

Propositional Relative Clauses in German

Manfred Sailer, University of Tübingen, sailer@sfs.uni-tuebingen.de

Introduction

In this study we will discuss a largely unnoticed type of relative clause which we will call *propositional relative clause* (PRC). PRCs manifest a mismatch between syntactic structure and interpretation: syntactically they are normal relative clauses, but in the semantics an NP which is modified by a PRC is interpreted as a proposition, even though the head noun in itself does not have a propositional interpretation. The clearest cases of PRCs will come from data with idiomatic expressions, in particular with bound words. We will show that the apparent mismatch can be resolved if we adopt a system of combinatorial semantics which exploits techniques of semantic underspecification such as *Underspecified DRT* (UDRT, Frank and Reyle (1995)), *Minimal Recursion Semantics* (MRS, Copestake et al. (2003)), *Constraint Language for Lambda Structures* (CLLS, Egg and Erk (2002)), or *Lexical Resource Semantics* (LRS, Richter and Sailer (2003a)).¹

A bona fide example of a German PRC is given in (1). The noun *Duzfuß* is a bound word which may only occur in the expression *mit X auf dem Duzfuß stehen* (*be on informal terms with X*). In (1) the noun heads a relative clause which contains the rest of the expression.²

- (1) um den anderen den Duzfuß ahnen zu lassen, [auf dem man mit den Spitzenkräften steht].
in order to the other the informal.foot suspect to let on which one with the top executives stand
'in order to make the other one suspect that one is on informal terms with the top executives.' (from the IDS corpora)

Examples such as (1) do not occur frequently in texts and many speakers consider them as "strange" or instances of creative language use. The English translation indicates the only possible interpretation: the NP with the relative clause is interpreted as a complement clause.

We will assume the following tentative definition of a PRC constellation: A sentence with an NP which is modified by a relative clause such that (i) the sentence is ungrammatical if the relative clause is removed, and (ii) the sentence is synonymous to a sentence in which the NP with the relative clause is replaced with a complement clause with the same lexical material. This is illustrated in (2): since *Duzfuß* is a bound word, it cannot occur freely (a). We can replace the NP with a complement clause to obtain a synonymous sentence (b).

- (2) a. * um den anderen den Duzfuß ahnen zu lassen.
b. um den anderen ahnen zu lassen, dass man mit den Spitzenkräften auf Duzfuß steht.
in order to the other suspect to let that one with the top executives on informal.foot stands

Data

To our knowledge PRCs have not been systematically studied so far. In the presentation we will provide a number of examples of PRCs from the WWW and from publically available corpora of German.

PRCs are not restricted to idioms, as illustrated in (3a). The sentence is ungrammatical without the relative clause (see (b)), and synonymous to a sentence with a complement clause instead of the NP (see (c)). Sentence (d) shows that the verb *bedauern* is in principle compatible with an NP complement under the intended reading, if this NP can denote a proposition.

- (3) a. Hans bedauerte das Vermögen, das er beim Spiel verloren hatte. b. * Hans bedauerte das Vermögen.
Hans regretted the fortune that he at the game lost had
'Hans regretted that he had lost the fortune at the game.'
c. Hans bedauerte, dass er beim Spiel das Vermögen verloren hatte. d. Hans bedauerte den Verlust (des Vermögens)
Hans regretted that he had lost the fortune at the game. Hans regretted the losing of the fortune

We predict ambiguity in cases such as (4), although the readings are admittedly hard to distinguish.

- (4) Hans bedauerte den Verlust, den er erlitten hatte. 'H. regretted the loss he had suffered.'
Hans regretted the loss that he suffered had 'H. regretted that he had suffered the loss.'

¹For reasons of space we only quote references which show the integration of the respective system with HPSG.

²The context for (1) is: Eine beliebte Variante ist das Bruderschafts-Dropping, bei dem man geschickt Vornamen wie Thomas, Viktor, Ioan, Otti etc. einflicht, ... ('A popular variant is the "friendship dropping" through which one drops first names such as Thomas, Viktor, ...')

Examples with bound words are easy to extract from corpora and provide particularly clear cases of PRCs since they are usually more restricted in their meaning potential than other words. There is a class of bound words in German which do not allow for modification or pronominalization. Nonetheless, the expressions in which they occur can be semantically decomposed. *Duzfuß* is of this type. A plausible meaning assignment to *Duzfuß* would be *informal terms*. In (5) we show that modification is not possible.

- (5) Der Parteivorsitzende steht mit dem Kanzler auf (*gutem/ *intimem) Duzfuß.
 the party leader stand with the chancellor on good/ intimate informal.foot

For unquestionably non-decomposable idioms such as *Maulaffen feilhalten* (*stand gaping*) PRCs are excluded.

- (6) * Mich erbosten die Maulaffen, die die Passanten feilhielten.
 Me made angry the mouth.monkeys which the passers-by kept for sale

We conclude that with free words a relative clause can have a PRC reading under certain conditions (see (3a)). For undecomposable idioms relative clauses are excluded altogether, (see (6)). For a number of bound words a PRC reading is the only way to interpret a relative clause, (see (1)).³

PRCs are reminiscent of the reinterpretation phenomena discussed in Egg (2002).

- (7) a. Amélie played the sonata for ten days. b. beautiful dancer

In (7a) an iterative operator must be introduced to make the VP “played the sonata” compatible with a durative adverbial. The NP in (b) has a reading in which the adverbial does not take scope over the entire noun but only over part of it (“person who dances beautifully”). In PRCs on the other hand the head noun is interpreted inside the relative clause and the entire NP has the meaning of a proposition. This means that in contrast to (a) no additional semantic material is inserted, and in contrast to (b) the syntactic modifier takes scope over the entire NP.

Analysis

PRCs do not differ structurally from other relative clauses. Within HPSG most publications on German relative clauses assume a variant of the analysis in Pollard and Sag (1994) rather than adopting the constructional approach in Sag (1997). For example Kiss (2004) and Holler (2003) postulate a functional head for relative clauses.

In German relative clauses often appear in extraposed position. This is also true for PRCs (see (1)). Kiss (2004) demonstrates that relative clauses are not “moved” into such an extraposed position — in contrast to complement clauses. In order to establish the semantic and agreement relation between an extraposed relative clause and the head noun, Kiss argues in favor of a system of combinatorial semantics which uses techniques of underspecified semantics, MRS in Kiss’ case, although other systems (UDRT, CLLS and LRS) are just as applicable. In Kiss’ analysis a relative clause can modify an index which is contributed by a noun deeper in the syntactic structure. In the present paper we will assume an LRS version of Kiss’ theory.

We will illustrate our analysis with the following two sentences, in which (8a) contains a normal relative clause, and (8b) contains a PRC.⁴ In (c) we will outline the syntactic structure of the sentences. In addition to the glossing we will also indicate the logical form. We will use a semantic representation language which allows us to differentiate between semantic objects of different type (entities, eventualities, propositions). We will ignore tense.

- (8) a. dass Peter das Photo betrachtete, das Maria knipste.
 that Peter the picture looked.at which Maria took
 $\lambda w.\exists e(\text{look-at}(w, e, p, \iota x(\text{picture}(w, x) \wedge \exists e'(\text{take}(w, e', m, x))))))$
- b. dass die Zahlen den Garaus belegen, den die Globalisierung dem Kleinbetrieb machte.
 that the numbers the ??? prove which the globalization to the small business made
 $\lambda w.\exists e(\text{prove}(w, e, Z, \lambda w.\exists e'(\text{make}(w, e', G, K, \iota x(\text{garaus}(w, x))))))$
- c. [COMP [S [S Subject [VP [NP Det Noun] Verb]] [RelP XP [Relativizer S]]]]]

³In Nunberg et al. (1994) relative clauses were used to motivate the meaningfulness of parts of idioms. Their examples are not instances of PRCs. Neither are the cases discussed as *Amount Relatives* in Carlson (1977).

⁴The expression *X den Garaus machen* (*kill/destroy X*) contains the bound word *Garaus* to which we can assign the meaning of *end* or *death*. An authentic PRC example is: *Einzig Vera Kutters ... Negativfoto der Wiener Secession bleibt als Hinweis auf den Garaus, den die Nazis der in ihren Augen "entarteten Kunst" machten.* ‘Only V.K.’s negative picture of the Vienna Secession remains as an indication of the fact that the Nazis destroyed what to their eyes was “degenerated art”.’ (<http://www.taz.de/pt/2001/11/30/a0123.nf/text>)

In the logical forms (lf) we use world indices (w) and events (e, e'). We also adopt the ι -operator for definites.⁵ We abbreviate the lfs of *die Zahlen*, *die Globalisierung* and *dem Kleinbetrieb* as **Z**, **G** and **K**. The challenge for the analysis of the PRC lies in the fact that the semantics of the NP ($\iota x(\text{garaus}(w, x))$) appear entirely inside the lf of the relative clause. We will demonstrate that this can be achieved using a semantic formalism such as LRS.

We will briefly introduce the necessary aspects of LRS. In LRS (Richter and Sailer, 2003a) the semantic representation of a sign consists of the following components: (i) The PARTS list of a sign contains all subexpressions of the overall logical form of a sentence which are contributed by this sign. (ii) The INTERNAL-CONTENT (INC) value is the main semantic constant contributed by a lexical element.⁶ (iii) The EXTERNAL-CONTENT (EXC) of an utterance is its logical form. It must consist of the elements of the utterance's PARTS list. (iv) The VAR value is needed for linking. Usually, the VAR value of a syntactic complement will appear as a semantic argument in the INC value of the syntactic head. (v) The INDEX value of Pollard and Sag (1994) is retained.

The SEMANTICS PRINCIPLE (SP) specifies how the semantic contributions of the daughters combine to form the semantics of the mother: Firstly the values of EXC, INC and VAR are shared between the head and the mother. Secondly the PARTS list of the mother contains exactly the elements of the daughters' PARTS lists. Thirdly specific subexpression requirements will be imposed depending on the semantic type of the daughters and on the syntactic construction. In this paper we need two such requirements, one for the combination of a noun with a definite article (DEF-N-RULE) and one for head-adjunct structures with an intersective modifier (INTMOD-RULE).

- (9) a. The DEF-N-RULE: When a noun N combines with the definite article D , then N 's INC value is a subexpression of D 's INC value and N 's VAR and EXC values are identical with D 's INC value,
 b. The INTMOD-RULE: When the nonhead is an intersective modifier M to a noun N , M 's EXC value is of the form $\alpha \wedge \beta$, N 's INC value is a subexpression of α and M 's EXC is a subexpression of N 's EXC.

In this paper we will use a shorthand notation for PARTS lists. For illustration consider (10) which is the LRS specification for the definite determiner and for nouns *Photo* and *Garaus*.

- (10) a. das/den: $\iota x(\alpha)$ b. Photo: $\text{picture}(w, x)$ c. Garaus: $\text{garaus}(w, x)$

The PARTS list of the noun *Photo* consists of the semantic constant picture , the two variables w and x , and the expression $\text{picture}(w, x)$. The PARTS list of the definite determiner contains the variable x and the ι -operator which binds x . However, it does not contain any specification of the scope of ιx . This is indicated by the use of lower case Greek letters. Thus in our shorthand notation, all subexpressions of the given formula are elements of the PARTS list, except for the lower case Greek letters.

The PARTS lists of the verbs *betrachten* and *belegen* are more complicated. In LRS a finite verb contributes a.o. a lambda abstract over the world index ($\lambda w.\beta$), and an existentially quantified eventuality variable ($\exists e(\gamma)$). We know that $\exists e(\gamma)$ is a subexpression of $\lambda w.\beta$. This is expressed in Richter and Sailer (2003a) with a subexpression relation (" \triangleleft "), written as $\exists e(\gamma) \triangleleft \beta$. In our shorthand notation we will symbolize this subexpression requirement more compactly as $\lambda w.\beta[\exists e(\gamma)]$. In general the notation $\alpha[\dots]$ means that the expression " \dots " must be a subexpression of α .

- (11) a. betrachten: $\lambda w.\beta[\exists e(\gamma[\text{look-at}(w, e, \delta[X], \epsilon[Y]])])]$ b. belegen: $\lambda w.\beta[\exists e(\gamma[\text{prove}(w, e, \delta[X], \zeta[Y]])])]$

The verb *betrachten* in (8a) introduces the constant look-at which requires a world index, an eventuality and two arguments of type e . Following Pollard and Sag (1994) argument identification is achieved lexically. A verb specifies how the VAR values of its complements relate to the INC value of the verb. In (11) we write X and Y for the VAR values of the subject and the direct object respectively. The specification indicates that the subject's VAR value must be a subexpression of the third semantic argument of the verbs ($\delta[X]$). In the case of *belegen* in (11b) the last semantic argument (ζ) is a proposition, i.e., of type $\langle s, t \rangle$.

We will derive the readings of the sentences in (8) step by step according to the syntactic structure in (8c). The INC values of the nouns are $\text{picture}(w, x)$ and $\text{garaus}(w, x)$ respectively. The INC value of the determiner is $\iota x(\alpha)$. The PARTS list of the NP nodes are outlined in (12). They show the effect of the SP, including the DEF-N-RULE. They contain all the expressions of the daughters' PARTS lists and they encode the conditions of the DEF-N-RULE, i.e., the noun's INC value is a subexpression of the determiner's INC value.

⁵For each variable x of type τ and for each ϕ of type t , $\iota x(\phi)$ is an expression whose denotation is an individual a of type τ such that $\llbracket \lambda x.\phi \rrbracket(a) = 1$ if there is exactly one such individual, otherwise the denotation is undefined.

⁶Richter and Sailer (2003a) distinguish between the INC and a MAIN value. For our purpose we can ignore the MAIN value.

- (12) a. das Photo: $\iota x(\alpha[\text{photo}(w, x)])$ b. den Garaus: $\iota x(\alpha[\text{garaus}(w, x)])$

For the NPs the VAR and the EXC values are identical to the determiner's INC, i.e., they express the ι -operator. Note that given the semantics of this operator the VAR values are of type e . Since *Garaus* does not allow for modification (see (5)) it must follow from its lexical specification that $\alpha = \text{garaus}(w, x)$. Thus the logical form of *den Garaus* must be $\iota x(\text{garaus}(w, x))$.

When these NPs occur as complements of the verbs in (11), their VAR values will appear in the indicated argument position of the verb (Y). The same holds for the subject (X). Since all complements in our examples are definite NPs the SP does not introduce new subexpression requirements, and the logical form of the saturated verbal projection is identical to that of the verb. This If can be described as in (13).

- (13) a. Peter das Photo betrachtet: $\lambda w.\beta[\exists e(\gamma[\text{look-at}(w, e, \delta[p], \iota x(\alpha[\text{picture}(w, x)]))])]$
 b. die Zahlen den Garaus belegen: $\lambda w.\beta[\exists e(\gamma[\text{prove}(w, e, \delta[Z], \zeta[\iota x(\text{garaus}(w, x)]))])]$

Next we must consider the relative clause (RelP). The INDEX value of the RelP is identical with that of the noun which it modifies. The RelP's VAR value is the variable bound in the noun's EXC value, i.e., x .⁷

According to the analysis of Pollard and Sag (1994) a constituent XP which contains a relative pronoun is extracted from the S-part of the RelP. This XP and its trace share the VAR value. This leads to the following description of the semantics of the S-part of the RelP.

- (14) a. (8a): $\lambda w.\eta[\exists e'(\theta[\text{take}(w, e', m, \kappa[Z])])]$ b. (8b): $\lambda w.\eta[\exists e'(\theta[\text{make}(w, e', \mathbf{G}, \mathbf{K}, \kappa[Z])])]$

The empty relativizer has two complements: the clause S and the XP which contains the relative pronoun. We will first discuss the semantics of relative clauses such as in (8a) and then address the semantic contribution of a PRC relativizer. We assume that the empty relativizer in (8a) contributes a logical conjunction, indicated in (15a).

- (15) a. Relativizer: $\mu[A, C] \wedge \nu[B]$ b. RelP for (8a): $\lambda w.\eta[\mu[x] \wedge \nu[\exists e'(\theta[\text{take}(w, e, m, \kappa[x])])]]$

The VAR value of the XP (A) appears as part of the first conjunct and that of the clause (B) in the second conjunct. The conjunction is the relativizer's EXC value. Its VAR value is identical with that of the relative pronoun, C . For our simple example the XP itself is the relative pronoun. the logical form of the relative clause in (8a) is given in (15b).

When the RelP combines with the rest of the sentence, the INTMOD-RULE requires the noun's INC value be a subexpression of μ . The coordination must be a subexpression of $\iota x(\alpha)$, the noun's EXC value. This allows us to specify the VAR values of the NP as $\iota x(\text{picture}(w, x) \wedge \exists e'(\text{take}(w, e', m, x)))$.

The relative clause also contributes an abstraction over the world index ($\lambda w.\eta$). $\lambda w.\eta$ must be outside the scope of the ι -operator. In sentence (8a) $\lambda w.\eta$ is identical with $\lambda w.\beta$.⁸ This leads to the following constellation:

- (16) $\lambda w.\beta[\exists e(\gamma[\text{look-at}(w, e, \delta[p], \epsilon[\iota x(\text{picture}(w, x) \wedge \exists e'(\text{take}(w, e', m, x))])])])]$

Since no more semantic material will be introduced into the sentence, the EXC value of the sentence must contain exactly the elements of its PARTS list. For our shorthand this implies that we can simply erase the Greek letters. The result is exactly the logical form given in (8a). Note that there is no other possibility of including $\lambda w.\eta$ in the logical form: $\epsilon[\lambda w.\eta[\iota x(\dots)]]$ is not well-formed since ϵ is of type e , $\lambda w.\eta$ of type $\langle s, t \rangle$. Similarly $\gamma[\lambda w.\eta]$ leads to a type conflict because γ must be of type t .

An analogous logical form is excluded for sentence (8b) because the logical form of *den Garaus* does not allow for the integration of semantic material into the scope of ιx . Instead, we must assume a different relativizer for PRCs. This relativizer does not contribute new semantic material into the logical form. Therefore the logical form of the RelP is identical to that of its S-part, as given in (14b). This reinforces our intuition that a PRC behaves semantically like a complement clause.

If we combine this RelP with the sentence in (13b) the INTMOD-RULE of the SP does not apply — nor does any other construction-specific constraint of the SP. Nonetheless, there is only one way to combine the logical forms of the matrix sentence and that of the PRC into a well-formed expression:

⁷In contrast to Richter and Sailer (2003a) we do not assume that VAR is defined on *index*.

⁸The possibility of identity of semantic operators has been explored in Richter and Sailer (2003a) for the analysis of Negative Concord.

(17) $\lambda w.\beta[\exists e(\gamma[\text{prove}(w, e, Z, \zeta[\lambda w.\eta[\exists e'(\theta[\text{make}(w, e', G, K, \kappa[\iota x(\text{garaus}(w, x))])])])])])])]$

If we eliminate the Greek letters in (17), it becomes clear that we have derived the PRC reading. In this reading $\iota x(\text{garaus}(w, x))$ appears as the semantic argument of **make**. This satisfies the condition that x , which is also the VAR value of the relative pronoun, must be a subexpression of this semantic argument. In fact the sentence does not contain any other slot of type e which would allow for the binding of this occurrence of x .

Another difference to the logical form of (8a) lies in the fact that embedded λ -abstract over the world index $(\lambda w.\eta)$ appears as the semantic argument of the matrix verb. This must be the case since **prove** requires an argument of type $\langle s, t \rangle$ and there is no other subexpression in the logical form which could appear in this position.

Conclusion

Data with bound words played an important role in our study of empirical motivation of PRCs. Nonetheless, as demonstrated by (3) PRC readings do not depend on bound words. Therefore we have not tied our account of PRCs to a particular analysis of expressions with bound words. As far as we can see our approach is compatible with any analysis of these expressions within HPSG (such as Riehemann (2001) or Richter and Sailer (2003b)).

The account presented seems to predict a general ambiguity of relative clauses. Note however that the availability of a PRC reading is correctly restricted to certain matrix predicates. For example we could not have derived a PRC reading for sentence (8a). To explain the often dubious grammaticality status of PRC readings we may assume that speakers tend to equate the VAR value of a complement with a semantic argument slot of a head, rather than allowing for the intervention of other material. Alternatively we might speculate that the use of the PRC relativizer is highly marked because it is semantically empty.

Recently the study of the semantics of restrictive relative clauses has led to a re-consideration of the “raising analysis” of Carlson (1977) (see Alexiadou et al. (2000); Bhatt (2002)), in which the head noun is moved out of the relative clause. We are optimistic that our analysis may naturally be extended to incorporate the data discussed there without imposing on ourselves an otherwise unmotivated head-internal syntactic structure. On the other hand PRC readings will contribute new insights to the ongoing discussion.

References

- Alexiadou, Artemis, Law, Paul, Meinunger, André, and Wilder, Chris (Eds.) (2000). *The Syntax of Relative Clauses*. Amsterdam and Philadelphia: John Benjamins.
- Bhatt, Rajesh (2002). The Raising Analysis of Relative Clauses: Evidence from Adjectival Modification. *Natural Language Semantics* 10, 43–90.
- Carlson, Greg N. (1977). Amount Relatives. *Language* 53, 520–542.
- Copestake, Ann, Flickinger, Dan, Pollard, Carl, and Sag, Ivan A. (2003, November). Minimal Recursion Semantics: An Introduction. Manuscript.
- Egg, Markus (2002). Semantic Construction for Reinterpretation Phenomena. *Linguistics* 40, 579–609.
- Egg, Markus and Erk, Katrin (2002). A Compositional Account of VP Ellipsis. In F. Van Eynde, L. Hellan, and D. Beermann (Eds.), *Proceedings of the 8th International Conference on HPSG*, pp. 162–179. Stanford: CSLI Publications.
- Frank, Anette and Reyle, Uwe (1995). Principle Based Semantics for HPSG. In *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, pp. 9–16. Association for Computational Linguistics.
- Holler, Anke (2003). An HPSG Analysis of the Non-Integrated Wh-Relative Clauses in German. In S. Müller (Ed.), *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, pp. 163–180. CSLI Publications.
- Kiss, Tibor (2004). Semantic Constraints on Relative Clause Extraposition. *NLLT*. To appear.
- Nunberg, Geoffrey, Sag, Ivan A., and Wasow, Thomas (1994). Idioms. *Language* 70, 491–538.
- Pollard, Carl and Sag, Ivan A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago/London: University of Chicago Press.
- Richter, Frank and Sailer, Manfred (2003a). Basic Concepts of Lexical Resource Semantics. Lecture notes for the course on *Constraint-based Combinatorial Semantics*, ESSLLI 2003, Vienna.
- Richter, Frank and Sailer, Manfred (2003b). Cranberry Words in Formal Grammar. In C. Beyssade, O. Bonami, P. C. Hofherr, and F. Corblin (Eds.), *Empirical Issues in Formal Syntax and Semantics*, Volume 4, pp. 155–171. Paris: Sorbonne.
- Riehemann, Susanne Z. (2001). *A Constructional Approach to Idioms and Word Formation*. Ph. D. thesis, Stanford University.
- Sag, Ivan A. (1997). English Relative Clause Constructions. *Journal of Linguistics* 33, 431–483.