In: Proceedings of the Third Conference on Formal Grammar. Aix-en-Provence, France. 1997.

Using lexical principles in HPSG to generalize over valence properties*

Walt Detmar Meurers
Sonderforschungsbereich 340 B4/B8, Universität Tübingen
Kleine Wilhelmstr. 113, 72074 Tübingen, Germany
dm@sfs.nphil.uni-tuebingen.de

1 Motivation

One of the key aspects of HPSG theories on Germanic and Romance languages is the lexical specification of verbs selecting a verbal complement. Hinrichs and Nakazawa (1989) showed how the idea of functional composition from categorial grammar can be expressed as part of the specification of a lexical entry, and versions of this argument raising specification have since been used in most work on Germanic or Romance languages. In these proposals, the discussion has focussed on the specification of exemplary lexical entries of verbs which are intended to be representative for certain lexical classes. These classes are not formally represented, but each verb in such an intuitively understood class is intended to bear the illustrated specifications.

A number of questions arise from this situation:

- How can the implicit notion of a lexical class actually be represented in the theory?
- How can generalizations (such as the argument raising distribution of valence) over all words in one lexical class be expressed?

Finally, in many current HPSG theories building on Pollard and Sag (1994, ch. 9), the argument structure is represented in every word in addition to the valence specification. Apart from serving as data structure for the binding theory, the precise role of the argument structure representation has yet to be

*The research presented here was carried out in the projects B4 and B8 of the Sonderforschungsbereich 340, financed by the Deutsche Forschungsgemeinschaft. I would like to thank Paola Monachesi, Kordula De Kuthy, Thilo Götz, Tilman Höhle, and Frank Richter for discussions on the topic presented here, and the three anonymous reviewers for interesting comments.

explored. One of the issues which needs to be investigated is the relationship between the elements in the ARG-ST attribute of a lexical entry and the values of the three valence attributes:

- How can valence properties, including special cases such as argument raising, be obtained as result of a principled mapping between ARG-ST and the valence attributes?

In this paper we provide one possible set of answers to these three questions by showing that the mechanisms for representing lexical classes and expressing generalizations over the elements of a class are already available in the HPSG architecture of Pollard and Sag (1994), and how they can be used to express lexical generalizations over the valence specification of different classes of Italian verbs. To situate the proposal, we start out with a brief overview of lexical generalizations and the methods used to capture them in linguistic theory.

2 Background: Expressing lexical generalizations

The need to organize the specification of the lexicon in a non-redundant fashion has long been recognized. The proposals range from the use of templates, macros or frames in implementation oriented proposals over work on special lexical representation formalisms such as DATR to extensions of feature-based grammars using defaults. Briscoe, Copestake, and De Paiva (1993) provide an excellent collection of papers reflecting the breadth of work on this topic. Here, we have to restrict ourselves to the two basic kinds of lexical generalizations and the methods used to express them in HPSG theories.

2.1 Two kinds of lexical generalizations

Flickinger (1987) distinguishes between two kinds of regularities within the lexicon: one is sometimes referred to as *vertical*, the other as *horizontal*.

Vertical generalizations express that certain properties are common to the words of a certain class or subclass. For example, in (Pollard and Sag, 1994) finite verbs are supposed to assign nominative case to their subject valence.

Horizontal generalizations, on the other hand, express a “systematic relationship holding between two word classes, or more precisely, between the members of one class and the members of another class” (Flickinger, 1987, p. 105). A common example for such a horizontal regularity is the relationship between active verbs and their passive counterparts.

2.1.1 Horizontal generalizations

Horizontal generalizations can be captured in the linguistic theory with the help of lexical (redundancy) rules, which have been used for this purpose at least since Jackendoff (1975). Two ways to formalize lexical rules in the HPSG architecture of Pollard and Sag (1994) were proposed in Calcagno (1995) and Meurers (1995). To allow for a clearer explanation of the relevant issue we focus on the latter formalization.

A lexical rule "$D \mapsto E$" in HPSG under the formalization of Meurers (1995) expresses an implicational statement of the form "If there is a grammatical word described by *D'*, then the corresponding grammatical word described by *E'* is also grammatical". The descriptions *D'* and *E'* are derived from D and E to reflect the fact that only the properties in which the output is intended to differ from the input is explicitly provided in the lexical rule specification.

Employing a lexical rule in a linguistic theory thus makes theoretical predictions, which can be seen from the fact that such predictions can be falsified if one observes that in a language a word described by a *D'* is grammatical but not its corresponding counterpart.

2.1.2 Vertical generalizations

Vertical generalizations are often expressed by some mechanism which allows the abbreviation of a lexical specification (macros, templates, ...). Once an abbreviation is defined, it can be used in the specification of each lexical entry in a class. By defining an abbreviation in such a way that it refers to an already defined one, it is possible to organize abbreviations for lexical specification in a hierarchical fashion, which has sometimes been referred to as "hierarchical lexicon".

Since the method allows for a compact specification of the lexicon, it is widely used for grammar implementation. From a theoretical perspective, though, we think it is not satisfactory. Abbreviations in the HPSG architecture of Pollard and Sag (1994)[1] have no formal status different from that of the set of descriptions which they abbreviate, i.e., macros are not really part of the theory, they only express it compactly. Using abbreviations, the notion of a lexical class that was at the basis of the original idea of vertical generalizations is neither formally reflected in the theory nor in the domain of linguistic objects denoted by the theory. Whether an abbreviation is actually used in the specification of lexical entries and where this is done is decided by the grammar writer on the basis of personal preference or some

[1]As formal basis for the HPSG architecture of Pollard and Sag (1994), we assume the formal setup provided by King (1989, 1994). We are not aware of other formal setups for HPSG, though, which contradict the claim about abbreviations made here.

kind of meta regime which he/she follows in writing the grammar but which is not present in the grammar itself.

That no proper generalization is expressed can be seen from the fact that no predictions which can be proven incorrect follow from such an encoding. If some word is not described by the abbreviation which is intended to capture the properties of that class, nothing in the grammar stops us from not using the problematic abbreviation in the specification of the lexical entry in order to license the problematic word.

Finally, due to the theory-external role of abbreviations, a possibly present hierarchical structure of the abbreviations is not reflected in the theory either. In particular, the hierarchical structure of abbreviations stands in no formal relationship to the hierarchical organization of types defined in the type hierarchy of an HPSG grammar.

2.2 Expressing vertical generalizations in the theory

We believe that vertical lexical generalizations should be expressed such that they have a predictive power just like the lexical rule mechanism mentioned above. A mechanism expressing vertical lexical generalization should capture statements of the form "If a word is described by D, then it is also described by E."

A mechanism to express such lexical generalizations is readily available in the HPSG architecture. The implicational constraints used to express the grammatical principles make statements of the form that all objects described by the antecedent must satisfy the consequent in order to be grammatical. This expresses a generalization over all objects described by the antecedent that can be falsified if one finds grammatical linguistic objects which satisfy the antecedent but violate the consequent.

The remaining question is which kind of antecedents are to be used as antecedents of the principles. If the lexical class to which a generalization is to be attached cannot directly be picked out on the basis of a property common to all and only the elements of that class, we need to make the lexical class explicit in the theory in some other way. Pollard and Sag (1987, ch. 8.1) introduced a 'hierarchy of lexical types' for a similar purpose. However, while there are clear intuitions behind these 'lexical types', their exact meaning and how they fit into the HPSG architecture has never been clarified. A straightforward solution is to introduce the relevant class distinctions as ordinary types. Since lexical classes often are subclasses of categorial distinctions, it seems reasonable to use subtypes of *head* for this purpose.[2]

[2]Riehemann (1993, p. 56) mentions an alternative to her 'lexical types' along the same formal lines, but she does not pursue this possibility.

Summing up, we propose to formalize classes of words and the vertical generalizations holding of the elements of such classes as

a) ordinary types, namely subtypes of *head*, to distinguish the classes (unless they are already distinguishable on the basis of independent properties)

b) ordinary implicational principles having descriptions of words as antecedents to express the properties of the classes. We will refer to these theory statements as *lexical principles.*

3 Generalizing over valence properties of verbs

Having addressed the two formal questions raised in the introduction, we can now make use of the formal setup obtained in order to express the generalizations over valence specifications which were the topic of the third question.

We start with a simple lexical principle which expresses that verbs only select via the COMPS and SUBJ attributes. It is shown in figure 1.

$$\begin{bmatrix} \textit{word} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \ \ \textit{verb} \end{bmatrix} \rightarrow \begin{bmatrix} \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{VAL} \mid \text{SPR} \ \ \langle\rangle \end{bmatrix}$$

Figure 1: Lexical principle enforcing an empty SPR value for verbs

The specification of the other two valence attributes SUBJ and COMPS is somewhat more interesting. To avoid the specification of redundant information and express generalizations over the valence specification of different verb classes, we propose to specify only the ARG-ST list in the lexical entries of verbs provided in the lexicon. The values of the valence attributes SUBJ and COMPS are then obtained as a result of generalizations over classes of words. To formulate the relevant lexical principles, we need to distinguish between the class of verbs enforcing argument raising and the class of verbs subcategorizing only their own arguments.

3.1 Argument raising and the specification of Italian verbs

In her detailed work on the grammar of clitics and the Italian verbal complex, Monachesi (1995) argues that data concerning clitic climbing, long NP-movement, tough constructions and auxiliary selection shows that the so-called *restructuring verbs* (modal, aspectual, and motion verbs) can select either a VP complement or a lexical verbal complement from which all arguments are raised. The *auxiliaries* are more restricted in that they only allow the second option.

In her analysis, Monachesi proposes to specify the lexical entries of restructuring verbs as selecting a VP complement. The argument raising variant of these verbs is then derived by a lexical rule, the Argument Composition Lexical Rule shown in figure 2.

$$\begin{bmatrix} \textit{restr-verb} & \\ \text{SUBJ} & \langle \text{NP}\rangle \\ \text{COMPS} & \left\langle \text{VP}\begin{bmatrix}\text{SUBJ} & \langle\text{NP}\rangle \\ \text{COMPS} & \langle\rangle\end{bmatrix}\right\rangle \end{bmatrix} \mapsto \begin{bmatrix} \text{SUBJ} & \langle\text{NP}\rangle \\ \text{COMPS} & \boxed{1} \oplus \left\langle \text{V}\begin{bmatrix}\text{CLTS} & \{\} \\ \text{SUBJ} & \langle\text{NP}\rangle \\ \text{COMPS} & \boxed{1}\end{bmatrix}\right\rangle \\ \text{ARG-ST} & \langle \text{NP, V}\rangle \oplus \boxed{1}\end{bmatrix}$$

Figure 2: The Argument Composition Lexical Rule (ACLR) of Monachesi (1995, p. 175)

The main task of the ACLR is to introduce an argument raising specification in place of the VP requirement for restructuring verbs. Which verbs belong to the class of restructuring verbs cannot in general be deduced from independent linguistic properties of verbs. Monachesi therefore resorts to marking the input of the lexical rule as *restr-verb*, without making explicit how this is to be formally interpreted.

We propose to rephrase the analysis of Monachesi (1995) such that the lexical rule is no longer needed and the classification assumed is made explicit in the theory. For this purpose, we introduce two subtypes of the head-type *verb* as shown in figure 3.

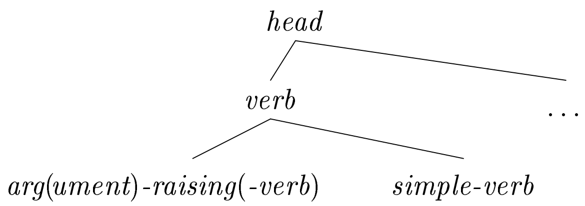

Figure 3: Head subtypes to reflect the lexical subclasses of verbs in the theory

Now that the different lexical classes of verbs are distinguishable in the theory, we can formulate the lexical principles expressing how the argument structure specified in the ARG-ST attribute of the lexical entries of verbs determines the value of the valence attributes SUBJ and COMPS.

The direct mapping for simple verbs is achieved by the lexical principle shown in figure 4.

$$\begin{bmatrix} \textit{word} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \;\; \textit{simple-verb} \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \boxed{2}\ \text{list}\left(\begin{bmatrix} \text{LOC} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS} \;\; \langle\rangle \end{bmatrix}\right) \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1} \mid \boxed{2} \rangle \end{bmatrix}$$

Figure 4: Lexical principle specifying the valence of simple verbs

The first element of the argument structure is assigned to be the subject valence, and the rest of the arguments are identified with the complement requirement. A relation[3] ensures that the complement requirements of the subcategorized complements have been saturated.

The more interesting case is the lexical principle shown in figure 5 which introduces the argument raising valence specification for the appropriate lexical class.[4]

$$\begin{bmatrix} \textit{word} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \;\; \textit{arg-raising} \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \boxed{2} \oplus \left\langle \boxed{3} \begin{bmatrix} \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{VAL} \mid \text{COMPS} & \boxed{2} \\ \text{CLTS} & \{\} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}, \boxed{3} \rangle \end{bmatrix}$$

Figure 5: Lexical principle specifying the valence of argument raising verbs

Note that we assume argument raising to take place only with respect to the valence attributes, not on ARG-ST. The intuition behind this is that the argument structure should be a direct syntactic reflection of the semantic roles of a predicate. Recently Abeillé, Godard, and Sag (1997) have proposed to employ two kinds of argument raising which include raising from/to the argument structure. An exploration of the consequences of the different proposals has to be left to future work.

Once the lexical principles shown in the figures above are included in the theory, the specification of the relevant lexical entries is very simple. The distinction Monachesi makes between auxiliaries, restructuring verbs, and

[3] An extension of the formal HPSG setup provided by King (1989, 1994) which includes relations is defined in (Richter, 1997).

[4] Following the ACLR of Monachesi (1995) we here only deal with ordinary control cases. Raising and ACI verbs can be dealt with in the same style, but require more complicated lexical principles.

other verbs is the result of the specification of one type in a lexical entry, namely that of the appropriate head subtype. The lexical entries of the auxiliaries are specified to have head type *argument-raising-verb*. This results in a mapping from ARG-ST to the valence attributes introducing argument raising. Restructuring verbs, on the other hand, are underspecified in the lexicon to have head type *verb*. Since all linguistic objects described by an HPSG theory are of exactly one most specific type, an actual occurrence of a restructuring verb will have one of the two subtypes of verb as head value. It thus either raises all its arguments like an auxiliary, or it subcategorizes for a VP complement. Finally, all other verbs are lexically specified to have head type *simple-verb*.

3.1.1 Related issues

Complex vs. type antecedents The approach presented here bears some similarity to the work of Sag (1996). While he subclassifies phrasal types and uses principles to express generalizations about nonlocal specification, we subclassify words by their head subtypes and express generalizations about their valence specification. One formal difference between the two approaches is that Sag (1996) only makes use of type antecedents, whereas we employ complex descriptions as antecedents of the lexical principles.

From a formal perspective, implicational constraints with complex antecedents and those with type antecedents are both well-formed expressions of the HPSG description language defined in King (1989, 1994) and they are interpreted in the same way as any other formula of that language: as the set of objects described by that formula.[5]

We believe that complex antecedents of implicational constraints have the advantage that they make it possible to use the articulate data structure of HPSG to refer to the relevant subset of objects for which some generalization is intended to be expressed. Being restricted to type antecedents, one needs to introduce new types for every set of objects to which a generalization applies, which duplicates specifications in case the information was already encoded under one of the feature paths for other linguistic reasons.

Underspecification vs. lexical rules In the way that the specification of the lexical entries of restructuring verbs interacts with the type hierarchy and the lexical principles restricting the occurrence of the verb subtypes, the theory sketched above is in the tradition of so-called underspecification

[5]In particular, implicational statements with complex antecedents do not need to be converted to some kind of a disjunctive normal form in order to be interpreted as one might assume.

approaches like (Kathol, 1994; Frank, 1994; Oliva, 1994; Sanfilippo, 1995).[6] Typical for an underspecification approach is that a relationship which is to be expressed between two classes of words is formalized by specifying lexical entries in which an attribute has a type value for which subtypes are defined in the type hierarchy. Principles in the theory then refer to those subtypes and impose requirements on the set of grammatical signs in which they occur.

With respect to the two classes of lexical generalizations discussed in the beginning, an interesting aspect of the proposal made above is that it recasts the horizontal regularity which Monachesi (1995) captured by a lexical rule as a vertical regularity expressed by a lexical principle.[7] To better understand the nature of lexical generalizations, in future work the discussion of lexical regularities should therefore be extended with an investigation of the differences and similarities between lexical rules and underspecification approaches.

4 Summary

We started out with an overview of the two kinds of lexical generalizations, horizontal and vertical ones, and the methods available in HPSG to capture them. Lexical rules manage to express the horizontal generalizations in a theoretically meaningful way. Commonly used abbreviation methods, on the other hand, turn out to be inadequate for capturing the vertical generalizations. As an alternative, we showed how classes of lexical entries can be formally represented and how vertical generalizations over such classes can be expressed by means of lexical principles. This provides a formalization of the intuitions behind the 'lexical types' of Pollard and Sag (1987) using simple, well-understood mechanisms readily available in the HPSG architecture of Pollard and Sag (1994) – a straightforward possibility, which nevertheless had not yet been pursued.

To illustrate the proposal, we reformulated part of the theory of the Italian verbal complex proposed by Monachesi (1995) in such a way that the lexical classes assumed but not formalized are made explicit in the theory. Based on these explicit classes, lexical principles were formulated to provide a principled mapping from argument structure to the valence attributes.

Finally, we raised two issues related to the lexical principles: the status of the antecedents in the implicational constraints, and the relationship be-

[6]However, unlike some of these approaches we do not need to introduce new attributes in which to store the specifications to be distributed: as common source for the different ways to specify the valence lists of restructuring verbs we make use of the ARG-ST attribute, which is independently motivated.

[7]Indeed, many of the underspecification proposals mentioned above were formalized with the goal of eliminating (certain) lexical rules.

tween proposals employing lexical rules and so-called underspecification approaches. Regarding the first, we argued that there are no formal reasons for preferring type antecedents and that complex descriptions as antecedents of constraints can make reference to the relevant class-distinguishing information without requiring their duplication at the sign level. Regarding the second, more discussion is needed to figure out the precise nature of the differences, and criteria for deciding which mechanism is best suited to express which linguistic generalizations.

References

Abeillé, Anne, Danieèle Godard, and Ivan A. Sag. 1997. Two kinds of composition in French complex predicates. In Erhard Hinrichs, Andreas Kathol, and Tsuneko Nakazawa, editors, *Complex Predicates in Non-derivational Syntax*, Syntax and Semantics Series. Academic Press, San Diego. In Press (Version: June 29, 1997).

Briscoe, Ted, Ann Copestake, and Valeria De Paiva, editors. 1993. *Inheritance, Defaults and the Lexicon.* Cambridge University Press.

Calcagno, Mike. 1995. Interpreting lexical rules. In *Proceedings of the Conference on Formal Grammar*, Barcelona. Also in: Proceedings of the ACQUILEX II Workshop on Lexical Rules, 1995, Cambridge, UK.

Flickinger, Daniel. 1987. *Lexical Rules in the Hierarchical Lexicon.* Ph.D. thesis, Stanford University.

Frank, Annette. 1994. Verb second by underspecification. In Harald Trost, editor, *KONVENS '94*, pages 121–130, Berlin. Springer-Verlag.

Hinrichs, Erhard, Detmar Meurers, Frank Richter, Manfred Sailer, and Heike Winhart. 1997. Ein HPSG-Fragment des Deutschen, Teil 1: Theorie. Arbeitspapiere des SFB 340 95, Universität Tübingen.

Hinrichs, Erhard and Tsuneko Nakazawa. 1989. Flipped out: Aux in German. In *Papers from the 25th Regional Meeting of the Chicago Linguistic Society*, pages 193–202, Chicago, Illinois.

Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language*, 51:639–671.

Kathol, Andreas. 1994. Passives without lexical rules. In John Nerbonne, Klaus Netter, and Carl Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, Lecture Notes 46. CSLI Publications, pages 237–272.

King, Paul. 1989. *A Logical Formalism for Head-Driven Phrase Structure Grammar.* Ph.D. thesis, University of Manchester.

King, Paul. 1994. An expanded logical formalism for Head-Driven Phrase Structure Grammar. Arbeitspapiere des SFB 340 Nr. 59, Universität Tübingen.

Meurers, W. Detmar. 1995. Towards a semantics for lexical rules as used in HPSG. In *Proceedings of the Conference on Formal Grammar*, Barcelona. Also in: Proceedings of the ACQUILEX II Workshop on Lexical Rules, 1995, Cambridge, UK.

Monachesi, Paola. 1995. *A Grammar of Italian Clitics.* Ph.D. thesis, University of Tilburg, Institute for Language Technology and Artificial Intelligence, Tilburg, The Netherlands. TILDIL and ITK Dissertation Series 1995–3.

Oliva, Karel. 1994. HPSG Lexicon without Lexical Rules. In *Proceedings of the 15th COLING*, Kyoto, Japan.

Pollard, Carl and Ivan A. Sag. 1987. *Information-based Syntax and Semantics, Vol. 1.* Number 13 in Lecture Notes. CSLI Publications, Stanford University. Distributed by University of Chicago Press.

Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar.* University of Chicago Press, Chicago.

Richter, Frank. 1997. Die Satzstruktur des Deutschen und die Behandlung langer Abhängigkeiten in einer Linearisierungsgrammatik. Formale Grundlagen und Implementierung in einem HPSG-Fragment. In (Hinrichs et al., 1997).

Riehemann, Susanne. 1993. Word formation in lexical type hierarchies: A case study of *bar*-adjectives in german. Master's thesis, University of Tübingen. Also published as SfS-Report-02-93, Seminar für Sprachwissenschaft, University of Tübingen.

Sag, Ivan A. 1996. English relative clause constructions. Ms., to appear in the Journal of Linguistics.

Sanfilippo, Antonio. 1995. Lexical polymorphism and word disambiguation. In *Proceedings of the AAAI-95.*