

Head or Non-head?

Semi-automatic procedures for extracting and classifying subcategorisation properties of compounds

Ekaterina Lapshinova-Koltunski, Ulrich Heid

IMS, Universität Stuttgart
Azenbergstr. 12
D-70174 Stuttgart, Germany
katerina, heid@ims.uni-stuttgart.de

Abstract

In this paper we discuss an approach to the semi-automatic classification of the compounds extracted from German corpora. Compound nominals are semi-automatically extracted from text corpora along with their sentential complements and classified. Our classification is based on the relationships between the subcategorisation of nominal compounds, e.g. *Grundfrage*, *Wettstreit* and *Beweismittel* and that of their constituent parts, *Frage*, *Streit*, *Beweis*, etc. We show that there are cases which do not match the commonly accepted assumption that the head of a compound is its valency bearer. Such cases should receive a specific treatment in NLP dictionary building.

1. Introduction

This study presents a set of semi-automatic extraction procedures for German nominal compounds which subcategorize for a subordinate clause; it is a part of a broader research programme about the semi-automatic classification of automatically extracted subcategorisation. As our aims serve symbolic NLP, especially large symbolic grammars for deep processing, such as HPSG (cf. work in the LinGO project (Copestake et al., 2004)) or LFG (cf. the PARGRAM project (Butt et al., 2002)), which need detailed subcategorisation data for their lexicons and grammars.

We review the subcategorisation of compounds and show that the commonly accepted assumption, that the head of a compound determines the predicate-argument structure of the whole construction (Zwicky, 1985) and (Bauer, 1988), is not always valid. There are at least two types of exceptions to this rule; this becomes apparent when the subcategorisation of compounds is compared with that of their complements.

We elaborate an architecture using linguistic knowledge about the phenomena we extract, and aim at answering the following questions: how can data about subcategorisation properties of nominal compounds be extracted from text corpora, and how can compounds be classified according to their subcategorisation properties?

In this paper we deal with German nominal compounds subcategorizing a subordinate *dass*, *wh-* or *ob* (that, *wh*- or *if*) clause, and focus on sentential complementation although our methods can be applied for other complements as well.

2. Data and existing approaches

2.1. Approaches to the linguistic description of compounds.

Many linguists call compounding an interface between morphology and syntax (Spencer, 1991). Compounds have a constituent structure and most of them have a compositional reading (Johnston and Busa, 1999). The relations

between their elements are similar to the relations of sentence parts. The head of a compound is characterized by the same properties as a phrase head: it determines the morphosyntactic categories and subcategorisation of the whole compound ((Zwicky, 1985) and (Bauer, 1978) or (Bauer, 1988)). We analyze nominal compounds followed by a subcategorized subordinate *dass*, *wh-* or *ob* clause. Nominal compounds were extracted along with the sentential complements which typically follow them in a sentence.

2.2. Data.

Our first extraction results show that there exist three types of nominal compounds: (a) those that share the subcategorisation with the head, (b) those that share their subcategorisation with the non-head, and (c) those that share their subcategorisation with neither the head nor the non-head constituents or that share it with both.

(a) The subcategorisation of the whole complex word is determined by its head:

- (1) *die Dauerfrage, ob...*
("the constant **question**, if...")
vs. *die Frage, ob...*
("the question if...").

Die Dauerfrage, ob die Gesellschaft oder der einzelne selber sein Schicksal stärker bestimmt, ist im republikanischen Amerika erst einmal beantwortet.

("The constant **question** if the community or individuals themselves make their own destiny has for now been answered in republican America".)

(b) The subcategorisation of the whole complex word is determined by its non-head:

- (2) *das Auswahlverfahren, wer...*
("the **selection** process, who...")
vs. *die Auswahl, wer...*
("the selection who...").

Das *Auswahlverfahren*, *wer Christine Frisinghelli ab dem Jahr 2000 an der Spitze des "steirischen herbstes" nachfolgen wird, geht ins Finale.*

("The **selection** process, who will follow Christine Frisinghelli at the top of the 'steirischer herbst' festival as of 2000, is about to be decided".)

(c) The subcategorisation of the compound is determined by both, its head and its non-head, or, alternatively, by neither of its elements:

(c1) both the head and the non-head can subcategorize for a subclause:

(3a.) *der Wettstreit*, *wer...*

("the bet battle (competition), who")

vs. *die Wette*, *wer* ("the bet, who")

or *der Streit*, *wer*

("the battle (argument), who").

In dem edlen Bonner Wettstreit, *wer die Kunst des Nebelwerfens und Schaumschlagens am besten beherrscht, rückt ein Mann unaufhaltsam nach vorn: Rudolf Scharping.*

("In the noble Bonn **competition**, who is the best smoke launcher and hot-air merchant, one man moves forward inexorably: Rudolf Scharping".)

(c2) neither the head nor the non-head can subcategorize for a subclause:

(3b.) *der Volksmund*, *dass...*

("folk mouth (common parlance) that...")

vs. *das Volk*, *dass...** ("folk that...")

or *der Mund*, *dass...** ("mouth that...").

Den Volksmund, *dass Kleinvieh auch viel Mist machen kann, bestätigen die Fluggesellschaften mit ihren Vielfliegerprogrammen.*

("The **common parlance** that many a little makes a mickle is confirmed by airlines with their frequent flyer programs.")

The type (c1) compounds are mostly borderline cases between the (a) and the (b) types. The subcategorisation of such cases as *Wettstreit* could be "inherited" from both the head and the non-head constituents (cf. (3a.)). For subtypes (c2), we assume that both parts of the compound do not have sentential complements at all, thus the compound has its own subcategorisation properties (cf. (3b.)).

More type (c) as well as type (b) examples extracted from German corpora are listed in the appendix.

3. Methods and tools for the extraction and classification of nominal compounds

3.1. Input.

We use a corpus of newspaper texts of German¹ which are annotated with part-of-speech tags, morpho-syntactic in-

¹We use texts from Germany, Austria and Switzerland, a total of ca. 950M words: Austrian ('AT', ca. 500M) and Swiss

formation, lemmatisation and partial syntactic analysis results², and can be queried in the form of regular expressions through the Stuttgart CorpusWorkBench (CWB³).

3.2. Linguistic assumptions and extraction architecture.

3.2.1. Predicate identification.

The first extraction tests were performed on nominal predicates in the "Vorfeld"-position (VF in figure 2, position left of the finite verb) cf. example (5) and the query in figure 1. In this case, we search for nouns or noun phrases at the sentence beginning, which are followed by a *wh-*, *ob-* or *dass-*clause. If a noun in "Vorfeld"-position is followed by a sentential complement, this complement can only be subcategorized by the noun ((Helbig and Buscha, 2005) and (Zifonun et al., 1997)).

(5) *Aber all die Erklärungsversuche, warum der Teufel sich an die Frau Doktor heranmacht, sind auf der Glätze gedrehte Locken.* ("But all the expansion-attempts (attempts to explain) why the devil chats up the female doctor are as futile as giving a bald man a comb").

The query in figure 1 has a noun phrase (or a PP) at the beginning of the sentence (lines 1 to 4), followed by a subordinate clause (lines 5 to 11) and the finite (main) verb (Lapshinova and Heid, 2007).

The list of nominal predicates extracted with the query in figure 1 is used for further classification. With the help of morphological tools⁴ we sort the list of nominal predicate candidates into two groups:

1) simplex predicates:

– *Erklärung, warum...*

Frage, ob...

Beweis, dass...

2) compound predicates:

– *Erklärungsversuche, warum...*

der Wahrheitsbeweis, ob...

die Grundfrage, dass...

The both lists resulting from the extraction and classification procedures shown in figure 2 are used for the subclassification of type (a) and (b) compounds with CQP procedures (cf. section 3.2.2.).

('CH', ca. 180M) texts are part of the German reference corpus DeReKo and have been made available to us by the Institut für deutsche Sprache, Mannheim, in a cooperative project. The corpora from Germany include extracts (1992-2000) from *die tageszeitung* ('taz', 111M), *Frankfurter Rundschau* ('FR', 40M), *Frankfurter Allgemeine Zeitung* ('FAZ', 71M), *Stuttgarter Zeitung* ('StZ', 36M), *DIE ZEIT* ('ZEIT', 86M) as well as literary texts from the 'Gutenberg' Archive ('DE Lit.', 138M).

²Annotations were obtained with Tree Tagger (Schmid, 1994) and (Schmid, 1999), YAC-Chunker (Kermes, 2003) and the morphology tool SMOR (Schmid et al., 2004)

³Cf. (Evert, 2005).

⁴e.g. SMOR (Schmid et al., 2004)

	Query building blocks	comments	extracted sentence
1.	<s>	sentence beginning	
2.	[pos!="NN V.FIN"]{0,3}	prenominal material	<i>Aber all</i>
3.	[pos="APPR APPRART"]?	optl. preposition or prep/art	
4.	(<np> ... </np>)	nominal phrase	<i>die Erklärungsversuche</i>
5.	“,”	comma	,
6.	(pos="PW.*"&word!="wobei womit")	relative pronoun, but not “wobei” and “womit” or	<i>warum</i>
7.	(word="ob")	conjunction “ob” or	
8.	(word="dass")	conjunction “daß”	
9.	[pos!="\$. V.FIN"]*	subclause: non-verbal part	<i>der Teufel sich an die Frau Doktor</i>
10.	[pos="V.FIN"]	finite verb of subclause	<i>heranmacht</i>
11.	“,”	comma	,
12.	[pos="V.FIN"]	finite main verb	<i>sind</i>
13.	within s;	rest of main clause	<i>auf der Glatze gedrehte Locken.</i>

Figure 1: Query for extraction of nouns in Vorfeld position subcategorizing a *wh*-, *ob*- or *daß*-clause

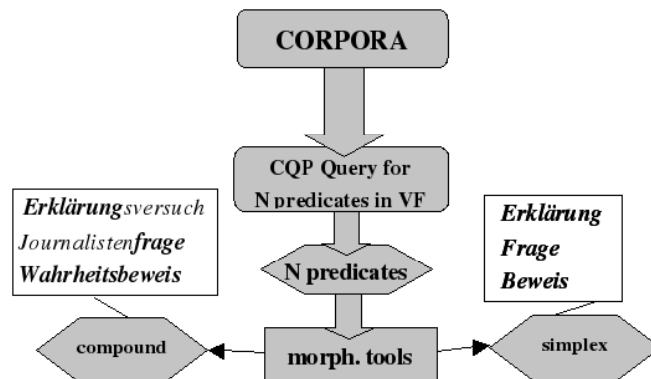


Figure 2: Extraction architecture

3.2.2. Subclassification.

To classify the extracted candidates according to the three groups mentioned in 2.2., we use compounds whose components are in a list of simplex nouns whose subcategorisation properties are “known” to our system. This knowledge comes from two sources: a list of extracted predicates subcategorizing for a sentential complement (cf. figure 1) and IMSLex, an NLP Lexicon for German containing subcategorisation information⁵:

- If the head of a compound is in the “known” list, and the non-head is not, the compound belongs to the type (a):
Paradebeispiel, Journalistenfrage;
- If the non-head of a compound is in the “known” list and the head is not, the compound belongs to the type (b):
Beweismittel, Erklärungsversuch;

- If both the head and the non-head are in the “known” list, the compound belongs to the type (c):
Wettstreit.

Cases where both elements are not contained in our lexical resources could either be new to these resources or of subtype (c2). We control these manually.

An overview of the cascaded classification procedures is given in figure 3.

A completely automatic separation of the compound types is still difficult. In example (6), the sentential complement is determined by the non-head noun *Zeit* (*Zeit, wann..*). The head noun *Punkt* can also subcategorize for a *wh*- clause: *Punkt, wo/warum...* (“the point where/why...”), but has not been found with *wann...* (“when”). So for a precise classification, we use the types of *wh*-pronouns preferred by the nominal predicates.

- (6) *Ein genauer Zeitpunkt, wann der Verkauf erfolgen soll, wurde ebenfalls nicht genannt.* (“The exact time

⁵(Lezius et al., 2000) and (Fitschen, 2004)

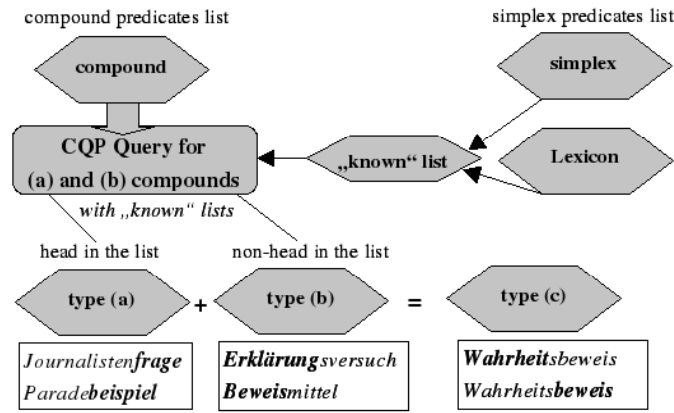


Figure 3: Subclassification of compounds

point (point in time) when the purchase should take place, was not mentioned either”.)

4. Results and interpretation

A full quantitative evaluation of the tools would require the preparation of a gold standard corpus. A partial evaluation of nominals with sentential complements (over 7,000 tokens) in Vorfeld (VF), on the 'FR', 'FAZ' and 'taz' (cf. section 3.), provided the precision shown in table 1: the precision of the extraction of nouns subcategorizing a subclause varies between 89 and 99%. The low figures for *wh*-subclauses have to do with the systematic ambiguity of *wh*-indirect questions, with headless relatives in German.

subclause type	precision
<i>dass</i>	98,50%
<i>ob</i>	99,00%
<i>wh-</i>	89,00%

Table 1: Extraction results for nominals subcategorizing for a *dass*-, *ob*- or a *wh*-clause in “Vorfeld”

We analyzed the morphological sorting procedure for a list of extracted nominal predicates (747 types) and achieved a precision of 94,2% and a recall of 95,7% in the automatic identification of nominal compounds. As the morphological tool provides us with the information about the part of speech of compound elements, their nominal (NN in table 2) or deverbal (V in table 2) nature, we outlined their evaluation in table 2.

type&example	proportion	precision	recall
NN.NN <i>Beweislast</i>	65,58%	97,03%	95,28%
NN.V <i>Völkbefragung</i>	12,99%	80,00%	95,24%
V.NN <i>Bedenkzeit</i>	21,43%	93,94%	97,06%
TOTAL	100,00%	94,20%	95,70%

Table 2: The morphological analysis of compounds: types of compounds, precision and recall

The information about the part of speech of compound elements also allows us to analyze the nature of elements of types (a) to (c). Most type (a) and type (c) compounds consist of two nouns (NN.NN). Most (b)-type compounds have a verbal non-head (V.NN), which is expectable (cf. table 3).

types	(a)	(b)	(c)
NN.NN	81,36%	25,00%	76,32%
NN.V	11,86%	20,00%	13,15%
V.NN	6,78%	55,00%	10,53%
TOTAL	100,00%	100,00%	100,00%

Table 3: The morphological analysis of compounds: types of compounds, precision and recall

In tables 4 and 5 we outline the frequency of compound nominal predicates, compared to the frequency of simplex nominal predicates extracted in “Vorfeld”. Both simplex and compound predicates were extracted from the newspapers 'FR', 'FAZ' and 'taz' along with the three types of subclauses under study.

As shown in tables 4 and 5, the phenomena we analyze are not frequent (about 12-15,7% of all nominal predicates in “Vorfeld” are compounds). The proportion of compound nouns varies significantly from ca. 2% for *ob*-clauses to ca. 20% for *dass*-clauses (compared by the number of extracted tokens), depending on the type of the subcategorized clause. In Table 6, we summarize the proportion of type (a) to (c) occurrences in the newspaper corpora mentioned above. The figures show that type (a) cases are the most frequent in text corpora. Compounds of type (b) and type (c), however, make over 40% of all compound cases in “Vorfeld”, which is a considerable amount.

Types	a	b	c
occurrences in corpora	56,6%	11,8%	31,6%

Table 6: Occurrence of (a) to (c) types in “Vorfeld”

The evaluation of the automatic identification of type (b) compounds extracted from 'FR', 'FAZ', 'taz', 'Stz' and 'ZEIT' provided a precision of 79%. In a naive string comparison, a part of the extracted noise are cases like *Zeitungs-*

subclause type	tokens				types			
	simplex	in %	compound	in %	simplex	in %	compound	in %
<i>dass</i>	4780	78,30%	1038	21,70%	863	88,40%	100	11,60%
<i>ob</i>	1479	97,80%	32	2,20%	153	79,30%	31	20,70%
<i>wh-</i>	771	95,30%	36	4,70%	290	91,00%	26	9,00%
TOTAL	7030	84,30%	1106	15,70%	1306	88,00%	157	12,00%

Table 4: Simplex and compound predicates according to the subcategorized subclause

subclause type	tokens				types			
	simplex	in %	compound	in %	simplex	in %	compound	in %
'FR'	1048	83,40%	174	16,60%	266	88,60%	33	12,40%
'FAZ'	2414	86,80%	319	13,20%	428	98,30%	46	10,70%
'taz'	3568	82,70%	613	17,30%	612	87,30%	78	12,70%
TOTAL	7030	84,30%	1106	15,70%	1306	88,00%	157	12,00%

Table 5: Simplex and compound predicates, subcategorizing for a *dass*-, *ob*- or a *wh*-clause.

bericht, *Zeitungsmeldung*, where the non-head constituent is not in the “known” list but it contains a string, which can subcategorize for a subclause. In *Zeitungsbericht* (“newspaper report”), the nominal predicate *Zeit* (“time”) is obviously not a valency bearer of the compound as it is just a graphical part of the non-head noun *Zeitung* (“newspaper”). Using the morphological analyser we can identify constituent parts of compounds and exclude the occurrence of this kind of noise. We are also able to achieve high recall (84,5%) in subclassifying type (b) compounds.

To further improve the morphology-based classification, we should include the extracted compound predicates in the “known” list. Some cases, e.g. *Grundsatzmißbilligung* (“deprecation of principle”) or *Rechtsgrundsatz* (“principle of law”), contain compound elements, e.g. *Grundsatz* (“principle”), which are built from elements that themselves can subcategorize for sentential complements, such as *Grund*, *dass...* (“reason that...”). The morphological analyser produces a list of all morphemes, but not a hierarchical analysis of word structure.

5. Compounds for symbolic NLP

As was mentioned above, we intend to use the results of our study for large symbolic grammars, such as HPSG or LFG (cf. section 1.). The treatment of compounds with the described set of procedures limits the need for listing all compounds in an NLP lexicon. On the basis of our automatic classification we can decide about which compounds to store in the subcategorisation lexicon. Compounds of type (b) and of type (c) should be included. There is, however, no need to store type (a) compounds. For compounds of this type, we could use a systematic method of matching the arguments of compounds to the subcategorisation frames of their heads (cf. work of O. Gurevich on the methods of mapping deverbal arguments onto those of their corresponding verbs (Gurevich et al., 2007)). In this case, a rewrite system would automatically map the subcategorisation frames of type (a) compounds to those of their head nouns, which are already stored in the lexicon. We have, however, not yet implemented this mechanism.

6. Conclusions and future work

Our experiments show that the subcategorisation of nominal compounds is not in all cases determined by the head component (types (b) and (c)). The data obtained from corpora allow us to classify nominal compounds into three groups according to their subcategorisation properties. This information can be semi-automatically retrieved from text corpora.

On the basis of extraction procedures we propose a semi-automatic extraction and classification architecture which can operate on tokenized, tagged, lemmatized and partially parsed texts. Our work is in an early stage, but the first results show that our tools can extract and classify compound candidates in terms of which of their components are valency bearers.

Future work will include the analysis of the nature of the non-head valency-bearer of type (b) and (c) compounds, as well as an extension of the kinds of extracted complements beyond subclauses.

7. Acknowledgements

This work is part of Ekaterina Lapshinova-Koltunski’s PhD work supported by the DFG-funded Research Training Group GK-609 “Linguistic Representations and their Interpretation”.

8. References

- L. Bauer. 1978. *The grammar of nominal compounding*. Odense University Press.
- L. Bauer. 1988. *Introducing Linguistic Morphology*. University Press, Edinburgh.
- M. Butt, H. Dyvik, T. King, H. Masuichi, and C. Rohrer. 2002. The parallel grammar project. In *Proceedings of COLING-2002 Workshop on Grammar Engineering and Evaluation*, pages 1–7.
- A. Copestake, F. Lambeau, B. Waldron, F. Bond, D. Lickinger, and S. Oepen. 2004. A lexicon module for a grammar development environment. In *Proceedings of the Linguistic Resources and Evaluation Conference 2004*, pages 1111–1114, Lisboa, Portugal.

- S. Evert, 2005. *The CQP Query Language Tutorial*. Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.
- A. Fitschen. 2004. *Ein computerlinguistisches Lexikon als komplexes System*. Ph.D. thesis, Universität Stuttgart: IMS. AIMS.
- O. Gurevich, R. Crouch, T.H. King, and V. de Paiva. 2007. Deverbal nouns in knowledge representation. *Journal of Logic and Computation Advance Access*, December 20.
- G. Helbig and J. Buscha. 2005. *Deutsche Grammatik: Ein Handbuch für den Ausländerunterricht*. Berlin, Langenscheidt.
- M. Johnston and F. Busa. 1999. Qualia structure and the compositional interpretation of compounds. In E. Viegas, editor, *Breadth and depth of semantic lexicons*. Dordrecht: Kluwer.
- H. Kermes. 2003. *Offline (and Online) Text Analysis for Computational Lexicography*. Ph.D. thesis, Universität Stuttgart: IMS. AIMS.
- E. Lapshinova and U. Heid. 2007. Syntactic subcategorization of noun+verb multiwords: description, classification and extraction from text corpora. In *Proceedings of the 26th International Conference on Lexis and Grammar*, Bonifacio, Corse, October 2-6.
- W. Lezius, S. Dipper, and A. Fitschen. 2000. Imslex - representing morphological and syntactical information in a relational database. In U. Heid, S. Evert, E. Lehmann, and C. Rohrer, editors, *Proceedings of EURALEX*, pages 133–139, Stuttgart, Germany.
- H. Schmid, A. Fitschen, and U. Heid. 2004. Smor: A german computational morphology covering derivation, composition, and inflection. In *Proceedings of LREC-2004*, Lisboa. LREC.
- H. Schmid. 1994. Probabilistic part-of-speech tagging using decision trees. In *International Conference on New Methods in Language Processing*, pages 44–49, Manchester, UK.
- H. Schmid. 1999. Improvements in part-of-speech tagging with an application to german. In Susan Armstrong, Kenneth Church, Pierre Isabelle, Sandra Manzi, Evelyn Tzoukermann, and David Yarowsky, editors, *Natural Language Processing Using Very Large Corpora*, volume 11 of *Text, Speech and Language Processing*, pages 13–26. Kluwer Academic Publishers, Dordrecht.
- A. Spencer. 1991. *Morphological Theory*. Cambridge, Blackwell.
- G. Zifonun, L. Hoffmann, and B. Strecker. 1997. *Grammatik der deutschen Sprache*, volume 2. Berlin/New York: de Gruyter.
- A.M. Zwicky. 1985. Heads. *Journal of Linguistics*, pages 1–29, 21.

9. Appendix: Examples

nominal compounds	Example sentences
Argumentationsblock	Der letzte große Argumentationsblock dafür, warum Rumänien in die Nato aufgenommen werden müsse, verweist auf die negativen Folgen, die eine Zurückweisung für das Land selbst und die Region hätte.
Argumentationskette	Auch die Argumentationskette, dass konservative Gemüter vom Verein dringend benötigte Spenden beziehungsweise deren Zusage zurückgezogen haben sollen, ist wenig schlüssig, da es Einzelspendern kaum zusteht, über den Verbleib einer Mitarbeiterin des Vereins zu befinden.
Bedenkzeit	Die einst gewünschte Bedenkzeit, ob er das Amt annehmen solle, dünkt ihm im Rückblick jedenfalls viel zu lang, fünf Minuten hätten gereicht.
Beweisführung	Seine abenteuerliche Beweisführung, dass Politiker eben auch Partei für einzelne Interessen ergreifen müßten, bringt einen der anwesenden Journalisten auf die naheliegende Frage, ob man das so verstehen müsse, dass Politiker nur dann ihre Aufgaben erfüllen könnten, wenn sie sich einladen ließen.
Beweislage	Die Beweislage, dass die verschiedenen Stationen des Schlafes unterschiedliche Implikationen auf die Gesundheit haben, ist sehr schwach.
Beweislast	Die Beweislast dafür, dass seine Aufhebung nicht zu einer neuen Rüstungsspirale führt, liegt jedoch bei den Amerikanern.
Beweismittel	Als Beweismittel, dass sie in dem Kennedy-Landsitz gewesen war, nahm sie ein Foto und einen Notizblock mit mehreren Telefonnummern aus dem Kennedy-Haus mit, die später der Polizei übergeben wurden.
Beweispflicht	Die Beweisspflicht, dass die neuen Techniken sich nicht sozial schädlich auswirkten, müßte von den Frauen, denen man sie zuschiebe, zurückverwiesen werden an die Ärzte und an die Industrie.
Denkumweg	Über den Denkumweg, dass ein Komplott stets nach einem Tod verlangt, wird dem armen Helden so die eigene Sterblichkeit, das unaufhaltsame Zusammenlaufen des Lebens auf sein Ende erst quälend bewußt
Druckmittel	Das Druckmittel, dass die Grundgesetzänderung nur mit einer Zweidrittelmehrheit zustande kommt, soll nach Auffassung der Grünen nicht aus der Hand gegeben werden.
Erfahrungssatz	Nach dem rechnerischen Erfahrungssatz, dass bei der Bahn rund 100000 DM einem Arbeitsplatz entsprechen, dürften etwa 400 Arbeitsplätze betroffen sein.
Erläuterungsbemühen	Das Erläuterungsbemühen des Wissenschaftsrates, wer aus welchen Gründen in die Evaluierungskommission berufen wurde, hielt sich bisher in engsten Grenzen.
Glaubenssatz	Den Glaubenssatz, dass es per se schwule Kunst oder Literatur gäbe, verwerfen die Freunde des schwulen Museums.
Grundsatzzerwägung	Seine Grundsatzzerwägung, dass ein Einsatz bewaffneter Kräfte zur Wahrung von Menschenrechten ein Widerspruch in sich sei, offenbart eine engelhafte Logik, nicht von dieser Welt.
Hoffnungsschimmer	Ein Hoffnungsschimmer, dass die Grünen noch gebraucht werden könnten, kommt derzeit ausgerechnet aus der Union.
Motivsuche	Die Motivsuche, warum Intellektuelle sich einer Diktatur zur Verfügung stellten, bleibt für die Betroffenen und für die Beobachter ein bislang kaum aufgearbeitetes Thema.
Rechenkunststück	Mit lobbyistischen Rechenkunststücken, wer denn nun die Kosten am meisten treibe, kann dem Patienten Gesundheitssystem nicht geholfen werden.
Rechtslage	Die unklare Rechtslage, ob Altreifen nun Abfall oder Wirtschaftsgut sind, hat weiter Auswirkungen auf den Kreis Groß-Gerau.
Schreckgespenst	Das Schreckgespenst, dass deutsche Physiker Hitler eine solche Bombe in die Hand geben könnten, trieb die deutschen Emigranten um.
Ursachenforschung	Die Ursachenforschung, woher im Einzelfall das im Körper nachgewiesene Quecksilber stamme, sei schwierig, da zum Beispiel jemand, "der den Mund voller Amalgam hat", an sich höhere Werte habe.
Ursachenpotential	Aber das Ursachenpotential, dass sie gewählt wurden, ist natürlich in der Stadt vorhanden.

Table 7: Type (b) examples

nominal compounds	Example sentences
Anhaltspunkt	Einen konkreten Anhaltspunkt, dass Schirinowski mit den transferierten Millionen finanziert wurde, gebe es allerdings nicht, so Dierdorf.
Angelpunkt	Dreh- und Angelpunkt dafür, dass die ostdeutsche Chemieindustrie nördlich von Halle wettbewerbsfähig wird, ist aber zunächst die Fertigstellung der Raffinerie bei LeunEhrgeiz
Denkmodell	Das Denkmodell, dass Männer so etwas machen, gibt es noch nicht lange.
Ehrgeiz	Den Ehrgeiz, dass sie die jungen Menschen zu "wetterfesten Persönlichkeiten" erziehen, erwartet der Bundespräsident von allen Lehrern und von allen Schulen.
Gewährsmann	Als Gewährsmann, daß es wirklich so entsetzlich gewesen sei, mußte immer wieder Manfred Schreiber erhalten, der frühere Polizeipräsident von München.
Grundsatz	Der Grundsatz, wer bestellt, zahlt auch, gelte nicht mehr, statt dessen heiße es, "wir bestellen, andere zahlen", erklärte Seib.
Meinungsstreit	Bei dem heftigen Meinungsstreit darum, ob Roma im Lande Bremen ein Bleiberecht erhalten oder abgeschoben werden sollen, finden zwei wichtige Aspekte nicht genügend Beachtung.
Naserümpfen	Und auf das obligatorische Naserümpfen darüber, dass solcher Film niemals ins Fernsehen gehört, verzichte ich, lege das Sofa mit Zellstoff aus und leide, leide, leide zum x-ten Mal.
Rätselraten	Das Rätselraten, wer in der Regierung sitzen soll, geht damit los.
Rechtsgrundsatz	Wegen des bei uns geltenden Rechtsgrundsatzes, dass alle Einkommen gleich zu besteuern seien, können wir diesen Weg nicht gehen.
Schlagzeile	Doch die Schlagzeile, dass mit ihm ein neuer Minetti gefunden sei, liest er nie.
Schlußfolgerung	Die Schlußfolgerung der Autoren, dass die Kernenergie das Hilfsmittel gegen weiteren CO2-Anstieg in der Atmosphäre ist.
Streitfrage	Auch mit der anfänglichen Streitfrage, ob eine Zusammenarbeit mit der Polizei legitim ist, wird heute bundesweit pragmatischer umgegangen. In der Streitfrage, wann Japan endlich sein Importverbot für Reis aufhebt, hält sich Miyazawa bisher bedeckt.
Verfahrensweise	Mit der Verfahrensweise, dass ein Gremium aus neun Leuten die Vorabentscheidung über die künftige Gestaltung des Marktplatzes fällte, ist die CDU überhaupt nicht einverstanden.
Volksmund	Den Volksmund, dass Kleinvieh auch viel Mist machen kann, bestätigen die Fluggesellschaften mit ihren Vielfliegerprogrammen.
Warnsignale	Warnsignale, dass das Radrennen unter den Einfluß skrupelloser Geschäftemacher geraten sei, habe es in der Vergangenheit immer wieder gegeben, sagt sie.
Werbemythos	Der Werbemythos, dass man in seinem Wagen einsam Freiheit und Freizeit genießen kann, hat sich ja wohl schon lange in den Staus erledigt, meint der Erfinder der ersten "Nachbarschaftsautovermietung" Deutschlands Markus Petersen.
Wettstreit	In dem edlen Bonner Wettstreit, wer die Kunst des Nebelwerfens und Schaumschlagens am besten beherrscht, rückt ein Mann unaufhaltsam nach vorn: Rudolf Scharping.
Wortspiel	Das Wortspiel, dass erst Pantani und dann Ullrich wie angestochen über die Gipfel rasten, verbietet sich.
Wunderglaube	Der Wunderglaube, dass mit der Parzellierung der Betriebe in viele GmbHs sich bereits der Erfolg einstelle, kann einen um den Schlaf bringen.
Wunschtraum	Der zionistische Wunschtraum, dass Juden ihr eigenes Land bebauen sollten, führte zu einer Landwirtschaftspolitik, die unter den in Israel naturgegebenen Umständen jeder Vernunft widersprach.
Wunschvorstellung	Die sportliche Wunschvorstellung, dass die Besten der winterlichen Wettkampfsreihe auch beim punktuellen Großereignis als Sieger gefeiert werden, ist keine logische Konsequenz.
Zeitplan	Einen Zeitplan, wann das größte Versandhaus Europas an die Börse gebracht werden soll, gebe es allerdings noch nicht.
Zeitpunkt	Einen präzisen Zeitpunkt, wann diese Ziele erreicht sein sollen, wollte Schneider noch nicht nennen. Der Zeitpunkt, dass ich meinen Rücktritt ausgerechnet am Samstag erklärt habe, hängt damit zusammen, dass im Weser-Kurier am Samstag ein Artikel erschienen ist mit der Lesart: Der Franke hört ja sowieso auf, da kann man doch mal mit ihm reden, ob er das nicht eher tut.

Table 8: Type (c) examples