

The Asymmetry of Optimality Theoretic Syntax and Semantics

Henk Zeevat

Abstract

This paper argues for a combination of semantics and syntax in an optimality theoretic framework that avoids the rat/rad problem and provides simultaneously a certain amount of bidirectionality, in the spirit of Blutner for an approach to ineffability. It can be succinctly described as taking the program of optimality theoretic syntax as basic, also as a theory of interpretation, and extending it with a bidirectional pragmatic component that is closely related to existing ideas about natural language interpretation. The paper argues for the priority of the direction from content to form, develops the pragmatic component, and argues for the bidirectionality of the pragmatic component on the basis of Grice's principle of cooperation. It applies the resulting theory to a small set of relevant examples. The asymmetry in the title is consistent with, but goes beyond, the asymmetry between syntax and semantics used in Smolensky (1996).

1 OT Semantics and Syntax

Optimality theoretic syntax (OT syntax) is the proposal to think of the knowledge of natural language syntactic structures as an ordered sequence of constraints that decide which are the best candidate sentences for expressing some given content¹ (the input). Optimal candidates are the ones that do better on the ordered constraints than all the other competing candidates. S_1 is a better candidate than S_2 if there is a strongest constraint C such that S_1 and S_2 do equally well on the constraints that are stronger than C but S_1 does better on C itself. Moreover, OT syntax makes the following assumptions. First, the set of constraints is the same for all languages, but languages differ in the ordering of the constraints. Second, constraint satisfaction is scored discretely. Both of these assumptions can be given up in principle without changing the essence of the theory as a descriptive device for a particular language, but they have an

¹Though this plays only a minor role in the argument, I wish to make clear my assumption that content is a semantic representation in some suitable logical formalism against the background of discourse context representing the common ground and the current discourse situation. The semantic features referred to by the constraints can therefore equally well be properties that the object identifiers have in virtue of their role within the discourse context. This goes against some proposals for the input, which favour underspecified representations or even quasi-syntactic inputs.

important methodological value since the first assumption militates against language particular constraints and the second keeps the theory formally simpler. Though there is as yet no consensus about a particular set of constraints for syntax, there is a lot of promising work going on in the area, like e.g. Grimshaw (1997), Choi(1998) and Bresnan(MS).

OT syntax suffers from a problem. The prediction —which arises from the formal conception itself— is that for any input there is a set of optimal candidates, i.e. any content can be expressed. This prediction is easily refuted by showing that some sentences are untranslatable. For example (1)

(1) Who ate what?

is a proper English sentence but does not have an Italian translation, like (2).

(2) *Chi ha mangiato che cosa?

It is a natural assumption that the input of the English sentence is also available to Italian language users. Yet, there does not seem to be an Italian form (except complicated paraphrases) that expresses this input. This problem is known as the *ineffability* problem. Contrary to what OT syntax predicts, not everything can be said in any language. The same problem has been noticed by Pesetsky(1997) using ungrammatical sentences that do not allow repair.

Optimality Theoretic Semantics (OT semantics) is a more recent enterprise in which the traditional methods for natural language interpretation are replaced by systems of ordered constraints. Given the problems that natural language semantics faces, this is a natural and wise move and has led to interesting approaches to *when*-sentences (De Hoop & De Swart (2000)) and to presupposition (Blutner (2000)). But there is a natural question to ask about the enterprise as such. If there is an OT semantics, how is it related to OT syntax? It is clear that we do not want a conflict: the OT semantics should not assign an optimal interpretation to a sentence for which the sentence is not optimal according to OT syntax. And also we do not want the OT syntax to assign a sentence to the input that does not have the input as an optimal interpretation. The problem is that both OT syntax and OT semantics are complete theories about the relation between form and content and it would therefore seem that they cannot be independent of each other.

Blutner has pioneered a first version of bidirectional OT which overcomes these problems. In his conception of superoptimality there is a single ordered set of constraints that regulates the relation between form and content. But the constraints are used twice: a pair $\langle Form, Content \rangle$ is superoptimal iff there is no better pair $\langle Form_1, Content \rangle$ and no better pair $\langle Form, Content_1 \rangle$. In weak superoptimality —the notion he really favours— we find also some recursion: A pair $\langle Form, Content \rangle$ is weakly superoptimal iff there are no weakly superoptimal better pairs $\langle Form_1, Content \rangle$ or $\langle Form, Content_1 \rangle$.

Both of these notions are highly interesting and lead to important results, like a solution of the ineffability problem and treatments of presupposition and lexical semantics. But superoptimality labours from its essentially symmetric character. One prediction that can be derived from weak superoptimality is that both synonymy and ambiguity are dying phenomena in natural languages: they tend to disappear. Now it is true that synonymy is not a stable phenomenon. It is a linguistic common place that “real” synonymy does not exist. Though debatable, the point about synonymy can certainly be defended and it seems the sort of fact that needs explanation in the kind of theory that we are discussing. But ambiguity seems ever on the increase. It is the major problem for computational linguistics and a remarkable ubiquitous and robust phenomenon. Moreover, it increases whenever a language loses phonological, morphological or configurational properties, i.e. almost whenever language change occurs.

The OT literature also contains a formal argument against the symmetric view: the rat/rad problem. The Dutch word *rat* (meaning rat) is homophonous with the Dutch word *rad* (meaning wheel) in its singular form. The pronunciation of *rad* (but not *rat*) is derived by a faithfulness violation: the underlying feature +voiced is lost at the end of Dutch words. In a treatment like Blutner’s, this has consequences for the interpretation of the sound /rat/. If it is interpreted as *wheel* there is a better form content pair, namely $\langle /rat/, rat \rangle$. According to both notions of superoptimality, this means that $\langle /rat/, wheel \rangle$ is thrown out of the competition, not just in interpretation but also in generation.

The rat/rad problem is a simple phonological problem, but it would arise in any ambiguity where in one of pairs $\langle Form, Content_1 \rangle$, $\langle Form, Content_2 \rangle$, *Form* is in one case derived by more serious syntactic constraint violations than the other. A simple case is perhaps² (3) assuming that (b) involves two violations of the constraint **STAY** enforcing constituents to stay in their canonical position rather than (at most) one as in (a).

- (3) a. Wie slaat Hans?
 a'. Who beats Hans?
 b. Wie slaat Hans?
 b'. Whom does Hans beat?

Superoptimality would predict not just that reading (a') is preferred but that it is the only reading, of course under the assumption of the analysis in terms of **STAY**. This does not match the facts of Dutch. There is a preference for reading (a'), but the other is also available. More serious than this particular example is the fact that given any particular syntactic system of constraints, examples of this kind can be found at will.

This paper is an attempt to develop a competing theory of the combination of syntax and semantics in optimality theory, which maintains as much as possible of the insights of Blutner, while avoiding the problems. It moreover aims at

²I want to remain strictly uncommitted to any syntactic analysis in this paper. Not in life.

being a naturalistic theory of these matters, i.e. a theory that can be interpreted as constraining actual processes of language production and interpretation.

It has been my view for a long time that the asymmetry between speaking and listening should be taken more seriously than theories generally do. Different parts of the body are involved and there can be vast differences between what people can say and what they can understand. Moreover listening and speaking differ in their very nature. Speaking is an active process in which the speaker has control, whereas listening is essentially a passive activity, in which the listener tries to make the most of the signal she receives. Equally important is the naturalistic character of an optimality theoretic account of speaking or understanding. OT was inspired by the consideration of processes in the brain and still derives much of its psychological plausibility from its interpretation as a theory about brain processes. A theory of the relation between form and content should therefore primarily be a theory of speaking and understanding, as these are the processes in which the brain uses the constraints. According to Smolensky(1996), the naturalistic interpretation still does not give a theory of the actual processes in performance (which would involve other mechanisms as well) but only a description of the grammatical norm. Therefore, naturalism here only means that we can think of the theory as a part of an overall account of the actual production and understanding mechanisms.

In the next two sections, the paper explores some general reasons for assuming that OT syntax is the basic theory. They are far from being conclusive arguments but they make that view plausible. That an OT syntax is needed at all (but possibly in conjunction with OT semantics) follows from the phenomenon of semantic blocking. For semantic blocking, see Zeevat (2000) and Bresnan (1998).

Section 4 tries to make it clear that interpretation cannot be handled by OT syntax on its own, because certain necessary constraints do not allow a proper reformulation in OT syntax. A minimal system of interpretation constraints that cannot be reduced to syntactic constraints is developed and defended. Sections 5 and 6 discuss the way in which syntax and semantics are connected. Section 6 applies the resulting theory to some key problems.

2 Chicken or Egg

What did evolution achieve when it created language? I think the right answer is the creation of a system of forms in which contents can be coded. Though the creation of the forms doubtlessly helped extend the richness of the contents that can be expressed by means of them, nothing suggests that the everyday thoughts we have and that we routinely transmit to our fellow humans are that different from the thoughts of somebody who lacks language or even from the thoughts of our closer biological cousins. After all, our basic drives are the same and so is the information we gather in order to satisfy these drives.

The wrong answer is surely that evolution created a stronger power of understanding that allows us to make sense of the complex contents expressed by the forms found in natural languages. This is the wrong answer if we assume that the new power of understanding is prior or independent of the creation of the system of forms. I do not think the system of understanding had to adapt very much. Already before language evolved, it was possible to interpret the behaviour of other humans and of animals and to interpret the environments. These are the hard problems, not language understanding³. Understanding limits the diversification of the production of acoustic signals: if a differentiation cannot be cashed in by a corresponding differentiation in understanding, it is not functional and will not become part of the language. The development of language use therefore can not be understood in isolation from the process of decoding the language tokens. But the biological achievement is the differentiation of the acoustic signals, which in combination with the recognitional and understanding capacities of the producing organisms use the differentiation to a biological advantage. Nothing rules out that the understanding powers grow as a consequence of the development of language, and that this growth then allows further differentiation in understanding for which new forms are developed. But the initiative is on the side of the forms.

This can be underpinned to some extent by physiological considerations. Whereas the ear is largely what it was before language as we know it, there are physiological changes in the larynx and in the way it is used.

The point of these remarks is that, as linguists interested in the nature of language, we should be primarily concerned with the production of language and develop theories of the production process. Producing language would not make sense without understanding, but it is not clear that the understanding needed to develop that much. It seems to follow that if we want to develop an empirical theory of the relationship between forms and meanings we find in natural languages we should be primarily concerned with the direction that goes from meaning to form. The other direction is like other perception problems where one reasons from a perceptual content to its causes.

The argument of this section is speculative and can at most underpin a certain bias towards the primacy of OT syntax and against an independent OT semantics. The argument in the next section has more substance, though it is not compelling either.

³Work on visual languages, especially Marriott & Meyer (1998) shows conclusively that going from the discrete to simple graphical diagrams leads to an immediate explosion of computational complexity. A tiger has to make predictions about the behaviour of its prey, birds need to orient themselves in their treks and all these tasks are seriously more complex than context-free parsing. Recursive structure, e.g. squares within squares, arises as much in the visual field as in natural languages.

3 Conflict in Production and Understanding

Following Boersma (1998), we can make the following observations. As in the production of speech, the production of sentences stands under two opposing principles. The first principle (expressiveness) is that the receiver of the sentence should be able to take out the message that the speaker has coded into the sentence. That is after all the purpose of language use. This goal is served by marking every semantically relevant property of the input by some syntactic feature, such as morphology, word order, lexical items etc.

At the same time, the speaker stands under a principle of minimal effort. There is no point in marking a feature that is inferable and often the available means of marking will be conflicting. The requirements conflict and the optimal realisation is a particular way of solving the conflict. The OT syntactic constraints reflect economy and expressiveness requirements and their ordering is the standard conflict resolution mechanism adopted by a language.

It is not clear that in interpretation the same conflict between different interests of the interpreter repeats itself. If the interpreter wants to minimise her effort, she runs the risk of not finding the speaker's intention. Of course, it does not pay off to put in more effort than is needed to recognise the speaker's intention, but economising on effort cannot go below the effort required, on the pain of disfunctioning.

There is of course the same principle of expressiveness: everything that is in the signal must be interpreted. But there does not seem to be a conflict between doing that to the maximal extent and the principle of not doing more than is required to find the speaker's intention.

∴ From this, I want to conclude that whereas there is a naturalistic interpretation of conflicting constraints in language production, there is no such naturalistic interpretation for conflicting constraints in interpretation. If there are conflicting constraints in language interpretation they must derive from constraints about language production⁴.

The situation can be fruitfully compared to the habit of hiding easter eggs for one's children on Easter Sunday. The parents engaged in hiding the eggs balance the amount of effort with the desired amount of difficulty in finding the egg. (They also picture the child looking for it and try to keep it possible for the child of finding the egg, without spoiling the fun.) For the child it is another matter. It just has to throw in the effort required for finding the eggs. Not more of course, but definitely not less. It is not a complicated balancing

⁴Later on I defend some defeasible interpretation constraints that are ordered with respect to each other. I accept the conclusion that their ordering is not a result of language users learning how their language resolves a conflict between opposing principles. This conclusion is also unavoidable given that nothing indicates that different languages could have them in different orderings. In fact, bizarre concepts of language use result if one tries alternative orderings, which seems to indicate that we should look for rational rather than empirical explanations of the ordering.

act.

This would be the argument that shows that the process of language production has to find a balance between conflicting constraints. Languages are an inventory and a conventionalised way of establishing the balance: the language particular ordering of the constraints. A similar argument for underpinning this balance in understanding cannot be given.

If it could be shown that the task of interpreting is in fact always an instance of an ongoing hermeneutic process of refinement — as some would perhaps argue—, the situation changes. Assuming the existence of a process of ongoing refinement, it is indeed possible to argue for a conflict between economy of effort on the one hand and the need for quick results on the other. It is however not easy to see how the semantic constraints that have been proposed can be seen as embodying a compromise between these conflicting needs.

My intuition also tells me that the hermeneutic circle is normally quickly closed. The communication of everyday thoughts (What time is it? Give me a coffee! Do you have something to eat?) quickly results in the grasp of the speaker intention. Negative feed-back can result in further reflection, but unprompted further reflection is pointless once a plausible and relevant speaker intention is found.

4 Proper Optimality Theoretic Semantics

The previous two sections may be read as arguments against assuming that there should be an OT semantics in addition to or side by side with OT syntax. My prejudice has in fact always been that there should not be a separate OT semantics. The proposed constraints of OT semantics and their ordering are really syntactic constraints in disguise and their ordering is the ordering of the disguised syntactic constraints. I tried to show the plausibility of this view by reconstructing the analysis of *when*-clauses of De Hoop & de Swart (2000) within OT syntax (Zeevat (2000)). But my plan of showing this ran up against the problem that there are some interpretation constraints that do important work and do not appear to allow a reformulation as syntactic constraints.

These are the ones I know about: ***ACCOMMODATION**, ***INVENT**, **STRENGTH**, **ANCHOR**, **CONSISTENCY** and **FAITH-INT**. I do not expect there to be many others and these ones also seem to form an interesting natural class, as I will try to show at the end of this section. I should also say at this point that my general solution does not depend on the question of which semantic constraints must be assumed or on the formulation of those constraints. The only requirement is that there should be some, otherwise the theory collapses into optimality theoretic syntax. Though the precise content of the system is not essential, I believe the system I present in this section has some independent merit.

The first constraint is ***ACCOMMODATION**. It (fallibly) prohibits accom-

modation of the antecedents of presupposition triggers. A presupposition trigger such as *regret* requires that its complement is already true in the context in which it is used. If it is not true, the content of the complement needs to be added to the context, a process called accommodation. Nothing should be added if the context (or one of the local contexts) already has the material and ***ACCOMMODATION** does just that.

I cannot imagine anything in syntax that has the effect of ***ACCOMMODATION**. It cannot be a prohibition against using the trigger in a context that does not have the antecedent: that occurs frequently and appropriately. If one wants, ***ACCOMMODATION** can be taken as a special case of a principle that forbids us to add material to the context of the utterance or to the content of the utterance without a proper reason (like external evidence or the material supplied by the sentence). ***INVENT** seems a good name for such a constraint. It is quite unclear how the speaker can rule out this bad behaviour of the listener by adding some feature to the sentence. For example, ***INVENT** forbids us to start thinking that John is ill, if all that the speaker said is that Mary had an icecream. It is the principle that asks us not to overinterpret.

STRENGTH expresses the preference for informationally stronger readings of the sentence. It is the odd man out here, because it does not seem to allow a discrete evaluation measure and also makes a couple of wrong predictions, as Geurts (2000) has pointed out. Nevertheless, a version of **STRENGTH** is needed for the interpretation of presupposition triggers and —as Peters et al. (1998) have argued— for the interpretation of reciprocals. It is obvious that there is no generation principle that can capture the effect of **STRENGTH**. From the generation perspective, it seems that the weaker inputs that **STRENGTH** rules out as an interpretation will nevertheless be optimally realised by the sentence.

ANCHOR is the principle that interpretations should be anchored. In essence, this means that all the pronouns, ellipses, tenses and topics should find proper antecedents and that a discourse relation must be constructed from the current sentence to the appropriate earlier element of the discourse or dialogue. Accommodation occurs because of the needs of **ANCHOR**. There is something in generation that corresponds to this: the principles that select the proforms, ellipsed versions, the presupposition triggers, the topic focus articulation and connectives based on the speaker's estimate of the context. It seems that **ANCHOR** can be reformulated as a syntactic principle that prevents the choice of a reduced form (a pronoun, an ellipsis, deaccented pronunciation, zero connective) when this is not appropriate. In principle, we could have a generation principle ***REDUCE** that prevents such reductions when the context does not licence them. (***REDUCE** would have to be ordered below the constraints that force the reductions.) But as will become clear, it suffices to have **ANCHOR** to get this effect and that seems the more natural choice. An additional argument is that reduced forms are not really required in the inventory of the language. E.g. the Latin *homo* (*a man* or *the man*) is not reduced with respect to the indefinite when it means *the man* and has a linguistic antecedent. Lan-

guages that have no reduced forms are just be less efficient for the generator. The interpreter would still be trying to identify as much material as possible in the preceding context or relate objects by bridging and discourse relations, by the principle **ANCHOR**. Without reduced forms, there is no effect of ***REDUCE** in generation and consequently there would be no reason for finding an antecedent.

CONSISTENCY prefers interpretations that do not conflict with the context. It plays a role in ambiguity resolution, selecting between different resolutions of anaphoric elements and in ruling out certain accommodations. It can be violated, since it is certainly possible to contradict the given context. Once more, there is no good generation constraint that rules out the expression of thoughts that contradict the context. It can just be done and the OT syntax tells us what is the best way of doing it. In certain cases, there is obligatory marking of inconsistency, using contrastive and concessive devices. Overt corrections have a number of syntactic features that make them recognisable. A language that would however not have such syntactic devices —or that does not always mark inconsistency with the context— does not seem impossible. Obligatory marking of inconsistency is therefore not an alternative to the assumption of **CONSISTENCY** as a principle.

For these violations of consistency, we need the principle of faithful interpretation **FAITH-INT**. This principle forces us to interpret all that the speaker has said. **FAITH-INT** could in principle⁵ be a generation constraint (“do not mark any features that are not in the input”) but the positive formulation is an interpretation constraint and that makes it more natural to think of it as one. In the scheme I am presenting in the next two sections, the principle is superfluous: it is captured by the first step of recovering the set of inputs that could lead to the sentence.

The ordering between the constraints is also fairly obvious. Readings can be inconsistent with the context if they are faithful and accommodation is only allowed because of the need to anchor. Accommodation is restricted to consistent additions to the context and selects the strongest reading when different ones are possible. This is just a rephrasing of the standard views on presupposition accommodation.

This gives us the following picture of what —if I am right — is the whole of OT semantics.

FAITH-INT >
CONSISTENCY, ANCHOR >
***INVENT, *ACCOMMODATION** >
STRENGTH

An example illustrating **FAITH-INT** > **CONSISTENCY, ANCHOR** is (4).

⁵As Bresnan does as part of the faithfulness constraint in Bresnan (MS).

- (4) A: John hates Bill.
B: He hates SUZY.

The second sentence, interpreted as a correction, violates consistency. Corrections would be impossible, if the ordering were reversed. The same example also illustrates that **ANCHORING** is not weaker than **CONSISTENCY**. If it were, the pronoun used by B. could not refer to *John*. So we have **CONSISTENCY, ANCHOR**.

***ACCOMMODATE** explains the contrast between (5a.) and (5b.) .

- (5) a. If John is in Berlin, he regrets that he is in Berlin.
b. If Mary is in Amsterdam, John regrets that he is in Berlin.

The (b.) example entails that John is in Berlin, but not the (a.) example, due to the presupposition trigger *regret*. In the (a.) example the presupposition is resolved to the condition of the implication, in (b.) that is not possible and the only interpretation is obtained by anchoring the trigger through the addition of the presupposition to the main context. This addition is ruled out in (a.) by ***ACCOMMODATE**. Addition to the context given by the condition in (b.) is ruled out by **STRENGTH**, as the resulting interpretation would be entailed by the addition to the main context. From this it follows that **ANCHORING > *ACCOMMODATE** and further that ***ACCOMMODATE > STRENGTH**.

It should be clear that without support from OT syntax the semantics given by these principles is unable to interpret any sentence whatsoever. But OT syntax exists and how it is integrated with the semantic constraints is the subject of the next two sections.

There is however one more aspect of the system that should be pointed out. It turns out to be no more than an OT reformulation of the essence of the received interpretation theory from the '70s of the last century. There we had the compositional semantics of Lewis(1970) and Montague(1974), supplemented with Karttunen(1973, 1974) and Stalnaker(1978)'s ideas about presupposition and assertion. In the '80s these have been supplemented by establishing that anaphoric resolutions and discourse relations can be best thought of as special cases of presupposition.

The combination of **FAITH-INT** and ***INVENT** restores important aspects of compositional semantics (not the full principle, but essential aspects). The combination of **CONSISTENCY** and **STRENGTH** are (a strengthening of) Stalnaker's principles of assertion and **ANCHOR** and ***ACCOMMODATION** together give a reconstruction of the field of discourse, including insights from discourse representation theory (e.g. Kamp(1982) and Heim(1982)) and the analysis of presupposition Heim(1983), Van der Sandt(1992)). The set of constraints itself is almost nothing more than the received theory. My proposal adds to the received theory by ordering the constraints and by allowing

exceptions. It is extremely unlikely that there would be reasons for changing the constraints and their ordering if one moves from language to language.

What is missing is not the rational argumentation for the constraints (that argumentation is just part of the literature) but the rational argumentation for their ordering. It is fairly clear from the empirical point of view that the ordering is as I sketched above. It seems that it is not hard to see that alternative orderings lead to problems. E.g. if **CONSISTENCY** would be weaker than **STRENGTH** we would be hunting for strong but false interpretations whenever that is possible, which does not seem a good idea. Or if **FAITH-INT** were weaker than **CONSISTENCY** we could not correct each other. Given the communication protocol that we seem to have adopted, alternative orderings would lead to a loss of functionality. A proper rational foundation is however a complicated matter. It should show in detail why each of the constraints is there, why each ordering statement must be there and, importantly, why it is rational to have defeasible constraints, etc. This task must be deferred to future work.

This section presented the case for preserving some OT semantics in the face of the criticism that OT semantics is not necessary or desirable given OT syntax. It is only a modest semantics that remains. In the next two sections I will only assume that OT semantics is a system of constraints that help us in deciding between the different readings predicted by OT syntax. I refer to that system as the interpretation constraints.

5 The Basic Connection

The prediction of OT syntax is that an optimal interpretation of a sentence S is any semantic input I that beats its competitors among the candidate set $\{ \langle S, J \rangle : J \text{ is a semantic input} \}$ by the system consisting of the normal syntactic constraints and their ordering. Smolensky (1996) points out that the winner of the interpretation competition for a sentence S is not necessarily going to be optimally generated as S by the same system and thereby explains observed asymmetries between production and generation in child language, since the competition in the other direction involves the different candidate set $\{ \langle S, I \rangle : S \text{ is a syntactic form} \}$.

Given what we have done so far, we can define the optimal interpretation of a sentence $Form$ in two steps. First we take our OT syntax system and determine the set $\{ Content : Form \text{ is an optimal form for the content } Content \}$. In a second step, we determine which of the elements of that set optimally satisfy the interpretation constraints. Those are then the best interpretations.

This can be understood as the evaluation of pairs $\langle Content, Form \rangle$ over two systems of constraints: the syntax constraints $G = CG_1, \dots, CG_n$ and the interpretation constraints $I = CI_1, \dots, CI_m$. The fact that we first take the set $\{ Content : Form \text{ is an optimal form for the content } Content \}$ orders

the interpretation constraints after the generation constraints, if we take both constraints as constraints on pairs.

In the table below, the evaluation starts with all pairs in which *Form* is the input. The optimal pairs are found before the evaluation by the semantic constraints begins and form the set *GEN* for semantic evaluation. The pairs that are optimal by the generation constraints give the optimal interpretations of *Form*.

| | | |
|-----------------------------------|---------------------|---------------------|
| $\langle Content_1, Form \rangle$ | CG_1, \dots, CG_n | CI_1, \dots, CI_m |
| . | | |
| . | | |
| $\langle Content_j, Form \rangle$ | | |
| . | | |
| . | | |
| $\langle Content_m, Form \rangle$ | | |

Since the generation and interpretation constraints form disjoint systems we have no problem with harmonizing between the interpretation and the generation process.

We can assume that an interpreter proceeds in this way (in an efficient implementation of it). But it is not wild to assume that the speaker does the same. Why say something knowing that it will be understood in the wrong way? It is also standard in natural language generation systems to check that the semantic representation from which generation started also comes out when the generated sentence is interpreted. One can even wonder whether a natural language speaker who —after all— is also a natural language understander can avoid interpreting her own words.

This basic system already suffices for an explanation of the ineffability problem: ineffable contents are those whose optimal realisation is misinterpreted by the interpretation constraints. I will give a more subtle account of ineffability later on.

6 Cooperativity

An important aspect of pragmatics we did not incorporate so far is Grice's principle of cooperation⁶. Language use is a special kind of cooperative behaviour and the speaker has a cooperative obligation when she speaks. In particular, the speaker has a responsibility for what the listener will make of her sentence.

⁶Charity of the interpreter is coded in the interpretation principle of consistency with the context and in the principle of going for the most informative reading. But this is only one aspect of cooperativity.

That makes it plausible to assume that the speaker goes through the interpreter's part of the process and makes sure that at least she would get the interpretation she intends. But there is something more to it. The speaker can make sure that interpretation is as painless as possible by avoiding violations of the interpretation constraints⁷. This gives us the following picture ($G + I$ is the system of generation constraints followed by the system of interpretation constraints):

Form is an optimal generation for *Content* iff

- a. $\langle \textit{Content}, \textit{Form} \rangle$ is optimal for $G + I$ in the set $\{\langle \textit{Content}, \textit{Form}_i \rangle: \textit{Form}_i \text{ an arbitrary form}\}$ and
- b. there is no pair $\langle \textit{Content}_j, \textit{Form} \rangle$ that is better by I than $\langle \textit{Content}, \textit{Form} \rangle$ is. ($\textit{Content}_j$ must come from the set of optimal inverses of the set of forms obtained in (a), but this is not essential).

And the definition of an optimal interpretation must be independent, but similar.

Content is an optimal interpretation for *Form* iff a. $\langle \textit{Content}, \textit{Form} \rangle$ is optimal for $G + I$ in the set $\{\langle \textit{Content}_i, \textit{Form} \rangle: \textit{Content}_i \text{ is an arbitrary form}\}$ and

- b. there is no pair $\langle \textit{Content}, \textit{Form}_i \rangle$ that is better by I than $\langle \textit{Content}, \textit{Form} \rangle$ is. (\textit{Form}_i must come from the set of optimal inverses of the set of contents obtained in (a), but this is not essential).

In generation, we carry out the basic combination first and then survey as interpreters the range of other interpretations of the form we found. In interpretation, we carry out the basic combination first and then survey as generators the range of other forms for the thought we found.

This is what people seem to do when they carry out the task of generating from a fixed content, like e.g. in literary translation. Real generation is probably better understood as a process starting from an only partially specified content.

A succinct formulation of the system is to say that we first do normal OT syntax and —after that— superoptimality over the interpretation constraints. The cooperativity of the speaker gives us superoptimality in the semantics.

The advantage of cooperativity is that we keep some of the effects of Blutner's bidirectionality. In particular, we preserve Blutner's theorem which offers revolutionary insights in the analysis of presupposition triggers, at least if you want to believe Zeevat (1999) or Zeevat (2000).

We also get a diagnosis for what is wrong with full superoptimality. In super-

⁷I am not sure of my equation of pain and constraint violation, but it is a natural idea. At least in syntax, it should be testable whether there is a relation between understanding times and the amount of constraint violation that goes on in sentences. Certainly the violations of the interpretation constraints that are the standard examples in the presupposition literature are not easy to understand.

optimality, it is not just the speaker that is cooperative, but also the listener. The listener must select a reading taking into account the effort of the speaker: the reading is deselected if the speaker has to violate a stronger constraint or the same constraint more severely for it than for another reading. But that does not make sense at all. The speaker will just spend the effort to express the content in question and the listener does not have the control necessary to reduce the speaker effort.

7 Applications

Rat and Rad

The last point of the last section is the solution to the rad/rat problem.

From the interpretation point of view *rad* (wheel) and *rat* (rat) are equally good interpretations for /rat/. Neither incurs a mark by any of the interpretation constraints. The mark occurred in the generation component is unimportant once /rat/ has become the optimal realisation of *rad* and *rat*.

The same applies to my syntactic version of the rad-rat problem. After *Wie slaat Hans?* has become the optimal realisation of both $?x \text{ beat}(x, Hans)$ and $?x \text{ beat}(Hans, x)$ the **STAY** violations become irrelevant.

Italian WH-phrases

Let us assume that Italian wants its *WH*-phrases fronted, i.e. it has strong constraint **FRONT-WH**, i.e. which is violated by *WH*-phrases that are not in the first position. Let us also assume that it wants to mark semantical *WH*-phrases (variables bound by the question operator) by the typical morphology of *WH*-phrases, but not as much as it wants to front them. This means that we have a constraint **PARSE-WH** that is weaker than **FRONT-WH**.

It then follows that the optimal candidate for $?xy \text{ eat}(e, x, y)$ is something like (6) (assuming *qualcosa* is the default NP of Italian).

- (6) Chi ha mangiato qualcosa?

The *WH*-constituent is fronted and the subject, but not the object is *WH*-marked. The object therefore violates **PARSE-WH**, but the damage is smaller than marking it and violating **FRONT-WH**. The generation competition gives—as always—an optimal candidate.

But in interpretation, by ***INVENT** the semantic correspondent of the *WH*-feature cannot be recovered. That means that the optimal candidate is in fact not a good expression of the input. It wins the syntactic competition, but its *WH*-interpretation always loses out the interpretation contest.

Killing and Causing to Die

We lose the ability to predict the semantic difference between *kill* and *cause to die* in this framework. A use of *kill* tends to be interpreted as a “standard killing” while *cause to die* indicates that the killing is indirect, or at least non-standard. Blutner⁸ explains this selection of meanings with weak superoptimality, using only the way generation and interpretation are combined.

It is a pity we lose this explanation, but there is no reason for despair because a simple alternative explanation is available. Let us assume that there is an **ECONOMY** constraint active in the OT syntax. This constraint militates against long and infrequent ways of expression. If the sheriff killed Bill in a normal way, **ECONOMY** will prevent the selection of *cause to die*.

For the interpreter, that means that the interpretation *kill* is not available for the form *cause to die*. That form is not a survivor, since for simple killing *kill* must be used due to the **ECONOMY** constraint.

Suppose that we also have a (stronger) constraint **PARSE-MARKED** which requires a marked way of expression when an input item is semantically marked, i.e. it belongs to the extension of a certain predicate, but it is an unusual member of that extension. Assume moreover that the use of long and/or infrequent expressions are marked ways of expression and so fulfill the constraint when the input is semantically marked. The interpreter can then only interpret *cause to die* as the expression of a marked way of killing. The generator would violate **PARSE-MARKED** by simply using *kill*, if there was something strange about the way the sheriff proceeded.

Though I appreciate the beauty of the explanation by weak superoptimality, I am worried by the fact that the interpreter actually overinterprets *cause to die* in Blutner’s account. As I see it, the interpreter would violate ***INVENT**. I avoid this problem by having an input feature that distinguishes the two readings.

Reflexives

Grice (1975) remarks that if you say (7), you imply that the woman is not his wife, his mother or his sister.

(7) I saw John in town yesterday with a woman.

We might add that the woman is also not the speaker or the listener or any other high salience item in the discourse situation. A natural explanation for this within OT is the assumption of a sequence of parsing constraints that force us to indicate in the output that the referent of an NP is the speaker⁹

⁸As Blutner points out, there is another problem. If there are not two possibilities, the prediction from superoptimality is that only the simple reading remains. That would predict that *make laugh* only has the direct interpretation, or that in Frisian, which has no reflexives, normal pronouns would only have reflexive meanings.

⁹A counterexample is Isherwood’s title (8), the disciple being the author himself.

(8) My Guru and his Disciple.

or the listener, c-commanded, currently in the discourse topic, in the visible surrounding of the utterance, has been mentioned before, is related to a highly salient discourse item by a relation expressed by the common noun of the NP, is uniquely described by the common noun of the NP etc. We further have to assume that first and second person pronouns express the person, reflexives c-commanding (or—in English— perspective), personal pronouns membership of the discourse topic, demonstratives the presence in the visible surroundings, the definite article either previous mention or a relation to an object in the discourse topic or uniqueness. The use of default rules for NP-selection is the standard technique in natural language generation and the only reason they have not found their way into linguistics is that most grammatical formalisms before OT syntax cannot accommodate them in natural way¹⁰.

In combination with ***INVENT** and **ANCHOR** the hierarchy of parse constraints give us precisely the effects that Grice predicts: that we can rule out all the properties higher up in the hierarchy.

8 Morals

In this paper, I have shown that a theory of semantic interpretation on the basis of OT syntax is feasible, if it is supplemented with some quite general semantic and pragmatic principles. The place of the Gricean maxims within this scheme has so far not been explored properly. It is clear that relevance and quantity must play a role at some point. Superoptimality (or weak superoptimality) and the speaker and listener games developed by Dekker & Van Rooy(2000) continue to be relevant, but do not penetrate syntax as such.

The treatment opens perspectives for the further development of the field of semantics as such. If I am right, compositionality does not need to be as much a straightjacket as it was in the heydays of the rule-to-rule hypothesis. A traditional problem is that of idiomatic expressions. The rule-to-rule hypothesis predicts that both sentences in (9) mean the same, i.e. that the speaker wants to know the time.

- (9) What time is it?
 How late is it?

Now the fact of the matter is that in English the second expression, though grammatical, is merely a source of wonder, while only the first actually expresses it. (This is reversed for the two Dutch equivalents.) It should be easy to configure the English OT syntax so that only the first is an optimal expression

Is this incorrect? Certainly there is suggestion of respect and modesty that would be absent in *My guru and Me*. Another literary effect seems that the topic of the book is neatly described: it is about the guru and Isherwood himself but only in his capacity as the guru's disciple.

¹⁰An exception should be made here for Panini, who by his general architecture and elsewhere principle is clearly a precursor of OT.

of the input (avoiding low frequency items would already seem to do that). The second sentence is then correctly predicted to be uninterpretable.

An important feature of OT syntax is that it can easily underspecify the full content of the semantic input. It is reasonable to assume that the representations in (10) are both optimally generated by *Every man likes a woman*.

$$(10) \quad \begin{aligned} &\forall x(\text{man}(x) \rightarrow \exists y(\text{woman}(y) \wedge \text{like}(x, y))) \\ &\quad \exists y(\text{woman}(y) \wedge \forall x(\text{man}(x) \rightarrow \text{like}(x, y))) \end{aligned}$$

The syntax parses the grammatical function of the two quantifiers and their quantificational force, but not their relative scope. The function of polarity sensitive items also becomes clearer: they parse a semantic feature of the environment of the semantic NP.

What we need is a weaker interpretation of the principle of compositionality. Frege does not say much more than that the meaning of a complex expression is a function of the meaning of its parts. What we need are slightly more liberal formulations. Parts must be taken to be the smallest meaningful part, which can include fixed combinations of words. And though we must admit that the meaning of a complex expression is determined by applying a function to the meaning of its parts, it does not follow that natural languages make it clearer what the precise logical content of that function is on a particular occasion than they make it clear what shade of blue is involved in my daughter's new blue dress. Though we can go for more precision in both cases, such precision is not required or desirable for everyday communication.

Author's Address

ILLC/Computational Linguistics
University of Amsterdam
Spuistraat 134
1012 VB Amsterdam
The Netherlands

Acknowledgments This paper comes out of remarks I made as a member of the panel at the OTS Conference in January 2000, Paul Smolensky's acute criticism of superoptimality, dinner time discussion after the conference and email discussion with Reinhard Blutner. I wish to thank especially Robert van Rooy, Reinhard Blutner, Marie Nilsenova and Anna Pilatova for their many useful comments. Many thanks also to the two anonymous reviewers for their excellent suggestions.

References

- Blutner, R. (2000) Some Aspects of Optimality Theory in Interpretation. (this volume).
- Boersma, P. (1998) Functional Phonology. PhD. Amsterdam.
- Bresnan, J. (MS B) Optimal Syntax. MS. Stanford.
- Bresnan, J. (MS A). The Emergence of the Unmarked pronoun: Chichewa Pronominals in Optimality Theory. BLS23.
- Choi, Hye-Won. (1998) Optimizing Structure in Communication: Scrambling and Information Structure. PhD Stanford.
- Dalrymple, M., M. Kanazawa, Y. Kim, S. Mchombo & S. Peters. (1998) Reciprocal Expressions and the Concept of Reciprocity. *Linguistics & Philosophy* **21**, p. 159-210.
- Dekker, P. & R. van Rooy. (2000). Bidirectional Optimality Theory: An Application of Game Theory. (this volume).
- Geurts, B. (2000) Buoyancy and Strength. (this volume).
- Grice, H.P. (1975) Logic and Conversation. In: P. Cole & J.L. Morgan. (eds.) *Syntax and Semantics 3: speech acts*. Academic Press, New York, p. 41-58.
- Grimshaw, J. (1997) *Projection, Heads and Optimality*. In: *Linguistic Inquiry*, 28, p. 373-422.
- Heim, I. (1982) *The Semantics of Definite and Indefinite Noun Phrases*. PhD University of Massachusetts, Amherst 1982.
- Heim, I. (1983) *On the projection problem for presuppositions*, *WCCFL* **2**. p. 114-26.
- P. Hendriks & H. de Hoop. (2000) Optimality Theoretic Semantics. *Linguistics and Philosophy*. (to appear)
- De Hoop, H. & H. de Swart. (2000) Temporal Adjunct Clauses in Optimality Theory. In H. de Hoop & H. de Swart (eds.) *OTS²*. OTS, Utrecht.
- Kamp, H. A Theory of Truth and Semantic Representation. In: Groenendijk, Janssen & Stokhof (eds). *Formal Methods in the Study of Language*. Mathematical Centre, Amsterdam 1981.
- Karttunen, L. Presuppositions of Compound Sentences. *Linguistic Inquiry* 1973, p. 167-193.
- Karttunen, L. (1974) Presupposition and linguistic context. *Theoretical Linguistics* **1**, p.181-94.
- Lewis, D. (1970) General Semantics. *Synthese* **22**: 18-67.
- K. Marriott & B. Meyer (1998) The CCMG Visual Language Hierarchy. In:K.

- Marriott & B. Meyer (1998) *Visual Language Theory*. Springer-Verlag, New York, Berlin and Heidelberg, p. 129-171.
- Montague, R. (1974) The Proper Treatment of Quantification in Ordinary English. In: R. Thomason (ed.) (1974) *Formal Philosophy*. New Haven and London, p. 247-271.
- Pesetsky, D. (1997) Optimality in Syntax. In: D. Archangeli & T. Langendoen (eds.) *Optimality Theory*. Blackwell.
- Van der Sandt, R. (1992) Presupposition projection as anaphora resolution. In: *Journal of Semantics* **9** , p. 333-377.
- Smolensky, P. (1996) On the Comprehension/Production Dilemma in Child Language. In *Linguistic Inquiry* **27** .
- Smolensky, P. Handout OTS Conference. Utrecht 3-1-2000.
- Stalnaker, R. (1978), 'Assertion', In: Cole, Peter (ed.): *Syntax and Semantics 9: Pragmatics.*, New York, p315 - 332.
- Zeevat, H. 2000. Optimal Semantics. In H. de Hoop & H. de Swart (eds.) OTS². OTS, Utrecht.
- Zeevat, H. 1999. Explaining Presupposition Triggers. In: P.Dekker(ed). AC 99 Proceedings. University of Amsterdam 1999.