Freezing and Marking

Henk Zeevat, ILLC, University of Amsterdam

Abstract The paper shows that word order freezing is not a good application of bidirectional optimality theory and that the facts it tries to cover in fact form a problem for bidirectional optimality theory. **max(X)** constraints allow moderate bidirectionality within a mono-directional production oriented optimality theory and give better results on freezing. Since freezing is the only argument for bidirectional syntax that is syntactic in nature, the paper also reviews the other arguments for full bidirectionality in syntax and concludes that they are not compelling.

1   Freezing

Jacobson (1984) observed that Russian *mat' ljubit doc'* (the mother loves the daughter) (*mat'* and *doc'* do not have different forms in the nominative and the accusative) only allows interpretations where *mat'* is the subject. This restriction on interpretation does not occur when case morphology distinguishes the two NPs. An accusative NP in the first position can then be interpreted as the object, or even an unmarked NP in the first position, provided the second NP is recognisably nominative. The same phenomenon can be found in German, Korean, Hindi, Polish, Latin and Dutch and the same description seems to apply

in all cases[1]. It has also been observed that, in all these languages, the excluded reading can be forced by parallellism. In the following German (1) *Peter* must be the loved one, because otherwise the question is not answered.

(1)   A: Wen liebt Maria? (Who does Maria love?)

         B: Peter liebt Maria. (Maria loves Peter.)

In optimality theory, the word order in these languages is controlled by a defeasible constraint that puts the subject before the object by default. The default may be overridden by other constraints that ask for priority for topics, *wh*-phrases and possibly other prominent NPs. In Dutch and German, the class of expressions that can override the default order of subject and object contains both contrastive topics and *wh*-phrases.

Lee (2001) starts from her version $\mathbf{CANON}_{GF}$ of $\mathbf{CN1}$ to derive the phenomenon of word order freezing in Korean and Hindi by bidirectional OT. $\mathbf{CANON}_{GF}$ requires that the syntactic functions are realised in the order subj > d-obj > i-obj > obl > adjunct. I will concentrate on Lee's Hindi example (2).

(2)   botal patt^{h}ar todegaa

         bottle-NOM stone-NOM break

         the bottle will break the stone

         (unavailable reading) the stone will break the bottle

(2) is a clear case of word order freezing. There is in Hindi a competition between the constraint $\mathbf{CANON}_{GF}$ and $\mathbf{TOP}_L$: the topic is realised in the leftmost position of the sentence. In Hindi, clearly $\mathbf{TOP}_L$ must be as strong

as $\mathbf{CANON}_{GF}$, given other examples like (3) where subject and object are distinguished by case and where the topical object moves to the first position.

(3) Ninaa_{TOP} Anuuko dik^{h}ii

        Nina-NOM Anu-DAT appear

        Anu saw Nina.

This ordering of the constraints however predicts that if the bottle is topic, (2) is a optimal realisation of the input "the stone will break the bottle" (roughly (4)) which is precisely the interpretation that is ruled out by freezing.

(4)   bottle:topic,object, stone:subject, break

This prediction is wrong, since as we saw in (2) , the sentence does not have that interpretation. But things change if we run the OT competition in reverse to obtain the optimal interpretation for the sentence given in (2) . The two interpretations[2] in (5) are then optimal, since -unlike the other interpretations- neither of them violates either of $\mathbf{CANON}_{GF}$ and $\mathbf{TOP}_L$.

(5)   bottle:topic,subject, stone:object, break

        bottle:subject, stone:object, break

Applying bidirectional OT, we can use the interpretational competition as a filter over the production competition. Winners that are not optimally interpreted as the original input are thrown out of the competition and the form of (2) becomes the unique winner. Where there is case marking, as in (3) faithfulness makes the nominative the object of the seeing and the non-canonical word order makes that object a topic, leading to a unique optimal interpretation for the optimal form given in (3) . Also agreement on the verb could

distinguish the subject from the object and that would also create an exception to freezing. This is a very elegant explanation of the facts of word order freezing and there is no reason to think that it would not work equally well for Russian, German, Korean or Dutch. But I think it has two serious problems, even in the case of Hindi.

2   Problems

The first problem is the following. The treatment as it stands rules out the interpretation (6) for (2) . The reason is that the interpretations of (5) will win from (6) since they do not break $\mathbf{TOP}_L$ in the interpretation direction.

(6)   bottle:subj, stone: object,topic , break

This is clearly wrong, since it is precisely the point of freezing that the word order is determined by the thematic roles, even if the normal trigger (topic) for optional inversion is present. There are two ways of defending this. One is to argue that topic is not a proper semantic feature, the other is going for different notions of bidirectionality. The first defence would involve limiting the bidirectional competition to the "proper" interpretation of the sentence, i.e. without the topic feature. This indeed gives the appropriate result. But it is hard to see what could be the reason for being so selective about features in the input. Relevance for the truth-condition of the sentence could not be the criterion because information structure has been clearly shown to influence truth-conditions.

For the other line of defence, notice that there is a large family of potential notions of bidirectionality. The notion that Lee uses is the one due to Smolensky (1996) (strong bidirectionality) and it is arguably the simplest: the input should be the winner for the output among the possible inputs and the output should be the winner for the input among the possible outputs. The same constraints with the same ordering decide in both cases who are these winners.

There are a number of alternative proposals in the literature for bidirectionality like Blutner's weak bidirectionality, Jaeger's conditional bidirectionality, Wilson's proposal or my own asymmetric proposal. Blutner's weak bidirectionality gives the same results in this case as strong bidirectionality: the pairs $<$ *botal patt^h ar todegaa*, the bottle (topic) will break the stone$>$ and $<$ *todegaa patt^h ar botal*, the stone (topic) will break the bottle$>$ are also weakly superoptimal and so will rule out interpretations where the object is the topic.

Jäger's proposal is to prefer superoptimal interpretations for forms that have a superoptimal realisation, but to allow interpretations for an optimal form if that interpretation does not have a superoptimal realisation. This in fact deals with the first problem: there is no superoptimal pair for the input *the bottle breaks the stone (topic)* and so it is an interpretation of its optimal realisation *botal patt^h ar todegaa*. Wilson's proposal restricts production optimisation to a competition between forms that are optimally interpreted as the input. That gives an account of freezing, but one that also suffers from the problem I am discussing: the missing meaning is not an optimal interpretation of the form.

My own proposal is to restrict bidirectionality to a limited set of pragmatic constraints and to be content with production optimisation for syntax and semantics. This makes the missing meaning possible but does not account for freezing.

The second problem is that there may be constraints that outrank **CN1**. In Dutch and German this holds for a constraint that puts *wh*-phrases first. What then happens to *wh*-objects that are not case-marked in the presence of a subject that is also not case-marked is that **CANON**$_{GF}$ inverts the interpretation. The interpretation that interprets the *wh*-phrase as the subject is preferred over the interpretation that interprets it as the object. This is wrong, since examples like (7) are simply ambiguous[4].

(7)   Welches Mädchen liebt Peter?

        Which girl loves Peter/does Peter love?

In fact, as noted in Zeevat (2001), these examples can be considered to be syntactic analogues of the *Rat/Rad* problem (Hale and Reiss, 1998). In the *Rad/Rat* problem, the devoicing of the final *d* of *Rad* in German lets the pronunciation /rat/ be the optimal one for both inputs since it overrides faithfulness for the feature voice in the case of *Rad*. In the interpretation direction however, that same faithfulness makes the interpretation *Rad* a clear loser, quite against the intuition that the pronunciation /rat/ is just ambiguous. Likewise, (7) is the optimal form for both the interpretation $?x(girl(x) \wedge love(p, x))$ and $?x(girl(x) \wedge love(x, p))$, but the extra movement required for the first form, punished by **STAY** or the requirements of **CANON**$_{GF}$) make it the case that

6

the only optimal interpretation is the one where Peter is the object. But there is a clear ambiguity.

The *Rat/Rad* problem and its syntactic analogues are a problem for any theory that is fully bidirectional and uses all the generation constraints also in the interpretive competition, such as strong and weak bidirectionality.

3   Abstract Solutions

One can try to come up with weaker versions of bidirectional OT to avoid the problem, like my own Zeevat (2001) and Jäger (2003). My proposal was to limit bidirectionality to pragmatics only and to be content with a monodirectional generative optimality theory in phonology and syntax. Interpretation happens through a mechanism that arrives at hypotheses about the meaning and tests them by the normal generative competition. Pragmatics favours consistent, plausible, maximally resolved interpretations which avoid extra material and maximise relevance. It filters out optimal forms for an interpretation if the optimal form allows pragmatically better interpretations.

The proposal avoids the *Rat/Rad* problems and saves bidirectional pragmatics. But it cannot handle freezing, since freezing is obviously not a question of pragmatics. *Mat' ljubit doc'* gets two readings that appear to be equally good in the respects that matter to the pragmatics.

Wilson's proposal runs into problems at once with the phonological and syntactic Rad-Rat problems since the initial filter omits the Rad and Wh-object interpretations.

In Jäger's proposal, a pair $< input, output >$ is hearer-optimal iff *input* wins the interpretation competition for *output*. Speaker-optimality can now be made dependent on hearer optimality: a pair is speaker-optimal iff it wins the generation competition and is hearer-optimal or there are no alternative outputs for the given input that are hearer-optimal. There are two possibilities for a pair to be speaker-optimal: either it is speaker-optimal and hearer-optimal, i.e. bidirectionally optimal or the form is just optimal for the input and there is no alternative optimal form for the input such that the input is hearer-optimal for that form.

If an interpretation I of a form F is defined to be good, iff I is the input of a speaker-optimal pair $< I, F >$, the syntactic and phonological *Rad/Rat* problems are solved. The stronger constraint that makes the *wh*-phrase come first, makes it impossible to put the subject before the object and so allows the form even though its interpretation is not hearer-optimal. Likewise, the *Rad*-interpretation of /rat/ survives, since voice on the last syllable is not an option in German.

Freezing however cannot be explained since both *subject* and *object* + *topic* are speaker-optimal interpretations of the first NP in *botal patt^{h}ar todegaa* without being hearer-optimal, assuming **CANON**$_{GF}$ and **TOP**$_L$.

There is also a conceptual problem with speaker-optimality. If interpretation is just reversing speaker-optimality, why bother about finding optimal interpretations by a competition? It does not help with finding good interpretations. If the interpretive competition is important, it should be the basis of interpreta-

tion and not something that has occasionally to be given up for an alternative method, like scanning possible inputs to see if they perhaps only have forms that are not hearer-optimal for that input. In this way, speaker-optimality supports a psychological prediction: that *Rad* is more difficult to recognise than *Rat*, that it is more difficult to recognise *Welches Mädchen liebt Peter (subj)* than *Welches Mädchen liebt Peter (obj)*. Such predictions are not plausible.

Speaker-optimality is similar to another notion that starts from monodirectional OT in the production direction. Again, interpretation is just looking for inputs that have the given form as optimal output. We add as a last constraint on $< input, form >$-pairs that there should not be both an alternative $input_1$ for *input* for which *form* is also optimal and that is indistinguishable from *input* except with respect to an important semantic dimension D (e.g. case or topic) and an alternative optimal form $form_1$ for *input* which is not optimal for $input_1$.

It is an ambiguity filter, like the constraint **UNAMBIGUOUS ENCODING** discussed in Gärtner (2003). If the form is ambiguous, it should be avoided. But if it is not possible to avoid the ambiguity, we do not bother. My formulation tries to limit the principle to certain aspects of meaning only, to avoid an implausible obligation to avoid any ambiguity. The version I gave is incomplete since the important semantic distinctions have to be stated. In principle, a list of important distinctions might be obtained by investigating which dimensions have to be obligatorily marked in different languages. It should then contain items like voice, case, tense, aspect, definiteness, etc. Sentences should not be

9

ambiguous in one of these important respects, unless it cannot be helped.

This way of proceeding overcomes the conceptual problem with Jäger's proposal, the misguided psychological prediction and can still explain the Rat/Rad problems including the syntactic versions.

Unfortunately, freezing is still not covered. The situation is just as in speaker-optimality.

The problem in the four general solutions I discussed seems to be that the solution is too abstract and general. Pragmatics does not help with freezing. There is no general notion of interpretation that can be used to condition generation appropriately, apart from reconstructing the input that the speaker had available, i.e. the Gricean concept of interpretation (and that one does not help for Jäger's purposes). There is also no *a priori* notion of ambiguity that can be made to work as part of a constraint **UNAMBIGUOUS ENCODING**. The important semantic dimensions seems to differ from language to language and typically one dimension is more important than another. What is important in one language can be far less important in another language. It should be possible to state that thematic role is more important than topic in the languages we have been considering. In short, we need to split up **UNAMBIGUOUS ENCODING** in a number of separate constraints.

In fact, that is part of standard optimality theoretic syntax, where there are **max(X)** constraints that are ranked with respect to other constraints and with respect to each other.

4 Marking

The concept of marking will give the solution. I am using it in a rather abstract way. An expression marks a semantic feature F if it lacks interpretations that do not have F. So case can mark an NP as binding a thematic role without "meaning" that thematic role. It should just be ruled that given the case, the verb and the syntactic structure of the sentence in which it occurs, it could not bind another thematic role. In the same sense, word order (e.g. NP1 before NP2) can mark NP1 as the agent, given the case, the verb and a constraint like **CANON**$_{GF}$, again without "meaning" agentivity.

Consider the three Dutch examples (8).

(8)   a. Wie slaat Jan? (Who hits John?, Who does John hit?)

          b. Piet slaat Jan. (Piet is hitting Jan.)

          c. Hem slaat Jan. (Jan is hitting him (and not somebody else X).)

(8a) is ambiguous, (8b) unambiguously subject object and (8c) again unambiguous with *hem* object and contrastive topic. The examples suggest the following view on word order. It is determined by *wh*-fronting, case or topic, but if word-order is determined by one of the dimensions, it becomes irrelevant for expressing anything other dimension.

We can capture this intuition by three constraints, ranked in the indicated order.

front(wh): *wh*-elements come first

max(θ): theta roles must be marked

max(topic): topic must be marked

The ambiguity in (8a) is explained by the fact that there are no means left for marking θ-roles after **front(wh)** has taken over the word order (and for topic, but perhaps the *wh*-element is always the topic), since there is no case morphology for the NPs and there is no agreement-morphology on the verb that could make a difference. In (8b) , θ-marking is succesfully achieved by word order, but the example is ambiguous for topic that can only be expressed by word order. In (8c) , there is case on the first NP which is sent to the first position by **max(topic)**. The three constraints seem in fact to capture the three examples, but it is not obvious how to formulate the two max(X)-constraints in a proper way.

To formulate the max-constraints one needs to have a notion of what marks the dimension X. For Dutch or German, **max(θ)** needs to refer to word order, case-marking and head-marking (agreement). Now case-marking is easily formulated for a verb that takes a subject and an object: one of the subject and the object needs to have a non-syncretic case form marking it unambiguously as subject or object. Head-marking is slightly more complicated: subject and object need to differ in an agreement feature expressed in the verbal agreement morphology. But word-order marking has a problem. The formulation: θ is word-order-marked iff the subject comes before the object is incorrect, since in (9) θ would come out as word-order-marked for one interpretation (who is Pete hitting?), while, in fact, it is not. The word order dimension has been taken over by **front(wh)** and therefore does not mark θ.

(9)   Wie slaat Piet?//Who is hitting Piet?/Who is Pete hitting?

Now in this particular case, it does not do any harm. But suppose there would be a fourth θ-marking device (e.g. directional morphology on the verb) that was available for (9). $\mathbf{max}(\theta)$ would force the use of the extra dimension in this case and it could be forced only by a definition of word-order-marking under which in (9), θ is not word-order marked.

The definition of *word-order marked* would therefore have to refer to possible stronger constraints on word order. But since it is a matter of ranking which constraints these are, this move would disqualify $\mathbf{max}(\theta)$ as a universal constraint and even as a possible OT constraint. An OT- constraint should not depend in its formulation on its ranking with respect to other constraints. It is also problematic that a universal constraint would refer to language particular marking devices, though it might be that an exhaustive listing is possible in principle.

I want to solve this problem by proposing a bidirectional definition of $\mathbf{max}(\mathbf{X})$ constraints, admittedly another breach of OT-orthodoxy. I believe it captures the essence of what it is to mark a certain distinction and it overcomes the problem of being dependent on rankings or on language particular marking devices. It also makes it clear that word-order marking is quite different from morphological marking, at least as long as there are no morphs which are ambiguous with respect to the dimension they mark. My definition is: a form marks a dimension X for an input I iff the form lacks interpretations which are exactly as I but have a different value for the dimension X. A mark is given for every

13

variation of the input for the dimension for which the form is also optimal.

This is a bidirectional formulation, since refer to other inputs than the given one. The conclusion however that processing max-constraints is inherently bidirectional is not warranted. Quite the contrary: one can always give language particular pure production versions of the constraint since there are only finitely many marking devices and stronger constraints in a particular language and checking marking can so be done without considering alternative meanings. It would be a compilation of the effects of the bidirectional max(X)-constraint given the stronger constraints and the particular marking devices of the language. The derived procedure is not itself an OT-constraint because it refers to other constraints and language particular marking devices and will be very different from language to language but it is still a plausible element of a processing model for the language.

I also do not think that formal hair-splitting leads to this reinterpretation of max-constraints. Instead a max-constraint for a dimension X as defined here is a straightforward formulation of the pressure to express content by conventional means, the sort of pressure that is opposed by economy and other markedness constraints. Any linguistic means can be harnessed to the task of expressing the dimension X and no means is inherently better suited than others. It follows that such a max-constraint must be expressed without reference to the means employed in the marking. Syntax is different in this respect from phonology where the relation between input and output is not merely conventional.

Let me give a detailed example for *Piet slaat Jan.*

From the perspective of a constraint **max(X)**, the input I is not alone, but comes with relevant other inputs I^X which can be obtained from I by changing I's value for dimension X to other values from X. I^{tense} can change the tense value of I from past to present, I^θ flips around the θ-roles of the participants in the input. I^{topic} gives another assignment to topic than in I. **max(X)** now just checks that the candidate form F receives a fatal error for every input in I^X. Otherwise, F for input I violates **max(X)**. I analyze the example (9b) using a constraint **PROM** saying that prominent things come first. i.e. a combination of **CANON**$_{GF}$ and **TOPIC**$_L$. Subjects and topics are both prominent and both should come first.

There are four relevant inputs.

Input: hit(piet,JAN)

topic(jan)

agent(piet)

goal(jan)

hit

Input$^\theta$: $hit(JAN, piet)$:

topic(jan)

goal(piet)

agent(jan)

hit

Input^{topic}: hit(PIET,jan)

topic(piet)

agent(piet)

goal(jan)

hit

Input$^{topic^\theta}$: hit(jan,PIET)

topic(piet)

agent(jan)

goal(piet)

hit

These four give all the new inputs one can make by flipping θ-roles and topics.

$hit(piet, JAN)$	**prom**	**max(θ)**	**max(topic)**
Jan slaat Piet	*	*!	
$\Rightarrow$ Piet slaat Jan	*		*

$hit(JAN, piet)$	**prom**	**max(θ)**	**max(topic)**
$\Rightarrow$ Jan slaat Piet			*
Piet slaat Jan	**!		

$hit(PIET, jan)$	**prom**	**max(θ)**	**max(topic)**
Jan slaat Piet	**!		
$\Rightarrow$ Piet slaat Jan			*

$hit(jan, PIET)$	**prom**	**max(θ)**	**max(topic)**
$\Rightarrow$ Jan slaat Piet	*		*
Piet slaat Jan	*	*!	

The difference between word order as a marking device for θ and topic and the morphological devices of head-marking and case-marking is in the constraints that play a role. The morphological devices are checked by the absolute constraint **FAITH** that checks whether the morpheme is consistent with the input. It gives an error to a form if its morphology expresses a property inconsistent with the input. If John is a subject, John should not have accusative morphology, if the subject is singular, the verb should not have plural agreement. The properties that the input should have to allow a particular kind of morphology

are coded in the lexicon. Word order does not have meaning in the same sense: it can be determined by grammatical factors (e.g. **front(wh)**), or can express prominence. It is the interaction with $max(X)$ constraints that sometimes gives word order meaning.

There is one more question that needs to be sorted out here. Suppose we have a language where subjects are unmarked and only objects may have an accusative case morpheme. In principle there should be an option between marking case by word order or by optional accusative morphology. According to Lee (2001) the interaction between freezing and optional case-marking in colloquial Korean is a case in point. But in many languages the case-marking is not optional. There are two equivalent ways of describing this situation. Obligatory case assigns a meaning to the unmarked form in the lexicon and **FAITH** should check that meaning just as it checks the case morpheme. But there could also be a constraint that makes the morphology obligatory, e.g. a constraint **CASE** that requires NPs to have case morphology.

One of these two ways is clearly needed in the description of English, Dutch, German or Russian for case, number and gender on NPs and for tense and agreement morphology on the finite verbs. Obligatory case morphology and agreement morphology is not explained by $\mathbf{max(\theta)}$ since it is also there if case is not needed to mark θ and if as in English θ is always completely marked by word order.

It has been observed that freezing is not a phenomenon without exceptions. Parallelism (as in answers to questions and in conjunctions and contrastive

pairs) can override freezing. The treatment I am giving can straightforwardly be extended to handle this phenomenon by assuming a word-order constraint **PARALLEL** stronger than **PROM** which says:

PARALLEL If the clause is parallel to another immediately preceding clause, the thematic functions appear in the order in which they occur in the parallel clause.

If **PARALLEL** instead of **PROM** determines the order of the NPs, it is **PARALLEL** instead of **PROM** that is sufficient for letting $\mathbf{max}(\theta)$ be satisfied.

(10)   Wen liebt Maria? (Who does Maria love?)

        Peter liebt Maria. (Maria loves Peter.)

In this example, $\mathbf{max}(\theta)$ is satisfied because the parallel constraint gives a fatal error to the form as an expression of the competing subject-object interpretation

5   Other Arguments for Bidirectionality

I argued that bidirectionality does not lead to an account of freezing and that $\mathbf{max(X)}$ constraints should do the job. Moreover, that these constraints in turn are best understood bidirectionally: a distinction X is expressed if the input without X is not expressible by the same form. This is a serious setback for bidirectional approaches: it is the only syntactic problem for which the assumption of bidirectionality seemed essential. How about the other arguments for bidirectionality? Is it possible to give a serious account of the syntax and semantics of natural language without going bidirectional all the way?

There are lots of technical problems with full bidirectionality ((Beaver and Lee, 2003) gives an overview for different versions) and not so many problems that it solves. There are five that used to convince me: ineffability, Blutner's theorem, Blutner's iconicity arguments, bidirectional learning and Smolensky's production/comprehension asymmetry. None of these really deals with a syntactic problem but still seem to establish a claim that bidirectionality is essential for saving OT syntax as a system (ineffability), for doing pragmatics (Blutner's theorem and the iconicity arguments), for dealing with language change and in accounting for language development. My view now is that while it is interesting and stimulating to apply bidirectionality in these cases, none of them is compelling.

In Zeevat (2000), I suggested that bidirectionality gives an interesting account of ineffability: ineffable inputs are inputs whose best expression is systematically understood as something else than the original input, by the interpretational competition.

In contrast, the mono-directional view on meaning is simply:

F means M iff F is optimal for M.

And that would appear to rule out the explanation, since there is always a winner in a production competition. However, the definition of meaning does not guarantee that every meaning of F is equally important. There is also the pragmatics of interpretation, preferring consistent, plausible interpretations where extra assumptions are maximally avoided and relevance is maximised. This can be enough for giving the same effect as bidirectionality. An ineffable

input can still be one that can only be expressed by forms that are plausibly interpreted as something else and only implausibly as the input. So the same explanation of ineffability may be given without bidirectionality. To use the same example as Zeevat (2001). If *Who ate what?* in Italian comes out as *Chi mangiava qualcosa?* it has to compete for plausibility with the interpretation *Who ate something?* and it is bound to lose that competition since the use of *qualcosa* expressing a secondary *wh*-elements would be limited to double *wh*-sentences and would lose out dramatically to the interpretation as an indefinite.

Blutner's Theorem (see (Zeevat, 2002)) says that presupposition triggers do not accommodate iff there is an alternative way of expressing the content that does not presuppose, and so offers a general explanation of why certain presupposition triggers do not accommodate. In the absence of other explanations this is very welcome. But special cases of non-accommodation can be explained by marking principles and economy, as shown in Zeevat (2003). Other cases can be perhaps be treated as suggested in Beaver and Zeevat (ta). And there are a number of apparent counterexamples (the triggers *manage, finish, stop*) that threaten the generalisation.

McCawley's case of Black Bart causing the sheriff to die is not really an argument for bidirectionality, since it is already explained as a flouting of Grice's maxim of manner. A flouting of this kind requires the hearer to find an intention of the speaker which in addition to the literal meaning contains a reason for her using the long and unusual form. Blutner's explanation in bidirectional OT can at best be seen as a contribution to the formalisation of the Gricean

reasoning, but it does not seem to be achieving that, because of two incorrect assumptions: (1) that the connection between "causing to die" and "indirect killing" is conventional (see e.g. van Rooy (ta) where the association is treated as the product of language evolution) and (2) the existence of a general economy principle that rules out longer (or less normal) expressions of some meaning. That (1) is wrong can be seen by noting that all the criteria that Grice gives for being a conversational implicature and not a part of the conventional meaning are fulfilled in this case. (2) follows from comparisons between: make a turn/turn, give a lick of paint to/paint, etc. where the longer form does not acquire a special interpretation.

The bidirectional learning algorithm of Jäger (2003) can be formulated without having recourse to bidirectionality in grammar. In fact it works for any representation of grammatical knowledge that can be learned by adjusting weights. A learning datum is a pair <m,f> (the hearer infers from context and situation that m is what the speaker intends with f, or the speaker infers from the hearer's reaction that the hearer understands f as meaning m). The datum can be used in two ways: the learner can take m and determine her optimal way f' of expressing m: learning takes place if $f' \neq f$. But the learner can also try to find out what meaning m' would have led to f given her grammar. And learning also happens if $m \neq m'$. Combining these directions leads to the same partially unstable learning that Jäger found and can likewise be used to underpin evolutionary processes, because functional pressure can push the weights away from the weights underlying the observations. Notice that the algorithm

learns in speaker and hearer mode and so still is genuinely bidirectional. If one of the learning directions is omitted the algorithm just reproduces the input weights.

A final argument for bidirection is Smolensky's production/comprehension asymmetry, though perhaps it is better seen as part of a theory of language development. The observation is that the child says "tata" when meaning Kate, but only understands "kate" as Kate. This can be explained by assuming that in the child's grammar markedness constraints are ranked high while faithfulness constraints are ranked low. The markedness constraints cause very unmarked structures to be produced, while in interpretation they do not play a role, causing only "kate" to be understood as Kate. It is not an argument for bidirectionality, but for the importance of an interpretive competition using the same system of constraints as underlies production. There are arguments against this assumption, e.g. the Rad/Rat problem (Hale and Reiss (1998) gives this and other counterarguments). But perhaps more importantly when we look at it as an argument for bidirectionality: it is not clear that this is the only way to account for the asymmetry. A less interesting, but not implausible explanation is to assume that the child lacks the psycho-motoric skills needed for reproducing the adult pattern, like a person having a cold or being under the influence of alcohol. That the child can produce the complicated forms when attempting to say even more complex words does not mean that it has the skills to imitate the complicated word itself, it hits it by accident trying to accomplish something more complex.

6 Conclusion

This paper argues against abstract conceptions of bidirectionality in syntax or semantics, but for a bidirectional interpretation of max-constraints. The explanation of freezing in this paper is substantially identical with Hanjung Lee's proposal, but avoids the two problems that I noted. Max-constraints as formulated here are also a good way of approaching the problems discussed in Gärtner (2003). The dimensions for which max-constraints are plausible are central in the sense that not marking the dimension would lead to systematic ambiguities against which there is considerable functional pressure, so that by bidirectional learning one expects language development to redress the balance in favour of marking. And that is what appears to be going on: morphology, articles, and particles form at those places where a max-constraint is present, i.e. where there is an expressive need. It is even possible to reason backwards and infer the functional pressure behind the max-constraint from the morphs it has been able to form. How else could we have those morphs?

Freezing and marking are clear indications that a theory of word order cannot be autonomous, but needs to be related to the other marking possibilities in the language and to the strengths of the expressive needs which it can help to meet.

Linguistics for their remarks which have helped me considerably in trying to make the

present paper more understandable.

Footnotes

1. Lee (2001) gives an overview of Hindi and Korean facts, freezing is a standard part of teaching German word order in Dutch schools, the similarity of Dutch to German in this respect is my own observation. Polish and Latin freezing have been investigated by my former students Piotr Labenz and Kelly Nedwick.

2. One could object that maybe there should be only one interpretation and maybe this the correct view for Hindi. In Dutch and German however, it is not topic as such that allows objects to come first, but contrastive topic and sentences do not need to have a contrastive topic.

3. Some have argued -but not in print as far as I know- that it is a mistake to have OT systems in which sentences end up as ambiguous. As they say, the fact of the matter is that people rarely see ambiguities because they always -in context- come up with a single reading. In an OT system this should be accounted for by a constraint **CONTEXT** that lets the winning reading win. I am in favour of a such a constraint, as part of pragmatics. It checks the consistency of the interpretation with the existing context of interpretation, the plausibility of the interpretation, given the context of interpretation and checks the relevance of the interpretation again in the context of interpretation. These are quite potent disambiguators, but it is easy to construct examples that remain ambiguous even if these factors are all taken into account. For example, the phonological ambiguity between Dutch *rad* (wheel) and *rat* (rat) does not get eliminated in a story like *We waren aan het fietsen zaterdag en toen zagen we een rat/d. (We were cycling last Saturday and then we saw a rat/wheel.)*

So unless better ways of using the context can be found, ambiguity is a fact of life and it seems that the single reading effect is much better attributed to processing: people just take the first interpretation that passes pragmatics and quite possibly not the one the speaker intended. The ambiguity notion in this paper abstracts from pragmatics and so merely opens the possibility of pragmatic ambiguities. The assumption of misunderstandings due to pragmatic ambiguity is also necessary in a functional theory of language evolution, it would be the force behind language change in a functional account.

References

Beaver, David and Lee, Hanjung (2003). Input-output mismatches in OT. In Reinhard Blutner and Henk Zeevat, editors, *Pragmatics and Optimality Theory*. Palgrave.

Beaver, David and Zeevat, Henk (t.a.). Accommodation. In Gillian Ramchand and Chris Reiss, editors, *The Oxford Handbook of Linguistic Interfaces*. Oxford University Press.

Gärtner, Hans-Martin (2003). On the OT status of Unambiguous Encoding. In Reinhard Blutner and Henk Zeevat, editors, *Pragmatics and Optimality Theory*. Palgrave.

Hale, Mark and Reiss, Chris (1998). Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry*, **29**, 656–683.

Jacobson, Roman (1958/1984). Morphological observations on Slavic declension (the structure of Russian case forms). In Linda R. Waugh and Morris Halle, editors, *Roman Jakobson. Russian and Slavic grammar: Studies 1931-1981.*, pages 105–133. Mouton de Gruyter, Berlin.

Jäger, Gerhard (2003). Learning constraint sub-hierarchies. the Bidirectional Gradual Learning Algorithm. In Reinhard Blutner and Henk Zeevat, editors, *Pragmatics and Optimality Theory*. Palgrave.

Lee, Hanjung (2001). Markedness and word order freezing. In Peter Sells, editor, *Formal and Empirical Issues in Optimality-Theoretic Syntax*. CSLI Publications, Stanford.

Smolensky, Paul (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry*, **27**, 720–731.

van Rooy, Robert (t.a.). Signalling games select Horn strategies. *Linguistics and Philosophy*.

Zeevat, Henk (2000). Semantics in Optimality Theory. In Helen de Hoop and Petra Hendriks, editors, *Optimality Theoretic Semantics*. OTS, Utrecht.

Zeevat, Henk (2001). The asymmetry of Optimality Theoretic syntax and semantics. *Journal of Semantics*, **17**(3), 243–262.

Zeevat, Henk (2002). Explaining presupposition triggers. In Kees van Deemter and Rodger Kibble, editors, *Information Sharing*, pages 61–87. CSLI Publications, Stanford.

Zeevat, Henk (2003). Particles: presupposition triggers, context markers or speech act markers. In Reinhard Blutner and Henk Zeevat, editors, *Pragmatics and Optimality Theory*. Palgrave.