Logical metonymies and qualia structures: an annotated database of logical metonymies for German

Alessandra Zarcone, Stefan Rüd

Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart
Azenbergstr 12, 70174 Stuttgart, Germany
alessandra.zarcone@ims.uni-stuttgart.de, ruedsn@studenten.ims.uni-stuttgart.de

Abstract

Logical metonymies like *The author began the book* involve the interpretation of events that are not realized in the sentence (**covert events**: → *writing the book*). The Generative Lexicon (Pustejovsky, 1995) provides a qualia-based account of covert event interpretation, claiming that the covert event is retrieved from the qualia structure of the object. Such a theory poses the question of to what extent covert events in logical metonymies can be accounted for by qualia structures. Building on previous work on English, we present a corpus study for German verbs (*anfangen (mit), aufhören (mit), beenden, beginnen (mit), geniessen*), based on data obtained from the deWaC corpus. We built a corpus of logical metonymies, which were manually annotated and compared with the qualia structures of their objects, then we contrasted annotation results from two expert annotators for metonymies (*The author began the book*) and long forms (*The author began reading the book*) across verbs. Our annotation was evaluated on a sample of sentences annotated by a group of naive annotators on a crowdsourcing platform. The logical metonymy database (2661 metonymies and 1886 long forms) with two expert annotations is freely available for scientific research purposes.

Keywords: logical metonymy, qualia, German

1. Logical metonymies and qualia structures

In **Logical metonymies** (e.g. *The author began the book*) event-subcategorizing verbs combine with entity-denoting objects; **covert events** (CE), not realized on the surface of the sentence, are involved in the interpretation of such constructions (→ *began writing the book*).

A classical account of logical metonymy was offered by the Generative Lexicon (Pustejovsky, 1991; Pustejovsky, 1995): the type-clash between the entity-denoting object and the event-subcategorizing verb leads to the recovery of a CE from the qualia structure (QS) of the object. QSs are complex lexical entries representing aspects of meaning of a word and the semantic relations involved in its understanding. The components of the QS (qualia roles) involved in logical metonymy interpretation are the **agentive quale** and the **telic quale** (Pustejovsky, 1995; Briscoe et al., 1990). The agentive quale (AQ) refers to the event which brings about the object; the telic quale (TQ) refers to the purpose of the object:

1. The author began the book → writing the book
   CE: *writing* (AQ)

2. The student began the book → reading the book
   CE: *reading* (TQ)

Consider the example 3 from Lascarides and Copestake (1998) and example 4:

3. My goat eats anything. He really enjoyed your book → eating the book
   CE: *eating* (AQ? TQ?)

4. The publisher began a series of books → publishing a series of books
   CE: *publishing* (AQ?)

The QS for *book* does not seem to accommodate for the interpretation *eating*. *Publishing* is involved in the creation of *book* as a physical object, but what is its relation with the AQ *writing*? Also, the agent plays a role in selecting one or the other qualia role:

5. The journalist enjoyed the movie → watching the movie
   CE: *watching* (TQ)

6. The director had enjoyed this movie the most → directing this movie
   CE: *directing* (AQ)

A number of issues seem to be at stake here, posing the question of whether QS are adequate to account for logical metonymies:

- **CE undergeneration**: QS accounts seem to undergenerate the set of CE interpretations (examples 3-4, see also Zarcone and Padó (2010)).

- **Role of non-lexical information**: Briscoe et al. (1990) and Verspoor (1997) have observed that the CE interpretation coming from the QS is used as a default, which can be overridden if a different interpretation is inferred from the context. On the other hand, Lascarides and Copestake (1998), though arguing that the lexicon must contain generalizations, claimed that world knowledge (pragmatic knowledge) has priority over these general rules. (examples 3-4).

- **Role of the agent**: Metonymic verbs select for the AQ or the TQ - for example, *enjoy* has a strong preference for the TQ of the object, (Briscoe et al., 1990; Verspoor, 1997) -, but the agent can override this selection, as in examples 5-6 (see also examples 1-2, Lapata et al. (2003) and Zarcone and Padó (2011)).

Verb	Total	NP					VP				
		all NP occurrences		artifacts (metonymies)			all VP occurrences		artifacts (long forms)		
				unique sent.		**metonymies**			unique sent.		**long forms**
anfangen	5463	2571	47.1 %	111	2.0%	**112**	2892	52.9%	446	8.2%	**472**
anfangen mit	4015	3691	91.9%	337	8.4%	**350**	324	8.1%	46	1.1%	**51**
aufhören	1223	13	1.1%	–	–	**–**	1210	98.9%	97	7.9%	**104**
aufhören mit	1223	1188	97.1%	46	3.8%	**47**	35	2.9%	5	0.4%	**5**
beenden	12014	12014	100.0%	228	1.9%	**231**	–	–	–	–	**–**
beginnen	41288	30111	72.9%	242	0.6%	**243**	11177	27.1%	1058	2.6%	**1110**
beginnen mit	36853	34858	94.6%	395	1.1%	**406**	1995	5.4%	94	0.3%	**110**
geniessen	20749	20477	98.7%	1052	5.1%	**1272**	272	1.3%	31	0.1%	**34**
4547 annotated metonymies						**2661**					**1886**

Table 1: Extracted and annotated sentences per verb. Please note that more than one object can occur in a sentence, therefore the number of metonymies and long forms (columns in boldface) is higher than the number of sentences.

- **Sets of interpretations**: Vendler (1968) and Lapata and Lascarides (2003) observe that metonymic expressions are usually interpreted by a set of interpretations (a cluster of meanings) instead of a single event (e.g. bringing about a book involves writing it, but also editing it, publishing it, and so on).

2. Qualia roles as clusters of meanings

For almost every noun there is a large number of verbs that can take it as an object. Consider the most frequent verbs for *Buch* ("book") in the deWaC corpus:

Buch (101241): 5006 lesen, 3468 schreiben, 1561 geben, 1092 kaufen, 1018 veröffentlichen, 893 empfehlen, 619 machen, 581 finden, 566 nehmen, 464 legen, 435 vorstellen, 385 kennen, 370 lassenI, 357 bestellen, 350 finden, 331 verfassen, 326 machen, 301 verkaufen, 267 führen, 261 halten, 259 ausleihen, 236 bringen, 233 erscheinen, 229 herausgeben, 215 bekommen, 200 abrunden, 186 sehen, 182 vorlegen, 176 herausbringen, 164 brauchen, 162 verschlingen, 161 aufschlagen, 160 nennen ...

All these verbs are in principle available for paraphrasing CEs. If we consider qualia roles to each contain one single predicate, then only *lesen* ("read") and *schreiben* ("write") would be considered to be its AQ and TQ, and many metonymies (e.g. example 4) could not be satisfyingly treated by qualia-based theories. If we instead assume that the QS is a model of conceptual knowledge, then the qualia roles can be considered prototypical concepts, bundling all these relations into the AQ or the TQ:

AQ: schreiben, veröffentlichen, machen, verfassen, vorlegen, herausbringen, ...

TQ: lesen, verschlingen, durchlesen, ...

In this study, we have considered the AQ or the TQ as concepts, and not as single predicates, following Vendler (1968) and Lapata and Lascarides (2003).

3. A corpus study of logical metonymy

In order to evaluate to what extent CEs in logical metonymies can be accounted for by QSs, we conducted a corpus study of German verbs (*anfangen (mit)*, "start (with)", *aufhören (mit)*, "stop (with)", *beenden*, "finish", *beginnen (mit)*, "begin (with)", *geniessen*, "enjoy"), based on data obtained from a dependency parsed version of the deWaC corpus, a very large collection of German sentences of about 1.7 billion words, gathered from the Internet (the WaCky project, Baroni et al. (2008)). The corpus was parsed with the BitPar (Schmid, 2004) and the FSPAR parser (Schiehlen, 2004).

Our method consisted of three steps:

Extraction of logical metonymies and long forms;
For each of these verbs, two types of instances were extracted:

- **logical metonymies (dependent NP)**, V(Obj)
  e.g. *Raucher können mit bestem Blick über die Stadt ihre Zigaretten geniessen.*
  "Smokers can enjoy their cigarettes with the best view of the city"

- **long forms (dependent VP)**, V(V2(Obj))
  e.g. *In dieser Zeit begann er, seine berühmten großformatigen Aquarelle zu malen.*
  "In this period he began to paint his famous large-format watercolors"

In order to avoid cases of alternation (e.g. *Der Film begann* - "the movie began") and non-metonymical uses, only transitive sentences with an animate subject and an artifact object were considered (semi-automatic labeling of subjects and objects was based on GermaNet 5.1, (Lemnitzer and Kunze, 2002)). Metaphorical uses were also excluded (e.g. *die schönen Seiten des Lebens geniessen* - "enjoy the beautiful pages of life"). See number of extracted sentences in Table 1.

Annotation of the extracted sentences;
Logical metonymies were annotated with a CE paraphrase; each CE (for logical metonymies) or subordinate event (for long forms) was annotated depending on its overlap with either the agentive quale or the telic quale of the object or to neither. The annotation was performed by one of the authors (ANN 1), a native speaker of German (Rüd and Zarcone, 2011), and at a later point also by a second native speaker of Ger-

QS coverage for CE interpretation in logical metonymies								
Verb	Tot.		AQ	TQ	SUM(AQ+TQ)	OTHER	CTXT_INSUFF	Q_UNDET
anfangen	112	ANN1	68 (60.71%)	23 (20.54%)	91 (81.25%)	4 (3.57%)	14 (12.5%)	3 (2.68%)
		ANN2	60 (53.57%)	46 (41.07%)	106 (94.64%)	6 (5.36%)	– (–)	– (–)
anfangen mit	350	ANN1	64 (18.29%)	120 (34.29%)	184 (52.57%)	80 (22.86%)	79 (22.57%)	7 (2%)
		ANN2	89 (25.43%)	219 (62.57%)	308 (88%)	34 (9.71%)	8 (2.29%)	– (–)
aufhören mit	47	ANN1	11 (23.4%)	29 (61.7%)	40 (85.11%)	3 (6.38%)	3 (6.38%)	1 (2.13%)
		ANN2	14 (29.79%)	31 (65.96%)	45 (95.74%)	2 (4.26%)	– (–)	– (–)
beenden	231	ANN1	118 (51.08%)	83 (35.93%)	201 (87.01%)	12 (5.19%)	18 (7.79%)	– (–)
		ANN2	122 (52.81%)	106 (45.89%)	228 (98.7%)	2 (0.87%)	1 (0.43%)	– (–)
beginnen	243	ANN1	214 (88.07%)	14 (5.76%)	228 (93.83%)	8 (3.29%)	7 (2.88%)	– (–)
		ANN2	210 (86.42%)	28 (11.52%)	238 (97.94%)	5 (2.06%)	– (–)	– (–)
beginnen mit	406	ANN1	144 (35.47%)	127 (31.28%)	271 (66.75%)	80 (19.7%)	55 (13.55%)	– (–)
		ANN2	162 (39.9%)	201 (49.51%)	363 (89.41%)	40 (9.85%)	3 (0.74%)	– (–)
geniessen	1272	ANN1	– (–)	1150 (90.41%)	1150 (90.41%)	26 (2.04%)	24 (1.89%)	72 (5.66%)
		ANN2	– (–)	1228 (96.54%)	1228 (96.54%)	37 (2.91%)	7 (0.55%)	– (–)
					2661 (100%)			

QS coverage for events in long forms							
Verb	Tot.		AQ	TQ	SUM(AQ+TQ)	OTHER	
anfangen	472	ANN1	189 (40.04%)	101 (21.40%)	290 (61.44%)	182 (38.56%)	
		ANN2	195 (41.31%)	136 (28.81%)	331 (70.13%)	141 (29.87%)	
anfangen mit	51	ANN1	– (–)	7 (13.73%)	7 (13.73%)	44 (86.27%)	
		ANN2	– (–)	22 (43.14%)	22 (43.14%)	29 (56.86%)	
aufhören	104	ANN1	29 (27.88%)	39 (37.50%)	68 (65.38%)	36 (34.62%)	
		ANN2	27 (25.96%)	50 (48.08%)	77 (74.04%)	27 (25.96%)	
aufhören mit	5	ANN1	– (–)	1 (20%)	1 (20%)	4 (80%)	
		ANN2	– (–)	3 (60%)	3 (60%)	2 (40%)	
beginnen	1110	ANN1	502 (45.23%)	217 (19.55%)	719 (64.77%)	391 (35.23%)	
		ANN2	472 (42.52%)	378 (34.05%)	850 (76.58%)	260 (23.42%)	
beginnen mit	110	ANN1	– (–)	23 (20.91%)	23 (20.91%)	87 (79.09%)	
		ANN2	1 (0.91%)	88 (80%)	89 (80.91%)	21 (19.09%)	
geniessen	34	ANN1	9 (26.47%)	12 (35.29%)	21 (61.76%)	13 (38.24%)	
		ANN2	6 (17.65%)	21 (61.76%)	27 (79.41%)	7 (20.59%)	
					1886 (100%)		

Table 2: Qualia roles coverage for events in logical metonymies and long forms

man (ANN 2 - for more details on the evaluation, see section 5.).

Table 1 shows the total amount of annotated sentences per each verb in the logical metonymy group and in the long form group, the tagset is described in section 4.1. At times more than one object occurred in a sentence (e.g. *Wir haben Kaffee und Kuchen genossen*, "we have enjoyed coffee and cake" - one sentence, two covert events: drinking coffee and eating cake), therefore the total number of metonymies and long forms (table 1, columns in boldface) is higher than the number of related sentences. A total of 1886 metonymies and 2661 long forms were annotated.

Computation of CE-QS matching;
  The matches between CE paraphrases in logical metonymies and QS and between explicit events in long forms and QS were computed.

The database of metonymies and long forms labelled by the two annotators is publicly available for scientific research purposes on the first author's website (http://www.ims.uni-stuttgart.de/~zarconaa/data.html.en).

4.   QS-event matching

4.1.   Tagset

Extracted sentences were annotated depending on whether their CE-paraphrases (for logical metonymies) or their explicit events (for long forms) matched the QS of the object. Possible tags were:

AQ: if the event corresponded to the agentive quale;

TQ: if the event corresponded to the telic quale;

OTHER: if the event did not correspond to either of them.

For logical metonymies, since the CEs were not explicit, it was sometimes problematic to find an appropriate paraphrase. For such cases, two more tagging options were introduced:

Q_UNDET: if the QS of the object was unclear;

CTXT_INSUFF: if the sentence context was not sufficient to find a paraphrase.

See percentages for such classes in Table 2.

4.2. Results

4.2.1. Tendencies of metonymic verbs

Metonymies with *anfangen* and *beginnen* yielded a strong preference for the AQ of the object, but a higher tendency towards the TQ for the corresponding constructions with *mit* (*anfangen mit* and *beginnen mit*). *Geniessen* showed a strong tendency for the TQ, whereas the (fewer) occurrences of *geniessen* in the long form allowed for AQ interpretation. Consider example 7:

> 7. *Ich habe es wirklich genossen, diesen Film <u>zu drehen</u> wenn man von den Szenen absieht, die ich bis zur Hüfte im Sumpf zubringen musste.*
> I really enjoyed <u>making</u> this film apart from the scenes I had to spend up to the hip in the swamp.

It seems that the preferred way to express an AQ activity for *geniessen* is to explicitly formulate it in a long form, in order to overcome the default TQ interpretation of logical metonymies with *geniessen*.

The majority of interpretations for the logical metonymies fall into the QS categories (AQ or TQ) (*anfangen, aufhören mit, beenden, beginnen, geniessen*), whereas the long form counterparts for these verbs, where the event is explicit, yielded higher percentages of non-QS interpretations. This is particularly interesting in the light of a "conversational maxim" account of logical metonymy: as observed by Lapata and Lascarides (2003), metonymy is strongly related to Grice's conversational maxims (Grice, 1975). If the QS captures a basic/default interpretation (*book → read* OR *write*), we tend to omit this information in a logical metonymy (*John began the book*). If, on the other hand, the event is a less typical one (e.g. *binding*), we need to make it explicit in a long form (*John began binding the book*).

4.2.2. Context-based interpretation

In some cases a broader context was needed to find the correct interpretation; since the sentences were collected from a web crawl corpus, it was sometimes possible to find the original source, as in example 8:

> 8. *Wir haben mit einem traditionellem Brett angefangen und es lief recht gut*
> We started with a traditional board and it went quite well.

From the original website it was apparent here that a mother wants to teach her 8-years-old son to play chess.

4.2.3. Comparison with previous work

For the English language, Briscoe et al. (1990) and Verspoor (1997) carried out corpus-based studies for metonymic verbs. The study by Briscoe and colleagues on the LOB (Lancaster-Oslo/ Bergen) corpus found that on average 17% of metonymies for the following verbs are pragmatic cases (not solvable with the QS): *enjoy, prefer, finish, start, begin, miss* and *regret*. Verspoor (1997) found in a BNC study that the CEs for the verbs begin and finish are determined by the QSs in about the 95% of the cases - again, a prevalence for AQ or TQ interpretations.

Our study was carried out for German and on a much larger corpus, and included long forms which are not included in previous work on English; also, our study only takes into consideration artifacts, whereas Briscoe et al. (1990) and Verspoor (1997) include non-artifact items (*sea* or *family*). Despite some methodological differences, though, lexical differences were noted by all three studies. *Enjoy* matched the low OTHER interpretations of *geniessen*. *Beginnen* and *anfangen* have a strong AQ preference, while the English *begin* has more TQ interpretations. *Begin* and *begin on* in Verspoor (1997) respectively showed a preference for QS-interpretations and for context interpretations, and a similar contrast holds for our analysis of *anfangen (mit), beginnen (mit)* and *aufhören (mit)*: the versions with *mit* have a larger proportion of non-QS interpretations.

The general claim from these studies, that QSs can account for up to 80% of the reported metonymic instances, seems to be quite consistent with the results of the present work, which yielded values above 80% for *anfangen, aufhören mit, beenden, beginnen, geniessen*.

5. Evaluation

Given the large number of annotated items (4547), the annotation was initially conducted by a single native speaker annotator, and only at a later stage by a second expert annotator. We computed agreement with regards to the assignment of QS categories (nominal scale): Krippendorff's α (Krippendorff, 1980) for the annotated database was .71 for the long forms and .6 for the metonymies (substantial agreement, table 3).

5.1. Crowdsourcing inter-annotator evaluation

In order to evaluate the annotation of the data set, an inter-annotator experiment was conducted on a sample of 100 randomly extracted sentences from the logical metonymy set and 100 from the long form set. The sample was used for a large-scale non-expert annotation experiment on the Amazon Mechanical Turk (AMT) platform for fast and affordable collection of native judgements (Snow et al., 2008).

5.1.1. Method

In the annotation study with the long form sample, annotators were asked if the underlined event (e.g. *Ich habe im Herbst 1997 angefangen, Gedichte zu schreiben.* - "I began writing poems in fall 1997") depicted the purpose of the object or its coming about (*Beschreibt es den Zweck des Objektes, oder seine Herstellung, oder etwas anderes?*). In the annotation study with the logical metonymies, annotators were told that an additional implicit event was involved (e.g. *Kleine Kinder sollten noch nichtmit großen Instrumenten anfangen → <u>spielen</u>* - "Little kids should not begin with big instruments $rightarrow$ <u>to play</u>"), and they were asked to provide the additional event (*Welches zusätzliche Ereignis verstehen Sie?*) and to say if it depicted the purpose of the object or its coming about (*Beschreibt es den Zweck des Objektes, oder seine Herstellung, oder etwas anderes?*).

24 unique annotators from the US took part in the annotation study with the long form sample (average of 10 annotations per item) and 22 annotated the metonymy sample (average of 10 annotations per item).

	Logical Metonymies	Long Forms
(db) ANN1 - ANN2	0.6	0.71
(sample) ANN 1 - ANN 2	0.57	0.65
AMT	0.5	0.42
ANN 1 - AMT	0.67	0.32
ANN 2 - AMT	0.56	0.52
ANN1 - ANN2 - AMT	0.59	0.5

Table 3: Agreement values from the evaluation (Krippendorff's α): pairwise annotations among annotator (ANN) 1 and 2 and majority vote from the AMT experiment; three-way agreement between ANN 1, ANN 2 and AMT majority vote; and agreement among all the AMT annotators.

5.1.2. Results

Agreement was computed among the AMT annotators (AMT in table 3), and between the majority vote per sentence yielded by AMT and, respectively, ANN1 and ANN2, and also a three-way agreement between ANN1, ANN2 and the AMT majority vote was computed. Interestingly, the crowdsourcing study with non-experts annotators yielded satisfactory agreement levels (fair to substantial agreement). On the same sample, the two expert annotators yielded a moderate to substantial agreement (.65 for the long forms and .57 for the metonymies).

Agreement was lower for the long form sample; recall that the long forms have a higher number of non-QS events than logical metonymies: non-expert annotators seem to have bigger difficulties in classifying less-clear cut events into the three categories of AQ, TQ and OTHER, and they provide much fewer OTHER responses for long forms (see table 4).

Logical Metonymies					
	AQ	TQ	OTHER	CTXT INSUFF	Q UNDET
ANN1	44	39	10	6	1
ANN2	44	49	0	0	0
AMT	44	53	0	1	0
Long forms					
	AQ	TQ	OTHER		
ANN1	26	24	50		
ANN 2	25	45	30		
AMT	37	54	9		

Table 4: Distribution of answers given for the long form sample and the logical metonymy sample by ANN1, ANN2 and by the AMT annotators (majority vote).

6. Conclusion

We have presented a corpus study for German verbs (*anfangen (mit), aufhören (mit), beenden, beginnen (mit), geniessen*), based on data obtained from the deWaC corpus. Covert events in logical metonymies (*The author began the book → reading*) were annotated and, together with events in long forms (*The author began reading the book*), were labelled by two expert annotators depending on their overlap with the QS of the object. Our annotation was compared and evaluated with a non-expert inter-annotator study conducted on a crowdsourcing platform, yielding fair to substantial levels of agreement.

Our study highlights different patterns of behavior for the analyzed verbs. Some have a clear tendency towards the TQ of the object, some for the AQ, others have a less clear-cut behavior. Despite methodological differences with previous work on English (Briscoe et al., 1990; Verspoor, 1997), some common patterns seemed to emerge: in particular, between *begin* and *begin on*, and our analysis of the constructions with *mit*: *begin on* and the *mit*-constructions both show a significantly bigger proportion of non-QS interpretations than their versions without *mit*. The general claim from Briscoe et al. (1990) and Verspoor (1997), that qualia roles can solve up to 80% of the reported metonymic instances, was mirrored by our results, which yielded percentages of QS-interpretations between 81% and 98% for logical metonymies with *anfangen*,*beginnen*, *beenden* and *geniessen*.

If on the one hand the vast majority of CE interpretations in logical metonymies falls into the QS range, on the other hand in cases where the event is explicit (long forms) chances are that this depicts a non-QS event, thus confirming the observations in Lapata et al. (2003) and Egg (2004) that metonymy is strongly related to Grices conversational maxims (Grice, 1975): implicit CE → basic/default QS-interpretation; explicit event → non-QS interpretation.

While English is often the language of choice for literature on logical metonymy (but see Lapata et al. (2003)), we are offering a study and an available resource for the German language: the study resulted in a database of 4547 annotated contexts (2661 metonymies and 1886 long forms), which is freely available for scientific research purposes.

8. References

Marco Baroni, Silvia Bernardini, Adriano Ferraresi, and Eros Zanchetta. 2008. The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation*, 43(3):209–226.

Ted Briscoe, Ann Copestake, and Bran Boguraev. 1990. Enjoy the paper: Lexical semantics via lexicology. In *Proceedings of the 13th COLING*, volume 2, pages 42–47.

Markus Egg. 2004. Metonymie als Phaenomen der Semantik-Pragmatik-Schnittstelle. *Metaphorik*, 6:37–53.

Herbert Paul Grice. 1975. Logic and conversation. In Peter Cole et al., editor, *Syntax and Semantics*, volume 3. Academic Press, New York.

Klaus Krippendorff. 1980. *Content Analysis*. Sage.

Mirella Lapata and Alex Lascarides. 2003. A probabilistic account of logical metonymy. *Computational Linguistics*, 29(2):263–317.

Mirella Lapata, Frank Keller, and Christoph Scheepers. 2003. Intra-sentential context effects on the interpretation of logical metonymy. *Cognitive Science*, 27(4):649–668.

Alex Lascarides and Ann Copestake. 1998. Pragmatics and word meaning. *Journal of Linguistics*, 34:387–414.

Lothar Lemnitzer and Claudia Kunze. 2002. Germanet – representation, visualization, application. In *Proceedings of the 3rd LREC*, pages 1485–1491.

James Pustejovsky. 1991. The Generative Lexicon. *Computational Linguistics*, 17(4):409–441.

James Pustejovsky. 1995. *The Generative Lexicon*. MIT Press, Cambridge, MA.

Stefan Rüd and Alessandra Zarcone. 2011. Covert events and qualia structures for German verbs. In *Proceedings of the Metonymy 2011 Workshop*, September.

Michael Schiehlen. 2004. Annotation strategies for probabilistic parsing in German. In *Proceedings of the 20th COLING*, pages 390–396, Geneva, Switzerland.

Helmut Schmid. 2004. Efficient parsing of highly ambiguous context-free grammars with bit vectors. In *Proceedings of the 20th COLING*, pages 162–168, Geneva, Switzerland.

R. Snow, B. O'Connor, D. Jurafsky, and A.Y. Ng. 2008. Cheap and fast - but is it good? Evaluating non-expert annotations for natural language tasks. In *Proceedings of EMNLP 2008*, pages 254–263, Honolulu, Hawaii.

Zeno Vendler. 1968. *Adjectives and Nominalizations*. Mouton, The Hague, The Netherlands.

Cornelia M. Verspoor. 1997. *Contextually-dependent lexical semantics*. Doctoral dissertation, University of Edinburgh.

Alessandra Zarcone and Sebastian Padó. 2010. "I like work: I can sit and look at it for hours" – Type clash vs. plausibility in covert event recovery. In *Proc. of the VERB 2010 workshop*, Pisa, Italy.

Alessandra Zarcone and Sebastian Padó. 2011. Generalized event knowledge in logical metonymy resolution. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, pages 944–949, Austin, TX. Cognitive Science Society.