

C-WEP—Rich Annotated Collection of Writing Errors by Professionals

Cerstin Mahlow

Institut für Deutsche Sprache
mahlow@ids-mannheim.de

Abstract

This paper presents *C-WEP*, the Collection of Writing Errors by Professionals Writers of German. It currently consists of 245 sentences with grammatical errors. All sentences are taken from published texts. All authors are professional writers with high skill levels with respect to German, the genres, and the topics. The purpose of this collection is to provide seeds for more sophisticated writing support tools as only a very small proportion of those errors can be detected by state-of-the-art checkers. *C-WEP* is annotated on various levels and freely available.

Keywords: error collection, multi-layer annotation, grammar, German

1. Introduction

Current state-of-the-art checkers for grammar and spelling are aimed to support experienced authors writing in their native language. For German, however, only a very small proportion of errors in texts produced by professional authors can be detected. Most often, these authors introduce errors by revising and editing texts, we almost find no spelling errors or typos.

With this paper, we close the gap on error collections by professional writers of German. We focus on grammar issues. *C-WEP*, the Collection of Writing Errors by Professionals Writers of German provides seed information for the development of more sophisticated writing support tools beyond spell checking. It currently consists of 245 sentences with grammatical errors.¹ All sentences are taken from published texts. *Published text* is understood in a broad sense and includes advertisements, formal e-mail messages, late drafts of conference papers, etc. Most of the errors were found in well-respected German and Swiss newspapers, which have a copy-editing stage. The texts are written by various authors and cover various topics; however, all authors are professional writers with high skill levels with respect to German, the genres, and the topics.

Almost none of the errors in *C-WEP* are detected by the state-of-the-art grammar checkers in standard word processors (we used the Microsoft grammar checker (Heidorn, 2000) as used in MS Word and Hunspell² as used in Apple TextEdit or OpenOffice). The sentences in *C-WEP* only contain a small number of typos, most errors are “grammatical errors.” However, both checkers mark a high number of “spelling issues,” which are listed in table 1. Hunspell and the Microsoft grammar checker correctly find 5 spelling issues each (one is a name). None of the grammar issues are detected by Hunspell. The Microsoft checker correctly identifies 7 instances of repeated words and also provides the correct suggestion (deleting one instance). However, this is only a very small portion of the errors.

We first present the state of the art in error annotation in various fields in section 2. and then report on *C-WEP* in detail in section 3. In section 4. we look at use case for *C-WEP*.

2. Related Work on Annotating Errors

2.1. What to Annotate

Most of the available error corpora focus on texts either produced by learners writing in their native language (here the authors are typically very young) or in a second language (here we also find texts from adults who are already experienced writers in their L1).

Annotating errors has a long tradition in various fields. In writing research and rhetoric and composition studies, we have long-term studies, for example the US-wide studies by Lunsford and Lunsford (2008) on error types in texts by college students. The authors compared their findings with a study from the late 1980s by Connors and Lunsford (1988). The writing medium had changed from pen and paper to computers, which resulted in a drastic decrease of spelling errors in students’ final drafts: from over 30% to only 6.5% of all errors. This shows that spelling checkers help to reduce the number of spelling errors—but spelling errors are not the only type of errors.

In applied linguistics and second language acquisition, the collection of learner data started in the late 1980s (see for example (Gries and Stefanowitsch, 2007; Reznicek et al., 2013)). Learner data is also of interest in computational linguistics; for an overview, see Leacock et al. (2010). Here, the goal of coding errors in learner corpora is to improve grammar checkers (Gamon, 2010), to develop applications for computer-assisted language learning (Meurers et al., 2010), or to make syntactic parsers more robust (Foster and Vogel, 2004).

In general, error analyses for professional or experienced L1 writers familiar with topic, genre, and tools, are rare (for example, (Napolitano and Stent, 2009)). There is much less research on writing in German than in English; the Falko corpus³ is one of the few examples, but it too focuses on foreign language learners. We are not aware of any resource for studying errors made by professional L1 writers of German; *C-WEP* aims to fill this gap.

2.2. How to Annotate

In writing research, there has been little research on errors in the last decades. Studies work towards systemizing errors by developing coding schemes based on the written evidence researchers get from subjects’ texts. However, there are no agreed-upon coding schemes. Coding is influenced by the

¹ All examples used in this paper come from *C-WEP*.

² <http://hunspell.sourceforge.net/>

³ <https://goo.gl/Qd7uMv>

Error Type	System	marked	true positive	appr. suggestion	no sugg.	wrong sugg.
name	MS Checker	73	1	0	11	64
	Hunspell	82	1	1	22	59
word	MS Checker	15	5	4	3	8
	Hunspell	47	5	3	29	15
compound issue	MS Checker	6	0	0	0	6
	Hunspell	1	0	0	0	1
grammar issue	MS Checker	59	13	0	0	59
	Hunspell	0	0	0	0	0
grammar as spelling	MS Checker	7	7	6	0	1
	Hunspell	0	0	0	0	0

Table 1: Error detection by MS Checker and Hunspell. (Names include names for persons, places, and products.)

goal of developing pedagogical approaches to teach writing. Categories are often unsuited for automatic classification. In (applied) linguistics and (second) language acquisition, we find more awareness of coding schemes and reproducibility of studies. Lüdeling et al. (2005) argue for annotation on various levels. In research on errors in computational linguistics, coding schemes are mostly developed from scratch (for an overview, see (Díaz-Negrillo and Fernández-Domínguez, 2006)) without reference to coding schemes developed in writing research.

Nicholls (2003) describes the coding system for the Cambridge Learner Corpus (CLC), which has been widely adopted. The focus is on coding as close to the original intention of the student as possible. As with other approaches, the schema does not take into account the source of the error. Coding the error involves assigning a label for the error class as well as a target hypothesis for the erroneous part of the sentence.

3. C-WEP

In this section we describe the annotation scheme used in C-WEP and some important characteristics of the collection.

3.1. Annotation

C-WEP is primarily a collection of *sentences*. As we are only interested in erroneous language, we do not store the entire texts, but we record information about the source texts as metadata, so the context can be looked up if needed. For every sentence, we also add further metadata such as genre, author, and time of publication or creation. Each sentence is labeled according to the main type of the error(s) and the POS involved. The error is marked in the sentence—except for missing words: due to relatively free word order in German, missing words could be inserted in more than one place. These annotations are done manually.

For all sentences, at least one target hypothesis (i.e., a grammatically well-formed version) of the *complete* sentence is manually added as a further annotation layer. To produce a grammatical sentence, mostly changes at the point of the observable error are required, but changes in other positions in the sentence may be needed as well. Note that in many

cases there are several target hypotheses. For 245 sentences, we have 434 hypotheses in total.

As the texts were written by professional writers, we can assume that they are familiar with grammar and spelling rules in general—in fact, we find very few spelling errors—and that the errors were typically introduced during editing and revising. For some of the sentences we have access to earlier versions, so we know the intended reading. For most of the sentences, the editing actions involved are very obvious and we can use this information to produce a grammatical version or to reproduce the sentence before the editing took place. When reconstructing the writing process, we often can tell that the author obviously tried various versions—resulting in residual words—but we have no clear indication which of the versions was the intended one. In such cases we aim to recreate all plausible versions.

Our reasoning is that an automatic checker should apply the same approach and create all grammatical versions and let the author decide which one to use. The versions could be ranked by edit distance measured in words; ranking could also take into account the position of editing operations—e.g., edits later in the sentence would result in a higher rank. However, an even better way would be to prevent errors in the first place by supporting the author *during* editing.

On additional layers, the original sentences and the target hypotheses are automatically annotated with POS by using TreeTagger. We also parsed all sentences (the erroneous ones and the hypotheses) with the Mate parser resulting in parses in the CoNLL2009 format and a broad coverage LFG parser for German from the ParGram Project in Stuttgart and added the resulting trees as further annotation layers.

Mate parser always produces parse trees including all word forms in a sentence, parse trees for erroneous sentences have to be inspected manually to detect whether or not Mate (as used out of the box) is suited to detect the errors by producing unusual trees or dependencies. For 118 of the 245 erroneous sentences, the LFG parser produced a coherent parse, the other sentences could only be parsed in fragments. For 322 of the 434 hypotheses, the LFG parser produced a parse, for 112 we still get fragmented parses. These results show, that state-of-the-art parsers can be used to detect more errors than state-of-the-art checkers in word processors do, but

Type	Number	POS involved	Example
missing words	62	all, including phrases and clauses	(1) * Die Römer hatten es im Laufe ihrer Geschichte mehrmals gegnerischen Kriegselefanten zu tun. (2) * Vor anderthalb Jahre starb Sohn des Ehepaars im Alter von 16 Jahren während eines Familienurlaubs. (3) * Außerdem muss der Klub im nächsten Spieler-Draft auf zwei Ziehungen – auf eine in der ersten und eine in der dritten Runde.
duplicate words (discontinuously)	46	all, including phrases	(4) * Wir fliegen wie geplant am 30. Oktober nach Kairo fliegen .
agreement	42	particle + verb, verb + noun, noun phrase	(5) * Der Verband unabhängiger Schweizer Hochzeitsplaner (VSUH) bietet einen Diplomlehrgang zum Hochzeitsplaner durch .
word order	28	complex phrases, noun phrases, verb constructions	(6) * Von der grauhaarigen Frau sieht Christin nur zuckenden den Rücken . (7) * Versuchen Sie nicht, Ihren zu Partner verändern .
duplicate POS (non-identical words)	25	preposition, verb	(8) * Sie sind in auf einem fremden Netz, es gelten die folgenden max. Gebühren (in CHF) (9) * Beide Parteien, heißt es aus Lausanne weiter, haben hätten bereits je einen Vertreter aus dem geschlossenen Schiedsrichterpool des Cas benannt.
duplicate words (continuously)	11	all, including phrases	(10) * Ob es den den Jungen wirklich gibt oder ob er nur eine Phantasiegestalt ist, spielt keine Rolle. (11) * Kurz nach dem nach dem Abitur traf ich meinen ehemaligen Direktor im Pressecafé in der Friedrichstraße.
semi-duplicate words	3	noun compounds, particle verbs (part reappears)	(12) * Der Brieffund bedeutet viel für die Reformations forschung Forschung .

Table 2: Error typology of C-WEP with frequency and examples.

also those are not able to detect all errors a human would recognize.

3.2. Characteristics

Mahlow (2015) argues for a typology of writing errors that takes the source of the error into account. In this paper, we focus on the surface only and present a preliminary typology of errors in texts from professional writers based on C-WEP, using the CLC coding scheme. The error classes are tailored towards automatic detection by NLP methods and are listed

in table 2⁴. Unfortunately, the assignment of error classes is compounded by the fact that sentences can contain multiple errors.

For some errors, it is easier to come up with grammatical variants than to classify the error:

- (13) a. * Laut Polizei hatte die Frau
According to the police did the woman
gegen 10.50 Uhr zu Fuß die Kreuzung in
at 10.50 hours on feet the junction at

⁴ Glosses and shallow English translations are given in the appendix.

Höhe Platz der Freundschaft überqueren
 height Platz der Freundschaft cross
 wollen, dabei aber nicht auf die
 want, in the process but not at the
 heranfahrende Straßenbahn der Linie 5
 approaching tram of the line 5
 nicht bemerkt.
 not noticed.

According to the police, the woman wanted to cross the junction near Platz der Freundschaft and didn't at the approaching tram 5 didn't notice.

- b. Laut Polizei hatte die Frau gegen 10.50 Uhr zu Fuß die Kreuzung in Höhe Platz der Freundschaft überqueren wollen, dabei aber nicht auf die heranfahrende Straßenbahn der Linie 5 geachtet.
- c. Laut Polizei hatte die Frau gegen 10.50 Uhr zu Fuß die Kreuzung in Höhe Platz der Freundschaft überqueren wollen, dabei aber die heranfahrende Straßenbahn der Linie 5 nicht bemerkt.

The cause for the error is most likely that the author tried several variants (the use of different verbs in the example above), focusing on one aspect and ignoring side effects (here one but not the other verb is used with a prepositional phrase). As for identifying the error type, there are too many words including the duplicate *nicht*, and the error involves both the POS verb and preposition. However, depending on the target hypothesis, the error is only in the verb or only in the preposition. Table 3 shows the distribution of involved POS in the errors contained in the current version of C-WEP.

POS	Number
Verb	108
Preposition	50
Determiner	36
Noun	29
Pronouns	25
Adjective/Adverb	20
Particles/Konjunctions	12

Table 3: POS involved in errors. Note that some errors involve several POS and that some sentences contain more than one error.

4. Effort for Fixing Errors

Detecting agreement errors and duplicate words is relatively easy when using state-of-the-art parsers—these sentences are not well formed. In principle, fixing such errors is not that hard: real duplicates occurring directly after one another can be removed; for duplicates with a longer distance, a parser should decide which of them could be removed to result in a well-formed sentence. It might be necessary to remove more than the duplicate word—a step-wise approach would be needed for those cases. For semi-duplicate words, the user should decide which version to keep: the more general one or the more specific one. For agreement errors, we could give more weight to the verb, so particles, prepositional phrases

and nouns have to be adjusted for agreement. When citations from direct speech are involved, preference should be given to the verb in this direct speech sequence and all other parts are governed by its tense and mood.

However, correcting sentences with missing words is much harder. Missing determiners and prepositions should be relatively easy to fix—there is also much research in those areas as these are typical error types for learners of German. For missing verbs or phrases, however, semantic information is needed. One not only has to guess the missing word as such, but also where to place it—the relatively free word order in German often allows more than one possibility. The semantic information might come from the sentence itself—i.e., the direct context of the missing verb—, from the text itself—i.e., the context of the sentence—, or from implicit information such as genre, author, or topic.

Cases like example 14 are almost impossible to solve when starting from the erroneous sentence:

(14) a. [*]

Demnach wurde Pastior, der unter
 Accordingly had Pastior, who amongst
 anderem wegen einiger
 other things because of some
 antisowjetischer Gedichte, die er während
 anti-soviet poems, that he during
 seiner Gefangenschaft geschrieben hatte, am
 his time in prison written had, on
 8. Juni die Zusage zur Mitarbeit
 8 June the agreement to collaboration
 abgenötigt.
 extorted.

Accordingly, Pastior, who [missing verb phrase], because of some anti-soviet poems he had written during his time in prison, was forced to agree to collaborate on June 8.

- b. Demnach war Pastior unter anderem wegen einiger antisowjetischer Gedichte, die er während seiner Gefangenschaft geschrieben hatte, ins Visier der Securitate geraten. Am 8. Juni 1961 wurde ihm dann die Zusage zur Mitarbeit abgenötigt.

Because of the structure of the sentence, the error is hard to detect in the first place. If we skip modifiers, it is example 15a. We can also skip the wellformed relative clause and get sentence 15b, then the error is obvious.

(15) a. [*]

Demnach wurde Pastior, der wegen
 Accordingly had Pastior, who because of
 einiger Gedichte, die er während seiner
 some poems, that he during his
 Gefangenschaft geschrieben hatte, die
 time in prison written had, the
 Zusage abgenötigt.
 agreement extorted.

Accordingly, Pastior, who because of some poems he had written during his time, was forced to agree.

b. [*]

Demnach wurde Pastior, der wegen Accordingly had Pastior, who because of einiger Gedichte, die Zusage abgenötigt. some poems, the agreement extorted.

Accordingly, Pastior, who because of some poems, was forced to agree.

In this particular case we actually know the correct version 14b, as both sentences (found in two different newspaper articles) are based on a single news agency report (by DPA). Taking also into account that the correct version was found in an article published one day before the article with the error, we can be confident that the error occurred while merging two sentences, probably triggered by space restrictions. We cannot fix this error without having access to the original version; instead it would have been better if the author had received some help while merging the sentences to prevent this error from happening in the first place.

5. Conclusion

In this paper we have briefly described C-WEP, the Collection of Writing Errors by Professionals Writers of German, and some of its characteristics and potential uses. One obvious use is as a gold standard for the development and improvement of grammar checkers. A useful grammar checker should be able to detect all of the errors in C-WEP and it should offer suggestions for most of the errors. C-WEP can also serve as starting point for deeper investigations on writing errors by professional writers.

C-WEP is made available under a CC license in XML format which allows easy transformation into other formats. We maintain the collection at www.lingured.info/linguistic-resources/cwep/ and are adding further annotation layers and further erroneous sentences. The license also allows contributions by others. We are also working to make C-WEP accessible via ANNIS.⁵

Acknowledgement

We thank Özlem Cetinoglu and Wolfgang Seeker from the IMS Stuttgart for helping with parsing C-WEP.

6. Bibliographical References

- Connors, R. J. and Lunsford, A. A. (1988). Frequency of Formal Errors in Current College Writing, or Ma and Pa Kettle Do Research. *College Composition and Communication*, 39(4).
- Díaz-Negrillo, A. and Fernández-Domínguez, J. (2006). Error tagging systems for learner corpora. *Revista Española de Lingüística Aplicada*, 19:83–102.
- Foster, J. and Vogel, C. (2004). Parsing Ill-Formed Text Using an Error Grammar. *Artificial Intelligence Review*, 21(3-4):269–291.

- Gamon, M. (2010). Using Mostly Native Data to Correct Errors in Learners' Writing. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 163–171, Stroudsburg, PA, USA, June. Association for Computational Linguistics.

- Stefan T. Gries et al., editors. (2007). *Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis*. De Gruyter Mouton, Berlin, New York.

- Heidorn, G. E. (2000). Intelligent writing assistance: Techniques and applications for the processing of language as text. In Robert Dale, et al., editors, *Handbook of Natural Language Processing*, pages 181–207. Marcel Dekker, New York, NY, USA.

- Leacock, C., Chodorow, M., Gamon, M., and Tetreault, J. (2010). *Automated Grammatical Error Detection for Language Learners*, volume 9 of *Synthesis Lectures on Human Language Technologies*. Morgan & Claypool, San Rafael, CA, USA.

- Lüdeling, A., Walter, M., Kroymann, E., and Adolphs, P. (2005). Multi-level error annotation in learner corpora. In *Proceedings of Corpus Linguistics 2005*.

- Lunsford, A. A. and Lunsford, K. J. (2008). "Mistakes Are a Fact of Life": A National Comparative Study. *College Composition and Communication*, 59(4).

- Mahlow, C. (2015). Learning from Errors: Systematic Analysis of Complex Writing Errors for Improving Writing Technology. In Núria Gala, et al., editors, *Language Production, Cognition, and the Lexicon*, volume 48 of *Text, Speech and Language Technology*, pages 419–438. Springer International Publishing.

- Meurers, D., Ziai, R., Amaral, L., Boyd, A., Dimitrov, A., Metcalf, V., and Ott, N. (2010). Enhancing Authentic Web Pages for Language Learners. In *Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 10–18, Stroudsburg, PA, USA, June. Association for Computational Linguistics.

- Napolitano, D. M. and Stent, A. (2009). TechWriter: An evolving system for writing assistance for advanced learners of English. *CALICO Journal*, 26(3):611–625.

- Nicholls, D. (2003). The Cambridge Learner Corpus: Error coding and analysis for lexicography and ELT. In *Proceedings of the Corpus Linguistics 2003 conference*, pages 572–581.

- Reznicek, M., Lüdeling, A., and Hirschmann, H. (2013). Competing target hypotheses in the Falko Corpus. a flexible multi-layer corpus architecture. In Ana Díaz-Negrillo, et al., editors, *Automatic Treatment and Analysis of Learner Corpus Data*, pages 101–123. John Benjamins, Amsterdam.

⁵ <http://corpus-tools.org/annis/>

A Glosses

This appendix contains the glossed examples and at least one target hypothesis for each sentence. The translations attempt to approximate the errors found in the German original as far as possible in English.

- (1) a. * Die Römer hatten es im Laufe ihrer Geschichte mehrmals gegnerischen Kriegselefanten zu tun.
The Romans had it in the course of their history several times enemy war elephants to do.
During their history, the Romans had to deal with enemy war elephants several times. (missing preposition)
- b. Die Römer hatten es im Laufe ihrer Geschichte mehrmals **mit** gegnerischen Kriegselefanten zu tun.
- (2) a. * Vor anderthalb Jahren starb Sohn des Ehepaars im Alter von 16 Jahren während eines Familienurlaubs.
Ago one and a half years died son of the couple at the age of 16 years during a family vacation.
One and a half years ago, son of the couple died at age 16 during a family vacation. (missing determiner (and adjective)).
- b. Vor anderthalb Jahren starb **ein** Sohn des Ehepaars im Alter von 16 Jahren während eines Familienurlaubs.
- c. Vor anderthalb Jahren starb **der** Sohn des Ehepaars im Alter von 16 Jahren während eines Familienurlaubs.
- d. Vor anderthalb Jahren starb **der älteste** Sohn des Ehepaars im Alter von 16 Jahren während eines Familienurlaubs.
- (3) a. * Außerdem muss der Klub im nächsten Spieler-Draft auf zwei Ziehungen.
Additionally must the club in the next player draft on two draws.
Additionally, the club must without two draws at the next drafts. (missing main verb)
- b. Außerdem muss der Klub im nächsten Spieler-Draft auf zwei Ziehungen **verzichten**.
- (4) a. * Wir **fliegen** wie geplant am 30. Oktober nach Kairo **fliegen**.
We fly as planned on 30 October to Cairo fly.
As planned, we fly on October 30 to Cairo fly.
- b. Wir **fliegen** wie geplant am 30. Oktober nach Kairo.
- c. Wir **werden** wie geplant am 30. Oktober nach Kairo **fliegen**.
- (5) a. * Der Verband unabhängiger Schweizer Hochzeitsplaner (VSUH) **bietet** einen Diplomelehrgang zum Hochzeitsplaner **durch**.
The union of independent Swiss wedding planners (VSUH) **bietet** a diploma course towards wedding planner **durch**.
out.
The union of the independent Swiss wedding planners (VSUH) offers a diploma course on wedding planning out. (offers vs. carries out)
- b. Der Verband unabhängiger Schweizer Hochzeitsplaner (VSUH) **bietet** einen Diplomelehrgang zum Hochzeitsplaner **an**.
- c. Der Verband unabhängiger Schweizer Hochzeitsplaner (VSUH) **führt** einen Diplomelehrgang zum Hochzeitsplaner **durch**.
- (6) a. * Von der grauhaarigen Frau sieht Christin nur **zuckenden den Rücken**.
Of the gray-haired woman sees Christin only shrugging the back.
Christin sees only shrugging the back of the gray-haired woman.
- b. Von der grauhaarigen Frau sieht Christin nur **den zuckenden Rücken**.
- (7) a. * Versuchen Sie nicht, Ihren **zu** Partner **verändern**.
Try you not, your to partner **verändern**.
change.
Don't try change to your partner.
- b. Versuchen Sie nicht, Ihren Partner **zu verändern**.
- (8) a. * Sie sind **in auf** einem fremden Netz, es gelten die folgenden max. Gebühren (in CHF).
You are in on a foreign net, it apply the following max. charges (in CHF).
You are in on a foreign net, the following charges apply (in CHF)
- b. Sie sind **in** einem fremden Netz, es gelten die folgenden max. Gebühren (in CHF)
- (9) a. * Beide Parteien, heißt es aus Lausanne weiter, **haben hätten** bereits je einen Vertreter aus dem geschlossenen Schiedsrichterpool des Cas benannt.
Both parties, say it from Lausanne further, **haben hätten** already each one representative from the closed pool of referees of the Cas named.
Lausanne says that both parties already **have would have** named one representative from the closed pool of referees of the Cas.
- b. Beide Parteien, heißt es aus Lausanne weiter, **haben** bereits je einen Vertreter aus dem geschlossenen Schiedsrichterpool des Cas benannt.
- c. Beide Parteien, heißt es aus Lausanne weiter, **hätten** bereits je einen Vertreter aus dem geschlossenen Schiedsrichterpool des Cas benannt.

- (10) a. * Ob es **den den** Jungen wirklich gibt
 Whether it the the boy really gives
 oder ob er nur eine Phantasiegestalt
 or whether he only one phantasy figure
 ist, spielt keine Rolle.
 is, plays no role.
 Whether the the boy really exists or whether he's
 only an imagined figure doesn't play a role.
- b. Ob es **den** Jungen wirklich gibt oder ob er nur
 eine Phantasiegestalt ist, spielt keine Rolle.
- (11) a. * Kurz **nach dem nach dem** Abitur traf
 Shortly after the after the Abitur met
 ich meinen ehemaligen Direktor im
 I my former director in the
 Pressecafé in der Friedrichstraße.
 Pressecafé in the Friedrichstraße.
 Shortly after the after the Abitur, I met my former
 director in the Pressecafé on Friedrichstraße.
- b. Kurz **nach dem** Abitur traf ich meinen ehemali-
 gen Direktor im Pressecafé in der Friedrichstraße.
- (12) a. * Der Brieffund bedeutet viel für
 The finding of the letter means much for
 die Reformations**forschung** **Forschung**.
 the research on reformation research.
 The letter means a lot for the research on refor-
 mation research.
- b. Der Brieffund bedeutet viel für die **Reformati-
 onsforschung**.
- c. Der Brieffund bedeutet viel für die **Forschung**.