

Attribute verbs within and across languages

Rainer Osswald and Thomas Gamerschlag

Institute for Language and Information

Department of General Linguistics

University of Düsseldorf, Germany

{osswald|gamer}@phil.uni-duesseldorf.de

Abstract

We describe a project that aims at a cross-linguistic resource of *attribute verbs*, which are stative verbs that encode attributes such as weight (*to weigh*) and price (*to cost*). To this end, we compiled a comprehensive lexical database of German attribute verbs, which have been classified with respect to the attributes they encode and a number of other relevant features. Based on these analyses, attribute verbs from other languages are added to the database in order to reveal typological differences in the encoding of attributes in the verbal domain.

1 Introduction

Language has different ways to encode attribute-value descriptions. If, as in English and many other languages, an attribute like weight or price can be expressed by a noun then the ascription of an attribute and its value can usually be expressed by a copula construction (1-a) or a *have*-possession construction (1-b).

- (1) a. The weight of the laptop is two kilos.
- b. The laptop has a weight of two kilos.
- c. The laptop weighs two kilos.

English also has the verb *weigh* for expressing such a description (1-c). An adjectival encoding such as *be two kilos heavy* is not possible in English, in contrast, for instance, to German:

- (2) Der Laptop ist zwei Kilo schwer.

The objective of the work reported in this paper is to explore the domain of verbs that encode attributes in the described way, with an eye on cross-linguistic variation. A related goal is to investigate the space of attributes encoded by verbs. As

a first step, we compiled a comprehensive list of such verbs in German by systematically exploring existing dictionaries. The verbs are manually classified with respect to the attribute they encode and certain other morpho-syntactic and semantic criteria, and the annotated entries are stored in a lexical database. After a careful revision and adjustment of the attribute space revealed during this process, data from other languages are added to the database, partly with the aim of full coverage, partly for contrastive purposes.

2 Attribute verbs

By an *attribute verb* (or *dimensional verb*) we mean a verb that, in one of its senses, characterizes an entity by specifying the value of an attribute of that entity. Attribute verbs are by definition stative. Standard examples are measurement verbs such as *weigh* (1-c), *cost* (3-a), and *last* (3-b):

- (3) a. The ticket costs two dollars. [PRICE]
- b. The lesson lasts two hours. [DURATION]

Each of the sentences in (3) describes the entity denoted by the subject with respect to the attribute expressed by the verb by specifying the value of that attribute through a “value phrase” (*two dollars*, *two hours*). The notion of attribute is taken here in a broad sense that subsumes also location, meaning, function, etc.

2.1 Empirical basis and coverage

We used the German standard dictionary *Duden Deutsches Universalwörterbuch* (DDUW) as a primary source for compiling a comprehensive list of German attribute verbs. Roughly 800 of the more than 13,000 verbs in the dictionary have been identified as attribute verbs. All of them are classified along the scheme described in Section 2.2 and are stored in a database (cf. Section 5). A representative set of French, Spanish, and English at-

tribute verbs is currently classified along the same lines and added to the database. In addition, a questionnaire has been used to gather data from native speakers and experts of further languages, which are mainly considered for contrastive purposes and with less ambition of lexicographic coverage.

2.2 Classification

The encoded attributes are by definition the most important features of the classification. Their formal properties will be discussed in Section 2.3. While the analyses and statistics presented in this section are based on the German sample introduced above, we often give corresponding English examples for ease of exposition (but see Figure 1). In addition to the encoded attribute, the verbs in the sample are characterized with respect to the following features:

Scalarity. Attributes like those in (3) are *scalar* in the sense that their value range is linearly ordered.¹ The attributes expressed by the verbs in (4), by comparison, are non-scalar.

- (4) a. The yoghurt tastes of bananas. [TASTE]
b. The path leads to a lake. [GOAL]

Although scalarity is a property of the attribute itself, and not of the verb, we used it as a verb feature in the first round of classification for reasons of consistency control.

Value incorporation. Some attribute verbs inherently restrict or specify the value of the attribute they encode. For example, the English verb *bulge* describes the shape of an object (5).

- (5) The bag bulges (with papers). [SHAPE]

Attribute verbs with inherent values can be seen as *descriptive verbs* in the sense of Snell-Hornby (1983), with the inherent value as the “modificand” of the “nucleus” given by the attribute.

Absolute use. Various attribute verbs that usually require an external value phrase can also be used absolutely, that is, without such a phrase. In this case, the implied value can depend on the typical properties of the entity denoted by the subject; viz. *The socks smell* vs. *The roses smell*.

¹Scalar structure has been recognized as an important factor for the analysis of gradation; see e.g. Kennedy and McNally (2005).

One vs. two attributes. Some attribute verbs are “two-dimensional” in that they encode two attributes. In our German sample, about 16% of the attribute verbs have been identified as two-dimensional. More than 70% of them encode LOCATION as an attribute with external value phrase and an additional attribute with inherent value. The two most frequent of these additional attributes are POSTURE and MANNER OF CONTACT; see (6) for English examples:

- (6) a. They squatted near the fire.
[LOCATION × POSTURE]
b. His wet shirt clung to his body.
[LOCATION × CONTACT]

There is moreover a small number of verbs such as English *hover* which combine LOCATION with an inherently specified “supporting medium” attribute:²

- (7) The hummingbird hovered over the flowers.
[LOCATION × MEDIUM]

Instead of LOCATION, two-dimensional attribute verbs can also encode “path-related” attributes such as SOURCE, GOAL, and COURSE. However, as illustrated by the examples in (8), the specific path aspect is usually provided by the preposition and not by the verb itself (Jackendoff, 1990; Kaufmann, 1995; Eschenbach et al., 2000).

- (8) a. The tree arched over the road.
[SHAPE × COURSE]
b. A male voice droned from the TV.
[SOUND × SOURCE]

It is nevertheless an inherent semantic property of these verbs to have an associated path. Moreover, word formation in German can give rise to attribute verbs with incorporated prepositions (see below).

Simplex vs. complex. A central concern in describing the lexical repertoire of attribute verbs in a language is to identify the *simplex* verbs, i.e., the monomorphemic lexemes within this class. Languages differ considerably in their morphological potential to form complex words. German has a rich system of verb prefixes comprising particles, prepositions, and adverbs, among others. The morphologically complex verbs in our sample are

²This analysis is based on Kaufmann (1995, Chap. 6.1).

subclassified into semantically transparent and intransparent formations. The first class consists mainly of regular incorporations of locative or directional prepositions, in which case the meaning of the resulting complex verb is compositionally derived from the meaning of the base verb and that of the preposition. For example, combining the directional preposition *über* with the attribute verb *sich wölben* results in the transitive complex verb *überwölben* (*arch over*), that is, *etwas überwölben* means *sich über etwas wölben*. In particular, the incorporated preposition reduces or eliminates the flexibility in the choice of the preposition usually available for locative or directional PP arguments. A special case of transparent formation is provided by verbs like *zurechnen* that are derived from proper prepositional verbs, here *rechnen zu* (*count among*).³ Despite the fact that the semantic contribution of the preposition is subregular at best in these cases, preposition incorporation only affects the syntactic valency and is hence transparent in the above sense. All in all, about half of the entries of the German sample are morphologically simple while approximately 25% have been classified as transparent and complex.

Primary vs. secondary. A considerable number of verbs used for attribute-value descriptions are “secondary” attribute verbs in that they are non-stative in their primary sense. The stative sense of *lead* in (4-b), for example, counts as secondary. The relation between this sense and the basic sense of *lead* can be described as a lexicalized *metonymic shift*. Other examples of this type are provided by change of direction verbs such as *ab-drehen* (*turn*) as in *Der Fluss dreht nach Westen ab*.⁴ *Argument alternations* are a further source for deriving secondary attribute verbs from non-stative verbs.⁵ Examples are the “characteristic property alternation” (e.g., *This knife cuts well*) and the middle alternation.

Approximately one out of three attribute verbs in the German sample have been analyzed as secondary. These verbs are furthermore classified with respect to the mechanism by which they are derived from the corresponding basic verb sense.

³Cf. Osswald et al. (2006) for more information on prepositional verbs in German.

⁴The stative sense of *turn* belongs to a class called “meander” verbs in Levin (1993) and “pseudo-motional locative” verbs in Dowty (1979).

⁵See Levin (1993) and Frense and Bennett (1996) for an overview of argument alternations in English and German.

Sense distinctions. We distinguish different senses of an attribute verb to the extent that they encode different attributes or attribute values. Under this regime, the attribute verbs in the German sample show an average polysemy of 1.2. About 13% of the entries are polysemous and the average polysemy within that set is 2.6.

Each sense is linked to the corresponding DDUW section or sections, if existent. For nearly 10% of the German entries in the sample, we added uses as attribute verbs that have no corresponding section in the DDUW. As to be expected, a good part of the missing readings are secondary in the sense introduced above, and thus related to non-stative senses in a more or less systematic way (cf. the discussion in Section 3).

Nominal and adjectival equivalents. We record if the attribute-value description expressed by an attribute verb can be expressed by a nominal (1) or an adjectival construction (2). Investigating the nominal and adjectival equivalents of attribute verbs is relevant for cross-linguistic comparison and also for questions concerning the diachronic development of attribute verbs.

A preliminary investigation of the simplex attribute verbs in the German sample has revealed that adjectival equivalents are rare. Nominal equivalents are fairly frequent, with less than half of them derivationally related to the verb. Within the latter class, deverbal nouns (e.g., *stinken* > *Gestank*) are considerably more frequent than denominal verbs (e.g., *Duft* > *duften*).

Valency. All attribute verbs in the sample are characterized with respect to their syntactic valency. In particular, the valency position of the value phrase has been explicitly marked. The optionality of this argument position corresponds to a possible absolute use of the attribute verb (see above).

2.3 The space of attributes

It is part of the project to explore the space of attributes encoded by attribute verbs on an empirical basis. The set of attributes used in the classification was not set up *a priori*, but is developed during the classification process and subsequent revision cycles.

Since attribute verbs can encode more than one attribute, there is no straightforward assignment of attributes to verbs. To put it the other way around,

one cannot expect to devise a simple taxonomy of attributes with attribute verbs uniquely attached to the nodes of the hierarchy. In fact, there seem to be at least three different ways of combining attributes that have to be taken into account: pairing, conjunction, and composition. *Pairing* (\times) is meant to refer to the two-dimensional case discussed before. *Conjunction* (+) is needed, for instance, to distribute the features TEMPORAL and SPACIAL over POSITION and DIRECTION. Conjunction thus realizes multiple inheritance. The attribute encoded by the German verb *datieren*, which is used to locate events in time, can then be written as POSITION + TEMPORAL. *Composition* (\circ), finally, is to be understood in the formal sense of functional composition, with attributes regarded as functions. Composition can be applied to characterize the attributes encoded by verbs like *begin* and *end*, when used to locate the begin or end of a trail, road and the like. For *The trail starts at the chapel*, the encoded attribute would then be expressed as BEGIN \circ LOCATION, or BEGIN \circ (POSITION + SPATIAL).

Figure 1 shows part of the system of classes of attribute verbs currently under development, with corresponding German examples. (Attribute composition is neglected in the figure.)

2.4 Cross-linguistic variation

The availability of verbs for encoding attribute-value descriptions differs considerably between languages. For instance, French has less verbs than English or German for encoding attributes of the SENSATION class. While the English attribute verbs *taste* (German: *schmecken*) and *feel* (German: *anfühlen*) can express TASTE and TOUCH, French uses constructions such as *avoir un goût* and *être ... au toucher* instead.

Another difference between Romance and Germanic languages shows up with two-dimensional verbs that involve LOCATION or PATH. In German and English, there are attribute verbs that combine LOCATION and POSTURE, whereas in French, a copula construction with past participle is required instead (Schwarze, 1993); compare *sitzen* (*sit*) and *liegen* (*lie*) vs. *être assis* and *être allongé*. The same is the case for the pairing LOCATION \times CONTACT, viz. *kleben* (*stick*) vs. *être collé*. This typological difference seems to be related to the distinction between verb-framed and satellite-framed languages proposed by Talmy (1985), according to

which verb-framed languages such as French tend to express manner by an adjunct in conflated constructions.

In order to investigate typological differences on a broad empirical basis, we developed a questionnaire to collect data about the verbal encoding of attributes in the languages Spanish, Korean, Russian, and Lakhota (Siouan). In addition, we are currently expanding the database to include French and English entries following the classification scheme described in Section 2.2.

3 The lexicographic perspective

A resource of attribute verbs as described in this paper can contribute to lexicography in various ways. For instance, it can be employed to improve the coverage of monolingual dictionaries with respect to stative uses of non-stative verbs. An overview of secondary attribute verbs and the underlying mechanisms of meaning shift can be helpful in this respect. This includes cases of valency alternation, which are often not systematically covered in dictionaries; see, e.g., Schwarze (2008) on the transitive and intransitive uses of the attribute verb *medir* (English: *measure*) in Spanish dictionaries.

The existence of sense gaps in the DDUW has already been mentioned in the discussion of sense distinctions in Section 2.2. Even the entries of frequent verbs such as *dreheren* and *wenden* (*turn*) do not make clear that they can be used to describe the change of direction of a road, river, etc. By comparison, the 5th edition of the ‘Longman Dictionary of Contemporary English’, is quite explicit about the corresponding sense of *turn*, which is listed in a section headed by ‘Direction’ and has the definition ‘if a road, river etc turns, it curves and starts to go in a new direction’.

Concerning the question of how to account for secondary attribute verbs in the dictionary, we therefore agree with Apresjan (2002), who requires that “all salient lexical classes should be fully taken into account and uniformly described in a dictionary in all of their linguistically relevant properties.” If applied to the classes of attribute verbs and the systematic relations between secondary attribute verbs and their non-stative base verbs, Apresjan’s “principle of systematic lexicography” would surely help to improve treatment of attribute verbs in the dictionary.

4 Comparison with existing verb classifications and resources

While some of the subclasses of attribute verbs have been studied before in work on stative verbs (Gerling and Orthen, 1979; Rothmayr, 2009), there has been no systematic investigation of such verbs in lexical semantics or lexicology up to now.

Existing lexical-semantic resources such as WordNet, FrameNet, or VerbNet do not pay attention to the systematic analysis and classification of attribute verbs (and stative verbs in general), with consequential gaps and inconsistencies in this domain. It is worth mentioning that the Brandeis Semantic Ontology of Pustejovsky et al. (2006) contains a number of subtypes of the type *Value Relation* such as *Amount*, *Cost*, *Height*, *Size*, *Temperature*, *Velocity*, and *Weight*, which are all to be located under QUANTITY/MEASURE in Figure 1. However, verbs are rather rare compared to nouns and adjectives in the set of lexical items associated with these ontological types.

5 Representation and implementation

The design of the lexical database and its implementation was driven by the requirement of a lean architecture that is easily modifiable and extensible and, furthermore, supports collaborative and platform-independent access and modification of the lexical data. We chose an XML database and implemented a web interface that allows editing and flexible browsing (including XQuery support).⁶ The chosen XML schema (specified via RELAX NG) has been kept as simple as possible for the moment. At a later point, we plan to migrate to a representational format more in line with existing standards for lexical data such as the Lexical Markup Framework (Francopoulo et al., 2006). It is planned to make the web interface accessible to the general public at the project end in summer 2011.

6 Ongoing work and prospects

The main focus of the project is currently on extending the lexical database with attribute verbs from other languages than German. Concerning the German sample, we plan to add more corpus-based examples.

⁶As to technical details, the implementation uses Berkeley DB XML, an embedded XML database, and Pylons, a Python-based web framework.

SENSATION/APPEARANCE	
SENSATION	
SMELL	<i>riechen, duften</i>
TASTE	<i>schmecken, munden</i>
TOUCH	<i>sich anfühlen, kratzen</i>
SOUND	<i>klingen, dröhnen</i>
LOOK	<i>aussehen, glänzen</i>
APPEARANCE	
POSITION	
+ SPATIAL	<i>sich befinden, wohnen</i>
× POSTURE	<i>knien, hocken</i>
× CONTACT	<i>kleben, stecken</i>
× MEDIUM	<i>schweben, schwimmen</i>
+ TEMPORAL	<i>datieren</i>
STRUCTURE	
SHAPE	<i>sich wölben</i>
COMPOSITION	<i>bestehen aus</i>
EXTENT	
QUANTITY/MEASURE	
WEIGHT	<i>wiegen</i>
DURATION	<i>dauern</i>
SPEED	<i>fahren, draufhaben</i>
CAPACITY	<i>fassen</i>
PATH	
COURSE	<i>führen, gehen</i>
DIRECTION	<i>verlaufen</i>
SOURCE	<i>wegführen</i>
GOAL	<i>hinführen, zeigen</i>
EXTREMAL	
BEGIN	<i>anfangen, entspringen</i>
END	<i>enden, münden</i>
SOCIOCULTURAL	
POSSESSOR	<i>gehören</i>
ROLE	<i>darstellen, verkörpern</i>
FUNCTION	<i>dienen, fungieren</i>
SEMIOTIC	
NAME	<i>heißen</i>
MEANING	<i>bedeuten</i>
REFERENCE	<i>sich beziehen auf</i>

Figure 1: Sketch of the system of attribute verb classes under development, with German examples.

Moreover, the formal characterization of the attribute space needs further investigation and might benefit from taking into account existing work on formal ontologies such as DOLCE (Borgo and Masolo, 2009).

It is furthermore planned to extend the coverage of the resource to non-stative attribute verbs, that is, to verbs which encode the change of attribute values.

Acknowledgments

The work presented in this paper has been financially supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) as part of the project ‘Dimensional Verbs’ of the DFG research group FOR600 ‘Functional Concepts and Frames’.

References

- Juri Apresjan. 2002. Principles of systematic lexicography. In Marie-Hélène Corréard, editor, *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins*, pages 91–104. Euralex, Grenoble.
- Stefano Borgo and Claudio Masolo. 2009. Foundational choices in DOLCE. In Steffen Staab and Rudi Studer, editors, *Handbook on Ontologies*. Springer, Berlin.
- David Dowty. 1979. *Word Meaning and Montague Grammar*. D. Reidel, Dordrecht.
- Dudenredaktion, editor. 2003. *Duden. Deutsches Universalwörterbuch*. Duden Verlag, Mannheim, 5th edition.
- Carola Eschenbach, Ladina Tschander, Christopher Habel, and Lars Kulik. 2000. Lexical specifications of paths. In Christian Freksa, Wilfried Brauer, Christopher Habel, and Karl Friedrich Wender, editors, *Spatial Cognition II*, Lecture Notes in Computer Science 1849, pages 127–144, Berlin. Springer.
- Gil Francopoulo, Monte George, Nicoletta Calzolari, Monica Monachini, Nuria Bel, Mandy Pet, and Claudia Soria. 2006. Lexical Markup Framework (LMF). In *Proceedings of LREC 2006*, pages 233–236, Genoa, Italy.
- J. Frense and P. Bennett. 1996. Verb alternations and semantic classes in English and German. *Language Sciences*, 18(1–2):305–317.
- Martin Gerling and Norbert Orthen. 1979. *Deutsche Zustands- und Bewegungsverben*. Studien zur deutschen Grammatik 11. Gunter Narr Verlag, Tübingen.
- Ray Jackendoff. 1990. *Semantic Structures*. MIT Press, Cambridge, MA.
- Ingrid Kaufmann. 1995. *Konzeptuelle Grundlagen semantischer Dekompositionsstrukturen. Die Kombinatorik lokaler Verben und prädikativer Argumente*. Niemeyer, Tübingen.
- Christopher Kennedy and Louise McNally. 2005. Scale structure, degree modification and the semantics of gradable predicates. *Language*, 81:345–381.
- Beth Levin. 1993. *English Verb Classes and Alternations*. University of Chicago Press, Chicago.
- Rainer Osswald, Hermann Helbig, and Sven Hartrumpf. 2006. The representation of German prepositional verbs in a semantically based computer lexicon. In *Proceedings of LREC 2006*, pages 461–464, Genoa, Italy.
- James Pustejovsky, Catherine Havasi, Jessica Littman, Anna Rumshisky, and Marc Verhagen. 2006. Towards a generative lexical resource: The Brandeis Semantic Ontology. In *Proceedings of LREC 2006*, pages 1702–1705.
- Antonia Rothmayr. 2009. *The Structure of Stative Verbs*. John Benjamins, Amsterdam.
- Christoph Schwarze. 1993. Primäre und sekundäre Lokalverben im Französischen. In Frank Beckmann and Gerhard Heyer, editors, *Theorie und Praxis des Lexikons*, pages 103–122. de Gruyter, Berlin.
- Brigitte Schwarze. 2008. Verbos dimensionales en el diccionario. In *Actas del II Congreso Internacional de Lexicografía Hispánica*.
- Mary Snell-Hornby. 1983. *Verb-Descriptivity in German and English. A Contrastive Study in Semantic Fields*. Carl Winter, Heidelberg.
- Leonard Talmy. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen, editor, *Language Typology and Syntactic Description. Volume III: Grammatical Categories and the Lexicon*, pages 57–149. Cambridge University Press, Cambridge.