

Chapter ?

Binding by Verbs: Tense, Person and Mood under Attitudes*

1. Introduction

This paper analyses some phenomena of the temporal and modal domain that have been worrying me for at least ten years. It deals with tense and mood in subordinate constructions, notably under attitudes.

The theory adopted here is a slight modification of the system proposed in Schlenker's dissertation (Schlenker, 1999). Recently, Schlenker has proposed other variants of his theory ((Schlenker, 2001a), (Schlenker, 2001c)), which do not work for the analysis of tense, however. That is the reason why I prefer to adhere to the first system. (Heim, 2001a) has proposed a very elegant modification of Schlenker's original system, which doesn't incorporate the analysis of tense either. I will only say a few words about that system at the end of the article.

The central principles of the system proposed are these:

1. Person, Mood and Tense are features of the verb that are checked by the features of the correspondent arguments of the verb, which are an individual, a world and a time variable.
2. At LF, the features of variables are interpreted. Variables that have them will be deictics. The features of verb ('checkees') are never interpreted.
3. Features of semantically bound variables are deleted and therefore not interpreted at LF.
4. Verbs of attitudes, modals and the future auxiliary **will** are variable binders that delete features under agreement.

The paper contains a full theory of attitudes that overcomes empirical problems of the classical theory of attitude due to (Kaplan, 1979). Nevertheless it defends Kaplan's celebrated **Prohibitions against Monsters**, viz. that the functors of natural language are at best intensional. Though I accept virtually everything that Schlenker says, I disagree with Schlenker on his claim that verbs of attitude are monsters in the sense of Kaplan. I will show that this claim cannot be maintained.

The organisation of the paper is as follows. Section 2 introduces the principle Feature Deletion under Binding. Section 3 lists the data to be analysed. Section 4 presents the theory. Section 4.1.3 gives a detailed account of the morphology/syntax interface of tense. 4.2 introduces the semantics. 4.2.1 reconstructs Schlenker's context theory and relates it to Kaplan's framework. It shows that the system has no quantifiers binding context variables. Section 4.1.2 gives a precise semantics of the features for person, tense and mood. Section 4.2.4 shows that *de se* attitudes are purely intensional; I sketch that this holds for *de re* attitudes as well. The semantics makes clear why these verbs are quantifiers that bind variables. Section 5 tests the theory against the data. 5.1 is a repetition of Schlenker's analysis of the first person, logophoric pronouns and controlled PRO. Amharic requires a special parameter. Section 5 is the central part of the paper. 5.2.1 shows how tenses are morphosyntactically realised in the different languages discussed. 5.2.3 derives the English Sequence of Tense Rules, and 5.2.4 the Russian ones. Russian requires a special parameter. 5.3 is about temporal adverbs; they are generated in the temporal argument position, move for type reasons and thereby bind the tense. The next 3 sections are devoted to the shiftable adverbial **in two days**. Schlenker's semantics is refined, and it is shown why this adverb must be embedded under a future operator. Section 5.4 discusses the interaction of mood and tense. It discusses the distributions of modals such as **ought** and shows that the German subjunctive must be a logophoric mood, i.e. bound by a verb of attitude. Section 6 summarises the binding principles, and section 7 explains why Kaplan's Prohibition against Monsters is not refuted. Finally, section 8 says that the theory of binding by verbal quantifiers should be literally related to LF movement.

2. LF-Binding and Feature Deletion

2.1. Feature Transmission/Deletion under LFMovement

One of the most important ideas of the system, which is attributed to (Heim, 1994), is that features of semantically bound variables are not interpreted. Heim's example is this¹:

- (1) SS: [Only I₅]^{1st}₈ did my₈ home work.
 LF: [DP only 5^{1st}]^{1st} λ₈ t₈^{1st} did 8^{1st}'s home work
 [by QR and Feature Transmission/Deletion under Binding
 (see below)]

I_5 is interpreted as a variable with the interpretable feature 1^{st} person, which restricts the denotation of the variable 5 to the actual speaker. This feature is projected to the **only**-DP, a generalised quantifier that must be QR-ed for type reasons. The moved DP λ -binds the trace t_8 and the variable 8 , which translates the possessive pronoun. The binding of the variables has the effect that the transmitted features are not interpreted. This should be obvious from the paraphrase of the intended meaning:

- (2) Everyone x such that x is different from me didn't do x 's homework.

The two occurrences of x in the **that**-clause correspond to the two variables 8 in the above LF. These variables range over the entire domain of individuals, which would be impossible if they had the interpretable feature 1^{st} . If we did represent **only** as a focussing sentence operator, the features of the bound **my** must be deleted as well.² Other sentences illustrating the same point are:

- (3) a. *The boys each think that they are alone in the room.*
(Sauerland, 2001)
'Each of the boys thinks that *he* is alone in the room'
- b. *Few men brought their children.* (Heim, 2001b)
'The number of those x such that x brought *his* children is small'
- c. *We brought our children.*
'Each of us brought *his* children'

In each case the bound pronoun is morphologically plural but semantically singular, though its 'antecedent' is semantically a plurality. To explain these facts, (Heim, 1994) proposes a principle of feature transmission, which reads as follows:

- (4) Feature Transmission under Variable Binding (Heim, 2001b)
Transmit features of a moved phrase to all variables it binds.

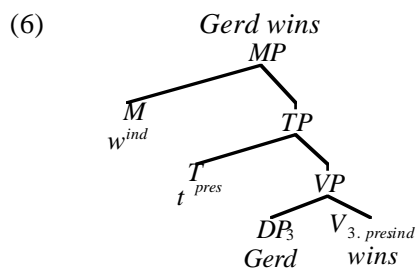
"The principle applies at the PF-branch. It affects the morphological shape of items, but not their interpretation. Only the features present prior to transmission are interpreted." This formulation seems to entail that we QR the antecedent of a bound variable at S-structure (= Spell-Out). Since I do not want to commit myself to that technical detail, I prefer the formulation that features are *deleted* under binding, as indicated in the notation (1). So my reformulation of Heim's principle is this.

- (5) Feature Deletion under Projection and Variable Binding.
- a. Binding requires agreement of φ -features at S-structure.
 - b. Delete the φ -features that are syntactically projected. (LF)
 - c. Delete the features of a moved phrase to all variables it binds. (LF)

One advantage of this formulation is that at the branching point all φ -features are present. But at LF only interpretable features are present – as the Minimalist Program wants to have it.³

2.2. Feature Deletion under Verbal Quantifiers

We will assume that verbs have a subject, a world and a time argument. Let us assume that the arguments are projected in that order. We will require that the features person, tense and mood match in the morphology by some appropriate agreement mechanism. Here is an example of the S-structure (Spell-Out) of a simple sentence

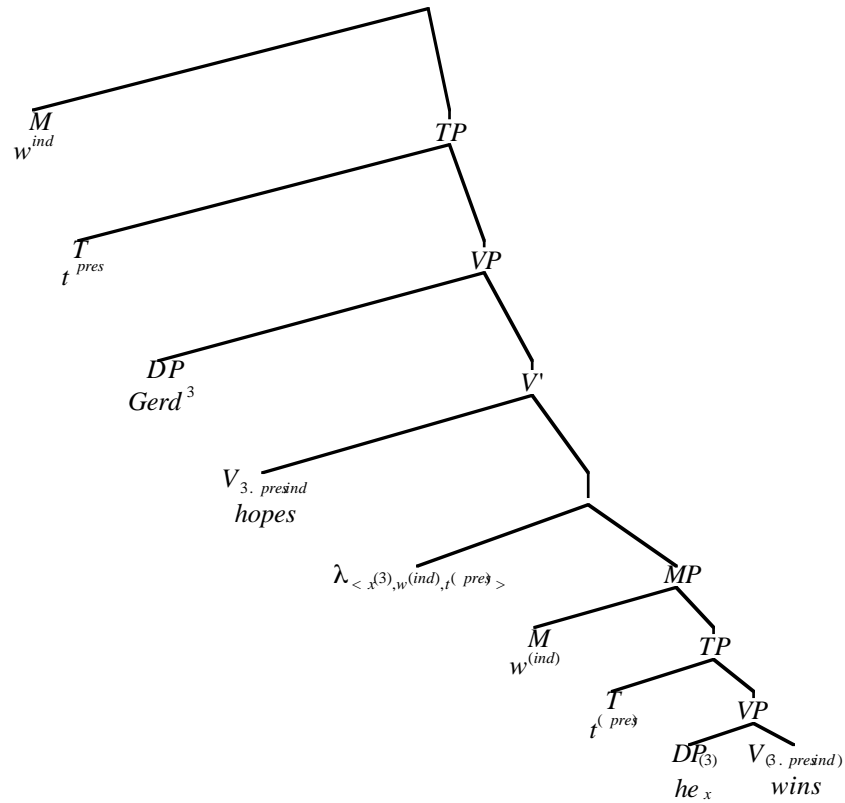


The feature 3rd (third person) is not interpreted. The features ind(icative) and pres(ent) are not interpreted when they occur at the verb, but they are when they are affixes of variables. Indicative restricts the denotation of the world variable to the actual world and present restricts the denotation of the time variable to the actual time. Look now at what happens if we embed the sentence under a verb of attitude:

- (7) *Gerd₁ hopes he₁ wins.*⁴

Arguably, the correct LF is something like the following (see (Stechow, 1984b) and (Stechow, 1995a) among many others):

- (8)



In this representation, only the indicative and the present features of the variables in the matrix sentence are interpreted. The features of the embedded variables are deleted under variable binding, as indicated by the brackets. Note in particular that the bound pronoun **he_x** is in no way anaphoric to the matrix subject **Gerd**.

If we compare this LF with that in (1), we discover that this kind of feature deletion cannot be subsumed straightforwardly under Feature Deletion as it is formulated in (5), because the principle presupposes LF-movement. If movement were involved here, too, this should be a sort of across-the-board movement. The verb of attitude should be a quantifier that binds a person variable, a world variable and a time variable simultaneously and transmits/deletes its features to the variables it binds:

- (9)
- a. Feature transmission/deletion under verbal quantifiers
 - S-structure
 - [_{VP} hopes^{ind.pres.3} _{<456>} [_{CP} ... w₄^{ind} t₅^{pres} he₆³ wins]]
 - |—————| morphological agreement
 - b. “LF-movement” with feature deletion
- [_{VP} hopes^{ind.pres.3} λ_{<456>}[_{VP} w₄^{ind} t₅^{pres} he₆³ wins]]

The analogy with movement might seem far-fetched because there is no obvious way in which the verb of attitude could occupy simultaneously the three positions word, time and subject. But we will see that semantically the analogy is fully justified: the verb **hopes** is a quantifier that simultaneously quantifies over individuals, world and times. That is the reason why we follow (Heim, 2001b) in calling verbs of this kind (verba sentiendi et dicendi) *verbal quantifiers*. Heim has developed a theory in which these verbs are literally moved from the argument position of the subordinate verb. This, however, requires a different ontological framework. Note that the matrix verb need not agree with the person of the subordinate verb as witnessed by a sentence such as **I hope that he will win**. But this only means that the pronoun **he** need not be bound by **hope**.

Temporal auxiliaries have different properties. They may be verbal quantifiers that quantify over times while keeping the subject and the individual argument constant. A classic example is the auxiliary **will/would**. It is an existential quantifier that introduces a relative future time while deleting the temporal features of its temporal argument.

- (10) a. *I will answer every e-mail that arrives.* (Abusch, 1998)
 b. $\text{will}(t_1^{\text{pres}}) \lambda t_1^{\text{pres}} \text{I answer every e-mail that arrives}(t_1^{\text{pres}})$

The embedded present variable in (10a) does not denote the speech time but ranges over future times. The analysis of **would** requires finer methods. This auxiliary must be bound by a predicate of attitude in the past, as we shall see.

Finally, let me mention modals. They may delete the indicative feature of subordinate world variables:

- (11) a. *We must have a secretary that speaks Amharic.*
 b. $\text{must}(w_1^{\text{ind}}) \lambda w_2 \text{we have a secretary that speaks}(w_2^{\text{ind}}) \text{Amharic.}$

The world variable in the relative clause does not refer to the actual world. It is not even said that there is such a secretary.

3. Data

This section gives an overview of some of the data the theory wants to cover. I start with pronouns, because these are the central topic motivating Schlenker's theory. We start with person, the best-studied topic in logical syntax. The idea is that tense and mood behave analogously.

3.1. Person

3.1.1. The first person

For (Kaplan, 1979), “I” is always directly referential; it invariably denotes the speaker.

- (12) *John thought that I was being attacked (by a bear). (de re)*
 John thought, “Arnim is being attacked.”

If we want to express a de se thought of the subject, we have to use a third person pronoun and encounter, in fact, an ambiguity between de se and de re.

- (13) *John thought he was being attacked. (de se/de re)*
 John thought, “I am being attacked.”
 John thought, “The man in the mirror is being attacked.”

If I speak of a de se/de re ambiguity, I mean that there are actually two different logical forms, which will be introduced below. One might dispute this for the reason that de se belief is a special case of de re belief (belief under the acquaintance relation of identity; cf. (Lewis, 1979)). (Percus and Sauerland, 2002) have given an argument, however, that this will not do in for all cases.

- (14) *Only John thought that he was being attacked by a bear.*

The scenario is Bill and Mary see themselves in the mirror without recognizing themselves and each one sees a bear in the mirror who is going to attack the person he is seeing. Each one thinks: “The bear is going to attack the person I am seeing in the mirror”. John sees a bear running toward him and thinks: “The bear is going to attack me”. Under this scenario, (14) is true intuitively, but not under a de re analysis.⁵ So let us assume that there is a special LF representing de se readings only.

We encounter the same ambiguity with a first person matrix subject:

- (15) *I thought that I was being attacked. (de se/de re)*
 I thought, “I am being attacked.”
 I thought, “The man in the mirror is being attacked.”

The generalisation seem to be that the de se/de re ambiguity arises when the matrix subject and the subordinate pronoun agree in their ϕ -features. This should follow from the binding principles for verbal quantifiers. Kaplan must assume the same LF for the ambiguous sentences, and (Kaplan, 1979: sect. XX) contains some ideas how to analyse de re readings, which could cover de se readings as a special case. Following (Zimmermann, 1991), (Stechow, 2001) developed a semantics for de re/de se (see the remarks in section 4.2.4).

A Kaplanian theory would have great difficulties with Amharic, where an embedded first person pronoun can refer to the speaker or it can be interpreted de se.

- (16) Amharic: (Schlenker, 2001a)
- | | | |
|-------------|-------------|--------------------------------|
| <i>j</i> | - | - |
| <i>John</i> | <i>hero</i> | <i>be.PF-ISO 3M.say-AUX.3M</i> |

Translation: ‘John says that he is a hero.’
 Situation: John says, “I am a hero.”

The next example shows that **I** is not quoted.

- (17) *alðttazzäzäNN* *alä* (Schlenker, 2001a)
 [Amharic, Leslau 1995, p. 779]⁶
I-will-not-obey-me he-said
 Translation: ‘He said he would not obey me.’
 Situation: He said, “I will not obey you.”

It is certainly not possible to derive this behaviour of “I” under Kaplan’s assumption that the first person is always directly referential to the speaker.

3.1.2. Logophors

The crucial property of logophoric pronouns is that they must occur in the scope of an appropriate verbal quantifier ((Clements, 1975), (Schlenker, 2001a)). Semantically, the logophoric pronouns encode the fact that the subject had an “I”-thought, i.e. where we find the logophor in indirect discourse, the subject had used “I” in his language (of thought). The following examples from Ewe illustrate the points.

- (18) a. *Kofi be ye~-dzo* [Ewe, (Clements, 1975)]
 Kofi say LOG-leave
 Translation: ‘Kofi said he had left.’

- Situation: Kofi said, "I left"
- b. *Kofi be e-dzo*
 Kofi say he/she left
 Translation: 'Kofi said he/she had left.'
 Situation: Kofi said, "He/she left."

Logophors must be interpreted de se, whereas ordinary pronouns need not have this interpretation.

(19) Ewe relative clauses [(Clements, 1975)]

- a. **ama do Nku nyOnuvi hi dze ye~*
gbO dyi
 ama set eye girl WH stayLOG
 side on
- b. *ama do Nku nyOnuvi hi dze e*
gbO dyi
 ama set eye girl WH stay pro
 side on
 'Ama remembered the girl who (had) stayed with her'
- c. *ama gblO be ye~-do Nku nyOnuvi hi*
dze ye~ gbO dyi
 ama say that LOG eye girl
 WH stay LOG side on
 Translation: 'Ama said she remembered the girl who (had) stayed with her.'
 Situation: Ama said: "I remember the girl who stayed with me"⁷

The last examples show that a logophoric pronoun may occur in a relative clause only if the clause is embedded under an attitude predicate.

3.1.3. Controlled PRO

From (Chierchia, 1989) we know that PRO must be interpreted as a de se pronoun. The following example is due to Schlenker, but Chierchia uses similar ones.

- (20) a. Situation: John is so drunk that he has forgotten that he is a candidate in the election. He watches someone on TV and finds that that person is a terrific candidate who should definitely be elected. Unbeknownst to John, the candidate he is watching on TV is John himself.

- True: John hopes that he will be elected.
 False: John hopes [PRO to be elected]
- b. Situation: John hopes: “I will be elected.”
 True: John hopes [PRO to be elected]

(21) *I hope Bill expects PRO to be elected.*

Possible situation: I hope: “Bill expects that he will be elected”.

Not possible: I hope: “Bill expects that I will be elected.”

Anticipating the semantics for attitudes to be introduced in the paper, the crucial property of controlled PRO will be that it must be bound by the nearest egocentric verbal quantifier.⁸ Note that the binder of PRO is not a DP. It is a verb.

3.2. Tense

One of the difficulties of an informal discussion of tenses is that they express relations. This is particularly true of the past tense. When we have a past under a verbal quantifier, it is not the past variable that is bound, but the present variable to which this variable is related. This subtle fact is a permanent source of confusion.

3.2.1. Sequence of tense

One of the central tasks of this paper is to derive the SOT facts and contrast for languages such as English and Russian. In analogy to de se I will use the term *de nunc* for “the time at which I am”.

- (22) English
- | | | |
|----|-------------------------------------|------------------------|
| a. | <i>Smith thinks Mary is sick.</i> | (simultaneous de nunc) |
| | Smith thinks, “Mary is sick now.” | |
| b. | <i>Smith thinks Mary was sick.</i> | (anterior de nunc) |
| | Smith thinks, “Mary was sick.” | |
| c. | <i>Smith thought Mary was sick.</i> | (simultaneous de nunc) |
| | Smith thought, “Mary is sick.” | |
| d. | <i>?Smith thought Mary is sick.</i> | (double access) |
| e. | Smith thought Mary had been sick. | (anterior de nunc) |
| | Smith thought, “Mary was sick”. | |

Our theory should entail the SOT-rules of English, which are the following.

- (23) SOT rules for English (without embedded future)
- a. When the main tense is present or future, then the subordinate tense is present in the case of de nunc simultaneity, past in the case of de nunc anteriority and future in the case of de nunc posteriority.
 - b. When the main tense is past, then the subordinate tense is past in the case of de nunc simultaneity, pluperfect in the case of de nunc anteriority and past future in the case of de nunc posteriority.

We will see that the rules follow from the assumption that a verbal quantifier deletes the semantic tense features of the time variable it binds. The theory does more than that: it gets the semantics right. Furthermore, it predicts that present under past can't have a de se interpretation.

People call Russian a non-SOT language, a misnomer, because Russian has SOT-rules as well. The difference is that (for the purpose of obtaining "de se" readings) Russian always treats subordinate clauses as if the matrix verb were in the present or future. With respect to the semantics and LF, the two languages do not differ.

- (24) SOT in Russian (folklore)
- a. *Petja skazal, čto Miša plačet*
(simultaneous de nunc/double access).
Petja said that Misha is-crying
'Petja said, "Misha is crying."'
 - b. *Petja skazal, čto Miša plakal*
(anterior de nunc).
Petja said that Misha is-crying
'Petja said, "Misha is crying."'

The tense in the relative clause can be bound by the verbal quantifier, but it need not be.

- (25) a. *Maša videla človeka kotoryj plačet.*
'Masha saw a man who is/*was crying.'
- b. *Maša skazala čto Ivan plačet.*
'Masha said that Ivan was crying.'
- c. **Kogda Maša videla Ivana, on plačet.*
'When Masha saw Ivan, he is crying.'
- (26) *Maša skazala čto ona videla človeka kotoryj plačet.*
'Masha said that she saw a man who was/is crying.'

According to Schlenker, the following example shows that not every verb that selects a CP triggers the Russian SOT rule:

- (27) *Chasto slučalos', čto Miša plakal/*plačet*
 [Janssen 1996]
 often happened, that Misha cried-IFPV / is-crying
 'It was often the case that Misha cried / *is crying.'
 The verb *slučat'sja* "to happen" behaves like its English counterpart.

3.2.2. Shiftable Aspectual Adverbs

(Schlenker, 2001a) observes that the local evaluation time of an adverbial of distance like **in two days** is either deictic or bound by a verbal quantifier. In other words, these adverbials behave like Amharic **I**.

- (28) **in two days** (in the sense of Russian **čerez** or Italian **fra/tra**).
 (Schlenker, 2001a)
- a. *John told me repeatedly over the years, 'I will finish my work in two days/the day after tomorrow.'*
 - b. *#John told me repeatedly over the years that he would finish his work the day after tomorrow.*
 - c. *John told me repeatedly over the years that he would finish his work in two days.*

Schlenker (p.c.) tells me that (Comrie, 1985: 107, fn. 7) makes the same observation for the adverbial **ten years ago**:

In English there is, incidentally, one set of adverbials that seems to behave anomalously with respect to the shift in deictic centre in indirect speech, namely adverbials with *ago*. Outside indirect speech, an expression like *ten years ago* invariably takes the here-and-now as deictic centre, and cannot be used with reference to any other deictic centre; cf. *I arrived here ten years ago*, but *I left that city last year, having arrived there ten years before/earlier*, where the ten years are to be counted from my leaving that city, not from the present moment. In indirect speech, however, one can have *John said that he had arrived ten years ago*, which allows both the interpretation where the ten years are counted from the here-and-now and the interpretation where the ten years are counted from the time of John's utterance.

German seems to behave similarly:

- (29) a. *#Ich traf Irene um 5. Sie rief mich in genau 10 Minuten an.*

- 'I met Irene at 5. She called me in exactly 10 minutes.'
- b. *Irene versprach mir, mich in 10 genau Minuten anzurufen.*
'Irene promised me to call me in exactly 10 minutes.'
 - c. *Irene sagte, sie würde mich in genau 10 Minuten anrufen.*
'Irene said she would call me in exactly 10 minutes.'

Possible exceptions come with statives:

- (30) *Ich rief Irene um 5 an. Sie war in genau 10 Minuten bei mir.*
'I called Irene at 5. She was at my place in exactly 10 minutes.'

3.2.3. Shiftable modals

The following data, which have been observed by (Abusch, 1993), suggest that the modal **ought** is either deictic or bound by a verbal quantifier.

- (31) **ought** (Abusch, 1993)
- a. **When John was a schoolboy, he ought to study more.*
 - b. *John thought that he ought to study more.*
 - c. *John ought to study more.*
 - d. *John will always be a student that ought to study more.*
(Heim)

The last example suggests that **will** behaves like a verbal quantifier. (Stechow, 1995b) observes that German **müsste**, **könnte**, and **sollte** behave exactly alike.

- (32) German **müsste** 'should', **könnte** 'could/might', **sollte** 'should/ought'
- a. **Als Karl zur Schule ging, müsste/könnte/sollte er mehr arbeiten.*
 - b. *Karl dachte, dass er mehr arbeiten müsste/könnte/sollte.*
 - c. *Karl müsste/könnte/sollte mehr arbeiten.*

The meanings are as in the previous examples. Obviously, the English modals **should**, **could**, and **might** belong to the same pattern.⁹

The past future auxiliary **would** seems to behave like the said modals in most contexts.

- (33) a. **Five minutes ago Irene would call me.*
b. Irene said that she would call me.

In some contexts, **would/würde** looks like an ordinary past future, however:

- (34) a. *1964 traf ich die Frau, die ich 5 Jahre später heiraten würde.*
'In 1964 I met the woman that I would marry 5 years later.'
b. *I child was born that would be king.* (Kamp)

I don't have a theory that accounts for all these facts.

3.2.4. Subjunctive in German

(Stechow, 1984a) observes that the German subjunctive gives rise to contradictions in certain contexts.

- (35) a. *I thought your yacht was longer than it is.*
(Russell, 1905)
b. *Ich dachte, ihre Yacht sei/wäre (subj.) länger als sie ist (ind.).*
(Stechow, 1984a)
c. *#Ich dachte, ihre Yacht sei/wäre (subj.) länger als sie sei/wäre (subj).*

(Schlenker, 2001a) gives similar examples. The facts can be explained by assuming that the German subjunctive is a logophoric mood: a world variable with the feature subjunctive must be bound by a verbal quantifier.

4. The Theory

As I said, I will adopt a version of the theory in Schlenker's dissertation, because among his several proposals this version offers the outlines of a theory of tense, the central interest of this article. Some technical details of his proposal are not entirely clear to me, and I will try to clarify the conceptual issues involved.

4.1. Syntax and Morphology

4.1.1. Features at S-structure

Like most semanticists, I will assume that lexemes have logical types: e (individuals), i (times), s (worlds) and t (truth values). The types determine the semantic type of the value of the expression. The language will consist of labelled trees, the terminals of which are labelled by lexemes, while the non-terminals are labelled by category symbols, logical types and features, which will be introduced. Some features are morphological, others will be purely syntactical and encode binding properties.

At the PF-branch of the grammar, the morphological features determine the pronunciation of the expression. I will assume that all features are still present at S-structure/Spell-Out. At LF, only interpretable features survive; non-interpretable features are deleted, and a very important deletion process is Deletion under Variable Binding, as we have seen. LF is generated from S-structure by scope shifting operations like Quantifier Raising (QR) or Quantifier Lowering (QL)¹⁰ and other operations like providing the verbal quantifier with the correct λ -binder.

With (Stechow and Sternefeld, 1988) and (Sternefeld, 2000), I will assume that there are two kinds of features, **checkees** and **checkers**.¹¹ The former are prefixed by a star.¹² As to semantics, there is an important difference. Here is a list of the morphological features that play a role in the paper.

(36) Morphological features

Conventional name	Checkee(s)	Checker(s)
i^{th} person, $i = 1, 2, 3$ ¹³	$*i^{\text{th}}$	i^{th}
Present	*pres	pres
(‘anterior’)	*<	<
Past	*<, *pres	<, pres
Pluperfect	*<, *<, *pres	<, <, pres
Indicative	*ind	ind

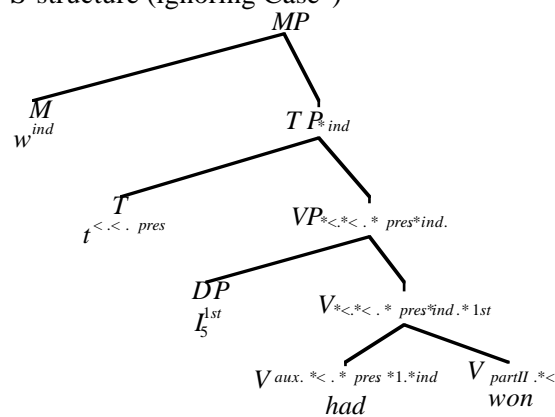
With the exception of present, the conventional tenses are decomposed into several features. I will assume that checkees are projected. Feature checking is done under agreement, i.e. when a checkee and a checker are adjacent, they are not projected further.

- (37) Feature checking (S-structure) Checkees are projected unless they are adjacent to a checker.

This formulation is ad hoc for the purposes of the discussion in this paper. Feature checking could be done as well by movement either by the bearer of the features or by abstract movement of the features. My approach is thus compatible with different variants of the ‘Minimalist’ tradition. Consider the official S-structure of the following sentence:

- (38) *I had won.*

- (39) S-structure (ignoring Case¹⁴)



The reader may verify that the checkee is projected according to the principle Feature Checking. A tree of this sort feeds the PF-branch. The notation should be even more abstract, as indicated by the following spell-out rules:

- (40) Morphological Spell-out rules
 a. $\zeta^1 \rightarrow /I/$, where ζ is any variable; Abbreviation: I_ζ^1
 b. $\text{win}^{\text{partIII}, * < } \rightarrow /won/$; Abbreviation: $\text{won}^{* < }$

It is often said that features are deleted under agreement. In the present system this means that features are not further projected. Features of lexical heads cannot be deleted since they are needed for phonetic interpretation.

4.1.2. Features in Logical Form

As a general principle of the LF interface we have it that non-interpretable lexemes and features are deleted. Non-interpretable features are: **3rd person (3)**, **aux**, **partII**, **infinitive** and others. Furthermore, checkees are non-

interpretable. Checkers are interpreted if they have semantic content. Interpretable checkers are the first and second person (1 and 2) and the temporal features **pres** and **past** (<), **ind** and **plural**. Here is an overview:

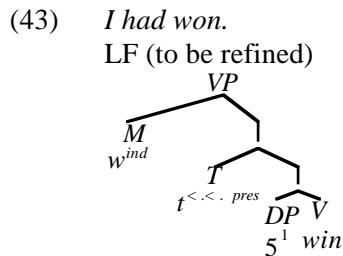
(41) Interpretable and non-interpretable features

Interpretable	Non-interpretable
1 st	*1 st
2 nd	*2 nd
pres	*pres
<	*<
ind	*ind
plural	*plural
	aux, partII,...

A special case of Chomsky's Principle of Full Interpretation is that non-interpretable material is deleted at LF. (Recall that features of bound variable are deleted at LF)

(42) Full Interpretation (LF)
Delete non-interpretable material!¹⁵

Given that the finite auxiliary **had** doesn't have a semantic interpretation, it follows that the LF of (39) is something like this:



4.1.3. LF of Tenses

There is still something missing. The temporal variable $t^{<<T}$ cannot be interpreted as it stands. Tenses are relations. Past means that the reference time is before the speech time, and Pluperfect means that the reference time is before a past time. In order to get this, Schlenker introduces the following well-formedness condition.

- (44) Syntax of the <-feature (LF) (Schlenker, 1999: 166)¹⁶
 A variable of form $\zeta^{<a}$, where a is a (possibly empty) set of tense features, is ill-formed unless followed by $[\zeta < \xi^a]$, where ξ is any variable.

The apposition to the variable will express a presupposition. I assume that the syntactic condition is met at LF, and I will assume that we simply add the condition to the variable there. So an official LF for (39) could be this:

- (45) *I had won.*
 Building an LF:
 $w^{ind} t_7^{<<pres} 5^{1st} win$
 $\Rightarrow w^{ind} t_7^{<<pres} [t_7 < t_6^{<pres}] 5^{1st} win$ (by (44))
 $\Rightarrow w^{ind} t_7^{<<pres} [t_7 < t_6^{<pres} [t_6 < t_0^{pres}]] 5^{1st} win$ (by (44))

A pluperfect variable such as $t_7^{<<pres}$ is always related to a past variable, here $t_6^{<T}$. The latter is related to a present variable, here t_0^{pres} . So modulo alphabetic variance the sentence (46a) has the LF (46a).

- (46) a. *I was happy because I had won.*
 b. $\exists t_6 [w^{ind} t_6^{<pres} [t_6 < t_0^{pres}] 5^{1st} happy \text{ because } \exists t_7 [w^{ind} t_7^{<<pres} [t_7 < t_6^{<pres} [t_6 < t_0^{pres}]]] 5^{1st} win]$

Like Schlenker, I am assuming that free time variables may be existentially bound at LF at the sentence or even discourse level. The same goes for the world variables. Variables with interpretable features may remain free, since the existential closure doesn't change their meaning, as we will see.

The LF correctly represents the fact that the interpretation of a simple past depends on two times whereas that of a pluperfect depends on three times. In the syntax, these variables are located at the same position, say SpecT, a somewhat puzzling feature of the approach.

4.2. Semantics

4.2.1. Context Theory

The point I want to make in this section is that there is no quantification over contexts in Schlenker's system, despite his claim to the opposite. A further point is that we need no such quantification. It follows that Schlenker's attitude predicates are not monsters.

Let α be any expression, let g be a variable assignment respecting the logical types, and let c be a context. Thus, $\|\varphi\|^{g,c}$ represents the semantic value with respect to g and c . This notion of context is, of course, not Stalnaker's context set but rather Kaplan's (1979) context of utterance.

But what is a context? Schlenker says that it is a sequence $c = \langle a, h, t, w \rangle$ such that a and h are individuals, t is a time, and w is a world. Schlenker calls a the **author** of c , h the **hearer** of c , t the **time** of c and w is the **world** of c . The context may contain other components. Schlenker says that we can quantify over contexts, and we can form abstracts of the form $\lambda\langle x, y, t, w \rangle \varphi$, which he calls **characters**. An **utterance** is a pair consisting of a (formal) context and a character.

- (47) Utterance (to be revised) (Schlenker, 1999: 79)
 $c \lambda\langle x, y, t, w \rangle \varphi$, where c is a formal context

Schlenker assumes that contexts may be represented in the language as special structured symbols $\langle x, y, t, w \rangle$ where x, y, t, w are of type e, e, i and s , respectively. Let us call such symbols **formal contexts**. An utterance of **I win** could be represented as:

- (48) a. *I win.*
 b. Utterance: $\langle \mathbf{Gerd}^{1st}, \mathbf{now}^{pres}, @^{ind} \rangle \lambda\langle \mathbf{x}^{1st}, \mathbf{t}^{pres}, \mathbf{w}^{ind} \rangle \mathbf{w}^{ind}$
 $\mathbf{t}^{pres} \mathbf{5}^{1st} \mathbf{win}$
 c. Its truth: $\mathbf{Gerd}^{1st} \mathbf{now}^{pres} @^{ind} \mathbf{win}$ (by λ -conversion)

The features **1st**, **pres**, and **ind** are transmitted/deleted under variable binding to the bound variables. (Schlenker calls these features **A**, **T** and **W** respectively.) Never mind that the expression **Gerd** doesn't (really) have the feature **1st** in the syntax. As constituent of a formal context it may have that feature. At S-structure the subject is the variable **5^{1st}**, which is pronounced as /I/. There are, however, conceptual problems with this approach. Suppose that the symbol $\langle \mathbf{x}, \mathbf{t}, \mathbf{w} \rangle$ were literally a context variable as assumed by Schlenker throughout. Then this variable should be bound in the λ -abstract in (47b). But it isn't. The complex symbol $\langle \mathbf{x}, \mathbf{t}, \mathbf{w} \rangle$ never shows up at some place within a formula of Schlenker's logical language. At an argument position of a predicate we only find the variables x, t, w . What is bound then are three different variables. I take it, therefore, that the notation $\lambda\langle \mathbf{x}, \mathbf{t}, \mathbf{w} \rangle$ is simply an abbreviation for three λ -abstractions.

- (49) $\lambda\langle x, t, w \rangle \dots := \lambda w \lambda t \lambda x \dots$

This redefinition shows that the complements of verbal quantifiers as in (8) are not characters in the sense of (Kaplan, 1979) but intensions of the property type. Note that I don't want to dispute the possibility of

quantifying over context like things. In Schlenker's dissertation, however, this is not done and need not be done. In what follows, the hearer coordinate y will always be omitted from context, because the hearer can be determined on the basis of x , t and w alone.

The second conceptual problem concerns the interpretation of the features **1st**, **pres** and **ind**. Schlenker's interpretation does not depend on a context but simply on a variable assignment. His semantic rules are these:

(50) Semantics of contextual features (to be revised) (Schlenker, 1999:

124)

Let **a**, **b**, **c** be constants denoting an individual, at time and world, respectively. For any assignment g :

- a. $\| \mathbf{a}^{1st} \|_g^g$ is defined only if $\| \mathbf{a}^{1st} \|_g^g$ is the speaker of the actual speech act.
- b. $\| \mathbf{b}^{pres} \|_g^g$ is defined only if $\| \mathbf{b}^{pres} \|_g^g$ is the time of the actual speech act.
- c. $\| \mathbf{c}^{ind} \|_g^g$ is defined only if $\| \mathbf{c}^{ind} \|_g^g$ is the world of the actual speech act.

It seems to me that the notion "actual speech act" is a disguised version of a Kaplanian context. So, after all, Schlenker's interpretation must depend on a context as well. Thus, my reconstruction of the notion of context is this:

(51) Context (official version)

A **context** is a triple $\langle x, t, w \rangle$ consisting of a person x , a world w and a time t . For any context c , a_c is the first component of c , t_c is the second component and w_c is the third component of c .¹⁷

For arguments to the extent that these three things are enough to determine the reference of every deictic, see (Lewis, 1980). As usual, we let the interpretation depend on a context. A Kaplanian character would then be a function of the following kind:

(52) Kaplanian Character

$\lambda c. \| \varphi \|_g^{g,c}$

A function of this kind never occurs as a complement of a verb of attitudes. So, contrary to Schlenker's claims, attitudes are not monsters. We may say that attitudes quantify over contexts, but they never shift the context of utterance. It remains true that the behaviour of Amharic **I** and logophors

cannot be treated in Kaplan's framework as it stands. This, however, has to do with Feature Deletion under Variable Binding, which requires a syncategorematic syntax (or movement) and cannot be reformulated in a compositional way.

4.2.2. Semantics of the features **1st**, **pres**, **past**, and **indicative**

These features are affixes to variables and restrict their interpretation, i.e. they operate on assignments.

(53) The egocentric features (official version)

Let μ , ν , ξ be variables of type e, i and s respectively. We define:

- a. **First person:** $\|\mu^{1st}\|^{g,c} = g(\mu)$, if $g(\mu) = a_c$. Undefined otherwise.
- b. **Present:** $\|\nu^{pres}\|^{g,c} = g(\nu)$, if $g(\nu) = t_c$. Undefined otherwise.
- c. **Indicative:** $\|\xi^{ind}\|^{g,c} = g(\xi)$, if $g(\xi) = w_c$. Undefined otherwise.

The truth of (48) with respect to the said context $c = \langle \text{Gerd, now, } @ \rangle$ can now be represented as:

(54) $\|\mathbf{w}^{ind} \mathbf{t}^{pres} \mathbf{5}^{1st} \mathbf{win}\|^{g,c}$
 $= 1$ if Gerd wins in @ now, provided $g(\mathbf{w}) = @$, $g(\mathbf{t}) = \text{now}$ and $g(\mathbf{5}) = \text{Gerd}$. If one of the conditions is not given, the utterance has no truth-value.

The meaning rule for **win** should be obvious:

(55) **win** is of type $e(i(st))$. $\|\mathbf{win}\| = \lambda x.\lambda t.\lambda w.x$ wins in w at t .

The semantics for the **past** feature is slightly more complicated. Recall that **past** is an abbreviation for **<pres**. So what really is needed is the interpretation of **<**.

(56) **Meaning of the <-feature**

Suppose we are given a term of the form $\zeta^{<a}[\zeta < \xi^a]$, where ζ and ξ are variables of type i and a is a (possibly empty) set of tense features.

$\|\zeta^{<a}[\zeta < \xi^a]\|^{g,c}$ is only defined if $\|\xi^a\|^{g,c}$ is defined and $g(\zeta)$ is prior to $\|\xi^a\|^{g,c}$. If this is fulfilled, $\|\zeta^{<a}[\zeta < \xi^a]\|^{g,c} = g(\zeta)$.

4.2.3. Existential Closure

While it makes little sense to λ -abstract over variables with semantic features, existential closure is needed and not problematic. Without loss of generality, I give the definition for temporal variables.

- (57) **Existential Closure of featured temporal variables**
 Let φ be any expression of type t and let ζ be any temporal variable, which may have a free occurrence in φ with temporal features.
 $\|\exists\zeta\varphi\|^{\text{g},c} = 1$ iff there is a ζ -alternative g' of g : $\|\varphi\|^{\text{g}',c}$ is defined and equal to 1.
 $\|\exists\zeta\varphi\|^{\text{g},c} = 0$ iff there is a ζ -alternative g' of g : $\|\varphi\|^{\text{g}',c}$ is defined and equal to 0.
 Otherwise, $\|\exists\zeta^a\varphi\|^{\text{g},c}$ has no truth value.

Here is the evaluation of **Gerd won** with the world argument disregarded.

- (58) $\|\exists t_1 t_1^{<\text{pres}}[t_1 < t_0^{\text{pres}}] \text{Gerd win}\|^{\text{g},c} = 1$ if there is a t_1 -alternative g' of g : $g'(t_1)$ is prior to $g'(t_0^{\text{pres}})$, where g' is defined for t_0^{pres} , and Gerd wins at $g'(t_0^{\text{pres}})$.
 Since g' is defined for t_0^{pres} , $g'(t_0^{\text{pres}}) = t_c$.

The conditions for falsehood and undefinedness are computed in the same way.

4.2.4. Attitudes

The semantics for attitude predicates follows the insights of (Lewis, 1979). Attitudes are universal quantifiers which take an egocentric proposition as their complement, i.e. they have the type pp with $p = e(i(s,t))$.

- (59) **Syntax of attitudes.** (English)
 Let ATT^{*a*b*c} be any verb of attitude, where $*a$, $*b$ and $*c$ are checkees of person, tense and mood features respectively.
 Let φ be a CP of type t . Let μ , ν and ξ be variables of type e , i and s respectively. Then
 $[\text{vpATT}^{*a*b*c}\varphi]$
 is a well-formed S-structure, whose LF is:
 $[\text{vpATT}^{*a*b*c}\lambda\langle\mu,\nu,\xi\rangle\varphi^-]$,
 where φ^- is obtained from φ by deleting the features a , b , c of the variables μ^a , ν^b and ξ^c that occur freely in φ .

The λ -abstraction obeys certain constraints for movement (see section 8).

The rule of feature deletion may be stated informally as follows:

(60) Feature Deletion under Attitudes (English)

A verb of attitude deletes the features of the variables it binds under agreement with its checkees. (A checkee **a** agrees with a checker **b** if **a** = ***b**.)

The syntax given in (Schlenker, 1999: 81) is slightly different. Schlenker transmits the features of the arguments of a verb of attitude to the variables it binds. Since these check the checkees of that verb, the result is the same. Here is an example of how a VP under an attitude behaves.

- (61) a. *I hope he wins.*
 b. SS: $[\text{VP hope}^{*ind.*pres.*1st} \lambda \langle x^{1st}, t^{pres}, w^{ind} \rangle [\text{CP } w^{ind} t^{pres} \text{he}^{3rd} \text{wins}^{*ind.*pres.*3rd}]]$
 c. LF: $[\text{VP hope } \lambda \langle x^{1st}, t, w \rangle [\text{VP } w t x \text{wins}]]$

The features of the time and the world variables are deleted under agreement. The feature 1^{st} cannot be deleted because it doesn't agree with the person checkee of the verb of attitude. (The semantically empty CP disappears.)

The meaning of attitudes is defined in two steps. First we assume for each predicate of attitude ATT a background function H_{ATT} that gives us for any index the relevant set of alternatives.¹⁸

(62) **ATT-alternatives**

$H_{ATT}(x,t,w) = \{ \langle x', t', w' \rangle \mid \langle x', t', w' \rangle \text{ is an ATT-alternative for } x \text{ in } w \text{ at } t \}$

For instance, if $ATT = \text{believe}$, then H_{believe} is the set of those $\langle x', t', w' \rangle$ that are not ruled out by the de se belief of x in w at t : for all x believes of himself in w at t , x could be x' in w' at t' .

(63) Semantics of attitudes (de se)¹⁹

Let ATT be any verb of attitude.

$\| ATT \| = \lambda P \in D_{e(i(s,t))}. \lambda x \in D_e. \lambda t \in D_i. \lambda w \in D_s.$

$\forall \langle x', t', w' \rangle \in H_{ATT}(x,t,w): P(x')(t')(w') = 1.$

This meaning is nothing but an intension. It is a mere restatement of the formulation given in (Schlenker, 1999: 81). So verbs of attitudes are not monsters in the sense of Kaplan. This holds for the derivation of de re

readings as well, i.e., for LFs with variables that are free in the complement of attitudes.

Here I give an informal discussion of what the semantics of *de re* attitudes could be. Following Kaplan's (1979: section XX)(see (Zimmermann, 1991) and (Stechow, 2001)) we might say that "John believes the proposition p at the index w, t " is true if there is a Kaplanian character ch such that John believes ch at w, t and $ch(\text{John}, w, t) = p$. The method presupposes that we can apply a character to an egocentric index, i.e. the index figures as a context. 'John believes the character ch in w at t ' means that for any doxastic alternative a, w', t' of John, w, t : $ch(a, w', t')(a, w', t') = 1$. In other words, John *se*-believes the diagonal of ch ; cf. (Haas-Spohn, 1995). John may believe the proposition "John is attacked" under the character "I am attacked" (*de se*) or "The man I see in the mirror" (*de re*). For the index (John, w, t) , both characters give the proposition "John is attacked". Thus, the account could yield a *de se* interpretation as a special case, even if there is no *de se* pronoun in the embedded clause. *De se* is a special case of *de re*. Normally, it is said that the subject must be acquainted with the *res*. Acquaintance is a gradual thing, depending on the character that is believed. For the character discussed, "the man I see in the mirror" must be John. It follows that John must be acquainted with John via the cognitive relation "seeing that man". A careful analysis of this account would show that we have a monster in the meta-language – the diagonal – but not in the semantics for "believe". Obviously, this matter is rather subtle and the diagonalisation approach can be trivialised, as it stands. So more has to be said about the matter. I leave that to another occasion.

Our theoretical tools are now complete and we can apply the theory against our data.

5. Analysis

5.1. Person

5.1.1. *The First Person*

The most important task is to explain the different behaviour of the first person in English vs. Amharic. We have to stipulate that the person checkee of an Amharic attitude verb deletes the feature **1st** of the person variable it binds.

- (64) **The Amharic parameter** [cf. (Heim, 2001b)]
 Amharic verbal quantifiers delete (LF) the feature **1st** of the person variable they bind, regardless of what their person checkee is.

Schlenker formulates the rule somewhat differently. Verbal quantifiers provide the variables they bind with subordination features a (author), t (time) and w (world), which simply mark subordination. If a bound subject variable has the feature a, then it is spelled out as /I/, regardless of what its number is; cf. (Schlenker, 1999: 134). Here is the analysis of an (our) example (17) with mood and tense disregarded.

- (65) Amharic **John says I am a hero**
 a. ...John³ says^{*3} λ<x¹...>...x¹ am^{*1} a hero
 (de se; Amharic parameter)
 John says, “I am a hero.”
 b. ...John³ says^{*3} λ<x¹...>...y¹ am^{*1} a hero
 (direct reference)
 John says, “You are a hero.”
- (66) Amharic **John says I don’t obey me**
 a. # ...John³ says^{*3} λ<x¹...>...x¹ don’t obey^{*1} x¹
 (de se/de se ; Amharic parameter)
 John says, “I don’t obey myself.”
 b. # ...John³ says^{*3} λ<x¹...>...y¹ don’t obey^{*1} y¹
 (direct ref./direct ref.)
 John says, “You don’t obey yourself.”
 c. ...John³ says^{*3} λ<x¹...>...x¹ don’t obey^{*1} y¹
 (de se/direct ref. ; Amharic parameter)
 John says, “I don’t obey you.”

We obtain a directly referential reading for the embedded **1st** person if the variable carrying that feature is not bound by the verb. The semantic rule (53a) ensures that the variable denotes the actual speaker in this case. Note that the verb **says** has the feature ***3rd**, and this feature is also projected at S-structure. Therefore, the matrix subject need not agree with the embedded subject. All these LFs are well-formed. The oddness of (66a/b) comes from the somewhat strange meaning. There would be nothing wrong to have a 3rd person embedded in Amharic, but it could not be bound by a verbal quantifier – if the theory is correct.

The theory makes the prediction that Amharic and English behave exactly alike if the matrix subject is the first person. This should be tested empirically.

- (67) Prediction: **I thought I was a hero** has the same readings in English and Amharic.
- a. I thought, “I am a hero.” (de se)
 - b. Seeing myself in the mirror without recognizing myself, I thought, “He is a hero.”

Here are the predictions for long distance de se-binding for English and Amharic.

- (68) *Bill thinks John thinks he is a hero.*
- a. Bill thinks John thinks, “I am a hero.”
 English:
 $\dots \text{Bill}^3 \text{ thinks}^{*3} \dots \lambda \langle \text{y}^3 \dots \rangle \dots \text{John}^3 \text{ thinks}^{*3} \dots$
 $\lambda \langle \text{z}^3 \dots \rangle \dots \text{z}^3 \text{ is}^{*3} \dots \text{a hero}$
 English: $z \rightarrow /he/$; Amharic: $z \rightarrow /I/$
 - b. Bill thinks, “John thinks I am a hero.”
 English:
 $\dots \text{Bill}^3 \text{ thinks}^{*3} \dots \lambda \langle \text{y}^3 \dots \rangle \dots \text{John}^3 \text{ thinks}^{*3} \dots$
 $\lambda \langle \text{z}^3 \dots \rangle \dots \text{y}^3 \text{ is}^{*3} \dots \text{a hero}$
 English: $y \rightarrow /he/$; Amharic: $y \rightarrow /I/$

5.1.2. Controlled PRO

PRO is a variable without case but with φ -features.²⁰ It is bound by the nearest verb of attitude. In order to insure this, PRO must have two syntactic features: **log** and **local**. The syntactic properties of the features are stated by the following rule:

- (69) The features **log(ophoric)** and **local**
- a. A variable with the feature **log** is bound by a verb of attitude.
 - b. A variable with the feature **local** is bound by a structurally adjacent verb of attitude.

Given that these features are not interpretable, they are deleted at LF. The definition of PRO is therefore this:

- (70) **Definition of PRO.** PRO is a variable without Case but with φ -features and the features **log** and **local**. It is not pronounced at PF.

An immediate consequence of this definition is that PRO can never be deictic. It is always de se. Here is the analysis of example (21).

- (71) a. $\dots x^1/I/ \text{ hope}^{*1} \dots \lambda \langle y^+ \dots \rangle \text{ Bill}^3 \text{ expects}^{*3} \lambda \langle z^3 \dots \rangle \dots z^{3, \log, local} / \text{PRO/ to enjoy } z^3 / \text{himself/}$
 b. $* \dots x^1/I/ \text{ hope} \dots \lambda \langle y^+ \dots \rangle \dots \text{ Bill}^3 \text{ expects}^{*3} \lambda \langle z^3 \dots \rangle y^{1, \log, local} / \text{PRO/ to enjoy } z^+ / \text{myself/}$

The error in (71b) is that the variable y in the most embedded clause is not locally λ -bound by a verbal quantifier.

For object control verbs we have to stipulate the following:

- (72) **Object Control:** Object control verbs are verbal quantifiers that delete the feature of the individual variable they bind if it agrees with their object.

Object control doesn't fit the pattern we find for other verbal quantifiers since we don't find agreement with the features of the verb but rather with a feature of an argument of the verb. Here is an example:

- (73) a. *John persuaded me PRO to kill myself.*
 b. $\dots \text{ John}^3 \text{ persuade}^{*3} \text{ me}^{1st} \lambda \langle x^{1st} \dots \rangle \dots x^{1st, \log, local} / \text{PRO/ to kill } x^{1st} / \text{myself/}$

We have to define the meaning rule for **persuade** appropriately, of course, i.e., the verb must have the logical type $p(e(p))$, with $p = (e(i(st)))$.²¹ Further complications arise with verbs of 'split control':

- (74) *Fritz proposed Susi to go to the movies.*

For a semantic treatment, see e.g. (Stechow, 2002a).

5.1.3. Logophors

Logophoric pronouns have Case and φ -features. They must be bound by some logophoric verb and therefore have the feature **log**.

- (75) **Definition of logophoric pronouns.**
 Logophoric pronouns are variables with Case and φ -features. They have the feature **log**.

Schlenker remarks that logophoric pronouns in Ewe or Gokana do not occur in the first person. If this is so, they must have the feature -1^{st} in addition. As an illustration, consider the analysis of the Ewe example (19c):

- (76) Ewe: Ama said she_{LOG} remembered the girl who stayed with her_{LOG}.
 ...Ama³ says³ $\lambda < x^3 \dots > \dots x^{3+log}$ remembers³ the girl who stayed with x^{3+log}

Since I am not interested in a full theory of logophors, I will ignore this possible complication.

5.2. Tense

5.2.1. Morphology and Syntax: Spell-Out Rules

Most languages I know encode the semantic features **pres**, **<pres** and **<<pres**, but they do it differently. This is so for even closely related European languages. Consider English, German and Russian. **pres** is more or less the same in all of these, and we have seen how it works. The variation starts with the encoding of **past** = **<pres**. The following trees ignore person and mood.

- (77) The syntax of **past** (= **<pres**): Engl./Germ./Russ. preterite morphology,
 Southern Germ./Yidd. present perfect morphology

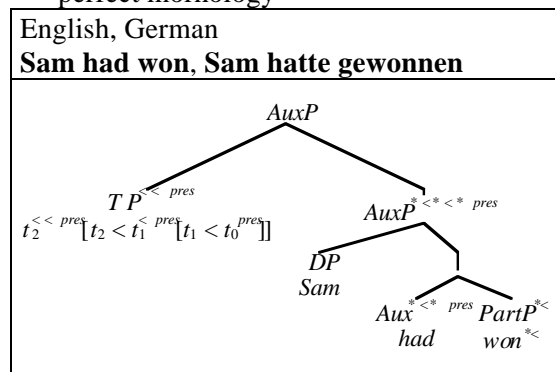
English: Sam won German: Sam gewann Russian: Sam vyigral	German, Southern German, Yiddish ²² Sam gewonnen hat
<p>A syntax tree for the English sentence "Sam won". The root node is VP, which branches into TP and VP. The TP node is labeled with a semantic feature <math>t_2^{< pres}</math> and contains the word "Sam". The VP node branches into a V node and another VP node. The V node is labeled with a semantic feature <math>V^{* < * pres}</math> and contains the word "won". The second VP node branches into a DP node and a V node. The DP node is labeled with a semantic feature <math>t_2^{< pres}</math> and contains the word "Sam". The V node is labeled with a semantic feature <math>V^{* < * pres}</math> and contains the word "won".</p>	<p>A syntax tree for the German/Southern German/Yiddish sentence "Sam gewonnen hat". The root node is VP, which branches into TP and V. The TP node is labeled with a semantic feature <math>t_2^{< pres}</math> and contains the word "Sam". The V node branches into DP and V. The DP node is labeled with a semantic feature <math>t_2^{< pres}</math> and contains the word "Sam". The V node branches into Part II and V. The Part II node is labeled with a semantic feature <math>V^{* < * pres}</math> and contains the words "gewonnen" and "won". The V node is labeled with a semantic feature <math>V^{* < * pres}</math> and contains the word "hat".</p>

--	--

In German two encodings of the semantic tense feature(s) **past** coexist, viz. as the preterite morphology and as the periphrastic present perfect form. With non-copulative verbs, the present perfect form is highly preferred in spoken language, and Southern German and Yiddish have only this form. In this representation, the auxiliary is semantically empty. It merely serves as a carrier of the feature ***pres**. It is a matter of dispute whether the participle II **gewonnen** ‘won’ has any temporal feature at all. We could equally well say that the auxiliary **hat** carries both ***<** and **pres**. In this paper, I am following (Schlenker, 1999: 126). The auxiliary **haben** has a second meaning, the perfect, to which we will come later. Note that the English present perfect doesn’t have the analysis we find in German.

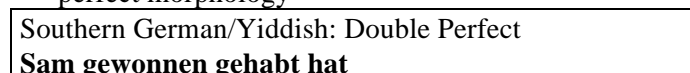
Having two different morphological outfits for the same meaning is a permanent source of confusion. Morphologically, the first tree has the verb in the preterite, whereas the second tree has it in the present perfect. Many researchers assume that therefore the two constructions mean something different.²³ This could be the case, but it need not be. I am following Schlenker here in assuming that these two different morphological forms encode the same meaning.

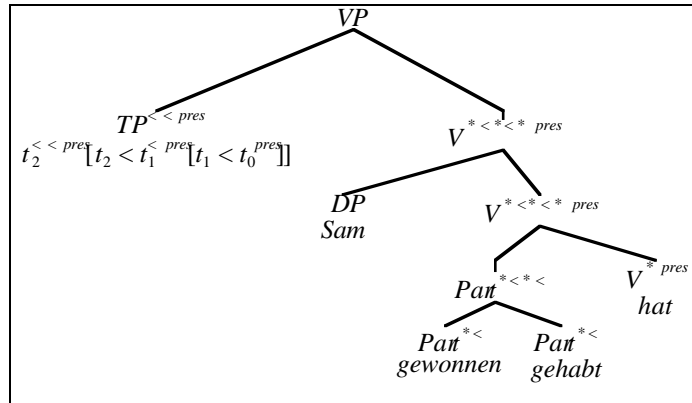
- (78) The syntax of **<past** in English and Standard German: preterite perfect morphology



This is what we expect from school grammar. The more complicated morphology is found in Southern German and Yiddish:

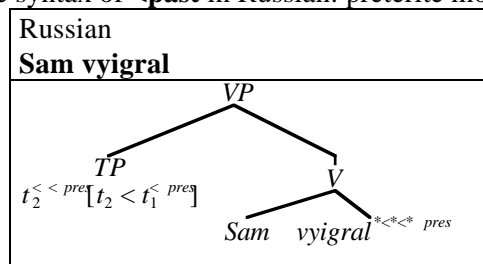
- (79) The syntax of **<past** in Southern German/Yiddish: present double perfect morphology





Russian doesn't have any verb morphology more complicated than the preterite. Therefore, this morphology is checked not only by **past** but by **<past** as well.²⁴ So the Russian preterite is ambiguous in some sense.

(80) The syntax of **<past** in Russian: preterite morphology



Here are the morphological rules deriving the differences between the languages:

(81) **Spell-Out Rules**

- a. ***past** (= ***<*<*<pres**)
win, ***3rd**, ***past** → /won/ (Engl.), /gewann/ (Germ.), /vyigral/ (Russ.)

No such rule in Yiddish/Southern German since there is no past morphology.

- b. ***<*<*<pres**
have, aux, *past → /had/ (Engl.), /hatte/ (Germ.)
 No such rule in Russian because there are no past auxiliaries.

No such rule in Yiddish/Southern German since there is no past morphology.
win, ***<*<*<pres** → /vyigral/ (Russian)

- c. Participle II
win, partII, * $\langle \rightarrow$ /won/ (Engl.), /gewonnen/ (Germ./Yidd.)
 No such rule in Russian given that there are no past participles there.
have, partII, aux, * $\langle \rightarrow$ /gehabt/ (Southern Germ./Yidd.)
 Dispreferred in Standard German/English.

These rules encode the typological differences that we find between the languages mentioned. And they connect semantic features with the morphology. They illustrate an important point: there is no 1-to-1 correspondence between semantic features and morphology. The same semantic feature can be encoded by two different morphologies even within one language, as we have seen for ***past** in German. And the same morphology may encode different semantic features, as we have illustrated for Russian.

5.2.2. Past and Pluperfect in Transparent Contexts

For the sake of illustration, I shall give some examples for the application of the theory.

- (82) *Mary was sick.*
 LF: $\exists t_7 \dots t_7^{\langle \text{pres}} [t_7 < t_2^{\text{pres}}] \text{ Mary was}^{*\langle \text{pres}} \text{ sick}$

At LF we may existentially close the referential time variable t_7 , whereas the egocentric **pres**-variable t_2 may remain free. Bear in mind that the auxiliary **was** is semantically empty and therefore deleted at LF. The temporal argument of **sick** is saturated by functional application of $\parallel \text{sick} \parallel$ to the time argument, though **sick** is not checked by the tense-features directly.²⁵

The pluperfect is an anaphoric tense: it is checked by a variable that is “prior” to a past variable. In isolation, pluperfect sentences are felt to be somehow incomplete:

- (83) *Mary had drunk.*
 LF: $\exists t_9 \dots t_9^{\langle \langle \text{pres}} [t_9 < t_7^{\langle \text{pres}} [t_7 < t_0^{\text{pres}}]] \text{ Mary had}^{*\langle \langle \text{pres}} \text{ drunk}^{*\langle$

The problem for the interpretation is the variable $t_7^{\langle \text{pres}}$, which must be anaphoric to a present variable, i.e., it cannot be existentially bound within the same sentence. Somehow we have to stipulate that only the referential variable of a tense is existentially bound in a sentence and this binding may

“dynamically” extend to temporal variables following in the text. When $t_7^{<pres}$ is introduced in a previous sentence, the sentence becomes perfectly acceptable:

- (84) *Mary was sick because she had drunk.*
 $\exists t_7 \dots t_7^{<pres} [t_7 < t_2^{pres}]$ **Mary was**^{*<pres} **sick** **because**
 $\exists t_9 \dots t_9^{<<pres} [t_9 < t_7^{<pres} [t_7 < t_2^{pres}]]$ **Mary had**^{*<<pres} **drunk**^{*<}
 True at t_c if there is a t_7 prior to t_c : Mary is sick at t_7
 because there is a t_9 prior to t_7 : Mary drinks at t_9 .

Homework for the reader: Compute this precisely! Compared with the complexity of other theories of the pluperfect (e.g. (Kamp and Reyle, 1993)), Schlenker’s theory is remarkably simple and works well.

5.2.3. Deriving English Sequence of Tense

In order to derive the SOT facts for English, recall that temporal variables have a well-formedness condition that must be checked at LF.

- (85) Temporal variables at LF. The well-formedness condition (44) for temporal features is checked at LF.

This principle should guarantee that verbal quantifiers always bind what intuitively is the egocentric temporal variable of the subordinate predicate. Recall the syntax (59) for verbs of attitude, viz. the stipulation that those features of the variables bound by the verb are deleted that agree with the morphology of the verb. Here are the LFs for the examples in (22).

- (86) a. *Smith thinks Mary is sick.* (de nunc simultaneous)
 $\dots t_1^{pres}$ Smith thinks^{*pres} $\lambda < \dots t_2^{pres} \dots > \dots t_2^{pres}$ Mary is^{*pres} sick
- b. *Smith thinks Mary was sick.* (de nunc anterior)
 $\dots t_1^{pres}$ Smith thinks^{*pres} $\lambda < \dots t_2^{pres} \dots > \exists t_3 \dots t_3^{<pres} [t_3^{<pres} < t_2^{pres}]$ Mary was^{*<<pres} sick
- c. *Smith thought Mary was sick.* (de nunc simultaneous)
 $\exists t_1 \dots t_1^{<pres} [t_1 < t_0^{pres}]$ Smith thought^{*<<pres} $\lambda < \dots t_2^{<pres} \dots > \dots t_2^{<pres}$ Mary was^{*<<pres} sick
- d. *Smith thought Mary is sick.* (double access!)
 $\exists t_1 \dots t_1^{<pres} [t_1 < t_0^{pres}]$ Smith thought^{*<<past} $\lambda < \dots > \dots t_2^{pres}$ Mary is^{*pres} sick
- e. *Smith thought Mary had been sick.* (de nunc anterior)

$\exists t_1 \dots t_1 \langle \text{pres} [t_1 < t_0^{\text{pres}}] \text{ Smith thought} \rangle \langle \text{pres} \dots \rangle \exists t_4 \dots t_4 \langle \text{pres} [t_4 < t_3^{\text{pres}}] \text{ Mary had been sick} \rangle$

(86a) is possible because the attitude verb deletes the feature **pres** of t_2 , thus satisfying (85).

The derivation of the LF in (86b) requires a further proviso. We first note that the referential variable t_3 of the embedded predicate must have the feature $\langle \text{pres}$, because the morphology of **was** requires so. Since the matrix verb deletes the feature **pres**, t_3 cannot be bound by the verbal quantifier, rather the variable t_2 , which figures in the presupposition is bound. Note that t_2 is not even present at S-structure. It comes into LF by a rule of construal. To obtain a correct LF, we have to delete the **pres**-feature of t_3 and leave the \langle -feature. In order to achieve that, we must accommodate the entire variable term.

(87) Accommodation of variable terms (LF)

Let $\zeta^{\langle a}$ be a temporal variable whose features are checked at S-structure, and let ξ^a be a variable with a deleted **pres** at LF. Then $\zeta^{\langle a} [\zeta < \xi^a]$ is a well-formed LF-term.

The principle entails that the term $t_3 \langle \text{pres} [t_3 \langle \text{pres} < t_2^{\text{pres}}] \text{ in the subordinate clause in (86b) is well-formed. Recall from definition (44) that the set of features } a \text{ following } \langle \text{ may be empty, i.e., } a \text{ is optional.}$

(86c) is obtained by a straightforward application of Deletion under Binding.

The derivation of (86e) requires again the accommodation of the variable term in the subordinate clause.

The theory predicts that the double access case (86d) is odd in several respects. The system requires that the **pres**-variable in the dependent clause be bound, for the matrix verb doesn't delete the feature **pres**. Therefore the variable must remain free and denote the speech time. So this variable must be interpreted de re. A problem with this is that Smith cannot be acquainted with the speech time t_c if this time doesn't extend to the time of the believing. But the standard assumption is that in English, **pres** denotes exactly the speech time.²⁶ So the meaning of t_2 has to be accommodated to cover this use: t_2 must denote a sort of Extended Now, a time extending from t_0^{pres} up to t_2 , and Smith must be acquainted with the beginning of that time.²⁷

The system works very well. Still, we have to add some qualifications. As it stands, the theory makes a false prediction for pluperfect under pluperfect:

(88) *Sam had thought that Mary had been sick.*

$\exists t_2 \dots t_2^{<past} [t_2 < t_1^{past} [t_1 < t_0^{pres}]]$ Smith had^{*past} thought^{*<}
 $\lambda < \dots t_3^{<past} \dots > \dots t_3^{<past}$ Mary had^{*<pres} been^{*<} sick
(simultaneous de nunc!)

The matrix verb deletes the pluperfect feature <past (= <<pres). So we should have a simultaneous de nunc reading. This doesn't seem available. Therefore, this derivation must be blocked by the following stipulation:

- (89) Only the features **pres** and **past** (= <pres) can be deleted under binding.

The question remains as to how we can derive the anterior de nunc reading for the sentence in (88). Treating the pluperfect as a tense won't do. Fortunately, **have** can also be a temporal shifter, and we can take that meaning to derive the reading in question. Sometimes past under past can express de se or de re anteriority.

- (90) *Mary said that Bill bought a car*
= Mary said that Bill had bought a car

As in Russian, the simple past can have pluperfect features with telic verbs, a sort of Past of Laziness.

5.2.4. Sequence of Tense in Russian

With respect to temporal binding, Russian (and other Slavic languages, e.g. Ukrainian) behave exactly as Amharic does with respect to the first person:

- (91) **The Russian Parameter.**
Russian verbs of attitude delete (LF) the feature **pres** at the temporal variable they bind, regardless what their tense (their temporal checkee) is.

Bear in mind that this deletion does not affect the temporal feature of the time argument of the verb. Again, Schlenker's account is slightly different. In his theory, verbal quantifiers transmit a subordination feature **t** to the temporal variable they bind, and the variables with that feature are pronounced as present, regardless of the tense of their verb; cf. (Schlenker, 1999: 175 ff.). Here is the analysis of the data in (24).

- (92) a. *Smith thinks that Mary is sick*
S. dumajet čto Maša bolna. (like
English)

- b. *Smith thinks that Mary was sick.*
S. dumajet čto Maša byla bolna. (like English)
- c. *Smith thought that Mary was (past) sick.*
S. dumal čto Maša (pres) bolna.
 $\exists t_1 \dots t_1^{<pres} [t_1 < t_0^{pres}]$ **Smith** **thought**^{*<pres}
 $\lambda < \dots t_2^{pres} \dots > \dots t_2^{pres}$ **Mary is**^{*pres} **sick**
- d. *Smith thought that Mary had been (plupf) sick.*
S. dumal čto Masha byla (past) bolna.
 $\exists t_1 \dots t_1^{<pres} [t_1 < t_0^{pres}]$ **Smith** **thought**^{*<pres}
 $\lambda < \dots t_3^{pres} \dots > \exists t_4 \dots t_4^{<pres} [t_4 < t_3^{pres}]$ **Mary was**^{*<pres} **sick**

The derivation of (92d) requires accommodation of the temporal variable term, of course. This system does not have the problem (Schlenker, 1999: 175 ff.) discusses on p. 161: Why can't past under past express simultaneity in Russian?

5.2.5. SOT in German

SOT in German is messy. It can follow either the Russian or the English strategy.

- (93) *Wolfgang sagte mir, dass die Gäste schon da sind/seien/were/wären.*
 W. said me that the guests already here are/be (Subj. I)/were(past ind.)/were (Subj. II)

They are all fine. The most thorough discussion of relevant data is found in (Fabricius-Hansen, 2002). It seems then that we have to describe the binding properties of German verbal quantifiers as a disjunction.

- (94) **German verbal quantifiers** delete (LF) the feature of the temporal variable they bind and they may always delete the feature **pres**.

If we disregard the subjunctive, this liberal rule seems to account for the variation we find in German. Still, not anything goes. If the main tense is a present, the SOT rules are as strict as in English or Russian. The facts remain to be investigated in greater detail.

5.3. Temporal Adverbs

5.3.1. Non-shiftable Deictics

Temporal adverbs such as **yesterday** and **tomorrow** are always interpreted with respect to the speech time. So their meaning obviously depends on the context c . The point I want to make in this section is that these adverbs have a temporal subject whose temporal features check the features of the finite verb. It follows that these adverbs are base generated in the temporal argument position of the main verb. For type reasons the adverbs undergo obligatory QR at LF. To my knowledge, this architecture was proposed for the first time in (Stechow, 2002b). It seems to me to be very important for an understanding of temporal adverbs. The syntax and semantics of these adverbs is described below.

- (95) $[\text{Adv } \mathbf{yesterday}]$ is a symbol of type $i((it)t)$. $\|\mathbf{yesterday}\|^{\text{g.c}} = \lambda t. \lambda P. t$ is on the day before t_c and $P(t) = 1$.

Here is the analysis of a simple sentence.

- (96) a. *Yesterday Gerd won.*
 b. SS: $w_0 [\text{AdvP } t_1^{<\text{pres}} \mathbf{yesterday}]^{<\text{pres}} \mathbf{Gerd won}^{* < * \text{pres}}$
 c. LF: $\exists t_1 \dots [\text{AdvP } t_1^{<\text{pres}} \mathbf{yesterday}]^{<\text{pres}} \lambda_1 w_0 t_1^{<\text{pres}} \mathbf{Gerd won}^{* < * \text{pres}}$

This analysis might look artificial, but I have argued at length in (Stechow, 2002b) that we need it anyway to get the interpretation of the following sentence right:

- (97) *John worked on every Sunday.*

If **yesterday** occurs under an attitude, the variable may be bound or it may be free. In both cases **yesterday** is described from outside, so to speak. Here is the de se reading.

- (98) a. *Smith thought Mary was sick yesterday/tomorrow.*
 b. $\exists t_1 \dots t_1^{<\text{pres}} [t_1^{<\text{pres}} < t_0^{\text{pres}}] \mathbf{Smith think}^{* < * \text{pres}} \lambda \dots t_2^{<\text{pres}} \dots \dots [\text{AdvP } t_2^{<\text{pres}} \mathbf{yesterday}] \lambda_1 \dots t_1 \mathbf{Mary sick}$
 Smith thought, "Mary is sick today."

From the fact that the embedded sentence has a de se/de dicto interpretation, it does not follow that Smith used any of the words of the **that**-clause to express the thought. It is the speaker who is responsible for the reference of the deictic adverb, not the subject of the sentence.

Schlenker's present shifter **now** is analysed exactly alike:

- (99) a. *Only now is the Concorde in Paris.* (Schlenker, 1999: 164)
 b. *Germ. Die Concorde ist erst jetzt in Paris.*
 "There is no time t earlier than t_c such that the Concorde is in Paris at t ."

The German translation makes it obvious that there is a temporal sense of **now** that means 'not earlier than'. Here is the semantics for the words.

- (100) a. $[\text{Adv } \mathbf{now}]$ is a symbol of type $i((it)t)$. $\| \mathbf{now} \|^{s,c} = \lambda t. \lambda P. t$ is at t_c and $P(t) = 1$.
 b. $[\text{Adv } \mathbf{only}_{erst}]$ is a symbol of type $i((it)t)$. $\| \mathbf{only}_{erst} \| = \lambda t. \lambda P. P(t) = 1. \neg \exists t' [t' < t \ \& \ P(t') = 1]$.

($P(t)$ is a presupposition.) Here is the derivation of the LF for (99).

- (101) SS: w $[\text{AdvP } [\text{AdvP } t_1^{\text{pres}} \mathbf{now}]^{\text{pres}} \mathbf{only}]^{\text{pres}}$ the Concorde is *pres in Paris
 $\implies [\text{AdvP } [\text{AdvP } t_1^{\text{pres}} \mathbf{now}]^{\text{pres}} \mathbf{only}]^{\text{pres}} \lambda t_2 w$ the Concorde is *pres in Paris
 (QR & Deletion of checkee and projected features)
 \implies LF: $\exists t_1 [\text{AdvP } t_1^{\text{pres}} \mathbf{now}] \lambda t_3 [\text{AdvP } t_3 \mathbf{only}] \lambda t_2 w$ **the Concorde is in Paris**
 (by QR and existential closure)

The reader may calculate for herself that this is exactly the correct meaning.

The following example by Ph. Schlenker shows that temporal adverbial clauses are semantic binders of tenses as well.

- (102) *Chaque fois que Pierre change d'emploi, il se querelle avec des gens qui étaient ses meilleurs amis un mois avant.*
 'Each time that Pierre changes jobs, he quarrels with some people who were his best friends one month before.'

A moment of inspection show that the sentence means something like this.

$$(103) \quad \forall t_1[(t_1[t_1 \subseteq t_0^{\text{pres}}] \& \text{Pierre changes jobs at } t_1) \rightarrow \exists t_2. \text{Pierre quarrels at } t_1 \text{ with people were his friends at } t_2^{\prec}[t_2 < t_1]]]$$

It is extremely tedious to derive this formally as you will see from my paper (Stechow, 2002b). Several things must be done. First, the present cannot denote the speech time here, but rather an interval that includes the speech time (cf. (Fabricius-Hansen, 1986)). Second, the entire subordinate clause has to be constructed as a quantifier that is base generated as a time argument of the main clause and QR-ed for type reasons, thus deleting the present feature of the time variable t_1 in the main clause. This has the side effect that the past in the relative clause loses its deictic character and becomes a relative past. The subordinate contains a bound present that is bound by the quantifier **each time**. While it takes me 30 seconds to write down a formula like (103), I would need one hour or more to generate a serious LF for sentences of this kind. Most work from the literature is on the 30 seconds level.

5.3.2. Schlenker on *in two days*

Here I take up adverbials of the type **in two days**. These adverbs are important witnesses for Schlenker's claim that attitudes quantify over contexts and are therefore monsters. He considers the following sentence and gives LFs for its two readings:

- (104) a. *John said a week ago that it would rain in two days.*
(Schlenker, 1999: 83 f.)
 b. $C^* \lambda C$ John said at time(C)-7 that [λc it would rain [in two days(C)]]] (de re)
 c. $C^* \lambda c$ John said at time(C)-7 that [λc it would rain [in two days(c)]]] (de se)

C and c are supposed to be context variables. The semantics for the adverbial is given on p. 86.

- (105) 'in two days' is a function that can take a matrix or a context variable as argument. Its value given an assignment s is ((Schlenker, 1999: 86)):
 $\| \text{in two days}(C) \|_s = \| \text{time}(C) \|_s + 2$
 $\| \text{in two days}(c) \|_s = \| \text{time}(c) \|_s + 2$

In section 4.2.1 I have argued that there are no context variables in this framework. Therefore, this semantics doesn't make sense. This raises the question of how we can express this semantics under our current

assumptions. Before we go on, we notice that Schlenker's analysis neglects the intervening future. According to his semantics, the following sentence should be grammatical:

- (106) **It is raining in two days right now.*
 C* λC . It is raining [in two days(C)]

Intuitively, the sentence is inconsistent. It entails that the speech time is two days later than it is. The semantics for **in two days** should bring that out. An intervening future somehow saves the statement.

5.3.3. Future as a Temporal Quantifier

According to (Ogihara, 1989) and many others, **will** is not a tense but a temporal quantifier. In the present approach, this verb is checked by the feature **pres** and it deletes this feature at the time variable it binds. The reasons for this treatment are examples such as the following ones:

- (107) a. *I will answer every e-mail that arrived.*
 (deictic = before the speech time; shifted before the future time) (Abusch, 1998)
 b. *John will say that he will be sick.*
 (not de se simultaneous)

The syntax for **will** should be formulated in analogy to that for verbs of attitudes (cf. (59)). Formally, **will** can be treated like a modal verb which abstracts over word and time and thereby deletes the features of the bound variables. The semantics ensures that this has no impact for the **ind**-feature.

- (108) **will** has the features **1st**, ***ind**, ***pres**. It is a symbol of type $is(is,t)$ and deletes the feature of the time and the world variable it binds at LF.
 $\| \text{will} \| = \lambda t. \lambda w. \lambda P. \exists t' [t' > t \ \& \ P(t')(w) = 1]$.

It follows that **will** is a raising verb with respect to the embedded subject because its person cannot be checked in the embedded position. Here is the analysis of Abusch's example(10b)

- (109) DS: $w_0^{ind} t_0^{pres} ()$ [_{VP} will^{*1st*pres*ind} [_{VP} x₁^{1st} answer every e-mail that arrived]]
 SS: $w_0^{ind} t_0^{pres} x_1^{1st} \lambda x_2^{1st} /I/$ [will^{*1st*pres*ind} [x₂^{1st} answer every e-mail that arrived]]
 (A-movement of /I/ for feature checking)

LF1: $w_0^{ind} t_0^{pres} x_1^{1st} \lambda x_2^{1st} \mathbf{will}^{*pres} \lambda < t_1^{pres}, w_1^{ind} > [w_1^{ind} t_1^{pres} x_2^{1st} \mathbf{answer\ every\ e-mail} \lambda_x \exists t_2. w_1^{ind} t_2^{<pres} [t_2 < t_1^{pres}] x \mathbf{arrived}^{*past*ind}]]$

(**will** must bind a world/time: movement across the board; t_1 bound by **will**; t_2 possibly after the speech time)

LF2. $w_0^{ind} t_0^{pres} x_1^{1st} \lambda x_2^{1st} \mathbf{will}^{*pres} \lambda < t_1^{pres}, w_1^{ind} > [w_1^{ind} t_1^{pres} x_2^{1st} \mathbf{answer\ every\ e-mail} \lambda_x \exists t_2. w_1^{ind} t_2^{<pres} [t_2 < t_0^{pres}] x \mathbf{arrived}^{*past*ind}]]$

(deictic: t_0 free; t_2 before the speech time)

Note that the features of the variables of infinitives need not be checked by the infinitive morphology. The checking is done by a finite verb and transmitted via binding. The embedded w_1 -variable is formally bound but the semantics makes sure that it is evaluated with respect to the actual world. So **will** looks like a genuine modal verb, but it doesn't really quantify over worlds.

5.3.4. The Logophoric Deictic **in two days** and the Future Effect

The distance adverbial **in two days** is a logophoric deictic, which means that its reference time is at a distance of two days into the future of the deictic or logophoric centre. From this semantics we derive what may be called the *future effect* of this adverbial: if the adverbial is not embedded under a future operator, we obtain a contradiction. Here is the semantics of the distance preposition **in**:

(110) Distance **in**.

The preposition of distance **in** has a logophoric variant **in**_{log}, which marks its first (nearest) argument with **log**, and a deictic variant **in**^{*}, which marks its first argument with the features **pres** and **M**, where **M** ("matrix") means that the variable cannot be λ -bound by a verb or operator. Both variants have the type $i(m(i(it,t)))$, where m is the type of time measure phrases. Let f be **log** or **pres**, **M**.

$\| \mathbf{in}_f \| (t)(m)(t')(P) = 1$ iff $P(t') = 1$ & distance(t' , t) = m .

The disjunctive definition shows that a generalisation is missing. So I agree with Schlenker that we want something like his (105). The framework of the dissertation doesn't allow us to express this, however, because it knows of no genuine context variables. Furthermore, Schlenker cannot derive the Future Effect.

(111) The Future Effect

- a. *It is raining in two days.
 $\dots \exists t_3 [_{PP} t_3^{pres} \text{in}(t_2^{pres.M}) \text{ 2 days }] \lambda t_1^{pres} t_1^{pres} \text{ is } ^{\#pres} \text{-raining}$
 Contradictory!
- b. ^{OK}It will rain in two days.
 $w_2^{ind} t_1^{pres.M} \text{ will } ^{\#pres\#ind} \lambda \langle t_2^{pres}, w_2^{ind} \rangle \lambda [_{PP} t_2^{pres} \text{in}(t_1^{pres.M}) \text{ 2 days }] \lambda t_3, w_2^{ind} t_3 \text{ rain}$

The inconsistency of (a) comes from the fact that the reference time and the deictic centre of the distance adverbial are interpreted with as the same time, viz. the speech time. But the speech time cannot be distant from itself. Note that the distance-PP is base generated at the position of its trace t_1 , and its tense is checked there at SS. At LF, it is QR-ed for type reasons.

would has, of course, exactly the same semantics as **will**. The only difference is that it marks its time argument with the features **past** and **log**. This means that the argument must be a variable bound by a logophoric verbal quantifier.

I leave it as homework for the reader to derive the following Schlenkerian contrast:

- (112) a. *John told me that he would finish his work in two days.*
 b. *#It often happened that John would finish his work in two days.* (OK in the sense: “It often happened that John would finish his work *within* two days.”)

Similarly for the future effect:

- (113) a. *#John told me that it was raining in two days.*
 b. *John told me that it would rain in two days*²⁸
 c. *#John told me that it had rained in two days.*

To the best of my knowledge there is no analysis in the literature that can derive these facts.

I leave it to the reader to convince herself that we have all then necessary tools for the analysis of Abusch’s version of the classical example by Kamp & Rohrer:

- (114) a. *John decided a week ago that in ten days at breakfast he would say to his mother that they were having their last meal together.*
 b. John decided ($t_1 \langle [t_1 < t_0^{pres}] \dots \lambda t_2 \text{ will}(t_2^{past}) \dots \lambda_3 \text{ John say}(t_3^{past}) \text{ to his mother } \lambda_4^{past} \text{ they were}(t_4^{past}) \text{ having their last lunch.}$

Omitting disturbing details, the temporal structure is as in (b), where the temporal variable of the verbs are written at the right side of the verb. The full analysis is exceedingly complex and is left as homework for the reader.

5.4. Mood and Tense

The following ideas for the analysis of mood suggest themselves.

5.4.1. Modals as Indicative Absorbers

All modals are world quantifiers and therefore delete the **ind**-feature of the world variable they bind. The tense feature of bound time variable is deleted as well, but since the time is not shifted, this doesn't have any semantic effect. Here is an entry for **must**:

- (115) First person **must** is of type $is(is,t)$. It deletes the features of the variables it binds at LF.
 $\| \mathbf{must} \|^c = \lambda t. \lambda w. \lambda p. \forall w' [w R_c w' \rightarrow p(w')(t)]$, where R_c is a contextually given relation of accessibility.

Here is a German example.

- (116) a. *Gerd musste gewinnen.*
 Gerd must-past win
 b. $w_1^{ind} t_1^{past} \mathbf{must}^{*ind*past} \lambda \langle w_2^{ind}, t_2^{past} \rangle w_2^{ind} t_2^{past} \mathbf{Gerd win}$

For historical reasons, most English modals have neither past forms, nor do they have non-finite forms. Therefore, we have to express the meaning by a different construction in English (e.g. **Gerd had to win**).

5.4.2. Subjunctive Modals

In this section I will take up Abusch's examples (31), i.e. the distribution of **ought**. The first thing we notice is that this modal cannot be checked by **pres** alone because it can be embedded under a verbal quantifier that deletes **past** without giving rise to a DA reading; cf. (31b). On the other hand, **ought** cannot be checked by the feature **past** because it cannot occur in matrix past-sentences; recall (31a). The puzzling fact is that **ought** is licensed under **will**; recall (31d). There is no elegant derivation of these facts in the present framework. So we stipulate the following syntactic features:

- (117) **ought** has a person feature one of the following two checkees: ***pres** or ***subj**.

***subj(junctive)** is the same feature as ***log**. The only difference is that it is checked by a world variable, whereas ***log** is checked by an individual variable. Given that **will** qualifies as a verbal quantifier, it follows that **ought** is licensed under **will**. I leave it to the reader to convince himself that this entry accounts for the data in (31). Most subjunctive finite modals behave in the same way:

- (118) Subjunctive Modals
- might, could, should** have the same features as **ought**.
 - Germ. **müsste, könnte, sollte, dürfte** behave alike.

Here is the analysis of just one of the data in (31):

- (119) a. *John will always be a student that ought to study more.*
(Heim)
- b. $w_1^{ind} t_1^{pres}$ **will** $\lambda < t_2^{pres}, w_2^{ind} > t_2^{pres} w_2^{ind}$ **John be a student**
 $\lambda_x t_2^{pres} w_2^{ind,log}$ **ought** $\lambda < t_3^{pres}, w_3^{ind} > t_3^{pres} w_3^{ind} x$ **study more**

The world variable that checks the feature ***subj** carries the feature **log**. Therefore, this variable must be bound by a verbal quantifier: **will** does the binding. Due to the fact that the said modals do not have non-finite forms in English, we find syntactic ambiguities that cannot be analysed compositionally in English, whereas in German (or Italian) they can.

- (120) a. *John might have won.*²⁹ (ambiguous)
- b. *John könnte gewonnen haben.* (unambiguous)
- c. *John hätte gewinnen können.* (unambiguous)
'*John had could win'

können is the so-called 'Ersatzinfinitiv', a suppletive form for the past participle **gekonnt**. Interestingly, semantically this form behaves like an infinitive. (c) has a counterfactual meaning, which obscures the point. Here are other examples:

- (121) a. *Du musst zwei Jahre gedient haben.*
you must two years served have
'You are required to have served for 2 years.'
- b. *Du hast zwei Jahre dienen müssen.*
you have two years serve must

‘You were required to serve for 2 years.’

In the first sentence, the modal accessibility is determined with respect to the speech time, in the second it is evaluated with respect to a past time. Even if we maintained the analysis (77)/(78) according to which **haben/have** is always semantically empty, **have** and **must** must have scope relative to each other. Reading (121a) is the one English can’t express compositionally.

5.4.3. German Subjunctive as a Logophoric Mood

The subjunctive facts in (35) are deduced from the assumption that a verb with the feature ***subj** checks a world variable with the feature **log**. This in turn means that the variable must be bound by a verb of attitude. Here is the analysis.

- (122) a. *Ich dachte, ihre Yacht sei/wäre (subj.) länger als sie ist (ind.).* [cf. (35)]
 $w_1^{ind} t_1^{<pres}$ I thought $\lambda < w_2 t_2 > w_2 t_2$ your yacht was *subj longer than $w_1^{ind} t_3$ it is *ind
- b. *#Ich dachte, ihre Yacht sei/wäre (subj.) länger als sie sei/wäre (subj).*
 $w_1^{ind} t_1^{<pres}$ I thought $\lambda < w_2 t_2 > w_2 t_2$ your yacht was *subj longer than $w_2^{log} t_3$ it was *subj

The contradiction comes from the constraint that both world variables in the comparative construction must be bound by the verbum sentiendi **thought**. (Schlenker, 1999: 198) discusses the issue of locality: should a subjunctive world variable not have an additional feature that enforces binding by the next verbal quantifier? Is the following indexing acceptable?

- (123) *²Fritz behauptet, Maria glaube (subj), meine Yacht sei (subj) länger als sie tatsächlich sei (subj).*
 $w_0 \dots$ Fritz claims $\lambda < \dots w_1 > \dots w_1 \dots$ Mary believe $\lambda < \dots w_2 > \dots w_2$ my yacht be longer than $\dots w_1$ it be long

I don’t get this reading, but better examples may exist.

Schlenker observes that the subjunctive interacts with the first person and with the present tense.

- (124) a. **Ich glaube (pres), dass Maria krank sei (subj).*
 I believe that Mary be sick.
- b. *OK Ich glaubte (past), dass Maria krank sei (subj).*

- c. I thought that Mary be sick.
^{OK}*Fritz glaubt* (past), *dass Maria krank sei* (subj).

Annette Farhan (p.c.) has observed that factive verbs (taken in their factive interpretation³⁰) don't allow for the subjunctive in their complement.

- (125) #*Fritz ärgerte sich darüber, dass Maria krank sei* (subj).
 Fritz was annoyed about the fact that Mary was sick.

I cannot explain the facts in (124). (Schlenker, 2001a) has explained the interaction of the the subjunctive with the speaker, but his theory doesn't work for factive verbs. I leave the matter for further research. In any case, the claim that the subjunctive is a logophoric mood (in one of its uses) seems justified.

6. Binding Principles for Person, Tense and Mood

Here is a summary of the most important theoretical ingredients of this framework. In most cases no special binding principles for the correct indexing of time, person and world variables are needed. Everything is done by Feature Deletion under Binding, a central principle of LF. We can index time, person and world variables freely. If a verbal quantifier binds them, their relevant feature is deleted and we obtain a de se reading. Generally, deletion is performed under agreement, but we have to stipulate exceptions in order to get the Amharic and the Russian Parameters:

- (126) Feature Deletion under Binding
- a. The Unmarked Option: Verbal quantifiers delete the features of the variables they bind under agreement with their features.
 - b. The Amharic Parameter: Verbal quantifiers delete the 1st person of the individual variables they bind, regardless of what their person is.
 - c. Object Control: OC-verbs delete the person feature of the individual variables they bind if the feature agrees with that of the object of the verb.
 - d. The Russian Parameter: Verbal quantifiers delete the feature present of the temporal variable they bind, regardless of what their tense is.

Binding of logophoric pronouns requires ad hoc syntactic features.

- (127) Binding of logophoric variables

- a. A variable with the feature **log** is bound by a verbal quantifier.
- b. A variable with the feature **local** is bound by an adjacent verbal quantifier.

Principle (a) regulates the distribution of logophoric pronouns and German subjunctive, whereas (b) regulates the distribution of PRO. There is some redundancy in these features: **local** implies the presence of **log**. As far as I can tell, any other theory variant known needs features of this kind as well. In this respect they are not ad hoc.

(128) Licensing of **ought**

ought is licensed by a time variable with the feature **pres** or (exclusively!) by a world variable with the feature **log**.

7. Prohibition against Monsters is Alive

The theoretical merit of Schlenker's theory is that he has given a principled account of the syntax and semantics of tense, person and mood. Verbs of attitude affect the interpretation of these features simultaneously, and he gives the outlines of a theory explaining their behaviour. Given that many languages encode these features in the morphology of the verb, the importance of the account for our understanding of the architecture of grammar should be obvious. The descriptive merit of the theory is that for the first time we can explain mysterious facts of language. In this paper I have been mostly focussing on temporal phenomena.

Let me comment again on the philosophical claim behind Schlenker's writings. He thinks to have refuted Kaplan's (1979) famous Prohibition against Monsters. I am not convinced that he has been successful in this respect. A monster is an operation that essentially operates on a character, i.e. on a function from contexts to intensions. As an example, Kaplan considers sentences like the following one:

(129) A monster:

At some contexts I am speaking.

True at context *c* iff there are contexts *c'* such that the speaker of *c'* is speaking at the time of *c'* in the world of *c'*.

Here, **at some contexts** would be a quantifier over contexts. Schlenker has given no argument that such operators exist in natural language. It is true that attitudes quantify over triples that are composed of the same components as contexts, but they are not contexts of utterances but rather indices. Attitudes are egocentric intensional operators; they never quantify

over the context of utterance. Schlenker's notation obscures this point, and I have gone to considerable effort to bring this out (see 4.2.4). As I have said, the framework of the dissertation doesn't even allow us to quantify over contexts. In recent versions of the theory ((Schlenker, 2001a) and (Schlenker, 2001b)), a formal language is proposed in which binding of context variables is introduced in a semantically sound way. But in these versions, too, no example for a Kaplanian monster is given. It remains true, of course, that many of Schlenker's data cannot be treated in Kaplan's theory as it stands, since that theory doesn't have egocentric propositions. This is a small point.

The crucial difference is the idea that features restrict the reference of variables, and the variables are bound by verbs under loss of their features. The idea that bound variables do not have features is found also in papers by Kratzer, e.g. (Kratzer, 1998). In Kratzer's theory, semantically bound variables obtain their φ -features at the PF-branch via coindexation with an appropriate antecedent. The possibility that variables may lose their features in the course of the derivation is explicitly rejected by Kratzer.

8. Loose Ends

There is a serious gap in the theory as it stands, viz. the locality principles that govern the relation between a verbal quantifier and the variables it binds. The syntax for verbal quantifiers (59) doesn't say anything about how distant the variables bound by $\lambda\langle\mu, \nu, \xi\rangle$ in the configuration $[\text{vpATT}^{*a*b*c} \lambda\langle\mu, \nu, \xi\rangle \varphi]$ are. As it stands, the system heavily overgenerates:

- (130) a. *Bill believes (w_0) λw_1 John is married (w_0) to a rich(w_1) woman.
 b. ^{OK}Bill believes (w_0) λw_1 John is married (w_1) to a rich(w_0) woman.
 c. ^{OK}Bill believes (w_0) λw_1 John is married (w_1) to a rich(w_1) woman.

(Percus, 2000) considers similar examples and claims that the λ -operator that abstracts a proposition from an open formula must be related to the variable it binds by a sort of superiority condition: it is not possible to bind a lower world variable if there is a higher one that could be bound. In recent work ((Percus and Sauerland, 2002)), the same point is made for de se pronouns.

- (131) a. *I λ_x x dreamt λ_y that x was getting married to y's granddaughter.

- I dreamt, “This man getting married to my granddaughter.”
- b. ${}^{OK}I \lambda_x x \text{ dreamt } \lambda_y \text{ that } y \text{ was getting married to } x\text{'s granddaughter.}$
- c. ${}^{OK}I \lambda_x x \text{ dreamt } \lambda_y \text{ that } y \text{ was getting married to } y\text{'s granddaughter.}$

Percus claims that the indexing (a) is not possible. The facts are very subtle and deserve further discussion.³¹ The hope is that we can explain the locality restrictions if we conceive of the relation between the λ -operator and the variable it binds as movement of a semantically empty operator. Heim’s original formulation of feature transmission (4) assumed movement anyway. Let us assume then that the λ -operators are indeed moved and that the superiority condition can be derived from the theory of movement. Then there would still be something missing, viz. the agreement relation between the person, tense and mood feature between the lambdas and the binding verb, which we need for feature deletion. Under current assumptions this relation has to be stipulated.

(Heim, 2001b) has proposed a radical movement approach in which a verbal quantifier is moved literally from the “word” argument of the embedded clause. In this approach, the sentence **Bill believes it is raining** would have roughly the following derivation:

- (132) DS: $[_{VP}([_{CP} \text{ Mary is sick}(I \text{ believe}(w_0)))]]$
 SS: $[_{VP} I \text{ believe}(w_0) \lambda w_1 [_{CP} \text{ Mary is sick}(w_1)]]$
 (by head movement of the verbal quantifier **I believe**(w_0))

This approach raises a number of questions. First of all, w_0 and w_1 are Lewisian individuals that live exactly in one world at one time. To make this precise, we need Lewis’s semantics of counterparts, which is not familiar to the great majority of linguists. The second question is of how tense and mood/modality are to be accounted for in this approach. Modal verbs must be treated like verbs of attitude, i.e., they must be QR-ed from the world argument of the verb they modify. Tenses must be features restricting the time of the individual. Presumably, the semantic perfect must be treated as a verbal quantifier as well. Heim has developed her proposal in great detail and it looks very promising, but the theory of tense has to be integrated yet, which I cannot do here. So this is another way to go.

Let me finally mention that the theory of tense I have proposed here needs to be better elaborated. This concerns the present in particular. Following Schlenker, I have treated **pres** as a feature that restricts its variable to the speech time. Presumably, this is not general enough. The present tense may very well be a relative tense as well, i.e. the variable should denote a time that overlaps the speech time and could be a rather large time in English. In German it could be a variable that denotes a time

not before the speech time. This would complicate analysis considerably but would be empirically more correct.³²

Another suspect point was that with past under attitudes, it was never the referential variable that was bound by the verbal quantifier but rather the “logophoric centre”, i.e. Abusch’s subjective now. This should follow from the theory. And I haven’t said anything about aspect and aktionsarten in this paper. Virtually every example is crudely simplified. So a lot remains to be done.

Notes

* I have to thank Philippe Schlenker and Orin Percus for helpful comments and Angela Cook for correcting the English of the first version.

1. (Heim, 2001a: 4)
2. Assuming the theory of (Rooth, 1992), representation of the sentence must be (a) and not (b) or (c):
 - a. only $[I_F \lambda x. x \text{ did } x\text{'s homework}] \sim F$
 - b. only $[I_F \text{ did my homework}] \sim F$
 - c. only $[I_F \text{ did my}_F \text{ homework}] \sim F$
 (b) entails “Fritz didn’t do my homework” and (c) entails “Fritz didn’t do Mary’s homework”. In (Schlenker, 2001a) we find some ad hoc devices for treating focused pronouns. They ignore the meaning of the person feature as well and do not work for Heim’s plural examples.
3. The example *Gerd hopes he will win* would be more natural. I choose the present to keep the exposition as simple as possible.
4. The de re analysis means: Each of John, Mary, and Bill bears a relation of acquaintance toward himself and thinks that the object toward which he bears the relation is being attacked by a bear. Clearly this condition is fulfilled in the scenario given. Those who would dispute that the sentence is true in the scenario given would reject Percus & Sauerland’s argument, of course.
5. Recently, Schlenker tells me that the second example is not very solid. A better investigated example would be this:

(Leslau 1995 p. 779)

what bring.IMPER-2M COMP-say.PF-3M-1SO NEG-hear.PF-1S-NEG

‘I didn’t hear what he told me to bring’ (lit. I didn’t hear that he said to me bring what)

The example is interesting for the fact that we have to interpret an embedded **you** as de se. Anticipating the semantics of the paper, we will have to say the following for the feature 2nd:

Let ζ be an individual variable. $\|\zeta^{2nd}\|^{g,c}$ is only defined, if $g(\zeta) =$ the person(s) addressed by s_c . If so, $\|\zeta^{2nd}\|^{g,c} = g(\zeta)$.

6. It follows that “remember” should not qualify as an attitude predicate in the sense discussed here, though the semantics certainly somehow involves an attitude. I don’t know what the correct semantics of “remember” should be. Neither do I know what the semantics of factive verbs is, which also behave rather differently from the verbal quantifiers considered here.
7. Of course, this generalisation presupposes theoretical notions like binding and binding by verbs.
8. Together with the present perfect, these forms show a scope ambiguity, which is syntactically resolved in German.
 - a. John might/could/should have studied more.

- b. Hans könnte/sollte mehr gelernt haben. might/could > have
- c. Hans hätte mehr lernen können/müssen. had > might/could

There is more involved than scope, since (c) is a past counterfactual, whereas (b) is not a counterfactual.

9. (May, 1977).
10. (Stechow and Sternefeld, 1988) use a subcategorisation hyphen, and (Sternefeld, 2000) puts an asterisk at both sides of the checkee.
11. I will also write 1st, 2nd, 3rd for the numbers.
12. In this system, the head of the mood phrase MP may have the feature *nominative. Assuming that the subject has it, it must move to SpecMP.
13. I leave it open how category labels ($\pm N$, $\pm V$) are treated. For convenience, I leave them at LF.
14. In Schlenker's system, the features <a occur with the second occurrence of ζ as well, i.e. the restriction is $[\zeta^{<a} < \xi^a]$. This would make the syntax circular, because this variable should satisfy the condition as well, so the LF would be $\zeta^{<a}[\zeta^{<a}[\zeta^{<a}[\zeta^{<a} \dots] < \xi^a] < \xi^a]$. Clearly this is not intended. My reformulation avoids this error.
15. For the reconstruction of Schlenker's theory of the historical present and free indirect discourse, we would need his idea of formal context. This would be a sequence of expressions of type e, i and s respectively, which by convention would have the features **1st**, **pres** and **ind**, respectively. The interpretation would not depend on the Kaplanian c anymore but only on the assignment g. A Kaplanian character would be a function from formal contexts into intensions. A proper reconstruction of Schlenker should proceed along these lines. I realised this possibility only after completion of the paper.
16. For this kind of semantic rule, see (Kratzer, 1978) A remark to the terminology: A Kaplanian character $ch(c)(i)$ is a two place function where both c and i are triples consisting of an individual, a time and world. The first argument, c, is a *context* and the second, i, is an *index*. $ch(c)$ is the intension expressed by ch at c, and $ch(c)(i)$ is the intension with respect to c and i. The Prohibition against Monsters says that the character of a functor α applied to an argument β can always be determined by the following principle of composition: $\|\alpha(\beta)\|(c)(i) = \|\alpha\|(c)(\|\beta\|(c)(i))$, i.e. the computation of a complex character never really depends on the entire embedded character.
17. The conventions underlying this formulation are the same as in (Heim and Kratzer, 1998)
18. Recall from the syntax folklore that the φ -features are needed for agreement reasons: *Mary thought it would be good to behave herself/*himself*.
19. Here is an approximation. $\|\text{persuade}\| = \lambda p.\lambda y.\lambda x.\lambda t.\lambda w.\forall w'\forall t'[y \text{ has realised in } w' \text{ at } t' \text{ what } x \text{ suggest to } y \text{ in } w \text{ at } t \rightarrow p(y)(t')(w')]$.
20. For Yiddish morphology, see (Schlenker, 1999: 126)
21. Cf. e.g. (Klein, 1998) or
22. The <past reading usually requires the verb to be in the perfective. This is ignored here. Cf. (Paszowska and Stechow, 2002)
23. (Klein, 2000) expresses this idea in saying that adjectives and participles must be linkable to finiteness (FIN-linkable). The theory here says that adjectives and participle have a time argument, but no temporal checkees.

24. See (Mittwoch, 1988), (Klein, 2000) and many others.
25. See (Abusch, 1991) for a proposal to that effect.
26. Susi Wurmbrand (p.c.) tells me that she cannot get this reading for the Germ. sentence *Hans sagte mir, dass es in zwei Tagen regnen würde*. I can say that. So Wurmbrand has only the deictic version of **in** her language, while I have Schlenker's logophoric variant as well.
27. I have observed the ambiguity in (Stechow, 1995b: 43) (Condoravdi, 2002) observes it as well.
28. (Fabricius-Hansen, 2002) gives many examples that show that that factive verbs can be reinterpreted as *verba dicendi vel sentiendi*:
Die Firma bedauerte, dass sie 2000 Angestellte entlassen müsse (subj).
The company regretted that it had to dismiss 2000 employees.
29. Even if Percus' observations were true for the verb **dream**, they are not valid for a verb such as **think**. The sentence
I thought I was being attacked by my son
surely has a de se reading, since my thought could have been "I am being attacked by my son".
30. The account of the present in (Stechow, 1995a) seems descriptively more adequate in this respect.

References

³ An alternative to this strategy has been proposed in (Kratzer, 1998). I became aware of it only one year after completion of this paper. In Kratzer's theory, semantically bound pronouns do not have φ -features at SS (spell-out). They obtain them on the PF branch via conindexing with an appropriate antecedent. I have compared Kratzer's and the present approach in (Stechow, 2003).

¹¹ I am always confused with this terminology. In the first version of this paper, I called the star-feature checker and the feature without star checkee. Orin Percus told me that a switch of the terminology would be more intuitive. At present, I find the former terminology more intuitive, because the element with the star-feature subcategorises the other feature. Since nothing hinges on that terminological issue I leave the terminology as it stands.

