

An Underspecified HPSG Representation for Information Structure

Jonas Kuhn

Institut für maschinelle Sprachverarbeitung

Azenbergstraße 12

D-70174 Stuttgart, Germany

Email: jonas@ims.uni-stuttgart.de

Abstract

Information structure can be of great use in linguistic applications, especially in those involving a speech component. However, focus marking by prosody is often ambiguous. Existing theories capture this by rules that produce alternative focus structures. This disjunction is hard to handle computationally. In this paper, a compact, graphically underspecified representation is proposed, along with composition principles and a resolution routine based on context information.

1 Introduction

This paper proposes an underspecified representation for information structure (IS) and HPSG¹ principles that build up this representation for German. The representation is designed as a basis for use in applications like text-to-speech and speech translation systems. It is obvious that for a non-tiring, natural output, especially the prosodic side of IS has to be taken into account. In this section, examples from sentence-based translation with semantic transfer visualize the role of IS in several empirical domains. For one thing, ignoring IS may result in confusing translations: (2) with default accenting,² despite being truth-conditionally equivalent, is not a suitable translation for (1B) in the given context. System users will probably presume that a serious mistranslation occurred.

- (1) A: Zu unserem Treffen wird also Frau Otto von der Firma Müller kommen?
(‘So, Frau Otto of Müller’s will participate in our meeting?’)
B: Nein, Frau SCHNEIDER wird Firma Müller vertreten.
No, Frau Schneider will company Müller represent
- (2) No, Frau Schneider will represent MÜLLER’S.

¹See (Pollard/Sag, 1994).

²SMALL CAPS are used to highlight pitch accents.

Moreover, IS can be exploited to choose between certain translation alternatives on sentence level (cf. (Eberle, 1996)). The particle *noch* has different translations depending on accentuation:

- (3) G. Maria sucht noch einen BRIEFKASTEN.
Maria looks-for still a post-box
E. Maria is still looking for a post box.
- (4) G. Maria sucht NOCH einen Briefkasten.
E. Maria is looking for another post box.

Concluding the set of initial examples, (5) shows that focus can play a role in simple inferences drawn from context to resolve lexical ambiguities (or similarly for ellipsis and anaphora resolution).

- (5) Frau Schulze kündigte an, die Firma im Herbst zu besuchen.
(‘Frau Schulze announced to visit the company in autumn.’)
G. Ich nehme an, daß die Managerin ihren *I assume that the manager her* Besuch BEGRÜSST.
visit/visitors welcomes
E. I assume that the manager will APPRECIATE her visit.
- (6) Ich kann Frau Müller nirgends im Gebäude finden.
(‘I can’t find Frau Müller anywhere in the building.’)
G. Ich nehme an, daß die Managerin ihren BESUCH begrüßt.
E. I assume that the manager is giving a welcome to her VISITORS.

Since the ambiguous *Besuch* in (5) is not focus-marked, it has to be contextually given. Even a simple context model will resolve the reference back to the *besuchen* event mentioned in the preceding sentence, so the noun *visit* is chosen as a translation.

The last example showed that certain ambiguities require inferences from context. However, the other examples make clear that focus can be effectively applied in sentence-based translation exploiting the existing linguistic focus marking *within* the sentence. Costly reasoning can thus be restricted to doubtful cases.

A sentence-based account of IS faces one problem: in the general case, a certain prosodic marking of a sentence does not correspond to a unique IS. These cases of *focus ambiguity* have been discussed extensively in the literature and will be briefly reviewed in sec. 2. The existing HPSG account of IS by Engdahl/Vallduví (1994) generates a disjunction of alternative sentence analyses for ambiguous sentences.³ Since focus ambiguity multiplies with all other kinds of ambiguity, a disjunctive treatment makes the use of IS in computational applications unattractive. Therefore, an underspecified representation of IS is proposed in this paper (sec. 3) along with a resolution routine on the basis of context information (sec. 4). Finally, the linguistic principles that compose the representation are worked out in HPSG (sec. 5). The data discussed are German, but English is very similar.

2 Focus marking

The basic data are well-known.⁴ A pitch accent on a direct object like *Buch* in (7), can serve to mark a number of constituents as focused: (8).⁵ The focus feature is usually assumed to project.

(7) Karl hat ein BUCH gelesen.
Karl has a book read

- (8) a. Karl hat [ein BUCH]_F gelesen.
b. Karl hat [ein BUCH gelesen.]_F
c. [Karl hat ein BUCH gelesen.]_F

A pitch accent on the subject *Karl* however cannot project focus (9), neither do adjuncts project focus (10).

(9) [KARL]_F hat ein Buch gelesen

(10) Maria hat [DRAUSSEN]_F geniest.
Maria has outside sneezed

These data can be explained by the following *F-Assignment Rules* from (Selkirk, 1995):⁶

- (11) *Basic F Rule*
An accented word is F-marked.

³The same is true for practically all accounts of focus projection that I am aware of – cf., e.g., (Selkirk, 1995; Schwarzschild, 1995).

⁴For German, cf., e.g., (Höhle, 1982).

⁵The standard means to determine the focus of an utterance is the question test, where the focus in the answer corresponds to the interrogative constituent in the question. However, this test can be misleading in cases where the question comes in a wider context (cf. (16) below and the discussion in (Kuhn, 1996)).

⁶Rule (12b) may need some refinement. It is not clear whether it is a syntactic condition that constrains indirect F-marking of a head. Jacobs (1994) argues that the notion of *integration* or *informational autonomy* is responsible instead, which is based on the complexity in semantic processing. Unfortunately, the formal nature of *integration* is still ill-understood.

(12) *F Projection*

- a. F-marking of the *head* of a phrase licenses the F-marking of the phrase.
b. F-marking of an *internal argument* of a head licenses the F-marking of the head.

According to these rules, the head itself can project focus, which appears to be refuted by data like the following.

(13) Karl hat ein Buch [GELESEN.]_F

However, (13) *can* have a wider focus if books are *contextually given* (this effect has been called *deaccenting*). Note the difference between (14) and (15), both marked according to (11) and (12): although their maximal focus domain is identical, *ein Buch* is F-marked only in (15).

(14) Karl hat [ein Buch [GELESEN.]_F]_F

(15) Karl hat [[ein BUCH]_F [gelesen.]_F]_F

This is reflected in Schwarzschild's (1995) interpretation of the F-feature. All constituents that are *not* F-marked need to be given, where givenness is defined as entailment by prior discourse. Entailment is extended to expressions of other type than propositions by existentially binding unfilled arguments.⁷ Even non-F-marked constituents embedded in an F-marked constituent, like *ein Buch* in (14) have to pass this givenness filter. So, deaccenting is no longer a special case for the theory.

To restrict the optional focus projection from (12) further, Schwarzschild assumes an additional pragmatic filter *Avoid F* that selects the tree with the least F-marking.⁸

Non-F-marked constituents that contain F-marked subconstituents need to be given as well. The context has to entail their *F-skeleton* – the

⁷The exact definition is as follows ((Schwarzschild, 1995), pp. 5-6.):

An expression T (for target) is *GIVEN* iff it has an antecedent, A, such that the existential closure of A entails the result of existentially binding F-variables in the existentially closed F-skeleton of T [where the existential quantifier binding F-variables quantifies over contextually salient values].

The *F-skeleton* of an expression is the result of replacing F-marked elements with variables (working top to bottom).

The *existential closure* of an expression is the result of existentially binding all unfilled arguments of the expression.

if α is type t, $\text{ExClo}(\alpha) = \alpha$.
if α is of conjoinable type $\langle a, b \rangle$,
 $\text{ExClo}(\alpha) = \exists x_a [\text{ExClo}(\alpha(x_a))]$.
if α is type e, $\text{ExClo}(\alpha) = \exists P_{\langle e, t \rangle} [P(\alpha)]$.

⁸I consider the status of this filter somewhat problematic. Why should a tree with less F-marking be pragmatically preferred? One could as well argue that since a sentence should be as informative as possible, given constituents should be avoided. The underspecified account developed here recasts Schwarzschild's ideas in a way that makes *Avoid F* redundant.

result of replacing embedded F-marked elements with variables.⁹ This condition allows to explain data like (16) – a puzzle for theories based on the question test for focus (cf. fn. 5 above).

(16) A: Paula hat eine rote Rose fotografiert.
Paula has a red rose photographed
 Was hat sie davor getan?
What has she before done

B: Sie hat eine [WEISSE]_F Rose fotografiert.
She has a white rose photographed

A's question leads to expect focus marking of the complete VP, but intonational marking plus projection rules produce a narrow focus on *weiße*. Schwarzschild (1995) predicts the indicated F-marking, since the *F-skeleton* of the NP *eine weiße Rose*, for instance, (*a X rose*) is actually entailed by the context.

While non-F-marked constituents have to be given, F-marked constituents need not necessarily be new. This is to account for data like the following, where *ihn* in (17B) is given:

(17) A: Wen hat Peters Mutter gelobt?
Whom has Peter's mother praised?

B: Sie hat [IHN]_F gelobt.
She has him praised

The indicated F-marking follows from the theory: there has to be *some* F-marking, since the meaning of the complete sentence (*Peter's mother praised Peter*) is not entailed by the context. The F-skeleton variant *Peter's mother praised X* is actually entailed by the question (17A), thus the F-marking of *ihn*.

3 An underspecified representation

In computational applications, a compact representation is a prerequisite for any successful treatment of IS. Take the following short sentence with two pitch accents.

(18) Die Direktorin der Firma MÜLLER BEGRÜSST
The director of company Müller welcomes
 ihren Besuch.
her visit/visitors

Even if functional elements are ignored, the rules in (12) produce nine alternative F-markings that have to be checked against the context for givenness. In order to resolve the lexical ambiguity of *Besuch* however (cf. the discussion of (5) above), some partial information about IS suffices.

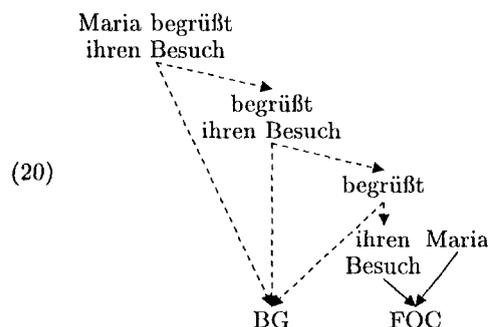
The representation developed here is relatively independent of the underlying semantic theory of focus. Two semantical partitions for focus (FOC) and background (BG) are assumed, each of them a set of semantic conditions. Underspecification is expressed in a graphical way. The interpretation¹⁰

⁹There is some affinity to the f-semantic value in Rooth's (1985) alternative semantics, although the specific details are different.

¹⁰The interpretation is given informally in the following examples.

of each syntactic constituent forms a node in the graph, while the directed edges express accessibility relations.

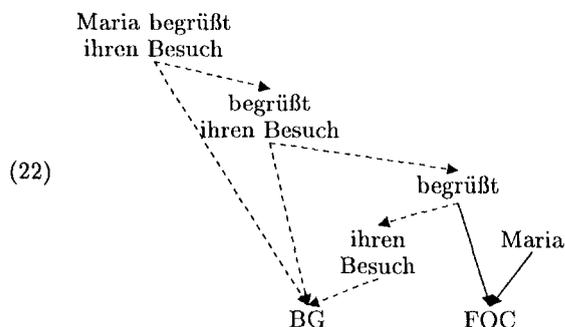
(19) MARIA begrüßt ihren BESUCH.
Maria welcomes her visit/visitors



The solid line arrows signify obligatory inclusion in the respective IS partition, the dashed line arrows defeasible inclusion. The VP can, e.g., either belong to FOC (via a chain of arrows) or to BG. The graph in (20) represents the amount of information that is encoded on sentence level without reference to context. Additional knowledge may introduce further solid arrows. If, e.g., a *begrüßen* event is contextually given (like in a question *Who greets whom?*), the arrow from *begrüßt* to BG will become an obligatory arrow. This arrow again will overrule the dashed arrow from *begrüßt* to *ihren Besuch*. Since the latter was the only path to access the FOC partition, the complete graph will collapse into a fully specified representation of IS. Resolution will be dealt with in more detail in sec. 4.

The principles composing the representation are worked out formally in sec. 5. Here is an intuitive overview: the arrows pointing directly to the FOC and BG partition originate from accenting or non-accenting of the single words, respectively. The additional arrows between the constituent nodes are introduced by the grammatical principles of F-projection – irrespective of the actual prosodic marking. This becomes clear when we regard the following alternative prosodic marking of (19):

(21) MARIA BEGRÜSST ihren Besuch.



Note that here, the dashed arrow from *begrüßt* to *ihren Besuch* is overruled right away, since the accented *begrüßt* is strictly tied to FOC. The

phrase *ihren Besuch* is forced into the BG partition, thus the utterance is correctly predicted to be restricted to contexts where *Besuch* is given. As anticipated above in the discussion of the lexical ambiguity involved, this conclusion can be drawn even if other points remain underspecified.

4 Resolution

An underspecified IS arising from the prosodic marking of a sentence can be resolved by information from the context. As noted above, entailment by context introduces additional solid line arrows. To cover more than just trivial cases, the counterpart of Schwarzschild's (1995) F-skeleton has to be kept track of in the underspecified representation. At first glance, this seems to be incompatible with the idea of underspecification, since the F-skeleton that is checked against the context for entailment requires settlement on what the actual F-marking is. For instance, to check the givenness of the VP in (23), *reads a book about X* has to be entailed, whereas on the basis of the marking in (24), *reads Y* has to be entailed.

(23) Franz liest ein Buch über [SCHIFFE]_F.
Franz reads a book about ships

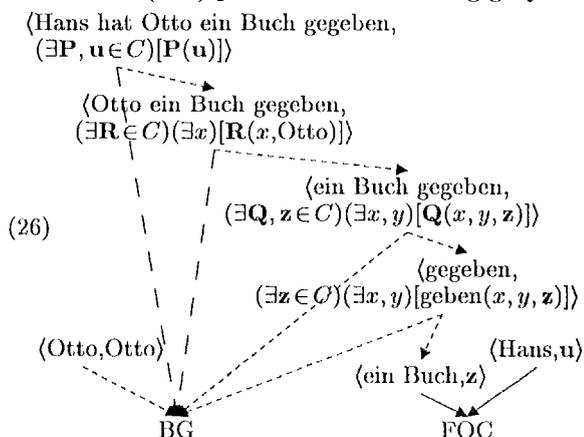
(24) Franz liest [ein Buch über [SCHIFFE.]_F]_F.

The solution lies in the observation that the latter F-skeleton entails the former. So, when a certain constituent (e.g., again the VP in the above examples) is checked for givenness, it suffices to assume F-marking of the maximal potentially F-marked subconstituents (I call this the *maximality assumption*). If it turns out that the actual F-marking is more restricted, this will be detected at a lower level. For example, if (23) is the right solution, this will be discovered even if *reads a book about X* is not checked, since in this case, *a book about X* will be contextually given as well.

To illustrate resolution in the graph representation, take the following example in context:

- (25) a. Anna hat Otto fotografiert.
Anna has Otto photographed
 b. HANS hat Otto ein BUCH gegeben.
Hans has to-Otto a book given

Sentence (25b) produces the following graph:



The nodes are now labelled by pairs. The first element of a pair is the semantic value of the constituent corresponding to the node (here again expressed informally in natural language); the second element is the F-skeleton based on the maximality assumption – the F-skeleton is obtained by (i) replacing the maximal F-marked subconstituents (or sister constituents, in the case of indirect F-marking of the head, following (12b)) by a variable, (ii) existentially binding unfilled arguments, and (iii) existentially quantifying over the variables from the F-skeleton (cf. also fn. 7). The latter quantification is restricted by the set of contextually salient values *C*.

In resolving underspecification from context, it is checked for each node with access to the BG partition¹¹ whether its F-skeleton is entailed by an antecedent in the context. For *gegeben* and the lowest verb projection *ein Buch gegeben*, there is no such antecedent in the context. However, for the VP *Otto ein Buch gegeben* the VP from (25a) *Otto fotografiert* is a suitable antecedent: after existential binding, $(\exists x)$ fotografieren(*x*, Otto) entails $(\exists R \in C)(\exists x)[R(x, Otto)]$. Similarly for the complete sentence: fotografieren(Anna, Otto) entails $(\exists P, u \in C)[P(u)]$. So, solid line arrows replace the two dashed arrows pointing to BG (suggested in (26) by the longer dashes).

The status of the lower verb projections in (26) is still underspecified. If a wider context was considered, antecedents could possibly be found, so it makes sense to end up with such a representation after processing the discourse (25). Once we know that all salient context has been considered, a rule of *Focus Closure* is applied. This strengthens the arrows pointing towards FOC for all unresolved constituents, predicting *ein Buch gegeben* as the focus for the example. In an application, heuristics may trigger Focus Closure earlier, to avoid unnecessary inferences.

The representation proposed here comes close to Schwarzschild's (1995) theory of F-marking in coverage,¹² avoiding the computationally expensive disjunction of alternative analyses in favour of a single graph representation that is underspecified when based on sentence internal infor-

¹¹ Constraining the givenness check to nodes with access to the BG partition makes sure that narrow (contrastive) focus on given entities (like in (17)) is treated correctly. Since an accent on a word introduces just an arrow towards FOC, narrow focus on a word survives the check even in cases where the word is given.

¹² For examples with several ambiguous accents, the modified account collapses some F-markings with minimal differences in interpretation into one. For instance, two of the twelve F-markings for (i) are indistinguishable: (ii) and (iii). A context that would enforce the latter in Schwarzschild's theory produces (ii) in the underspecified account.

mation only. Furthermore, Schwarzschild's pragmatic condition *Avoid F* that selects the analysis with the least F-marking (cf. fn. 8) is no longer needed.

5 HPSG principles

The representation proposed in sec. 3 and 4 can be implemented directly in a sign-based formalism like HPSG. In this section, the central composition principles for German are worked out. A binary branching structure is assumed. Again, the representation will be fairly open as to which particular semantic formalism is chosen. This is reflected in the Semantics Principle I assume, which specifies the semantic value of a phrase as the application of a two-place function *compose* to the semantic values of both daughters. The function can be spelt out in different ways depending on the choice of a semantic theory.¹³ Furthermore, a function *variable* is assumed that maps a semantic object to a new variable of the same type.

The HPSG type *cont* (the value of the $\text{CONT}(\text{ENT})$ feature) has the following four new features: O-SEM ('ordinary semantics') and F-SKEL (F-skeleton) of the type of a semantic object, the set-valued IS-CSTR (IS constraints) and the binary MAX-F (for potential maximal focus).

The phonological information is enriched by a feature PROM (prominence) with values *accented* and *unaccented*.

The following principles specify the IS-CSTR set for a sign, introducing arrows, or links, between

- (i) Der Direktor der SPARKASSE begrüßte Frau
The director of-the savings-bank greeted Frau
MAIER.
Maier
- (ii) Der Direktor der [SPARKASSE]_F begrüßte [Frau
MAIER.]_F
- (iii) [Der Direktor der [SPARKASSE]_F]_F begrüßte
[Frau MAIER.]_F

However, such a context is quite intricate to construct. Here is an example:

- (iv) Der Direktor der Firma ABC saß im Café.
The director of company ABC sat in-the cafe
Der Kellner begrüßte eine Bekannte.
The waiter greeted a friend

Context (iv) does not entail *der Direktor der X begrüßte Y*, so (ii) is ruled out according to Schwarzschild's system. The resolution process proposed in this paper is based on the maximality assumption, and thus checks givenness for the complete sentence only once, with the complete subject F-marked. To rescue the difference between (ii) and (iii), it would have to be enforced that resolution of the subject NP takes place before the resolution of the focus projected from the object. This is a conceivable solution – however, declarative perspicuity would be sacrificed for a very moderate benefit, considering the main point of this paper.

¹³In the example below, simple predicate-argument semantics is used for illustration.

semantic objects. (The IS partitions *foc* and *bg* are here also treated as semantic objects.) The *h.link* relation corresponds to the obligatory (solid line) arrows in the graphs, *s.link* reflects defeasible (dashed line) arrows. In the following, the link relations are expressed verbally. (27) is the counterpart to (11); (28) and (29) reflect clauses (12a) and (12b).

- (27) *Focus Linking Principle*
The O-SEM value of a pitch accented word is *h.linked* to *foc*.
- (28) *Head F-Projection Principle*
The O-SEM value of a phrase is *s.linked* to the O-SEM value of its head daughter.
- (29) *Indirect Head F-Marking Principle*
In a head-complement-structure, where none of the head daughter's arguments have yet been saturated,¹⁴ the O-SEM of the head daughter is *s.linked* to the O-SEM value of the complement daughter.

These three principles establish direct or indirect links towards *foc*. For the constituents that are not obligatorily focus marked, the underspecified representation requires additional defeasible links to *bg*:

- (30) *Background Linking Principle*
The O-SEM value of every sign that is not accented is *s.linked* to *bg*.

The principles just presented compose the representation introduced in sec. 3, with the nodes in the graph corresponding to the O-SEM values. To provide the input for the resolution routine, the representation was enriched in sec. 4 by the F-skeleton. This is kept track of in the F-SKEL feature (assuming independent existential binding of unfilled arguments and free variables).

- (31) *F-Skeleton Instantiation*
The F-SKEL value of a word with $[\text{O-SEM } \alpha]$ is
 - (i) the function *variable*(α), if α is *h.linked* to *foc*;
 - (ii) α , if the word is marked $[\text{MAX-F } -]$;¹⁵
 - (iii) *composition*(α, β) (where β is the F-SKEL value of the word's complement), if the word is marked $[\text{MAX-F } +]$, but α is not *h.linked* to *foc*.
- (32) *F-Skeleton Principle*
The F-SKEL value of a phrase is the function *compose* applied to
 - (i) the F-SKEL value of a daughter with $[\text{MAX-F } -]$; or
 - (ii) the function *variable*(α) of a daughter with $[\text{MAX-F } +]$ and $[\text{O-SEM } \alpha]$;
 for both daughters.

¹⁴This formulation will be subject to changes once there is a clear concept of integration (cf. fn. 6) – (29) applies to integrated constituents.

¹⁵The feature MAX-F is actually redundant. It has been introduced for convenience, signifying that a chain of links to *foc* exists – a condition that could be checked directly in the graph. But it does not hurt much to carry the feature along in the principles (27)–(29).

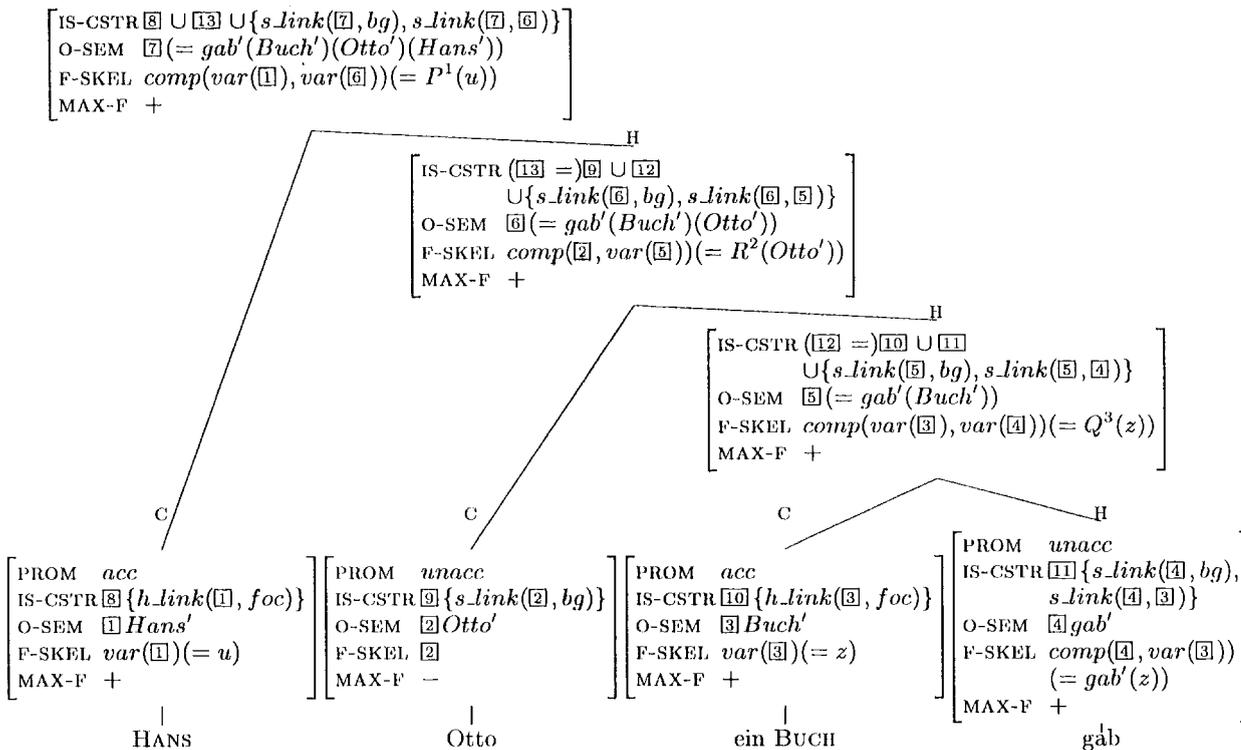


Figure 1: Sample HPSG analysis

A sample analysis for (33), a slight simplification of (25b), is given in fig 1. The graph produced by the linking constraints is the one in (26).

(33) (Ich weiß, daß) HANS Otto ein BUCH gab.
I know that Hans to-Otto a book gave

Let us briefly see how the principles interact to produce the phrase *ein Buch gab* (For simplicity, the NP is treated as if it was a word). Since *ein Buch* bears a pitch accent, the Focus Linking Principle (27) applies, introducing an obligatory link to FOC ($h.link(\boxed{3}, foc)$). For the unaccented *gab*, the Background Linking Principle applies, giving rise to a defeasible link ($s.link(\boxed{4}, bg)$). At phrase level, the Indirect Head F-Marking Principle (29) and the Head F-Marking Principle apply (introducing $s.link(\boxed{4}, \boxed{3})$ for the head, and $s.link(\boxed{5}, \boxed{4})$ for the phrase, respectively). In addition, (27) applies again: $s.link(\boxed{5}, bg)$.

As for the F-Skeleton, subclause (i) of (31) applies at *ein Buch*, subclause (iii) at *gab*, causing the function *compose* to apply to *gab*'s own semantic value and to its sister's F-SKEL value. The phrase is covered by (32), where both daughters are marked [MAX-F +] and thus fulfil subclause (ii).

6 Conclusion

This paper shows that a fully expressive underspecified representation of IS can be effectively composed by linguistic principles, circumventing the computational problems that the disjunctive analyses of existing theories pose. Also, a resolution routine was presented. The idea is to leave

the representation underspecified in applications, unless resolution is required for a specific reason.

References

- Eberle, Kurt (1996): "Disambiguation by information structure in DRT," COLING-96 (this conference).
- Engdahl, Elisabet, Enric Vallduví (1994): "Information packaging and grammar architecture: A constraint-based approach," in: Engdahl, E. (ed.): *Integrating information structure into constraint-based and categorial approaches* (DYANA-2 Report R.1.3.B). ILLC, Amsterdam. 39-79.
- Höhle, Tilman (1982): "Explication für 'normale Betonung' und 'normale Wortstellung'," in: Abraham, W. (ed.): *Satzglieder im Deutschen*. Tübingen: Gunther Narr Verlag.
- Jacobs, Joachim (1994): "Informational Autonomy." In: Bosch, P. and R. van der Sandt (eds.): *Focus and Natural Language Processing*. Working Paper 6, ILL, IBM Heidelberg.
- Kuhn, Jonas (1996): "Context effects on interpretation and intonation". Ms. Institut für maschinelle Sprachverarbeitung, Universität Stuttgart. (<ftp://ftp.ims.uni-stuttgart.de/pub/papers/kuhn/>)
- Pollard, Carl, Ivan Sag (1994): *Head-Driven Phrase Structure Grammar*. University of Chicago Press.
- Rooth, Mats (1985): "Association with Focus," PhD thesis. University of Massachusetts, Amherst.
- Selkirk, Elisabeth (1995): "Sentence Prosody: Intonation, Stress and Phrasing," In: Goldsmith, J. A. (ed.): *Handbook of Phonological Theory*, Basil Blackwell. London.
- Schwarzschild, Roger (1995): "On Givenness," Handout of talk presented at University of Massachusetts, Amherst.