 \gg C_2 \gg \dots \gg C_n$. This condition is crucial for the convergence of the proposed learning mechanism (Tesar & Smolensky, 2000). It can be shown that the iterative procedure of constraint demotion converges to a set of totally ranked constraint hierarchies in this case, each of them accounting for the learning data.

What is the status of these conditions? One way to look at these conditions is to see them as oversimplifications that are made mainly for didactic and practical reasons. Oversimplifications may be needed to allow one to concentrate on a central problem

and to sweep aside many problems that are less critical for understanding the central one (i.e., the problem of learning ‘hidden’ structure.) Moreover, oversimplifications may be necessary to achieve interesting mathematical results that simply are not possible without them. But it is not necessary to see them as simplifications, we can also see them as conditions reflecting the true nature of the domain under discussion and are taken to be empirically justified.

It must be admitted that it is not always simple to find out which position really is taken by the representatives of OT. For example, concerning the condition (C), we find the following statement in Tesar & Smolensky (2000):

From the learnability perspective, the formal results given for Constraint Demotion depend critically on the assumption that the target language is given by a totally ranked hierarchy. This is a consequence of a principle implicit in Constraint Demotion. This principle states that the learner should assume that the description is optimal for the corresponding input, and that it is the *only* optimal description. This principle resembles other proposed learning principles, such as Clark’s Principle of Contrast and Wexler’s Uniqueness Principle. (p. 47 ff)

It appears likely to us that learning languages that do not derive from a totally ranked hierarchy is in general much more difficult than the totally ranked case. If this is indeed true, demands of learnability could ultimately explain a fundamental principle of OT: UG admits only (adult) grammars defined by totally ranked hierarchies. (p. 50)

Taking condition (C) as a kind of principle that indicates when language learning is simple, however, is a different idea than taking it as a strict demand on theories of

learning. In our opinion, the first idea is right and the second wrong. There are many examples where the target language produces synonymies (scrambling data in German and Korean may provide a case in point). We agree that this can delay learning in one case or the other. In this vein, the suggestion is to take (C) as a kind of oversimplification, the acceptance of which is justified only for doing the first significant research steps. Notably, Anttila & Fong (2000) take a similar view (cf. also **Beaver & Lee**). As a consequence, the condition (C) should be given up in an advanced stage and a more general theory should be developed, a theory that *explains* (C) as a principle about the complexity of language learning. In our opinion, Paul Boersma's learning theory (Boersma, 1998; Boersma & Hayes, 2001) is on the right track for doing this job.

With regard to the condition (B), Smolensky himself sees it as a "regimentation and pushing to extremes of the basic notion of Harmonic Grammar" (Prince & Smolensky, 1993, p.200). And Gibson & Broihier (1998) argue that this restriction does not appropriately characterize the manner in which parsing preferences interact.

What about condition (A)? Many representatives of OT seem to consider it as a *conditio sine qua non*. Boersma's work on functional phonology (Boersma, 1998), however, puts forward convincing arguments exposing principle (A) likewise as a kind of oversimplification.

These questions about the status of the conditions (A)-(C) become highly relevant when we try to extend the domain of applications for OT, especially when we try to apply the OT framework to the domain of pragmatics. Hence, for pragmatics in OT debating and clarifying the status of the condition (A)-(C) is an opportunity and challenge. Most papers in this volume are directly or indirectly concerned with this task.

2 Pragmatics in OT

The idea of optimization was present in the pragmatic enterprise from the very beginning. Much more than in other linguistic fields optimality scenarios are present in most lines of thinking: Zipf's (1949) balancing between effect and effort, the Gricean conversational maxims (Grice, 1975, 1989), Ducrot's argumentative view of language use (e.g., Ducrot, 1980), the principle of optimal relevance in Relevance theory (Sperber & Wilson, 1986). However, in the course of the development of OT, the area of OT semantics and pragmatics was developed after everything else. This appears rather puzzling, and the reasons for that are not very clear. There may be stylistic aspects that may frighten a serious semanticist or logician: the curious tables with shadows, and the famous little hands. A more serious reason may have to do with an unfortunate 'dynamic turn' which was directed against Kamp's (1981) programmatic outline of a cognitively oriented approach to language. In contrast to Kamp's original paper, which is based on the tenets of 'radical pragmatics', much research which falls under the rubric of the 'dynamic turn' is in the spirit of the conservative view of language which radical pragmatics set itself against.

While the compositionality assumption underlying the 'dynamic turn' has strengthened the methodology of semantics, it has also led to a mechanistic approach at points where pragmatics and semantics are difficult to keep apart. The habit of interpreting trees with fully resolved pronouns fails to make the distinction between rule-based grammar and the complicated salience weighting of different antecedents required for pronoun resolution, a process that leads to preferences at best. In treatments of presupposition, sometimes presupposition boils down to a single logical operator and so obscures the distinction between the semantic role of

presuppositions for their triggers and their function as a sign that the speaker is making an assumption, a distinction that also shows up as the distinction between treating a presupposition by resolving it to the context or by accommodating it. We would submit that an OT theory has an advantage over logical or grammatical treatments in that the ideal of rational cooperative communication can be almost directly captured by constraints that directly derive from Grice's analysis of this cooperative behavior (cf. **van Rooy, Zeevat**)

What can be called with more justice 'radical pragmatics' (cf. Cole, 1981) is to hypothesize a division of labor between (i) a linguistic system determining the semantic representation of a sentence (Grammar including the lexicon) and (ii) a pragmatic system constituting the interpretation of the corresponding utterance in a given setting (contextual information, encyclopedia). The pragmatic system is taken as realizing Grice's (1975) idea of conversational implicature, and it is modeled with the instruments of OT. As a consequence, many linguistic phenomena which had previously been viewed as belonging to the semantic subsystem, in fact can be explained within the pragmatic subsystem of OT.

Before we enter the discussion in which way optimality theory may help to close the gap between formal (linguistic) meaning and interpretation, we have to consider this distinction more closely. For Grice (1975) the theoretical distinction between what the speaker explicitly said and what he has merely implicated is of particular importance. What has been said is supposed to be based purely on the conventional meaning of a sentence, and is the subject of semantics. What is implicitly conveyed (scalar and conversational implicatures) belongs to the realm of pragmatics. It is assumed to be calculable on the basis of the setting – a notion already introduced by Katz & Fodor (1963), and referring to previous discourse, socio-physical factors and any other use of "non-linguistic" knowledge. Fruitful as

this theoretical division of labor may have been – especially as a demarcation of the task of logical semantics –, it has inherent problems. More often than not, what is said by a speaker's use of a sentence already depends on the context. Even for Griceans, propositional content is not fully fleshed out until reference, tense, and other indexical elements are fixed. However, propositional content must be *inferred* in many cases – going beyond the simple mechanism of fixing indexical elements.

Proponents of relevance theory (see, for example, Carston, 2002; Carston, 2003a, 2003b; Sperber & Wilson, 1986) have pointed out that the pragmatic reasoning used to compute implicated meaning must also be invoked to fill out underspecified propositions where the formal meaning contributed by the linguistic expression itself is insufficient to give a proper account of truth-conditional content. A similar point was made in lexical pragmatics (e.g., Blutner, 1998, 2002). Both relevance theory and lexical pragmatics agree in assuming a Gricean mechanism of pragmatic strengthening in order to fill the gap between formal, linguistic meaning and the propositional content (i.e., the *explicit* assumptions communicated by an utterance – called explicature in relevance theory; cf. Sperber & Wilson, 1981, p. 182).

In a similar vein, de Hoop & de Swart (2000) and Hendriks & de Hoop (2001) argue with regard to the theory of interpretation, what compositional semantics gives us is a radically underspecified notion of meaning represented by a possibly infinite set of interpretations of a well-formed syntactic structure. In addition, these authors were the first who propose to use the framework of optimality theory in order to select the optimal interpretation associated with a particular syntactic structure. For that purpose, they propose a particular set of constraints and a rankings between those constraints, based on general principles of rational communication. The interpretive perspective on optimization provides insights into

different phenomena of interpretation, such as the determination of quantificational structure and domain restriction (Hendriks & de Hoop, 2001), nominal and temporal anaphora (de Hoop & de Swart, 2000), and the interpretational effects of scrambling (de Hoop, 2000).

Stimulated by Horn's (1984) theory of conversational implicature and related ideas in relevance theory, Blutner (2000) argued 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 (*Gen*) can pair *different* forms with one and the same interpretation. The existence of such alternative forms may lead to *blocking* effects which strongly affect what is selected as the preferred interpretation. The phenomenon of blocking has been demonstrated in a number of examples where the appropriate use of a given expression formed by a relatively productive process is restricted by the existence of a more "lexicalized" alternative to this expression. One case in point was provided by Householder (1971). The adjective *pale* can be combined with a great many color words: *pale green*, *pale blue*, *pale yellow*. However, the combination *pale red* is limited in a way that the other combinations are not. For some speakers *pale red* is simply anomalous, and for others it picks up whatever part of the pale domain of red *pink* has not preempted. This suggests that the combinability of *pale* is fully or partially blocked by the lexical alternative *pink*.

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.⁴ As mentioned in section 1, bidirectional optimality can be seen as describing the equilibrium that results from language learning in the limit.

In the domain of pragmatics, the 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 Hearer's effort (cf. Horn 1984). 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. Conversational implicatures which are derivable essentially by appeal to the Q-principle are called Q-based implicatures. Standard examples are scalar implicatures and clausal implicatures. I-based implicatures, derivable essentially by appeal to the I-principle, can be generally characterized as enriching what is said via inference to a rich, stereotypical interpretation (cf. Atlas & Levinson, 1981; Gazdar, 1979; Horn, 1984; Levinson, 2000)

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 with 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.

The so-called strong version of bidirectional OT – it conforms to the equilibrium established during OT learning – can be formulated as given in (8). Here, pairs (f, m) of possible (syntactic) forms f and utterance meanings (= interpretations) m are related by means of an ordering relation $<$, *being less costly (more harmonic)*. At the moment, the precise metric underlying this ordering relation is still open, and the sign $<$ is not much more than a place holder for such a metric. In OT, the ordering relation $<$ can be constituted by a system of ranked constraints, as discussed in many contributions to this volume. Another option would be to work with a single, graded markedness constraint such as RELEVANCE (see **Van Rooy**).

(8) **Bidirectional OT** (*Strong Version*)

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

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

It should be mentioned that the I-principle is very much in line with the mono-directional view on optimality theoretic interpretation as proposed by de Hoop & de Swart (2000) and Hendriks & de Hoop (2001), which exclusively, adopts the hearer's

perspective on disambiguation. What is interesting in (8) is that it also implements the Q-principle, thereby also taking the speaker's perspective into account. Hence, a proper treatment of interpretation in OT has to take into account both the perspective of the hearer and the perspective of the speaker. Because this framework of bidirectional OT can be characterized in game-theoretical terms (Dekker & van Rooy, 2000), optimality theoretic pragmatics can be given a proper formal interpretation.

One of the main advantages of the optimality theoretic framework is that it allows to isolate three substantial components of the overall mechanism: (i) the generator, which provides the potential form interpretation pairs, (ii) the underlying metric, possibly constituted by a system of ranked constraints, and (iii) the two perspectives of optimization. In relevance theory it is *relevance* that constitutes the underlying metric; in other frameworks notions of information, efficiency, and salience are more important (cf. **van Rooy**).

There are however several old problems with assuming full symmetric bidirectionality to phonological and syntactic processing in both directions. In phonology, the problem is mostly discussed as the *Rad/Rat* problem (cf. Hale & Reiss, 1998). The German word *Rat* (council) is pronounced as [rat] without any change from the underlying form to the surface form. The word *Rad* (wheel) is pronounced in the same way but here two constraints come into play: the DEVOICING constraint that prefers the pronunciation [rat] to [rad] and FAITHFULNESS that would prefer the pronunciation [rad] and that is outranked by DEVOICING in German. If we want to apply the same constraints in the direction from pronunciation to optimal underlying form, *Rat* is always preferred because of FAITHFULNESS in interpretation. The same problem can arise in syntactic ambiguities. Again in German, the sentence *Welches Mädchen mag Reinhard?* is ambiguous between *Which girl likes Reinhard?*

and *Which girl does Reinhard like?* The wh-object has a longer road to go from its canonical position to its sentence initial position than the corresponding wh-subject. The constraint STAY (= *Do not move*) adopted by most OT syntacticians then prefers the reading with the wh-subject. Since there is general agreement that there is a proper ambiguity in these cases, full bidirectionality needs to be restricted by some principle which makes the system less symmetric than the Tesar & Smolensky learning algorithm assumes. In this volume, **Jäger** uses an asymmetric bidirectional system for his learning algorithm, **Vogel** restricts his OT-syntax by powerful pragmatic principles and **Beaver & Lee** consider different ways to avoid the Rat/Rad problem in their survey of bidirectionality.

Another problem has to do with the specific features of blocking we find in natural languages. The scenario of strong bidirection describes the case of **total** blocking where some forms (e.g. **furiousness*, **fallaciousness*) do not exist because others do (*fury*, *fallacy*). However, blocking is not always total but may be **partial**, in that only those interpretations of a form are ruled out that are pre-empted by a “cheaper” competing form. 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 restricted by the existence of a corresponding lexical causative.

- (9) (a) Black Bart killed the sheriff
(b) Black Bart caused the sheriff to die

Whereas lexical causatives – e.g. (9a) – tend to be restricted in their distribution to the stereotypic causative situation (direct, unmediated causation through physical action), productive (periphrastic) causatives tend to pick up more marked situations of mediated, indirect causation. For example, (9b) could be used appropriately when

Black Bart caused the sheriff's gun to backfire by stuffing it with cotton. 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, p. 26) – a tendency that Horn terms the *division of pragmatic labor*.

There are two principal possibilities for avoiding the fatal consequences of total blocking that are described by strong bidirection. The first possibility is to make some stipulations concerning *Gen* in order to exclude equivalent semantic forms. The second possibility is to weaken the notion of (strong) optimality in a way that allows us to derive Horn's division of pragmatic labor in a principled way by means of a sophisticated optimization procedure.

In Blutner (1998; 2000) it is argued that the second option is much more practicable and theoretically interesting. A recursive variant of bidirectional optimization was proposed (called *weak* bidirection) which was subsequently simplified by Jäger (2002):

(10) **Bidirectional OT** (*Weak Version*)

A form-meaning pair (f, m) is called super-optimal iff $(f, m) \in Gen$ and

- (a) there is no other super-optimal pair $(f, m') : (f, m') < (f, m)$
- (b) there is no other super-optimal pair $(f', m) : (f', m) < (f, m)$

Under the assumption that $<$ is transitive and well-founded, Jäger (2002) proved that (10) is a sound recursive definition and is equivalent with the formulation in Blutner (1998; 2000). In addition, he proved that each pair which is *optimal* (strong

bidirection) is *super-optimal* (weak bidirection) as well, but not vice versa. Hence, weak bidirection gives us a chance to find additional super-optimal solutions. For example, weak bidirection allows marked expressions to have an optimal interpretation, although both the expression and the situations they describe have a more efficient counterpart. Hence, this formulation is able to describe Horn's division of pragmatic labor. The notion of weak bidirection is discussed more detailed by **Mattausch** (this volume, section 3.2).

The existence of two notions of bidirectionality raises a conceptual problem: Which conception of bidirectionality is valid, the strong or the weak one? Obviously, this question relates to the foundation of bidirection in an overall framework of cognitive theory. As we have seen already, the *strong mode* of optimisation in (8) – *what we produce we are able to understand adequately and what we understand we are able to produce adequately* – corresponds to the equilibrium established by the OT learning algorithm. Hence, the strong conception of bidirectionality can be seen as a kind of synchronic law describing the results of language learning.

Weak bidirection gives a chance to find additional solutions. Is it possible to give a natural interpretation for these additional solutions? We want to propose the idea that these additional solutions are due to the ability and flexibility of self-organization in language change which the weak formulation alluded to. In other words, we propose to take these additional solutions as describing the possible outcomes of self-organization before the learning mechanism has fully realized the equilibrium between productive and interpretive optimisation.

Jäger (2002) and Dekker & Van Rooy (2000) have proposed algorithms that update the ordering (preference) relation $<$ such that (i) optimal pairs are preserved and (ii) a new optimal pair is produced if and only if the same pair was super-optimal at earlier stages. Consequently, we can take the solutions of weak bidirection to be

identical with the solutions of strong bidirection considering all the systems that result from updating the ordering relation. Recently, Van Rooy (forthcoming) and **Jäger** (this volume) have reconsidered this problem and have proposed algorithms within an evolutionary setting – realizing a mechanism of self-organization in language change. This point may be clarified when we (re)consider Horn's *division of pragmatic labor* and relate it to the principle of *constructional iconicity* in the school of "natural morphology" (for references cf. Wurzel, 1998)

Constructional iconicity: A semantically more complex, derived morphological form is unmarked regarding constructional iconicity, if it is symbolized formally more costly than its semantically less complex base form; it is the more marked, the stronger its symbolization deviates from this. (Wurzel, 1998, p. 68).

In this school the principle plays an important role in describing the *direction of language change*. In fact, constructional iconicity and Horn's division of pragmatic labor can be proven to be a consequence of weak bidirection. This observation gives substance to the claim that *weak bidirection* can be considered as a principle describing (in part) the direction of language change: super-optimal pairs are tentatively realized in language change. This relates to the view of Horn (1984) who considers the Q principle and the I principle as diametrically opposed forces in inference strategies of language change. Of course, the idea goes back to (Zipf, 1949), and was reconsidered in van Rooy (forthcoming).⁵ Arguing that Horn's division of pragmatic labor is a *conventional* fact about language, this convention can be explained in terms of equilibriums of signaling games introduced by Lewis (1969) – making use of an evolutionary setting (cf. van Rooy, forthcoming).

But is it really the case that weak bidirectionality does not play a role in synchrony? The Horn example is the pair *Black Bart shot the sheriff* / *Black Bart*

caused the sheriff to die. A similar example is Grice's *Mrs. T produced a series of sounds closely resembling the score of "Home Sweet Home"*, which contrasts with: *Mrs. T. sang "Home Sweet Home"*. Horn's and Grice's point is that the long and unusual form are used to convey that there was something special with the killing and the singing and that this is not accidental. The process by which this special interpretation is arrived at cannot be diachronic language change: the long and unusual forms are so unusual that it is not possible to assume a special conventionalisation process that associates the special meaning with the special form.

Grice's explanation from his maxim *Be Brief* can almost directly be translated in OT pragmatics. The relevant constraint is ECONOMY, which we can reinterpret as the requirement that there is no correct form interpretation pair that is more economical (or more standard?) in either dimension. This immediately leads us to reject the association of the complex (unusual) forms with the standard meaning: for that we have a simpler and more usual form. It likewise rules out the association of the simple form with the non-standard meaning. The result is that we obtain an underspecified special meaning for the special forms which must be interpreted further with respect to the context and the situation to give us the concrete interpretations (kill in a bizarre way, sing rather badly) that we seem to obtain. Notice however that the speaker has not said any of this, she has merely suggested that there is something special going on. There is no convention that fixes the meaning. The vagueness and cancellability of the extra interpretation suggests that we are dealing with an implicature and not with part of the truth-conditional content.

There are three points to be made about this reinterpretation of Grice's stylistic maxim. In the first place, it is a very low constraint which can be overridden by any grammatical or semantical constraint that one needs to assume. It is the lowest of the low. Second, it is obviously weakly bidirectional for it to work. If the

standard-form/standard-meaning or the marked-form/standard-meaning were in competition with marked-form/standard-meaning, that last pair would not survive. And third, it seems that – with some charity – all other pragmatic principles can be related to it. As Blutner & Jaeger show, the constraint DO NOT ACCOMMODATE can be seen as a special case of semantic economy, minimizing the number of new discourse referents. The constraint RELEVANCE can also be seen as a kind of semantic economy: irrelevant information is information that the interlocutor is not seeking for and requires the accommodation of new questions or interests of the interlocutor. Information that is consistent or consistent with the context is pragmatically less complex than information that is inconsistent in itself or inconsistent with the context. The whole of pragmatics would be weakly bidirectional under this interpretation.

If this would be the case, it would also give us an indication of why weak bidirectionality is such a powerful explanatory principle in diachronic linguistics. Pragmatic weak bidirectionality creates special interpretations that can become conventionalized. Assume that a marked form is used with some frequency to indicate the same marked meaning. It will then become a conventional device to indicate the marked meaning and the marked meaning will no longer be derived by weak bidirectionality but by a lexical or grammatical convention. Think about Hebrew optional object case marking conventionally meaning that the referent is definite. Or about the Dutch *wijf* – originally the standard word for woman, but pushed away by *vrouw* (originally mistress) – that can now only be used for the purpose of expressing contempt for the referent in question.

Summarizing, we suggest to take the strong conception of bidirectionality as a synchronic law and the weak one as conforming to diachrony (with the reservation and clarification just sketched). In addition, the present conception conforms to the

idea that synchronic structure is significantly informed by diachronic forces. Further, it respects Zeevat's (2000) acute criticism against super-optimality as describing an online mechanism (see also **Beaver & Lee**). From the perspective of *grammaticalization*, we are very close to Hyman's (1984) dictum of seeing grammaticalization as the harnessing of pragmatics by a grammar. And there are connections to a recent proposals by Haspelmath (1999) for an OT-based theory of language change.

3 Overview

The aim of this book is to demonstrate that OT also finds fruitful applications in the domain of pragmatics and can contribute in overcoming the gap between linguistic meaning and utterance meaning. This section contains an overview of the different topics treated in the book and it explains in which respects the single contributions aim to satisfy our cooperative goal: giving the tradition of radical pragmatics a new impulse.

The promise of OT pragmatics is that by using the OT architecture some order can be brought to the seemingly unrelated approaches that constitute pragmatics. There have been a series of studies that try to reformulate treatments of pragmatic phenomena to optimality theory. De Hoop & De Swart (2000) study the determination of quantifier restrictions, a classical challenge to compositional semantics, since that determination is only partially determined by the syntactic tree, and can involve interactions with the context, the information structure and the linear order of the quantifier. One of the factors in the solution is relating the interpretation to given material, either in the topic or in the context. This problem area comes back in studying pronoun syntax and resolution (Beaver, to appear; Bresnan, 2001),

presupposition (Jäger & Blutner, 2000; Zeevat, 2000), the binding theory (Burzio, 1991, 1998; Levinson, 1987, 2000). Other areas of pragmatics where OT has been attempted are intonation and information structure (Beaver & Clark, 2002; Schwarzschild, 1999), scalar implicatures (Blutner, 2000; Van Rooy, 2001).

In the present volume, Helen **de Hoop** provides an in depth discussion based on real data of the Complementary Preference Hypothesis as an account of stressed pronouns in English and formulates an alternative account in terms of two interpretive constraints: Contrastive Stress and Continuing Topic to overcome the problems with the earlier account.

Petra **Hendriks** combines a semantic analysis of only (*only(A)(B) = all(B)(A)*) with an OT account of how intonation and syntax conspire in determining the scope and restrictor of determiners and focus-sensitive particles. The account builds on earlier work of De Hoop & De Swart (2000) and Hendriks & De Hoop (2001) using optimality theoretic semantics.

Jason **Mattausch** introduces the influential work of Levinson on the origin and typology of binding theory and reformulates the different historical stages assumed by Levinson in bidirectional optimality theory. The reformulation is able to avoid and solve a number of problems in Levinson's proposal and can avoid the M-principle altogether which comes out as a theorem in bidirectional optimality theory.

Henk **Zeevat** reviews an earlier attempt to treat 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.

A proper framework of OT is also the correct platform for asking foundational questions. Given that we have violable principles and reliability ranking between them (let's assume this can be decided on empirical grounds), what follows about the representations on which the constraints have to work, can a rational foundation be found for each of the constraints and can the order between the constraints be founded in some rational principle? The notions of relevance and economy have particularly been in focus here. Another foundational issue concerns the nature of bidirection and the symmetry assumption (e.g., Zeevat, 2000). Further questions concern the division of labor between semantics and pragmatics in particular, and the modularity stipulation in general. And what is the proper architecture of an overall system integrating elements from syntax, prosody, semantics and pragmatics?

In the present volume, David **Beaver** and Hanjung **Lee** give an overview of various proposals in bidirectional optimality theory where crucial tests are total and partial blocking, the *Rat/Rad* problem and some other problems. They show conclusively that weak superoptimality cannot be combined with standard proposals for optimality theoretic syntax with a larger number of constraints.

Gärtner's analysis of Icelandic object-shift and differential marking of (in-)definites in Tagalog addresses the issue of disambiguation in natural languages. In the first part he suggests a family of OT-constraints called "Unambiguous Encoding", which can be understood as a correlate of Gricean "Avoid Ambiguity". In the second part he points out some shortcomings of this approach, and he suggests that the OT-status of "Unambiguous Encoding" is epiphenomenal. Two ways of reduction are explored which bolster the way for a functionalist understanding of the phenomenon – viewing grammars as "harnessed" or "frozen" pragmatics (cf. Hyman,

1984). In addition, and not unrelated to the contribution of **Beaver & Lee**, he points out some serious problems for Blutner's version of bidirectional OT.

Robert **van Rooy** is arguing that the general framework of optimality theoretic pragmatics is able to include basic insights from relevance theory. Starting from the bidirectionality of Blutner (1998; 2000) in terms of the Q- and I-principle, he develops a decision-theoretic notion of relevance to take – in the first instance – the place of the Q-principle in this scheme for pragmatics. Though this leads to improvements, further problems then force the tentative adoption of a relevance based exhaustivity operator as a basis for reconstructing the Q-principle, the I-principle and Blutner's bidirectionality. Horn's M-principle is then derived by minimisation of effort.

Ralf **Vogel** is addressing the problem of OT architecture. Following Jackendoff (1997) he is assuming three levels of representation: a semantic (= conceptual), a syntactic and a phonological level. The *correspondence* between these levels is modeled by a (bi-directional) OT grammar. Arguing that syntax is much less encapsulated and 'autonomous' than generative grammar usually assumes, Vogel's model is able to restrict OT-syntax by powerful pragmatic principles. In addition, there is a methodological point that deserves particular attention. The proposed architecture is not only motivated by its ability to account for certain intriguing linguistic phenomena. It is also justified by its compatibility with current OT learning theory.

OT pragmatics is a theory of pragmatic competence that invites to cross the boundaries of traditional pragmatics and to relate it to psycholinguistic theories of natural language performance (both production and comprehension) on the one hand, and to theories of language learning and language evolution on the other hand. This volume contains two contributions that explicitly conform to this challenge.

Jennifer **Spencer's** psycholinguistic investigation concerns the choice between two demonstrative forms in Swedish (one simple, the other compound). A multitude of factors influence the choice of one referential form over another, such as abstractness, animacy, and the level of activation of the referent. The general finding is that the simple form typically is used with more accessible and salient referents while the compound form is used for referents with a lower level of activation. Spencer argues that stochastic optimality theory is capable to model the subtle, yet statistically significant differences between the two demonstrative forms – making use of constraints that are independently motivated.

The contribution of Gerhard **Jäger** can be seen as the first step in a long research agenda which derives from the view that many syntactic and semantic facts are frozen pragmatics. It should be possible to show how particular languages emerge from pragmatics assuming the fairly standard account of the evolution of phonological forms. Even the advantages involved in moving from a purely pragmatic language to a language with partial conventionalisation can be studied from this perspective. The potential contribution of OT here is twofold. OT can inspire learning algorithms and it can provide the framework for the representation and evolution of grammatical knowledge. The diachronic perspective here offers a far more sophisticated picture of the mode of existence of a language. It is not just a conventional association between form and meaning, happening on some rather poorly understood hardware and offering a window on the nature of that hardware, it is one of the possible conventional associations that has a certain degree of stability due to the conditions under which language is transferred to ever new speakers, their ways of organizing these data, and the frequencies with which the various elements making up the association are used.

In particular, **Jäger** applies a bidirectional generalisation of Boersma & Hayes' (2001) learning algorithm to the formalisation and simulation of the grammaticalisation processes underlying case systems. He is able to show that structural case is the natural outcome of pragmatic case marking, that split ergative systems naturally evolve into nominative accusative systems and that some other systems are stable whereas others are either unlearnable or very unstable. The account also explains and underpins Aissen's (2000) treatment of differential case marking.

4 Problems and perspectives

The OT approaches to pragmatic phenomena seem to gain empirical advantages with respect to their non-OT predecessors, but that is not the only advantage. Important is the fact that we gain a different way of talking about these things in which uniformities can appear across the description of the different phenomena and that we have the prospect of a single theory of pragmatics where all the phenomena come together. This unification is still a prospect but there a number of issues that can already be distinguished.

The first issue is the existence of a pragmatic factorial typology. If there is a factorial typology, then it would fly in the face of the pragmatic tradition that has always maintained that pragmatics is universal and consists of a few principles that can be founded in the conceptual analysis of linguistic communication, as in Stalnaker (1999), Grice (1989), Sperber & Wilson (1986), Levinson (1983; 2000), Horn (1984; Horn, 2003) and others.

Is it really possible that a constraint CONSISTENT (sometimes treated as part of GEN) could be outranked by a constraint like ECONOMY OF EXPRESSION or RELEVANCE? This would mean that there could be communities where it is more

important to be economical than to be consistent with the context, or more important to be relevant than to be consistent with the context. In the first case, it would be not be possible to mark corrections, in the second, the interpretation process would maximize relevance without bothering about what we know already. It seems though that there is a general functional case for keeping corrections apart from consistent updates since the changes that have to be made to the knowledge of the interpreter are of quite different. Rerankings of this kind have to our knowledge not been found in the language communities of the world or only marginally (e.g., politeness can override sincerity).

If we succeed in agreeing on a universal system of ranked pragmatic constraints, there arises a second issue – a foundational one. It concerns the need not just for explaining why there are these constraints and no others and why they are ranked in this way. Because of the lack of variation, the factorial typology does not help to support an empirical argument that our system is correct. Possible strategies are the classical one of deriving the pragmatic system from pure reason, other strategies might try to use an evolutionary argument, which establishes that the pragmatic system is an evolutionary stable state by showing that any mutations (rerankings, small changes to the individual constraints) are eliminated and moreover that it is the only evolutionary stable state among a range of competitors. Of our contributors, **Van Rooy** and **Jäger** are following these different strategies, and it is one of the important questions of future research how to relate these different approaches (see van Rooy (forthcoming) for a first step in answering this direction)

The third issue is how to reconcile a universal pragmatics with the obvious fact that there is a great deal of variation in the syntactic, lexical and phonological expression of pragmatic properties in the languages of the world. It is an important insight that even if we have a pragmatic system, this does not mean that pragmatics

is purely universal. Languages exhibit enormous differences in their inventory of pragmatically relevant items, like in pronouns (for a basic typology, see Bresnan, 2001), tense and aspect, definite and indefinite markers, presupposition triggers, elliptical constructions, discourse particles. They also differ widely in their marking strategies for information structure. The richness of the data here is still largely unexplored especially in their interaction with the pragmatic treatments that have been in the focus of OT pragmatics. It is unclear to what extent these typological variations have a reflection on the abstract semantics. In Bresnan (2001), we see that Chichewa free pronouns (i.e. the closest analogon to English pronouns) do not allow antecedents that are topical, unlike English, where the pronoun predominantly refers to topical elements. The difference is that, in Chichewa, there is a class of bound pronouns realized in verbal agreement morphology that are used whenever the antecedent is considered to be a topic. Chichewa is not so different from French: he French clitic pronouns are used for topic, the free pronouns are used for the other cases. (These cases are not so easy to delineate.) The morphological distinction between zero, bound, clitic and free pronouns is not realized in all languages but seems to align in different ways with a prominence hierarchy on the antecedents. Whether this hierarchy is universal cannot be decided on the current state of research. The hierarchy itself may be universal, but it is clear from data in Gundel, Hedberg & Zacharski (1993) that zero pronouns do not align with the same property in the different languages that have them in their inventory.

For example, it cannot be decided yet whether pronoun resolution can be split up in a part to be treated in OT syntax and general pragmatic constraints on pronoun resolution. If one follows **Van Rooy**, the general principle is relevance. It would seem that for a particular treatment of e.g. Chichewa free pronouns, resolution would

need additional facts about the Chichewa inventory and the preference of bound pronouns for topical discourse referents.

A fourth issue is the nature of pragmatic constraints. One of the special features of the constraints that seem useful in pragmatics is that they seem like small OT competitions on their own. Consider a neutral, as far as we know original example, that is reasonably well understood, the resolution of ellipsis.

(11) Jan heeft een rode wollen trui gekocht en Piet drie blauwe.

“Jan has a red woolen sweater bought and Piet three blue”

The resolution process maximizes the similarity between the antecedent sentence and the ellipsed sentence. In a syntactic copying perspective, it copies the verb, the auxiliary, the object noun and one of the object adjectives. It does not copy the color adjective, the subject and the object determiner. It is clear that higher order unification, a tree assimilation algorithm, computation of the most specific common denominator, and source reconstruction – to mention only some of the techniques that have been applied to ellipsis – all attempt to make the ellipsed sentence as similar as possible to its antecedent. This can be naturally described as an OT competition⁶. The point is that constraint violations to a constraint MAXIMISE SIMILARITY must be scored by the existence of more similar candidates and that there is no alternative to that, since correctness of the resulting sentence misses out on the presence of optional material in the antecedent sentence, predicting e.g. that (12) is a correct interpretation even though the adjective *wollen* (“woolen”) is not taken along.

(12) Piet heeft drie blauwe truien gekocht.

“Piet bought three blue sweaters”

This seems the correct way to score DO NOT ACCOMMODATE, ECONOMY and RELEVANCE and STRENGTH, the main pragmatic constraints that people have come up with. We see whether there are otherwise correct interpretations with less discourse referents, otherwise correct sentences with less nodes and words that have the same interpretation in the context, or interpretations that deal with more questions that the interlocutor can be assumed to entertain.

In concluding these introductory remarks we want to stress once more that OT gives us a powerful instrument for implementing basic pragmatic mechanisms. However, one should not forget that having a hammer in one's hands may seduce one into seeing everything as a nail. For that reason methodological considerations for restricting the proper domain of OT applications in the area of pragmatics are important, and the significance of the three general conditions (A)-(C) of Section 1 deserves special attention in the area of pragmatics.⁷ On the other hand, we are at the beginning of deeper understanding of our instrument, which – unlike a real hammer – has proven to be helpful in quite different respects. It possibly will facilitate the integration of syntax, prosody, and pragmatics. It may allow to develop an evolutionary perspective showing that particular language traits emerge from pragmatics. And it may well provide a new framework research in psycholinguistics.

With any luck, the present volume helps to give a start.

Notes

-
- ¹ Roughly, this condition states that the subject position must not be empty.
- ² In this Introduction, names in bold type without a date refer to contributions to this volume.
- ³ See, for instance, Hayes & Hayes (1989) and Green (1990). Studies with chimpanzees have shown that they typically fail the symmetry test, but children older than two years pass it (Dugdale & Lowe, 2000).

It should be noticed that the first half of the equilibrium's condition – *what we produce we are able to understand adequately* – follows from the assumed initial state of the OT Grammar (the markedness constraints outrank the faithfulness constraint) plus the assumed mechanism of constraint demotion. In contrast, the second half of the condition – *what we understand we are able to produce adequately* – is independent of the initial state and an immediate consequence of the learning mechanism. In the more general case of learning arbitrary codes it needs extra requisites to ensure the symmetry condition. For example, it requires a particular asymmetry between expressive and productive optimization (see Zeevat, 2000; Jäger).

- ⁴ The origin of these ideas goes back to Blutner, Leßmöllmann, & van der Sandt (1996) and Blutner (1998).
- ⁵ A very similar point was made in functionalist phonology (e.g., Boersma, 1998). Most 'phonetically-driven' or functionalist theories of phonology propose that two of the fundamental forces shaping phonology are the need to minimise effort on the part of the speaker and the need to minimise the likelihood of confusion on the part of the listener. The need to avoid confusion is hypothesised to derive

from the communicative function of language. Successful communication depends on listeners being able to recover what a speaker is saying. Therefore it is important to avoid perceptually confusable realisations of distinct categories; in particular distinct words should not be perceptually confusable. The phonology of a language regulates the differences that can minimally distinguish words, so one of the desiderata for a phonology is that it should not allow these minimal differences, or contrasts, to be too subtle perceptually. There is nothing new about the broad outlines of this theory and it very closely relates to Zipf's (1949) two opposing economies (see also Lindblom, 1986, 1990; Martinet, 1955).

- ⁶ with different flavours of resolution arising by different data structures that have to be made as similar as possible, and the possibility of having different maxima to account for ambiguities. This is not the place to take a stance on the empirical and computational issues involved here.
- ⁷ Concerning the condition B, for instance, an interesting and new hypothesis is that the hierarchical encoding of constraint strengths is correlated with the effect of automaticity in psychological processes. Perhaps it is the area of pragmatics where this hypothesis can be tested in the most effective way.

References

- Aissen, J. (2000). *Differential object marking: Iconicity vs. economy*. Unpublished manuscript, Santa Cruz.
- Anttila, A., & Fong, V. (2000). The partitive constraint in optimality theory. *Journal of Semantics*, 17, 281-314.

-
- 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.
- Beaver, D. (to appear). The optimization of discourse anaphora. *Linguistics and Philosophy*.
- Beaver, D., & Clark, B. (2002). The proper treatment of focus sensitivity. In C. Potts & L. Mikkelsen (Eds.), *Proceedings of WCCFL 21* (pp. 15-28): Cascadilla Press.
- 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. (2002). Lexical semantics and pragmatics. *Linguistische Berichte*, 10, 27-58.
- Blutner, R., Leßmöllmann, A., & van der Sandt, R. (1996). *Conversational implicature and lexical pragmatics*. Paper presented at the AAAI Spring Symposium on Conversational Implicature, Stanford.
- Boersma, P. (1998). *Functional phonology*. The Hague: Holland Academic Graphics.
- Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, 32, 45-86.
- Bresnan, J. (2001). The emergence of the unmarked pronoun. In G. Legendre & J. Grimshaw & S. Vikner (Eds.), *Optimality-Theoretic Syntax*. Cambridge MA: MIT Press.
- Burzio, L. (1991). The morphological basis of anaphora. *Journal of Linguistics*, 27, 81-105.

-
- Burzio, L. (1998). Anaphora and soft constraints. In P. Barbosa & D. Fox & P. Hagstrom & M. McGinnis & D. Pesetsky (Eds.), *Is the best good enough?* Cambridge, Mass.: The MIT Press.
- Carston, R. (2002). Linguistics meaning, communicated meaning and cognitive pragmatics. *Mind and Language*, 17(1/2), 127-148.
- Carston, R. (2003a). Explicature and semantics. In S. David & B. Gillon (Eds.), *Semantics: A Reader*. Oxford: Oxford University Press.
- Carston, R. (2003b). Relevance theory and the saying/implicating distinction. In L. Horn & G. Ward (Eds.), *Handbook of Pragmatics*. Oxford: Blackwell.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Cole, P. (Ed.). (1981). *Radical pragmatics*. New York: Academic Press.
- de Hoop, H. (2000). Optimal scrambling and interpretation. In H. Bennis & M. Everaert & E. Reuland (Eds.), *Interface Strategies* (pp. 153-168). Amsterdam: KNAW.
- de Hoop, H., & de Swart, H. (2000). Temporal adjunct clauses in optimality theory. *Rivista di Linguistica*, 12(1), 107-127.
- Dekker, P., & van Rooy, R. (2000). Bi-directional optimality theory: An application of game theory. *Journal of Semantics*, 17, 217-242.
- Ducrot, O. (1980). *Les Echelles argumentatives*. Paris: Minuit.
- Dugdale, N., & Lowe, C. F. (2000). Testing for symmetry in the conditional discriminations of language-trained chimpanzees. *Journal of the Experimental Analysis of Behavior*, 73(1), 5-22.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition*, 28, 3-71.

-
- Gibson, E., & Broihier, K. (1998). Optimality theory and human sentence processing. In P. Barbossa & D. Fox & P. Hagstrom & M. McGinnis & D. Pesetsky (Eds.), *Is the Best Good Enough. Optimality and Competition in Syntax* (pp. 157-191). Cambridge, Mass.: The MIT Press.
- Green, G. (1990). Differences in development of visual and auditory-visual equivalence relations. *Journal of the Experimental Analysis of Behavior*, 51, 385–392.
- 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.
- Grice, P. (1989). *Studies in the way of words*. Cambridge Mass.: Harvard University Press.
- Grimshaw, J., & Samek-Lodovici, V. (1995). Optional subjects and subject universals. *University of Massachusetts Occasional Papers in Linguistics*, 18, 589-605.
- Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69(2), 274-307.
- Hale, M., & Reiss, C. (1998). Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry*, 29, 656-683.
- Haspelmath, M. (1999). Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft*, 18, 180-205.
- Hayes, S. C., & Hayes, L. J. (1989). The verbal action of the listener as the basis of rule governance. In S. C. Hayes (Ed.), *Rule governed behavior: Cognition, contingencies, and instructional control* (pp. 153-190). New York: Plenum Press.

-
- Hendriks, P., & de Hoop, H. (2001). Optimality theoretic semantics. *Linguistics and Philosophy*, 24, 1-32.
- 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. (2003). Implicature. In L. Horn & G. Ward (Eds.), *Handbook of Pragmatics*. Oxford: Blackwell.
- Hyman, L. (1984). Form and substance in language universals. In B. Butterworth & B. Comrie & Ö. Dahl (Eds.), *Explanations for Language Universals* (pp. 67-85). Berlin: Mouton.
- Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, Massachusetts: MIT Press.
- Jacobson, R. (1962). *Selected writings 1: Phonological studies*. The Hague: Mouton.
- 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.
- Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk et al. (Ed.), *Formal methods in the study of language*. Amsterdam: Mathematisch Centrum.
- Kaplan, D. (1979). On the logic of demonstratives. *Journal of Philosophical Logic*, 8, 81-89.

-
- Katz, J., & Fodor, J. A. (1963). The structure of semantic theory. *Language*, 39, 170-210.
- Levinson, S. (1983). *Pragmatics*. Cambridge: CUP.
- Levinson, S. (1987). Pragmatics and the grammar of anaphora. *Journal of Linguistics*, 23, 379-434.
- Levinson, S. (2000). *Presumptive meaning: The theory of generalized conversational implicature*. Cambridge, Mass.: MIT Press.
- Lindblom, B. (1986). On the origin and purpose of discreteness and invariance in sound patterns. In P. J. S & D. H. Klatt (Eds.), *Invariance and Variability in Speech Processes* (pp. 493-510). Hillsdale, N.J.: L. Erlbaum.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H & H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403-439). Dordrecht: Kluwer.
- Martinet, A. (1955). *Économie des Changements Phonétiques: Traité de Phonologie Diachronique*. Berne: Éditions A. Franke.
- Montague, R. (1970). Universal grammar. *Theoria*, 36, 373-398.
- Prince, A., & Smolensky, P. (1993). *Optimality theory*. Rutgers Center for Cognitive Science: Technical Report RuCCSTR-2.
- Schwarzschild, R. (1999). GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics*, 7, 141-177.
- Sperber, D., & Wilson, D. (1986). *Relevance*. Oxford: Basil Blackwell.
- Stalnaker, R. (1999). *Context and content: Essays on intentionality in speech and thought*. Oxford: Oxford University Press.
- Tesar, B., & Smolensky, P. (2000). *Learnability in optimality theory*. Cambridge Mass.: MIT Press.

Van Rooy, R. (2001). Conversational implicatures. In J. v. Kuppevelt & R. Smith (Eds.), *Proceedings of the 2nd SIGdial Workshop on Discourse and Dialogue*. Aalborg.

van Rooy, R. (forthcoming). Signalling games select Horn strategies. *Linguistics and Philosophy*.

Wurzel, W. U. (1998). On markedness. *Theoretical Linguistics*, 24, 53-71.

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

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge: Addison-Wesley.