On Bidirectional Optimality Theory for Dynamic Contexts

Anton Benz*

benz@zas.gwz-berlin.de

Abstract

In this paper we study context–sensitive versions of bidirectional Optimality Theory (OT) which can be used to model online communication. Our guiding examples are taken from anaphora resolution. We discuss a puzzle presented by Jason Mattausch which shows that context–sensitivity may lead into circularity. In order to represent it, we have to introduce more expressive mathematical structures for BiOT. We call the fundamental structures *Blutner structures*. A core problem is to account for the epistemic asymmetry between speaker and hearer in online communication. This leads us to Blutner structures which combine bidirectional OT with Dynamic Semantics.

1 Introduction

Bidirectional Optimality Theory (Blutner, 2000, BiOT) was originally introduced as a theory that models the pragmatic reasoning in online communication. Although many early proponents of the theory now prefer a diachronic interpretation, e.g. with applications to blocking phenomena, there remain substantial applications which ask for an online model, e.g. the resolution of pronouns and other anaphoric expressions.

In this paper, we are concerned with the mathematical structure of BiOT models of online communication. Online communication raises a number of questions in connection with the *bidirectionality* of BiOT. In contrast to *unilateral* OT, which concentrates on the production, i.e. the speaker's perspective (Kager, 1999; Smolensky and Legendre, 2006), bidirectional OT assumes that the choice of linguistic forms and their interpretation is the result of a two–sided optimisation process which involves the speaker's and the hearer's perspective. We argue that these perspectives must be kept distinct, both with respect to the constraint hierarchies which define the preferences of speaker and hearer, and with respect to the speaker's and hearer's knowledge.

Anaphora resolution is especially interesting for our purposes as here the difference between the speaker's and hearer's perspectives becomes clearly visible.

*The core of this paper dates back to Spring 2001 when I was employed on the DFG project LA 633/5–1 on *Dialogue Semantics* located at the Humboldt–Universität Berlin. Since then, the paper has been revised and updated substantially. It profited especially from discussions at the BIOT07 conference at the ZAS in Berlin.

Our guiding example is due to *Jason Mattausch* (2000, pp. 33–36). The puzzle about it is that it seems to show that bidirectional optimisation can lead into a circle without solution:

(1) Assume that Marion is a male person, and Jo a female one. The speaker wants to express with the second sentence that Jo was pulling Marion's hair out:

- **a)** Marion was frustrated with Jo. She was pulling his hair out.
- **b)** Marion was frustrated with Jo. He was pulling her hair out.
- **c)** Marion was frustrated with Jo. Jo was pulling Marion's hair out.

Intuitively, c) is the right way to put it. Mattausch considers three constraints. Firstly, pronouns have to agree with the natural gender of the person referred to. Secondly, pronouns are preferred over names. Thirdly, the hearer prefers an interpretation where *Marion* is female, and *Jo* male. If the speaker wants to express the meaning *pulling–hair–out*(*Jo*, *Marion*), he will choose *She was pulling his hair out.* This is the meaning of the horizontal arrow in the top row of the following graph. The hearer will interpret this form according to his preferences as *pulling–hair–out*(*Marion*, *Jo*). This is represented by the vertical arrow in the right column of the graph. But this meaning should be expressed by the speaker as *He was pulling her hair out.* For this form the hearer should prefer the interpretation *pulling–hair–out*(*Jo*, *Marion*). And here the circle closes.

$$
\begin{array}{ccccl}
\langle \text{He}, \text{her} \rangle & & \langle \text{She}, \text{his} \rangle & & \\
 & & & & \\
\bullet & \longrightarrow & \bullet & & \textit{pulling–hair–out}(\textit{Jo}, \textit{Marion}) \\
\uparrow & & \downarrow & & \\
\bullet & \longleftarrow & \bullet & & \textit{pulling–hair–out}(\textit{Marion}, \textit{Jo})
\end{array}
$$

In the standard accounts of BiOT, (Blutner, 2000; Blutner and Jäger, 2000; Jäger, 2002), circular structures are excluded by definition. There, the preferences are always well-founded. For a graph as that shown above this means that, wherever we start, by following the arrows our path will end in a node which has no arrows leading away from it. In order to guarantee well–foundedness, strong a priori assumptions have to be made about the ranking of the OT–constraints and the epistemic states of speaker and hearer.

Our discussion of the foundations of BiOT will concentrate on Jäger's algorithm and the mathematical structures, called *OT–systems*, on which he builds BiOT (Jäger, 2002). We address two issues related to the basic mathematical structures in which from–meaning selection in BiOT can be described. The first issue concerns the assumption that all OT-constraints are comparable and ranked in a single scale independent of whether they are speaker's constraints that select optimal forms for given meanings, or whether they are hearer's constraints which select optimal meanings for given forms. We argue that, in general, this assumption is cognitively implausible. Moreover, if hearer's and speaker's constraints are ranked on the same scale, then this is an empirically meaningful statement which should be carefully argued for from application to application. But it should not be an assumption built in a priori into the mathematical structures.

The second issue concerns the epistemic asymmetry between speaker and hearer. We will see that this asymmetry is the main reason why a circularity seems to arise with Mattausch's example. Whereas the speaker has full knowledge about Marion and Jo, the hearer knows only their names. In BiOT structures, this asymmetry can not be represented. Instead to offer a solution to Mattausch's puzzle, this restriction merely reduces the expressibility of BiOT.

In order to answer the raised questions and to solve Mattausch's puzzle, we provide mathematical structures for BiOT which separate speaker's and hearer's constraints and which allow us to represent the epistemic asymmetry between them. This means that we finally have to embed bidirectional reasoning in contexts known from dynamic semantics (Kamp and Reyle, 1993; Groenendijk and Stockhof, 1991). It also follows that our model improves BiOT as a framework for analysing production and interpretation of linguistic forms in *online* processes, in contrast to diachronic processes or language acquisition models.

In Section 2, we discuss the structure of OT–systems as used by Jäger (2002). Especially, we use Mattausch's example in order to highlight their shortcomings in dynamic contexts. In the following section, we will step-by-step extend these structures.

In Section 3, we first will introduce structures that separate the speaker's and the hearer's constraints in two separate scales. We call them *Blutner structures.* They are generated by two OT–systems over the same set of grammatical form–meaning pairs. We show that central notions of BiOT, like *optimality* and *weak optimality*, can be generalized to these structures. We will also see that in Blutner structures it is no longer guaranteed that the combined preferences are well–founded. Furthermore, we will show how we can handle arbitrary context–sensitive constraints.

In Section 4, we address Mattausch's example and show how to solve it in our framework. It is essential for the example that the interpreter has only a limited knowledge about the actual context. In a dynamic setting, we can assume that he has *less* information than the speaker. Hence, he has more form–meaning pairs to consider — all the form–meaning pairs which are grammatical in any of his epistemically possible contexts. We assume that these epistemic possibilities are given by the information state defined by the previous discourse. This leads to two additions to context sensitive Blutner structures: first we have to represent the epistemic asymmetry, and, second, we have to add a new principle to the algorithm for calculating optimal form–meaning pairs. As the hearer lacks certain information, the speaker has to take care not to *mislead* him. In our graphical representations, which we will introduce later, this will mean that certain paths run into a *dead end.* Hence, Jäger's algorithm has to be modified so that the speaker avoids dead ends, i.e. misleading forms. This modification solves the puzzle.

In Section 4, we provide our motivations and outline the solution. The precise framework in which this solution can be formulated is then presented in Section 5. We call the central structure *Blutner structures for dynamic contexts.* They provide the foundation for a system which allows the integration of bidirectional OT and Dynamic Semantics.

2 OT–Systems

Radical Pragmatics holds that semantic meaning is highly underspecified and needs to be completed by pragmatic mechanisms in order to obtain the specific contextual meaning of an utterance. This completion can be understood as a two-sided optimisation process in which the speaker has to choose an optimal form F for the meaning M he wants to express, and the hearer an optimal meaning M for the form F presented to him by the speaker.

In Optimality Theory (OT), it is assumed that producers and interpreters of language use a number of constraints which govern their choice of forms and meanings. These constraints may get into conflict. OT proposes a mechanism for how these conflicts are resolved. It assumes that the constraints are ranked in a linear order. If they get into conflict, then the higher-ranked constraints win over the lower ranked. This defines preferences on forms and meanings. In BiOT, it is further assumed that speaker and hearer co-ordinate on form–meaning pairs which are preferred from both perspectives. As mentioned before, we concentrate on the standard account of BiOT as introduced in (Blutner, 2000; Blutner and Jäger, 2000; Jäger, 2002). Beaver and Lee (2004) provide a useful overview of versions of optimality theoretic semantics prior to 2004. They discuss seven different approaches. In particular, they compare them according to whether they can explain partial blocking and Horn's principle of division of pragmatic labour (Horn, 1984). According to Beaver & Lee, the only approach which can fully explain them is Blutner's BiOT (Beaver and Lee, 2004, Sec. 7 & 5). This success is the main reason for Blutner's BiOT becoming the dominant version.

In (Jäger, 2002), the algorithm by which optimal form–meaning can be calculated is discussed in greater detail.[1] The speaker starts by choosing an optimal form F_0 for a given meaning M_0, then the hearer a meaning M_1 which is an optimal interpretation for F_0 according to his ranking of meanings. Then again, the speaker chooses the most preferred form F_1 for M_1, then again the hearer the most preferred meaning M_2 for F_1. A form–meaning pair is optimal, if speaker and hearer ultimately choose the same forms and meanings. If $\langle F, M\rangle$ is optimal in this technical sense, then the choice of F is the optimal way to express M such that both speaker's and interpreter's preferences are matched.

Preferences can be identified with transitive relations $\preceq$, where we read $F \prec F'$ as F' *is preferred over* F, and $F \approx F'$ as F *and* F' *are ranked equally.* Before we proceed, we first fix some terminology concerning transitive relations:

Definition 2.1 *Let M be a set, and $\preceq \subseteq M \times M$ a relation. We say that:*

1. $\preceq$ *is a* pre–order*, iff* $\preceq$ *is reflexive and transitive:* $m \preceq m$*, and* $m \preceq m' \wedge m' \preceq m'' \Rightarrow m \preceq m''$.
2. $\preceq$ *is an* order*, iff it is a pre–order and in addition* $m \preceq m' \wedge m' \preceq m \Rightarrow m = m'$.
3. *An order* $\preceq$ *is linear, iff for all* m, m'*:* $m \preceq m' \vee m' \preceq m$.
4. *An order is well founded, iff for every set* $X \subseteq M$ *there is an* $m \in X$ *such that* $\forall m' \in X\,(m' \preceq m \Rightarrow m' = m)$.

[1] We describe the procedure which provides for a *strong z–optimal* form–meaning pair. (Blutner, 1998, 2000) introduced in addition *weak* optimality, also called *superoptimality*.

5. *If $\preceq$ is a pre–order, then the sets $[m] := \{m' \in M \mid m \preceq m' \,\&\, m' \preceq m\}$ are equivalence classes. If we set $[m] \preceq [m'] \Leftrightarrow m \preceq m'$, then $\preceq$ is an order relation on the set of equivalence classes $[m]$.*

6. *We call a pre–order well–founded or linear, iff the associated order on the set of equivalence classes $\{[m] \mid m \in M\}$ is well–founded or linear.*

A non-linguistic example of a well founded pre–order $\preceq$ is the cost relation between goods, i.e. if we read $a \preceq b$ as b *is at least as expensive as* a. Then, every good a is at least as expensive as itself (*reflexivity*), and from b being at least as expensive as a, and c being at least as expensive as b, it follows that c is at least as expensive as a (*transitivity*). Furthermore, for each set G of goods, there is at least one item $a \in G$, such that every other item $b \in G$ is at least as expensive as a (*well foundedness*). If $[a]$ is the set of all goods with the same price as a, then the *at–least–as–expensive–as* relation becomes a well founded order on the set of all equivalence classes $[a]$.

In (Jäger, 2002), specific mathematical structures are proposed in which to develop BiOT. They consist of a set Gen, the so-called *generator*, and a set C of constraints. Gen is a set of form–meaning pairs $\langle F, M \rangle$ which are considered to be grammatical according to an underlying un-defeasible grammar. In general, this grammar will leave the form–meaning relation underspecified, i.e. for a form F there my exist several forms M such that $\langle F, M \rangle \in \text{Gen}$. Hence, if the addressee hears F, then Gen tells him what are the possible interpretations of F. Gen is also the place where to represent indefeasible *faithfulness* constraints. Faithfulness constraints state conditions which depend on form and meaning simultaneously. For example, the constraint that a proper name (form) has to be interpreted by the bearer (meaning) of the name is a faithfulness constraint. If it is indefeasible, then we can directly represent it in Gen by restricting Gen to name–referent pairs in which the referent must be the bearer of the name.

A constraint can be most naturally represented by a formula $\varphi(v_1, \ldots, v_n)$ with free variables. This is equivalent to a representation as a binary valued function c with n argument positions, where $c(a_1, \ldots, a_n) = 0$ iff $\varphi(a_1, \ldots, a_n)$ holds, and $c(a_1, \ldots, a_n) = 1$ iff $\varphi(a_1, \ldots, a_n)$ does not hold. If we assume in addition that a constraint does also *rank* the tuples $\langle a_1, \ldots, a_n \rangle$, then we can identify it with a function c which maps $\langle a_1, \ldots, a_n \rangle$ into the set of natural numbers $\mathbb{N}$. If we restrict constraints to form–meaning pairs, then c is a function from Gen into $\mathbb{N}$. We can bring these parts together. The resulting structure is called an *OT–system* (Jäger, 2002, Def. 4):

Definition 2.2 (OT–System) *An OT–system is a pair $\mathcal{O} = \langle \text{Gen}, C \rangle$, where* Gen *is a relation, and $C = (c_\alpha)_{1 \leq \alpha < \beta}$ is a sequence of functions from* Gen *to* $\mathbb{N}$*, β an ordinal number.*

1. *Let $a \approx_{\mathcal{O}} b$ iff for all $\alpha < \beta$ $c_\alpha(a) = c_\alpha(b)$.*

2. *Let $a, b \in$ Gen. $a <_{\mathcal{O}} b$ iff there is an γ with $1 \leq \gamma < \beta$ such that $c_\gamma(a) < c_\gamma(b)$ and for all $\alpha < \gamma :$ $c_\alpha(a) = c_\alpha(b)$.*

We write $a \leq_{\mathcal{O}} b$ for $a \approx_{\mathcal{O}} b \vee a <_{\mathcal{O}} b$.

For technical reasons the preference relation is read here in inverse order, i.e. $a <_{\mathcal{O}} b$ means that a is preferred over b. This implies that c_1 is the strongest constraint, then comes c_2, etc. In the context of our paper, the elements $a, b \in$ Gen are form–meaning pairs $\langle F, M\rangle$.

It is important to notice that OT–systems, which underlie BiOT, are structurally identical to the systems that underlie the relevant part of *unilateral* OT (Smolensky and Legendre, 2006). In order to capture the two–sidedness of BiOT, Jäger introduces the distinction between *input* and *output*–markedness constraints. Both types of constraints are constraints that only depend on one argument: input markedness constraints on the forms, and output markedness constraints on the meanings. They are defined as follows (Jäger, 2002, Def. 9,11):

Definition 2.3 *Let $\mathcal{O} = \langle \mathrm{Gen}, C\rangle$ be an OT–system. A constraint c is an* input *markedness constraint, iff $\langle F, M\rangle, \langle F, M'\rangle \in \mathrm{Gen} \Rightarrow c(\langle F, M\rangle) = c(\langle F, M'\rangle)$. A constraint c is an* output *markedness constraint, iff $\langle F, M\rangle, \langle F', M\rangle \in \mathrm{Gen} \Rightarrow c(\langle F, M\rangle) = c(\langle F', M\rangle)$.*

An application of BiOT to anaphora resolution shows that these constraints are too restrictive. It suffices here to consider the constraints involved in Mattausch's example. For this example, we only need the following constraints:

(Agr) pronouns agree with the natural gender of the person referred to;

(Pro) pronouns are preferred over names;

(Dft) pronouns agree with the default interpretation of the gender of names.

Neither (Agr) nor (Dft) are input or output markedness constraints. We consider (Agr). First, let us fix a pronoun, say *she*, and choose two different meanings, one for which the referent is *male*, and one for which it is *female*. Then, (Agr) is met by $\langle she, female\rangle$ but not by $\langle she, male\rangle$. Hence, (Agr) is not an input markedness constraint. Second, let us fix a meaning, say *female*, and choose the forms *she* and *he*. Then (Agr) is met by $\langle she, female\rangle$ but not by $\langle he, female\rangle$. Hence, (Agr) cannot be an output markedness constraint either. In the same way, we can see that (Dft) is neither an input nor an output markedness constraint. Nevertheless, they are natural constraints and should not be excluded from bidirectional OT.

The first lesson, we can learn, is that the two–sidedness of bidirectional OT should not be regarded as a restriction on the type or number of arguments. It follows that, in OT–systems, the bidirectionality of BiOT is only present in the algorithm for calculating optimal form–meaning pairs. It is an important property of OT–systems that this algorithm always terminates, i.e. there is always a form–meaning pair that is optimal from both perspectives. This result holds for any type of constraints, i.e. it is not restricted to systems with only input and output markedness constraints (Jäger, 2002, Lem. 2):

Lemma 2.4 *Let $\mathcal{O}$ be an OT–system. Then $\leq_{\mathcal{O}}$ is a linear and well–founded pre–order.*

Proof: It is clear that $\leq_{\mathcal{O}}$ is linear, and that it is a pre–order. Assume that there exists $(a_n)_{n\in\omega}$ such that $\forall n\, a_{n+1} <_{\mathcal{O}} a_n$. Let $\gamma'_n := \min\{\gamma < \beta | c_\gamma(a_{n+1}) < c_\gamma(a_n)\}$ and $\gamma_n := \omega\gamma'_n + c_{\gamma'_n}(a_n)$. It is clear that all γ_n exist and that $\gamma_{n+1} < \gamma_n$. But this contradicts the well–foundedness of the ordering of ordinals. □

It is instructive to represent Mattausch's Example **(1)** by an OT–system and to see why the circle doesn't emerge. We consider only (Pro) and (Dft), and assume that (Agr) is part of the underlying grammar and represented in Gen. The algorithm starts with the speaker searching for a form to express that M_0 = *Marion was pulling Jo's hair out*. Due to (Pro), he prefers *he/she* to *Marion/Jo*. Then, (Dft) seems to force the hearer to choose M_1 = *Jo was pulling Marion's hair out* for the form *he/she*: both pairs, $\langle he/she, M_0\rangle$ and $\langle he/she, M_1\rangle$, satisfy (Pro), but only the second pair satisfies (Dft). Then, the algorithm continues again with the choice of the speaker, and it seems we run into a circle again. Where is the mistake? As indicated, it occurred when we considered the hearer's choice. In an OT–system, the choice is restricted to grammatical form–meaning pairs $\langle F, M\rangle \in$ Gen. As Marion is male and Jo female, $\langle he/she, M_1\rangle$ is not grammatical, hence not a possible choice for the hearer. His only possible choice is M_0, and therefore the algorithm terminates in an optimal form–meaning pair.

The problem about this solution to Mattausch's puzzle should immediately be clear: as the hearer does not know that Marion is male and Jo female, he has no possible way of knowing that only $\langle he/she, M_1\rangle$ is grammatical. This points to a deeper problem of OT–systems, and BiOT in general: it doesn't pay sufficient attention to the epistemic asymmetry between speaker and hearer. Implicit in BiOT is the assumption that both speaker and hearer have complete knowledge about what are the grammatical form–meaning pairs in the actual state of the world.

The second problem that arises with OT–systems in connection with our example relates to the ranking of all OT–constraints in a single order. Let us slightly change the representation of Mattausch's example. Let us assume that Gen is completely unrestricted and that (Agr), (Pro), and (Dft) are three constraints. Again, Lemma 2.4 predicts that Jäger's algorithm must terminate. Again, the speaker chooses F_0 = *he/she* for M_0 = *Marion was pulling Jo's hair out*. Then, the hearer chooses M_1 = *Jo was pulling Marion's hair out* for *he/she*; then, the speaker will choose F_1 = *she/he*, which in turn induces the hearer to choose M_0, and the circle seems to close. Again, where is the mistake? There are two possibilities depending on the relative ranking of the constraints. The pair $\langle F_0, M_0\rangle$ satisfied (Agr) but violates (Dft), hence, in order to win over $\langle F_1, M_0\rangle$, (Agr) must be assigned a lower index than (Dft), (Agr) < (Dft). But then, if the hearer has to choose between M_0 and M_1 for F_0, the simultaneous ranking of all constraints forces him to evaluate (Agr). Besides the problem of epistemic asymmetry, which arises again, it is counterintuitive to assume that the hearer evaluates (Agr) at all. But this assumption is necessary in order to break the circle as evaluating (Agr) will force him to choose M_1. If we assume that (Dft) < (Agr), a similar problem arises for the speaker and constraints (Pro) and (Dft). If he chooses between $\langle F_0, M_0\rangle$ and $\langle Marion/Jo, M_0\rangle$, then he has to decide whether (Pro), i.e. his preference for pronouns, is stronger than (Dft), i.e. the hearer's preferences for stereotypical interpretations. Intuitively, these aspects are incommensurable, and a comparison should not be forced upon the speaker. We therefore conclude that the speaker's and hearer's constraints should be ranked separately.

In the following sections, we will introduce more general mathematical structures which generalise OT–systems step-by-step. In dynamic semantics, the meaning of a sentence is identified with an update function on information states. The updated information states are implicitly assumed to represent the hearer's knowledge. The actual state of the world is only known to the speaker. This asymmetry is an essential feature of discourse and it must play a central role if we are to apply BiOT to discourse interpretation. The integration of this epistemic asymmetry and BiOT will lead us to develop mathematical structures for BiOT in dynamic contexts. The separation of preferences will be captured in *Blutner Structures*, Section 3. Later, in Section 5, we will introduce Blutner structures for dynamic contexts.

3 Blutner Structures

In the last section, we saw that it is desirable to divide constraints into one ranked group which provides for the speaker's preferences on forms, and a second ranked group which provides for the hearer's preferences on meanings. We call the resulting structures *Blutner structures*. If these structures shall be acceptable, we have to show that we can represent the core concepts of BiOT. Hence, we show that Jäger's definitions of *optimality* and *weak optimality* from (Jäger, 2002) generalise to Blutner structures. We then show that Blutner structures can always be generated by two OT–systems over a shared set of grammatical form–meaning pairs, one OT–system for the speaker and one OT–system for the interpreter. Finally, we show how to represent arbitrary context–sensitive constraints within our framework.

Blutner Structures

OT–systems are pairs $\langle \text{Gen}, C\rangle$, such that Gen is a set of form–meaning pairs which is *generated* by a given grammar. We use the same underlying structure. For convenience, we make the set of *forms* $\mathcal{F}$ and *meanings* $\mathcal{M}$ explicit, i.e. we represent the underlying structure by triples $\langle \mathcal{F}, \mathcal{M}, \text{Gen}\rangle$. The following definition of *Blutner structure* can be motivated by the way how speaker and interpreter find an optimal form–meaning pair. As explained in the previous section, the speaker starts with a meaning M and searches a grammatical form F which fits best to his preferences. This means that, for a fixed M, he has to search the set

$$R(M) := \{F \mid \langle F, M\rangle \in \text{Gen}\}.$$

His preferences must rank the elements in this set, i.e. they must define a binary relation $\preceq_M$ on $R(M)$. We assume that $\preceq_M$ is at least a linear pre–order. A symmetric assumption has to be made for the interpreter. If we collect all the preference relations $\preceq_M$ and $\preceq_F$, then we end up with structures of the following type:

Definition 3.1 (Blutner Structure) *A tuple* $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \text{Gen}, \preceq\rangle$ *is a* Blutner structure *if*

1. Gen *is a subset of* $\mathcal{F} \times \mathcal{M}$.
2. $\preceq$ *is a family* $(\preceq_p)_{p\in P}$ *where* $P \subseteq \mathcal{F} \cup \mathcal{M}$, *and:*
    - $\preceq_F$ *is a linear pre–order on* $\{M \mid \langle F, M\rangle \in \text{Gen}\}$, $F \in \mathcal{F}$.
    - $\preceq_M$ *is a linear pre–order on* $\{F \mid \langle F, M\rangle \in \text{Gen}\}$, $M \in \mathcal{M}$.

We use the following terminology: We call $\mathcal{B}$ a two–sided *Blutner structure, iff $P = \mathcal{F} \cup \mathcal{M}$. If $P = \mathcal{F}$, then we call $\mathcal{B}$ $\mathcal{M}$–sided, and if $P = \mathcal{M}$, then $\mathcal{F}$–sided. We call $\mathcal{F}$ a set of* forms*, and $\mathcal{M}$ a set of meanings.*

Of course, $\mathcal{F}$ and $\mathcal{M}$ may denote any sets of objects. In case of Mattausch's Example we will define $\mathcal{F}$ as the set of sentences of natural language, and the set $\mathcal{M}$ as the set of their translations into a language of formal logic.

We claimed that Blutner structures make the idea explicit that constraints divide into a group which provides for the speaker's preferences on forms, and one group which provides for the hearer's preferences on meanings. It may seem that Blutner structures are much more general, and, hence, that this is not captured by the definition. But we will see in Lemma 3.6 that Blutner structures are really generated by *two* ranked sequences of constraints provided that the inverse relations of $\preceq_F$ and $\preceq_M$ are all well–founded.

Optimality and Weak Optimality

We first show that the notions of strong and weak optimality generalise to Blutner structures. As in the case of OT–systems, the preferences on forms and meanings, represented by $\preceq$, induces a pre–order $\leq$ on Gen. It is the (inverse) counterpart to the order $\leq_{\mathcal{O}}$ induced by an OT–system.

Definition 3.2 *Let $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \mathrm{Gen}, \preceq \rangle$ be a Blutner structure. Then we can define a pre–order $\leq$ on* Gen $\times$ Gen *if we build the transitive closure of the following relation: $\langle F, M \rangle \leq \langle F', M' \rangle$ iff*

$$(F = F' \;\&\; M \preceq_F M') \vee (M = M' \;\&\; F \preceq_M F').$$

Clearly, we can find a Blutner structure which allows us to represent circular structures. This implies that, in general, Blutner–structures cannot be generate by an OT–system.

We next turn to the Blutner–Jäger definitions of optimality and weak optimality.

Definition 3.3 (Optimality) *Let $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \mathrm{Gen}, \preceq \rangle$ be a Blutner structure, and $\langle \leq, \mathrm{Gen} \rangle$ the associated pre–order. A pair $\langle F, M \rangle \in$ Gen is* optimal*, iff it is a maximal element in $\langle \leq, \mathrm{Gen} \rangle$, i.e. iff, for all $\langle F', M' \rangle$, $\langle F, M \rangle \leq \langle F', M' \rangle \Rightarrow \langle F, M \rangle = \langle F', M' \rangle$.*

What we defined here as optimality is sometimes called *strong* optimality. Optimality is a central notion of BiOT as it assumes that speakers and interpreters agree only to choose strongly or *weakly* optimal form–meaning pairs. That they choose optimal form–meaning pairs can be expected from general considerations about rationality. Weak optimality is the empirically more interesting notion because its implications are not immediately explainable by general rationality assumptions. But we will make no use of it in our future considerations. Therefore, we just show that its definition can be generalised to Blutner structures. For more motivation we refer to (Blutner, 2000). The crucial condition for weak optimality is the following:

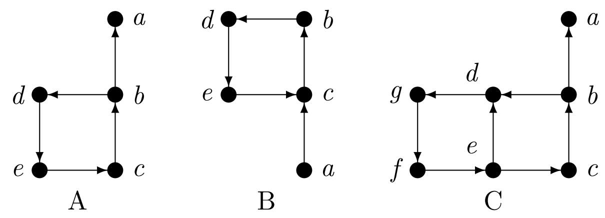

Figure 1: Three examples of 'circular' graphs.

Definition 3.4 (Weak Optimality Set) *Let $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \text{Gen}, \preceq \rangle$ be a Blutner structure. A Set $S \subseteq \text{Gen}$ is a* weak optimality set*, iff, for all $\langle F, M \rangle \in \text{Gen}$, $\langle F, M \rangle \in S$ implies that the following two conditions hold:*

1. *there is no pair $\langle F', M \rangle \in S$ such that $F \prec_M F'$,*
2. *there is no pair $\langle F, M' \rangle \in S$ such that $M \prec_F M'$.*

Intuitively, a pair $\langle F, M \rangle$ is weakly optimal, iff there is no weakly optimal $\langle F', M \rangle$ or $\langle F, M' \rangle$ which is preferred by one of the interlocutors. As it stands, this definition is circular, and, in fact, there are many sets S which have the property of Definition 3.4. We show that we can construct a weak optimality set if we start with the set of optimal form–meaning pairs, then take away all pairs which are immediately *dominated* by an optimal form–meaning pair (the first part of the disjunction in the following definition of D_α), and add then the new set of optimal pairs. This process can be repeated until no new optimal form–meaning pairs emerge. It is possible to extend the process even beyond this point (here we need the second part of the disjunction as well), but the choice of new elements for S is absolutely arbitrary. Clearly, this construction is a straightforward generalisation of Jäger's algorithm.

In order to get a better intuition for the following construction of weak optimality sets, let us consider the graphs in Figure 1. Assume that we have constructed a weak optimality set $S_{<\alpha}$ for Graph A with $a \in S_{<\alpha}$. Then we remove the dominated nodes b, c. In the remaining graph $\text{Gen}_\alpha = \{d, e\}$, e is optimal. Hence, we add it to the set of weakly optimal nodes and set $S_\alpha = S_{<\alpha} \cup \{e\}$. In Graph B, there is no optimal node. Hence, we can choose any node and add it to the set of already constructed weakly optimal nodes. Let us assume we choose b. In this case, we must remove all dominated nodes, i.e. a and c, and the dominating node d. This explains the disjunction in the following definition of D_α. Graph C combines the situations of graphs A and B. After adding a to the set of weakly optimal nodes and removing b, c, we can choose any of the nodes d, e, g, f and add them in the next construction step to the set of weakly optimal nodes.

Lemma 3.5 *Let $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \text{Gen}, \preceq \rangle$ be a Blutner structure, and $\langle \leq, \text{Gen} \rangle$ the associated preorder. It holds: 1.) The set of optimal elements is a weak optimality set. 2.) Every weak optimality set S can be extended to a maximal weak optimality set S^* by the following construction: Let $S_0 := S$ and $\text{Gen}_0 := \text{Gen}$. For $\alpha > 0$ we set:*

1. $S_{<\alpha} := \bigcup_{\beta<\alpha} S_\beta$,

2. $D_\alpha := \{\langle F, M\rangle \in \text{Gen} \,|\, \exists \langle F', M'\rangle \in S_{<\alpha} : (F = F' \vee M = M') \wedge (\langle F, M\rangle \leq \langle F', M'\rangle \vee \langle F', M'\rangle \leq \langle F, M\rangle)\}$,

3. $\text{Gen}_\alpha := \text{Gen} \setminus D_\alpha$.

4. *If* $\text{Gen}_\alpha \neq \emptyset$, *let* $Op_\alpha :=$ *the set of (strongly) optimal elements of* Gen_α.

5. *If* $Op_\alpha \neq \emptyset$, *set* $S_\alpha := S_{<\alpha} \cup Op_\alpha$. *If* $Op_\alpha = \emptyset$, *set* $S_\alpha := S_{<\alpha} \cup \{\langle F, M\rangle\}$ *for some arbitrary* $\langle F, M\rangle \in \text{Gen}_\alpha$.

6. *If* $\text{Gen}_\alpha = \emptyset$, *set* $S^* := S_{<\alpha}$.

3.) If we start the construction with the set S *of all (strongly) optimal elements of* Gen *and set* $S^* = S_{<\alpha}$ *for the first* α *such that* $Op_\alpha = \emptyset$ *or* $\text{Gen}_\alpha = \emptyset$, *then the maximal set* $WOT := S^*$ *is a weak optimality set. It is the intersection of all weak optimality sets which extend* WOT.

Proof: 1.) That the set of optimal elements is a weak optimality set follows trivially by definition.

2.) The union $\bigcup_{\beta<\alpha} S_\beta$ and all S_α are weak optimality sets. This follows directly by definition. Furthermore, $\beta < \alpha$ implies $S_\beta \subseteq S_\alpha$, hence, it follows with Zorn's Lemma that we ultimately find a maximal weak optimality set.

3.) That WOT is a weak optimality set follows by induction over α. Assume that S is a weak optimality set which extends WOT. Let $\langle F, M\rangle \in S \setminus WOT$. Then, there is no $\langle F', M'\rangle \in WOT$ such that $(F = F' \vee M = M') \wedge (\langle F, M\rangle \leq \langle F', M'\rangle \vee \langle F', M'\rangle \leq \langle F, M\rangle)$, and $\langle F, M\rangle$ can't be an optimal element for $\langle \preceq, \text{Gen}_\infty\rangle$, where we set $D_\infty := \{\langle F, M\rangle \in \text{Gen} \,|\, \exists \langle F', M'\rangle \in WOT : (F = F' \vee M = M') \wedge (\langle F, M\rangle \leq \langle F', M'\rangle \vee \langle F', M'\rangle \leq \langle F, M\rangle)\}$, and $\text{Gen}_\infty := \text{Gen} \setminus D_\infty$. Hence, there must exist a $\langle F', M'\rangle \in \text{Gen}_\infty\, (F = F' \vee M = M') \wedge (\langle F, M\rangle \leq \langle F', M'\rangle)$. As $\langle F', M'\rangle \in \text{Gen}_\infty$ we can start with WOT, choose $\langle F', M'\rangle$, and construct a maximal weak optimality set S'. But $\langle F, M\rangle \notin S'$ by definition of $D_{\infty+1}$, hence, $\langle F, M\rangle \notin S \cap S'$. □

From the third part of the lemma, it follows that the set WOT is the unique non–arbitrary maximal extension of the set of all optimal form–meaning pairs. Hence, we call a form–meaning pair $\langle F, M\rangle$ *weakly optimal* in an absolute sense iff it is an element of WOT.

Blutner Structures and OT–Systems

Now we show how Blutner structures are related to OT–systems. We will see that every Blutner structure $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \text{Gen}, \preceq\rangle$ which is such that the inverse of $\preceq_F$ and $\preceq_M$ are well–founded relations can be generated by two OT–systems over the same set of grammatical form–meaning pairs, i.e. they are generated by pairs $\langle \mathcal{O}_\mathcal{F}, \mathcal{O}_\mathcal{M}\rangle$ of OT–systems $\mathcal{O}_X = \langle \text{Gen}, C_X\rangle$. As both systems share the same set Gen, they differ only with respect to the sequence C of constraints. The construction below shows that $C_\mathcal{F}$ is responsible for the definition of the preferences $\preceq_M$ on $\mathcal{F}$, and $C_\mathcal{M}$ for the preferences $\preceq_F$ on $\mathcal{M}$. This means that the interpreter's preferences over meanings are defined by one sequence of constraints, and the speaker's preferences over forms by another sequence. Hence, Blutner structures really incorporate the idea that preferences are defined by *two* separate sequences of constraints, one for the speaker and one for the interpreter. We will see in the next section that we have to weaken the condition that both OT–systems are defined over the *same* set Gen if we want to capture dynamic contexts.

We first show how OT–systems over a fixed set Gen of form–meaning pairs generate Blutner structures. Let $\mathcal{O} = \langle \text{Gen}, C \rangle$ be an OT–system over a set of form–meaning pairs $\text{Gen} \subseteq \mathcal{F} \times \mathcal{M}$ with $C = \langle c_\alpha \rangle_{\alpha<\beta}$. We can then Definition 2.3 in order to acquire for each $M \in \mathcal{M}$ and $F \in \mathcal{F}$ associated linear well–founded pre–orders $<_{\mathcal{O}_M}$ on $R(M) := \{F \in \mathcal{F} \mid \langle F, M \rangle \in \text{Gen}\}$ and $<_{\mathcal{O}_F}$ on $R(F) := \{M \in \mathcal{M} \mid \langle F, M \rangle \in \text{Gen}\}$: For $\langle F, M \rangle, \langle F', M' \rangle \in \text{Gen}$ we set

1. $F <_{\mathcal{O}_M} F'$ iff there is an γ with $1 \leq \gamma < \beta$ such that $c_\gamma(F, M) < c_\gamma(F', M)$ and for all $\alpha < \gamma : c_\alpha(F, M) = c_\alpha(F', M)$;
2. $M <_{\mathcal{O}_F} M'$ iff there is an γ with $1 \leq \gamma < \beta$ such that $c_\gamma(F, M) < c_\gamma(F, M')$ and for all $\alpha < \gamma : c_\alpha(F, M) = c_\alpha(F, M')$.

Then, we define $\preceq_F$ as the inverse of $<_{\mathcal{O}_F}$, and $\preceq_M$ analogously. If we set $\preceq = (\preceq_p)_{p \in P}$ for $P = \mathcal{F}$ or $P = \mathcal{M}$, then $\langle \mathcal{F}, \mathcal{M}, \text{Gen}, \preceq \rangle$ is a one–sided Blutner structure.

It is easy to see that two one–sided Blutner structures $\mathcal{B}_\mathcal{F}$ and $\mathcal{B}_\mathcal{M}$, where $\mathcal{B}_\mathcal{F}$ is $\mathcal{F}$–sided, and $\mathcal{B}_\mathcal{M}$ is $\mathcal{M}$–sided, generate a two–sided Blutner structure $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \text{Gen}, (\preceq_p)_{p \in P} \rangle$ if we set $P = P_\mathcal{F} \cup P_\mathcal{M}$. Hence, we see that any pair $\langle \mathcal{O}_1, \mathcal{O}_2 \rangle$ with $\mathcal{O}_i = \langle \text{Gen}, C_i \rangle$ generates a two–sided Blutner structure.

Now we also show the inverse. Again it is easy to see that we can separate a two–sided Blutner structure into a $\mathcal{F}$–sided and a $\mathcal{M}$–sided structure. Hence, we can restrict our considerations to one–sided structures. We show that we can find for any $\mathcal{F}$–sided Blutner structure $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \text{Gen}, \preceq \rangle$ where for all $M \in \mathcal{M}$ the inverse of $\preceq_M$ is a well–founded pre–order an underlying OT–system $\langle \text{Gen}, C_\mathcal{F} \rangle$ which generates $\mathcal{B}$.

Lemma 3.6 *For every $\mathcal{F}$–sided Blutner structure $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \text{Gen}, \preceq \rangle$ which is such that the inverse $\leq_M$ of $\preceq_M$ is a well–founded pre–order for all $M \in \mathcal{M}$, there exists an OT–system $\langle \text{Gen}, C_\mathcal{F} \rangle$, such that for all $\langle F, M \rangle, \langle F', M \rangle \in \text{Gen}$:*

$$F \leq_M F' \iff F \leq_{\mathcal{O}_M} F'.$$

Proof: Let $\leq_M$ be the inverse relation of $\preceq_M$ on $R(M)$. By assumption, $\leq_M$ is well–founded. For each $F \in R(M)$ we can define equivalence classes $[F]_M := \{F' \in R(M) \mid F \leq_M F' \,\&\, F' \leq_M F\}$, where $[F]_M \leq_M [F']_M \Leftrightarrow F \leq_M F'$. Hence $\langle \{[F]_M \mid F \in R(F)\}, \leq_M \rangle$ is well–ordered for each $M \in \mathcal{M}$. Let $f(M, F)$ be the associated order type of $[F]_M$. Then we set:

$$\begin{aligned} c_\alpha(M, F) := 0 \quad &\iff \quad f(M, F) = \alpha \\ c_\alpha(M, F) := 1 \quad &\iff \quad f(M, F) \neq \alpha. \end{aligned}$$

We set $\mathcal{O}_\mathcal{F} := \langle \text{Gen}, C_\mathcal{F} \rangle$ with $C_\mathcal{F} := \langle c_\alpha \rangle_{\alpha<\beta_\mathcal{F}}$, where $\beta_\mathcal{F} := \sup_{M \in \mathcal{M}}(\beta_M + 1)$, and $\beta_M := \sup_{F \in R(M)}(f(M, F) + 1)$. Let $F, F' \in R(M)$, $\gamma := f(M, F)$ and $\gamma' := f(M, F')$. Then we find: $F \leq_M F'$ *iff* $\gamma \leq \gamma'$ *iff* $(\forall \alpha < \beta_M \, c_\alpha(M, F) = c_\alpha(M, F')) \vee \forall \alpha < \gamma \, (c_\alpha(M, F) = c_\alpha(M, F') \wedge c_\gamma(M, F) < c_\gamma(M, F'))$ *iff* $F \leq_{\mathcal{O}_\mathcal{F}} F'$. □

The same result holds for $\mathcal{M}$. Hence, we can sum up our result as:

Lemma 3.7 *Let $\mathcal{B} = \langle \mathcal{F}, \mathcal{M}, \text{Gen}, \preceq \rangle$ be a Blutner structure such that the inverse of $\preceq_F$ and $\preceq_M$ are well–founded relations. Then there is a pair $\langle \mathcal{O}_\mathcal{F}, \mathcal{O}_\mathcal{M} \rangle$ of OT–systems $\mathcal{O}_\mathcal{F} = \langle \text{Gen}, C_\mathcal{F} \rangle$ and $\mathcal{O}_\mathcal{M} = \langle \text{Gen}, C_\mathcal{M} \rangle$ which generate $\mathcal{B}$.*

Blutner Structures with Context–Sensitive Constraints

Our structures represent situations where the constraints depend on arguments of the form $\langle F, M\rangle$. In case of Mattausch's Example, we applied the constraint (Agr), which says that the gender of a pronoun must agree with the gender of the person referred to. The real person is neither part of a natural sentence nor of its translation into a sentence of formal logic. Hence, this constraint cannot be represented as a function which only depends on a form–meaning pair. We want to extend Blutner structures in order to be able to handle arbitrary contexts given by a set $\mathcal{C}$. Later, in Sections 4 and 5, we define structures where the real world and the information state of the interpreter are possible arguments of constraints. Now, we consider only constraints with arguments $\langle c, F, M\rangle$. But we will show that this extension does not introduce really new structures, i.e. they can still be considered to be Blutner structures in the sense of Def. 3.1.

Definition 3.8 (Blutner Structure with Contexts)
A Blutner structure with contexts *is a tuple* $\mathcal{B} = \langle \mathcal{C}, \mathcal{F}, \mathcal{M}, \mathrm{Gen}, \preceq\rangle$ *where*

- $\mathcal{C}$, $\mathcal{F}$, *and* $\mathcal{M}$ *are sets.*
- Gen *is a subset of* $\mathcal{C} \times \mathcal{F} \times \mathcal{M}$.
- $\preceq$ *is a family* $(\preceq_p)_{p\in P}$ *with* $P \subseteq \mathcal{C} \times (\mathcal{F} \cup \mathcal{M})$ *where*
  - $\preceq_{c,F}$ *is a linear pre–order on* $\{M \mid \langle c, F, M\rangle \in \mathrm{Gen}\}$.
  - $\preceq_{c,M}$ *is a linear pre–order on* $\{F \mid \langle c, F, M\rangle \in \mathrm{Gen}\}$.

We call $\mathcal{F}$ *a set of* forms*,* $\mathcal{M}$ *a set of meanings, and* $\mathcal{C}$ *a set of contexts.*

We now show that every Blutner structure $\mathcal{B} = \langle \mathcal{C}, \mathcal{F}, \mathcal{M}, \mathrm{Gen}, \preceq\rangle$ with contexts can be represented as a Blutner structure in the sense of Definition 3.1. Let $\langle \mathcal{C}, \mathcal{F}, \mathcal{M}, \mathrm{Gen}, \preceq\rangle$ be given. Then we set:

$$\begin{aligned} \mathcal{M}' &:= \{\langle c, M\rangle \mid \exists F \in \mathcal{F}\ \langle c, F, M\rangle \in \mathrm{Gen}\}, \\ \mathcal{F}' &:= \{\langle c, F\rangle \mid \exists M \in \mathcal{M}\ \langle c, F, M\rangle \in \mathrm{Gen}\}, \\ \mathrm{Gen}' &:= \{\langle \langle c, F\rangle, \langle c, M\rangle\rangle \mid \langle c, F, M\rangle \in \mathrm{Gen}\}. \end{aligned}$$

Then we can define $\preceq$:

$$\begin{aligned} \langle c, M\rangle \preceq'_{\langle c,F\rangle} \langle c, M'\rangle \quad &\textit{iff} \quad M \preceq_{c,F} M' \\ \langle c, F\rangle \preceq'_{\langle c,M\rangle} \langle c, F'\rangle \quad &\textit{iff} \quad F \preceq_{c,M} F'. \end{aligned}$$

$\preceq'_{\langle c,F\rangle}$ and $\preceq'_{\langle c,M\rangle}$ are linear pre–orders on the sets $\{\langle c, M\rangle \mid \langle c, F, M\rangle \in \mathrm{Gen}\}$ and $\{\langle c, F\rangle \mid \langle c, F, M\rangle \in \mathrm{Gen}\}$ respectively. Hence, $\mathcal{B}' = \langle \mathcal{F}', \mathcal{M}', \mathrm{Gen}', \preceq'\rangle$ is a Blutner structure in the sense of Def. 3.1.

We now show that they both induce the *same* preference relation on Gen and Gen$'$. The map

$$e : \mathrm{Gen} \longrightarrow \mathrm{Gen}', \langle c, F, M\rangle \mapsto \langle \langle c, F\rangle, \langle c, M\rangle\rangle$$

is a bijection. Let $\leq$ be the pre–order on Gen induced by $\mathcal{B}$, and $\leq'$ the pre–order on Gen$'$ induced by $\mathcal{B}'$. Then, for all $a, b \in \mathrm{Gen}$ it holds that:

$$a \leq b \iff e(a) \leq' e(b).$$

This follows by:

$$\begin{aligned}
&\langle c, F, M\rangle \leq \langle c', F', M'\rangle \;\Leftrightarrow \\
&\quad \Leftrightarrow c = c' \;\&\; (M = M' \wedge F \preceq_{c,M} F' \vee F = F' \wedge M \preceq_{c,F} M') \\
&\quad \Leftrightarrow c = c' \;\&\; (M = M' \wedge \langle c, F\rangle \preceq_{\langle c,M\rangle} \langle c, F'\rangle \vee F = F' \wedge \langle c, M\rangle \preceq_{\langle c,F\rangle} \langle c, M'\rangle) \\
&\quad \Leftrightarrow \langle \langle c, F\rangle, \langle c, M\rangle\rangle \leq' \langle \langle c', F'\rangle, \langle c', M'\rangle\rangle
\end{aligned}$$

This allows us to generalise all results to Blutner structures with contexts.

4 Mattausch's Example

In this section we show how to solve Mattausch's Problem (2000) within a Bi-OT approach using Blutner structures. Beaver (2004) applied a version of a two–sided OT to predict choice and explain resolution of anaphoric expressions. Mattausch (2000) improves on this approach. He introduces his crucial example in a discussion (Mattausch, 2000, pp.33–36) of bidirectional OT in the version of (Blutner, 2000), (Blutner and Jäger, 2000) or (Jäger, 2002). For convenience, we repeat it here as **(2)**:

(2) Assume that Marion is a male person, and Jo a female one. The speaker wants to express with the second sentence that Jo was pulling Marion's hair out.

- **a)** Marion was frustrated with Jo. She was pulling his hair out.
- **b)** Marion was frustrated with Jo. He was pulling her hair out.
- **c)** Marion was frustrated with Jo. Jo was pulling Marion's hair out.

As we have seen before, with the constraints stated on page 6, we seem to arrive at a circle as depicted in Figure 2.

$$\begin{array}{cccccccc}
 & F_3 & & F_1 & & F_2 & \\
 & \langle Jo, Marion\rangle & & \langle He, She\rangle & & \langle She, He\rangle & \\
(jo, ma) & \langle F_3, M_1\rangle & \longrightarrow & \langle F_1, M_1\rangle & \longrightarrow & \langle F_2, M_1\rangle & M_1 \\
 & & & \uparrow & & \downarrow & \\
(ma, jo) & & & \langle F_1, M_2\rangle & \longleftarrow & \langle F_2, M_2\rangle & M_2
\end{array}$$

Figure 2: Mattausch's Circle with meanings $M_1 :=$ *pulling–hair–out*(jo, ma) and $M_2 :=$ *pulling–hair–out*(ma, jo).

Our discussion at the end of Section 2 showed that the apparent circularity is an effect of the epistemic asymmetry between speaker and hearer. The situation where speaker and hearer have to co–ordinate their choice of an optimal form–meaning pair is given by the context generated by the utterance of *Marion was frustrated with Jo.* The second sentence contains pronouns, hence, it is natural to look for a dynamic framework. In a first step we will combine Dynamic Semantics with bidirectional OT using Blutner structures for (static) contexts.

The underlying idea will be that Dynamic Semantics accounts for the contextual meaning of formulas, and OT provides for the pragmatic constraints which govern anaphora resolution.[2]

Intuitively, the continuatiuon '*She was pulling his hair out*' is misleading if Marion is male and Jo female. We solve Mattausch's puzzle by postulating a new principle which rules out the choice of misleading forms. In order to formulate it, we first need a precise characterisation of *misleading forms.* This characterisation will be formulated in terms of the generator sets Gen. The epistemic asymmetry has the effect that for the hearer there exists many more grammatical form-meaning pairs than for the speaker. In Dynamic Semantics it is implicitly assumed that the speaker knows the true state of affairs, hence, he knows Gen. The hearer's grammatical form meaning pairs will be collected in a set Gen_H. We say then that a form F is misleading in a context c iff the hearer prefers an interpretation M such that $\langle c, F, M\rangle \in \text{Gen}_H \setminus \text{Gen}$.

We now set up our general framework and analyse Mattausch's example. This will lead us to the searched characterisation of Gen_H. We introduce two structures: the first $\langle NL, \mathcal{L}, ^* \rangle$ tells us how to translate natural sentences $F \in NL$ into first order formulas $\varphi \in \mathcal{L}$, the second $\langle \mathcal{L}, \mathcal{W}, [\![\ ]\!]\rangle$ represents the dynamic semantics of $\mathcal{L}$.

Let $\mathcal{L}$ be a set of well–formed formulas of a first order language. We assume that for all sentences of natural language there exists a translations in $\mathcal{L}$ which is unique up to substitutions of free variables. For a given sentence $F \in NL$, we denote the set of all translations by F^*. Let W be a set of first order models for $\mathcal{L}$ representing the possible states of affairs. In Dynamic Semantics, the *meaning* of a sentence is identify with its *update potential*, i.e. with a relation between world–assignment pairs. A world–assignment pair is a pair (w, f) in which f is a partial assignment function of variables and $w \in W$. We collect all world–assignment pairs in a set $\mathcal{W}$ and denote denote the set of variables by Var. With these structures given, we can say what counts as a dynamic interpretation: For every formula $\varphi \in \mathcal{L}$, $[\![\varphi]\!]$ is a set of pairs $\langle \sigma, \tau\rangle$ with $\sigma, \tau \subseteq \mathcal{W}$. Furthermore, we assume that $[\![\varphi]\!]$ is a function. This means that the update effect of a sentence of natural language is given once we have resolved anaphora. This allows us to define the *meaning* $[\![\varphi]\!]_\sigma$ *of a formula* φ *in context* σ: It is defined if $\sigma \in \text{dom}[\![\varphi]\!]$ and $[\![\varphi]\!]_\sigma = \tau \Leftrightarrow \langle \sigma, \tau\rangle \in [\![\varphi]\!]$. The subsets σ, τ of $\mathcal{W}$ are called *information states.* They represent — and this is important for the following representation of contexts — the *hearer's* knowledge. This completes the definition of $\langle NL, \mathcal{L}, ^* \rangle$ and $\langle \mathcal{L}, \mathcal{W}, [\![\ ]\!]\rangle$. We next introduce Blutner structures with contexts.

In general, the set of forms $\mathcal{F}$ is given by the set of all syntactically correct sentences of natural language. For the present purpose we can restrict our considerations to the set:

$$\mathcal{F} := \{\underbrace{\text{She was pulling his hair out}}_{F_1}, \underbrace{\text{He was pulling her hair out}}_{F_2}\}$$

The set $\mathcal{M}$ of meanings is the set of all first order formulas. Hence, $\mathcal{M} = \mathcal{L}$.

By the *epistemic asymmetry* between speaker and hearer, we mean that the speaker knows the actual state of the world w but the hearer only the common

[2]That OT should be seen in context with Dynamic Semantics has been suggested by Blutner (2000). The idea that the two theories should work together in anaphora resolution goes back to (Beaver, 2004), and, later, Mattausch (2000).

ground which is represented by an information state σ. Hence, we consider contexts of the form $c = \langle w, \sigma \rangle$ such that the set of all context C is:

$$\mathcal{C} := \{\langle w, \sigma \rangle \mid w \in W \,\&\, w \in \sigma \subseteq \mathcal{W}\}.$$

Now we have to say which form–meaning pairs can be generated in which contexts. The elements in Gen have the form $\langle c, F, \varphi \rangle$ such that c is a pair $\langle w, \sigma \rangle$. We have to consider the constraints imposed by our dynamic semantics and the fact that φ must be a possible translation for F. This leads to the following definition:

$$\text{Gen} := \{\langle \langle w, \sigma \rangle, F, \varphi \rangle \mid \varphi \in F^* \,\&\, \sigma \in \text{dom}[\![\varphi]\!] \,\&\, \exists f\, (w, f) \in [\![\varphi]\!]_\sigma\}.$$

This means, Gen is the set of all $\langle c, F, \varphi \rangle$ such that 1) φ is a possible translation of the natural sentence F ($\varphi \in F^*$), 2) the information state of the hearer can be updated with φ ($\sigma \in \text{dom}[\![\varphi]\!]$), and 3) the existential closure of φ must be a true sentence in w ($\exists f\, (w, f) \in [\![\varphi]\!]_\sigma$). $\text{dom}[\![\varphi]\!]$ denotes the domain of φ. Hence, if $\sigma \in \text{dom}[\![\varphi]\!]$, then $[\![\varphi]\!]_\sigma$ is defined. Gen is the set of all proper *grammatical* form–meaning pairs.

As the hearer does not know the actual state of the world, he cannot distinguish between contexts with equal information states. Hence, his set of *potentially* grammatical form–meaning pairs is, in general, larger that the set of actually grammatical form–meaning pairs. In a context $\langle w, \sigma \rangle$, he has to consider all form–meaning pairs $\langle F, \varphi \rangle$ for wich there is a context $\langle w', \sigma \rangle$ such that $\langle F, \varphi \rangle \in$ Gen. This means that the hearer's set of alternatives is defined by:

$$\text{Gen}_H := \{\langle \langle w, \sigma \rangle, F, \varphi \rangle \mid \exists w' \in W\ \langle \langle w', \sigma \rangle, F, \varphi \rangle \in \text{Gen}\}.$$

The speaker, on the other hand, has full knowledge about the utterance situation, i.e. the set of forms where his choice takes place is determined by the set Gen. This asymmetry between speaker and hearer cannot be represented by Blutner structures with contexts as introduced in Definition 3.8. We will provide more general structures in the next section. But first, we provide a detailed analysis of Mattausch's example.

We follow Jäger's algorithm for finding optimal form–meaning pairs. We will see that it leads us into what we call a "dead end", i.e. a meaning which cannot be expressed by the speaker. Let us denote *Marion* by m, and *Jo* by j. As short forms we use:

$$\begin{aligned} \varphi(v_1, v_2) \quad &:\Leftrightarrow \quad \textit{pull–hair–out}(v_1, v_2), \\ \psi(v_1, v_2) \quad &:\Leftrightarrow \quad \textit{Marion}(v_1) \wedge \textit{Jo}(v_2) \wedge \textit{frustrated–with}(v_1, v_2). \end{aligned}$$

We assume that the actual world is such that

$$w_0 \models \psi(m, j) \wedge \varphi(j, m) \wedge \textit{male}(m) \wedge \textit{female}(j)$$

We assume that the hearer chooses a variable x for *Marion* and a variable y for *Jo* when processing the sentence *Marion was frustrated with Jo*. His information state is then represented by the set

$$\sigma_0 = \{(w, f) \,|\, x, y \in \text{dom} f \,\&\, (w, f) \models \textit{Marion}(x) \wedge \textit{Jo}(y) \wedge \textit{frustrated–with}(x, y)\}.$$

The context where the speaker has to choose the best form and where the hearer has to find the correct translation is given by the pair $c = \langle w_0, \sigma_0 \rangle$. This implies that the hearer has to choose the variable x as translation of a pronoun if it should refer to Marion, and y if it should refer to Jo. For convenience, we repeat the three constraints introduced on page 6:

(Agr) pronouns agree with the natural gender of the person referred to;

(Pro) pronouns are preferred over names;

(Dft) pronouns agree with the default interpretation of the gender of names.

The algorithm starts with the meaning which the speaker wants to express. It is his goal that the hearer updates his information state with:

$$\tau_1 := \{(w, f) \mid x, y \in \mathrm{dom} f \;\&\; f(x) = m \;\&\; f(y) = j \;\&\; (w, f) \models \varphi(y, x)\}.$$

This is the meaning of the formula $\varphi(y, x)$ in σ_0, i.e. $[\![\varphi(y, x)]\!]_{\sigma_0} = \sigma_0 \cap \tau_1$. By (Agr) and (Pro), the best choice for the speaker are the pronouns '*he*' for Marion and '*she*' for Jo. This means that, in context $c = \langle w_0, \sigma_0 \rangle$, he prefers F_1 over F_2, i.e. $F_2 \prec_{c,\varphi(y,x)} F_1$.

Now we have to take the perspective of the interpreter and choose the best translation for F_1. We assume here that it is part of the semantics of a pronoun whether or not its referent is male or female. It follows that the set of translations for F_1 is given by:[3]

$$F_1^* = \{\underbrace{\varphi(v_1, v_2) \wedge \mathit{female}(v_1) \wedge \mathit{male}(v_2)}_{\mu(v_1, v_2)} \mid v_1, v_2 \in \mathrm{Var}\}$$

The set of alternatives where the choice takes place in context $c = \langle w_0, \sigma_0 \rangle$ is defined as:

$$R_c(F_1) = \{\lambda \in \mathcal{L} \mid \exists w \in W \; \langle \langle w, \sigma_0 \rangle, F_1, \lambda \rangle \in \mathrm{Gen}_H\}$$

Hence, $R_c(F_1) = \{\mu(v_1, v_2) \mid v_1, v_2 \in \mathrm{Var} \;\&\; \sigma_0 \in \mathrm{dom}[\![\mu(v_1, v_2)]\!] \;\&\; \exists w \exists f \, (w, f) \in [\![\mu(v_1, v_2)]\!]_{\sigma_0}\} = \{\mu(v_1, v_2) \mid (v_1 = x \wedge v_2 = y) \vee (v_1 = y \wedge v_2 = x)\} = \{\mu(x, y), \mu(y, x)\}$. Here we have to assume that x and y are the only variables interpreted so far by the hearer.

With (Dft), it follows that the hearer has to choose y for '*He*' and x for '*She*' as $\sigma_0 \models \mathit{Jo}(y) \,\&\, \mathit{Marion}(x)$. Hence, he will translate F_1 into $\mu(x, y)$, and update his information state with the set:

$$\tau_2 := \{(w, f) \mid x, y \in \mathrm{dom} f \;\&\; (w, f) \models \varphi(x, y) \wedge \mathit{female}(x) \wedge \mathit{male}(y)\}.$$

[3]Of course, our decision to make the gender of a pronoun part of the meaning, and therefore of its translation, has some repercussions with regard to our previous definitions because now F_1 is no longer a possible choice for the producer. This follows because $\varphi(y, x)$ is not a translation for F_1, hence $\langle \langle w_0, \sigma_0 \rangle, F_1, \varphi(y, x) \rangle \notin \mathrm{Gen}$. This problem is due to the strong condition for grammaticality which implies that the formula has to be an exact translation. But we can weaken this condition: The exact translation has to imply the formula. This is without influence on our subsequent discussion. The definition of Gen then is as follows: $\langle \langle w, \sigma \rangle, F, \varphi \rangle \in \mathrm{Gen}$ iff

$$\sigma \in \mathrm{dom}[\![\varphi]\!] \wedge w \in [\![\varphi]\!]_\sigma \wedge \exists \psi \in F^* \, (\sigma \in \mathrm{dom}[\![\psi]\!] \wedge w \in [\![\psi]\!]_\sigma \wedge [\![\psi]\!]_\sigma \subseteq [\![\varphi]\!]_\sigma).$$

Now we have to switch to the perspective of the speaker again. He has to choose the optimal sentence in the context $c = \langle w_0, \sigma_0 \rangle$ for $\mu(x, y)$. Hence, in accordance with Definition 3.8, he has to choose the preferred form from the set

$$R_c(\mu(x,y)) = \{F \in \{F_1, F_2\} \mid \langle c, F, \mu(x,y) \rangle \in \text{Gen}\}.$$

This is the set of all $F \in \{F_1, F_2\}$ such that

$$\mu(x,y) \in F^* \;\&\; \sigma_0 \in \text{dom}[\![\mu(x,y)]\!] \;\&\; \exists f\, (w_0, f) \in [\![\mu(x,y)]\!]_{\sigma_0}.$$

$[\![\mu(x,y)]\!]_{\sigma_0}$ is the set of all $(w, f) \in \sigma_0$ such that:

$$x, y \in \text{dom} f \;\&\; (w, f) \models \varphi(x,y) \wedge \mathit{female}(x) \wedge \mathit{male}(y).$$

Hence, it is the set of all (w, f) such that $x, y \in \text{dom} f$ and

$$(w, f) \models \mathit{Marion}(x) \wedge \mathit{Jo}(y) \wedge \psi(x,y) \wedge \varphi(x,y) \wedge \mathit{female}(x) \wedge \mathit{male}(y).$$

But for all (w_0, f) with $x, y \in \text{dom} f$ and $f(x) = \mathit{Marion}$ and $f(y) = \mathit{Jo}$ it holds that $(w_0, f) \models \neg\varphi(x,y) \wedge \neg\mathit{female}(x) \wedge \neg\mathit{male}(y)$. Hence

$$R_c(\mu(x,y)) = \emptyset.$$

This means that, in context c, there are no grammatical forms which the speaker could choose from. Hencve, the algorithm has reached a *dead end.* $\langle \langle w_0, \sigma_0 \rangle, F_1, \mu(x,y) \rangle$ is hearer optimal in Gen_H but it is not an element of Gen. Clearly, whenever the hearer chooses an interpretation M for a form F in context c such that $\langle c, F, M \rangle$ is an element of $\text{Gen}_H \setminus \text{Gen}$, Jäger's algorithm ends in a dead end. Hence, we postulate the following principle:

> We call the elements of $\text{Gen}_H \setminus \text{Gen}$ *dead ends*, and require for (strong) *optimality* that the considered form–meaning pairs are maximal in the set Gen minus the set of all form–meaning pairs for which the hearer's preferences would lead into a dead end.

Don't force the interpreter to choose an ungrammatical meaning!

We show now how Mattausch puzzle can be solved. As mentioned before, we supply a precise description of the necessary modifications to Blutner structures in the next section.

According to our definition, the pair $\langle$*She was pulling his hair out*$, \mu(x,y) \rangle$ is a dead end in context $c = \langle w_0, \sigma_0 \rangle$. Hence, the triple $\langle c, F_1, \mu(y,x) \rangle$ is a maximal element in Gen but it must be avoided as $\mu(y,x) \prec_{c,F_1} \mu(x,y)$ and $\langle c, F_1, \mu(x,y) \rangle \notin \text{Gen}$. The principle stated above implies that we have to exclude all elements in Gen which necessarily lead into a dead end, i.e. we have to remove all triples of the form $\langle c, F_1, \,.\, \rangle$. In this way we reduce the set of possible choices for the speaker and get new optimal form–meaning pairs. Especially, F_1 is no longer a possible choice.

We now consider Mattausch's problem in its full version with originally three choices for the speaker; i.e. $\mathcal{F} := \{F_1, F_2, F_3\}$ with $F_1 :=$ 'She was pulling his hair out,' $F_2 :=$ 'He was pulling her hair out,' and $F_3 :=$ 'Jo was pulling Marion's hair out.'

Our considerations showed that the choice of F_1 for $\varphi(y,x)$ in situation $\langle w_0, \sigma_0 \rangle$ necessarily leads into a dead end, hence our mechanism removes it

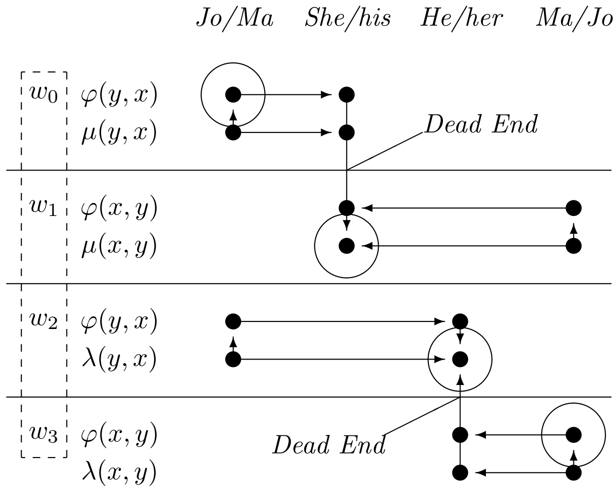

Figure 3: A graphical representation of the Mattausch's Problem. The worlds w_i are the worlds where $Marion(x) \wedge Jo(y) \wedge frustrated\text{–}with(x,y)$, and where the formulas listed in the second column hold.

from $\mathcal{F}$. F_2 is not a possible choice either because $\langle\langle w_0,\sigma_0\rangle, F_2, \varphi(y,x)\rangle$ is not an element of Gen. Hence, it only remains F_3 as a possible choice. Of course, $\langle\langle w_0,\sigma_0\rangle, F_3, \varphi(y,x)\rangle$ is grammatical, and it is easy to see that $\varphi(y,x)$ is also the preferred translation of F_3. Hence, it turns out that

$$\langle\langle w_0,\sigma_0\rangle, \text{Jo was pulling Marion's hair out}, \mathit{pull\text{–}hair\text{–}out}(y,x)\rangle$$

is optimal.

Figure 3 provides a graphical solution to Mattausch's Problem. The first row lists the different forms the speaker can choose. In the first column we list the different contexts. As the hearer's information state σ_0 is fixed, we listed only the worlds $w_i \in \sigma_0$. All elements of σ_0 are indiscernible for the hearer. This is indicated by the dashed box around them. In the second column we list the different formulas which represent the possible translations of the forms. We use the following abbreviations:

- $\varphi(v_1,v_2)$ *iff* *pull–hair–out*(v_1,v_2),
- $\mu(v_1,v_2)$ *iff* $\varphi(v_1,v_2)$ & *female*(v_1) & *male*(v_2),
- $\lambda(v_1,v_2)$ *iff* $\varphi(v_1,v_2)$ & *male*(v_1) & *female*(v_2).

The big dots represent the form–meaning pairs which can be generated in the context listed in the first column. The horizontal arrows show the preferences of the speaker, the vertical ones the preferences of the interpreter. The vertical arrows which cross the lines separating the different contexts indicate that they lead into a dead end. Of course, the dead ends itself are not listed in the picture. The circles around the big dots indicate the optimal form–meaning pairs.

5 Bidirectional OT for Dynamic Contexts

It remains to describe the structures which we used to solve Mattausch's Problem. First, we notice that it was essential that the set of meanings from which the interpreter can choose the best meaning may depend on a set Gen_H of possible form–meaning pairs which is different from the set $\mathrm{Gen}_{(S)}$ which restricts the speaker's choice of forms. We saw that Gen_H is generated by an equivalence relation on contexts. This has consequences for the speaker's and the hearer's preference relation on meanings and forms: The speaker's preference relation on forms is defined for the original contexts, whereas the interpreter's preference relation on meanings is defined for equivalence classes of contexts.

We saw in Section 3 that we can identify a Blutner structure with a pair of OT–systems $\langle \mathcal{O}_\mathcal{F}, \mathcal{O}_\mathcal{M}\rangle$, or equivalently with a pair $\langle \mathcal{B}_\mathcal{F}, \mathcal{B}_\mathcal{M}\rangle$ of two one–sided Blutner structures. Both systems are built over the same set of grammatical form–meaning pairs Gen. We use this representation and generalise it in order to cover dynamic contexts.

Definition 5.1 (Blutner Structures for Dynamic Contexts)
Let $\langle \mathcal{B}_\mathcal{F}, \mathcal{B}_\mathcal{M}, [\ ]\rangle$ *be a triple where*

1. $B_\mathcal{F} = \langle \mathcal{C}, \mathcal{F}, \mathcal{M}, \mathrm{Gen}, \preceq\rangle$ *is a* $\mathcal{F}$*–sided Blutner structure.*
2. $\mathcal{B}_\mathcal{M} = \langle \mathcal{C}_H, \mathcal{F}, \mathcal{M}, \mathrm{Gen}_H, \preceq\rangle$ *is a* $\mathcal{M}$*–sided Blutner structure.*
3. $[\ ] : \mathcal{C} \longrightarrow \mathcal{C}_H$, $c \mapsto [c]$, *maps* $\mathcal{C}$ *onto* $\mathcal{C}_H$.
4. $\mathrm{Gen}_H = \{\langle [c], F, M\rangle \mid \langle c, F, M\rangle \in \mathrm{Gen}\}$

We write $c' \in [c]$ *iff* $[c'] = [c]$.

As $[\ ]$ is a function, it follows that $c \sim c' \Leftrightarrow [c] = [c']$ defines an equivalence relation. This definition is slightly more general than is needed for dynamic contexts. There, contexts had the form $\langle w, \sigma\rangle$ in which w represents the real utterance situation, and σ the common ground. If we set $[\langle w, \sigma\rangle] := \{\langle v, \sigma\rangle \mid v \in \sigma\}$, or simply $[\langle w, \sigma\rangle] := \sigma$, then Gen_H is of the form:

$$\mathrm{Gen}_H = \{\langle [\langle w, \sigma\rangle], F, \varphi\rangle \mid \langle \langle w, \sigma\rangle, F, \varphi\rangle \in \mathrm{Gen}\}.$$

Hence, for a fixed $F \in NL$ we find that the set of meanings $R_{[\langle w,\sigma\rangle]}(F)$ from which the interpreter can make his choice is given by:

$$\begin{aligned} R_{[\langle w,\sigma\rangle]}(F) &= \{\varphi \in \mathcal{L} \mid \langle [\langle w, \sigma\rangle], F, \varphi\rangle \in \mathrm{Gen}_H\} \\ &= \{\varphi \in \mathcal{L} \mid \exists v \in \sigma\ \langle \langle v, \sigma\rangle, F, \varphi\rangle \in \mathrm{Gen}\}. \end{aligned}$$

This means that he has to consider a formula φ as a translation for a sentence F, iff form him there is any epistemically possible context in which $\langle F, \varphi\rangle$ is grammatical. This is exactly as intended.

A two–sided Blutner structure for dynamic contexts again induces a pre–order $\leq$ on Gen:

Definition 5.2 *Let* $\langle \mathcal{B}_\mathcal{F}, \mathcal{B}_\mathcal{M}, [\ ]\rangle$ *a Blutner structure for dynamic contexts with* $\mathcal{B}_\mathcal{F} = \langle \mathcal{F}, \mathcal{M}, \mathrm{Gen}, \preceq\rangle$. *Then we define a pre–order* $\leq$ *on* $\mathrm{Gen} \times \mathrm{Gen}$ *by:* $\langle c, F, M\rangle \leq \langle c', F', M'\rangle$ *iff*

$$(c = c' \ \&\ M = M' \ \&\ F \preceq_{c,M} F') \vee (c' \in [c] \ \&\ F = F' \ \&\ M \preceq_{[c],F} M').$$

Let us for now assume that the sets of forms and meanings are finite, and that the preference relations $\preceq$ are proper linear orders. Jäger's algorithm defines a sequence $(a_n)_{n\in\mathbb{N}}$ of forms and meanings for which $a_{2n} \in \mathcal{M}$ and $a_{2n+1} \in \mathcal{F}$. If the induced order $\leq$ on Gen is well–founded, then the sequence will eventually always choose the same forms and meanings, i.e. $\exists M \in \mathbb{N} \forall m \geq M\, a_{m+2} = a_m$. If $\langle \leq, \mathrm{Gen}\rangle$ is not well–founded, then the sequence $(a_n)_{n\in\mathbb{N}}$ may end in a larger circle, i.e. it may be that there exist $M, N \in \mathbb{N}$ such that for all $m \geq M$ and for all $k \in \mathbb{N}$ $a_{m+k\cdot N} = a_m$ but not $a_{m+2} = a_m$. Now, if we consider Blutner structures for dynamic contexts, then a third case can occur: the sequence can run into a dead end, i.e. it may happen that the interpreter chooses a meaning M such that there is no F for the speaker to choose. Hence, it may happen that there is an $m \in \mathbb{N}$ such that the sequence can not be extended beyond a_m. Exactly this happens if, in context c, the interpreter prefers a meaning M for which there is no grammatical triple $\langle c, F, M\rangle \in \mathrm{Gen}$. In order to avoid dead ends, the speaker has to avoid all forms for which the hearer would choose an ungrammatical meaning. This leads to the following definition of dead end and the modified definition of (strong) optimality in which we incorporate the principle of avoiding dead ends.

Definition 5.3 (Dead Ends) *Let* $\langle \mathcal{B}_{\mathcal{F}}, \mathcal{B}_{\mathcal{M}}, [\ ]\rangle$ *a Blutner structure for dynamic contexts with* $\mathcal{B}_{\mathcal{F}} = \langle \mathcal{F}, \mathcal{M}, \mathrm{Gen}, \preceq\rangle$*.* Dead ends *are a set of ungrammatical form–meaning pairs:*

$$DE := \{\langle c, F, M\rangle \mid \langle [c], F, M\rangle \in \mathrm{Gen}_H \;\&\; \langle c, F, M\rangle \notin \mathrm{Gen}\}$$

We denote by DE^+ *the set of all form–meaning pair that* lead into *dead ends, i.e.* DE^+ *is the set of all* $\langle c, F, M\rangle \in \mathrm{Gen}$*:*

$$\exists M' : M \prec_{[c],F} M' \;\&\; \forall M''(M' \preceq_{[c],F} M'' \Rightarrow \langle c, F, M''\rangle \in DE).$$

Definition 5.4 (Optimality) *Let* $\langle \mathcal{B}_{\mathcal{F}}, \mathcal{B}_{\mathcal{M}}, [\ ]\rangle$ *a Blutner structure for dynamic contexts with* $\mathcal{B}_{\mathcal{F}} = \langle \mathcal{F}, \mathcal{M}, \mathrm{Gen}, \preceq\rangle$*, and* $\langle \leq, \mathrm{Gen}\rangle$ *the induced order. We call an element* $\langle c, F, M\rangle \in \mathrm{Gen}$ optimal for $\langle \mathcal{B}_{\mathcal{F}}, \mathcal{B}_{\mathcal{M}}, [\ ]\rangle$*, iff it is maximal in* $\langle \leq, \mathrm{Gen} \setminus DE^+\rangle$*.*

Now we can specify how the preferences of speaker and interpreter enter into the resolution problem of pronouns, and therefore how it determines the information update. We can single out three structures which determine the optimal choice of pronouns and their translation:

1. $\langle NL, \mathcal{L}, ^*\rangle$ in which $^* : NL \longrightarrow \mathcal{P}(\mathcal{L})$ maps natural sentences $F \in NL$ to a set of formal sentences $\varphi \in \mathcal{L}$. The translation part.
2. $\langle \mathcal{L}, \mathcal{W}, [\![\ ]\!]\rangle$ in which $\mathcal{W}$ is a set of world–assignment pairs, and $[\![\ ]\!]$ a function from $\mathcal{L}$ into sets of pairs $\langle \sigma, \tau\rangle$ with $\sigma, \tau \subseteq \mathcal{W}$. The dynamic semantics part.
3. $\langle \mathcal{B}_{\mathcal{F}}, \mathcal{B}_{\mathcal{M}}, [\ ]\rangle$ the Blutner structure with contexts $\mathcal{C} = \{\langle w, \sigma\rangle \mid w \in \sigma \subseteq \mathcal{W}\}$. The Blutner structure is determined by $\langle NL, \mathcal{L}, ^*\rangle$ and $\langle \mathcal{L}, \mathcal{W}, [\![\ ]\!]\rangle$ up to preferences, i.e. $\mathcal{F} = NL$, $\mathcal{M} = \mathcal{L}$, and

$$\mathrm{Gen} := \{\langle \langle w, \sigma\rangle, F, \varphi\rangle \mid \varphi \in F^* \;\&\; \sigma \in \mathrm{dom}[\![\varphi]\!] \;\&\; \exists f\, (w, f) \in [\![\varphi]\!]_\sigma\}.$$

Together, these structures define transition systems $\langle \mathcal{C}, NL, \longrightarrow \rangle$ such that for all $c = \langle w, \sigma \rangle, c' = \langle w', \sigma' \rangle \in \mathcal{C}$:

$$c \xrightarrow{F} c' \iff w = w' \;\&\; \exists \varphi \in \mathcal{L}\; (\langle c, F, \varphi \rangle \text{ is optimal} \wedge \sigma' = [\![\varphi]\!]_\sigma).$$

Now, if we assume that the preferences induced by the OT–constraints are linear orders (not only pre–orders), then there can only be one φ such that $\langle c, F, \varphi \rangle$ is optimal. In this case, we can derive the *update potential* of a sentence $F \in NL$ as follows:

$$\sigma[\![F]\!]\tau \iff \exists w\; \langle w, \sigma \rangle \xrightarrow{F} \langle w, \tau \rangle\,.$$

6 Summary

This paper grew out of an dissatisfaction with the foundations of BiOT in dynamic contexts, i.e. with application of BiOT to contexts which typically arise in online applications of BiOT, and which are characterise by an epistemic asymmetry between speaker and hearer. Standard BiOT as developed by Blutner and Jäger (Blutner, 2000; Blutner and Jäger, 2000; Jäger, 2002) has no representation for the interlocutors' information states. Hence, the algorithm for finding strongly optimal form–meaning pairs is, as we have shown, only applicable if both interlocutors can decide which form–meaning pairs are grammatical in the given utterance context. As an interlocutor can only decide which form meaning–pair is grammatical if he knows the actual world, this assumption is practically always violated in ongoing communication. Hence, Jäger's algorithm may be justifiable in diachronic but not in dynamic contexts. Mattausch's example is making this especially clear.

As an answer to our dissatisfaction, we outlined a framework which allows us to combine bidirectional OT and Dynamic Semantics. We introduced Blutner structures which are more flexible than Jäger's OT–systems. Blutner structures are generated by two OT–systems, one for the speaker and one for the interpreter. This means that we are not forced to decide about the relative strength of production and interpretation constraints, in particular, constraints which are only relevant to one process cannot outrank constraints which are relevant to the other process. Furthermore, it allows us to account for the epistemic asymmetry between speaker and hearer. We saw that for the evaluation of some production constraints it is necessary to know the true state of the world. As the addressee does not know the true state, he cannot check these constraints. We also saw that there is a price to pay for separating production and interpretation constraints: it is not anymore guaranteed that preferences regarding form–meaning pairs are non-circular. Finally, we formulated our models such that we can handle arbitrary context–sensitive constraints.

Our model of Mattausch's example shows that the circularity is apparent only. If we take utterance context and epistemic context into account, then the circularity vanishes. But, in contrast to using OT–systems, this is an empirical finding, no longer a restriction imposed a priori by the underlying mathematical structure. The consideration of the role of epistemic perspectives leads us to postulate a new restriction on the set of form–meaning pairs. The epistemic asymmetry makes it necessary to remove misleading forms before calculating optimal form–meaning pairs. As the interpreter has only limited information about the utterance context he may prefer a meaning M for a form F such that the resulting form–meaning pair $\langle F, M \rangle$ is not grammatical in the given context. We called such a pair a *dead end*. We saw that only those form–meaning pairs can count as optimal for which F does not lead into a dead end, i.e. the speaker is not allowed to choose a form which misleads the interpreter. This additional rule applies equally to the calculation of strongly optimal form–meaning pairs as well as to weakly optimal formmeaning pairs. Hence, the principle of avoiding dead ends does not depend on Blutner's version of weak BiOT but applies to all versions which use strong optimality.

References

Beaver, D. (2004). The optimization of discourse anaphora. *Linguistics and Philosophy*, 27(1):3–56.

Beaver, D. and Lee, H. (2004). Input-output mismatches in ot. In Blutner, R. and Zeevat, H., editors, *Optimality Theory and Pragmatics*, pages 112–153. Palgrave Macmillan, Basingstoke.

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. and Jäger, G. (2000). Against lexical decomposition in syntax. In *Proceedings of IATL 15*, pages 113–137, University of Haifa.

Groenendijk, J. and Stockhof, M. (1991). Dynamic predicate logic. *Linguistics & Philosophy*, 14:39–100.

Horn, L. R. (1984). Towards a new taxonomy of pragmatic inference: Q-based and r-based implicature. In Schiffrin, D., editor, *Meaning, Form, and Use in Context: Linguistic Applications*, pages 11–42. Georgetown University Press, Washington.

Jäger, G. (2002). Some notes on the formal properties of bidirectional optimality theory. *Journal of Logic, Language, and Information*, 11(4):427–451.

Kager, R. (1999). *Optimality Theory.* Cambridge University Press, Cambridge.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory.* Kluwer, Dordrecht.

Mattausch, J. (2000). On optimization in discourse generation. Master's thesis, Universiteit van Amsterdam.

Smolensky, P. and Legendre, G. (2006). *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar.* The MIT Press, Cambridge, MA.