

*Sandra Hansen, Ralph Dirksen, Martin Küchler,
Kerstin Kunz, Stella Neumann**

**Comprehensible legal texts
– utopia or a question of wording?
On processing rephrased German court decisions**

Abstract

This paper presents a study on the comprehensibility of rephrased syntactic structures in German court decisions. While there are a number of studies using psycholinguistic methods to investigate the comprehensibility of original legal texts, we are not aware of any study looking into the effect resolving complex structures has on the comprehensibility. Our study combines three methodological steps. First, we analyse an annotated corpus of court decisions, press releases and newspaper reports on these decisions in order to detect those complex structures in the decisions which distinguish them from the other text types. Secondly, these structures are rephrased into two increasingly simple versions. Finally, all versions are subjected to a self paced reading experiment. The findings suggest that rephrasing greatly enhances the comprehensibility for the lay reader.

1. Introduction

In the course of the last thirty years we have seen many contributions to the question of how to make the legal language more comprehensible. In the Anglophone world the plain language movement is quite influential, in Germanic research – on which we will mainly concentrate here since we are concerned with the German language – the work of the interdisciplinary working group “Analyse der Juristischen Sprache” (cf. Rave et al. 1971) from the beginning of the 1970s onwards has

* The authors would like to thank the reviewer for valuable comments on the paper. Of course, he cannot be held responsible for any remaining weaknesses.

* *Sandra Hansen, Ralph Dirksen, Martin Küchler, Kerstin Kunz, Stella Neumann*
Universität des Saarlandes
FR 4.6
Universitätscampus Saarbrücken A2 2
D-66123 Saarbrücken

to be mentioned and most recently the working group at the Berlin-Brandenburgische Akademie der Wissenschaften (cf. Lerch 2004a). Lexical and syntactic specificities of legal language are described such as nominal style, longer than average sentences which are said to contribute to the incomprehensibility of the language of law (cf. Wagner 1981, Oksaar 1988). In this paper we concentrate on the language of court decisions whose linguistic characteristics are described in an example-based way by Altehenger (1983), contrasting German and Danish decisions by Engberg (1997) and using corpus-linguistic methods by Hansen-Schirra & Neumann (2004). This latter study confirms the specificities of longer sentences, nominalisations and highly complex noun phrases. These syntactic elements previously only described on an exemplary basis were analysed in a quantitative way in comparison to a reference corpus consisting of samples from 15 text types.

Building on work like this, several studies investigated legal language with the help of methods for researching comprehensibility. Basedow (1999) reports on a word-based study using the Flesch-test (Flesch 1948). This readability formula developed in the 1940s allows quantifying criteria like word and sentence length. Nevertheless, the readability approach is problematic because it is limited to counting syllables and words per sentence and completely ignores the semantic content of the text (a short sentence may still be difficult to understand because it consists of short, but rare words). In the framework of research on comprehensibility it is therefore regarded as obsolete (cf. Rickheit 1995; see also Lerch 2004b for a review of the readability approach for measuring the comprehensibility of the law).

The above mentioned working group “Language and Law” at the Berlin-Brandenburgische Akademie der Wissenschaften did a psycholinguistic experiment aiming at the investigation of the general knowledge background of the participant. The complex research design used various methods like Thinking Aloud Protocols to look into the transparency rule in the German legislation on standard terms and conditions. It investigated deeper coherence structures in standard terms and conditions of insurance contracts. Unfortunately, the results of this highly interesting study are not published yet. Therefore it is not possible to draw any consequences from the findings.

Neumann & Hansen-Schirra (2004) used an acceptability judgment experiment (Gernsbacher 1994) in a pilot study to test the comprehensibility of syntactic peculiarities in German court decisions. The German verb mood *Konjunktiv*¹ was chosen as a register-specific feature of court decisions as described by Altehenger (1983). This pilot study showed the use of the combination of corpus-linguistic methods establishing the grammatical characteristics of a given register with psycholinguistic experiments testing the comprehensibility of the specific features. The described results served as a basis for the research discussed in the present paper.

So far, we have discussed legal language as such. However, this is a cover term which has to be specified for our purposes. We have to distinguish between at least legislative texts, administrative texts and court decisions. These text types differ in view of the authors, the addressees as well as the function of the texts. Jaspersen (1998) discusses the comprehensibility of legislative texts. There are a number of linguistic guidebooks for German administrations (cf. „Bürgernahe Verwaltungssprache“ by the Bundesverwaltungsamt² and publications resulting from the project „Bürgerfreundliche Verwaltungssprache“³). Furthermore, Grönert (2004) reports on research geared towards improving the communication between administrations and citizens.

Court decisions are written by jurists and are not addressed at a clearly identifiable recipient but – apart from the parties involved in the case – rather at unspecific recipients subdivided into two groups. Court decisions are addressed both at legal experts who work with the texts and at the lay citizen who is supposed to accept the decision and abide by it. The needs of the two recipient groups partly diverge: Experts expect to be supplied with concise and precise information efficiently packed into typical terms and syntactic constructions. On the other side, the lacking specialised knowledge of citizens not legally trained requires a more elaborate – in the view of the expert probably lengthy – presentation of information. As this type of legal texts particularly draws public attention (cf. Jaspersen 1998) it should be comprehensible to a large

¹ We intentionally use the German term for the subjunctive mood to make the functional difference to the English subjunctive clear.

² http://www.bva.bund.de/imperia/md/content/bbb_win/allgemeines/149.pdf

³ <http://www.ruhr-uni-bochum.de/vt/>

group of recipients, i.e. both, experts and citizens. For this reason, our main interest lies in investigating the comprehensibility of court decisions by tracking down their linguistic complexity.

Our study focuses on how readers process syntactic specificities found to be typical of German court decisions. The aim of the study is to quantify syntactic specificities of German court decisions, to rephrase the complex structures into simplified versions and finally to test these version psycholinguistically. Consequently, the research setup combines corpus-linguistic and psycholinguistic methods against the background of legal knowledge. This interdisciplinary research design is reflected in the authors' affiliations with not only linguistics but also law and psychology.

The remainder of the paper is organised as follows. In the following section 2, we will present the design of the study with the preparatory steps for the psycholinguistic test, namely the corpus-linguistic analysis of the texts as well as the process of rephrasing three types of syntactic specificities conforming to the legal content of the original wording. In section 3, we will present the psycholinguistic experiment in detail and discuss the results. Finally, we draw conclusions from this study and look at consequences for future research in section 4.

2. The study

2.1. Corpus-based analysis

Corpus design

As mentioned above, the first step to analysing the syntactic complexity of German legalese consists in a corpus-linguistic analysis. We intend to identify and quantify the syntactic features causing this complexity. For this purpose, we compare intralingual versions of German legal texts. Hence, our corpus of investigation consists of decisions of the German Federal Constitutional Court and press releases and newspaper reports on these decisions. One main interest for looking into the three versions resides in the assumption that, serving different purposes, they may display different linguistic properties, in particular, that they vary in their syntactic complexity: We assume the decisions to be the most complex version as they contain language for specific purposes. They constitute a written version of the oral pronouncement in court. They

display a complexity of legal content in combination with linguistic complexity that may manifest on every linguistic level.

The newspaper reports are assumed to be the version displaying the least complex syntactic structures. They are expected to exhibit more general language features as they are designed to be the citizen's everyday reading. Further on, we expect the press releases to be of medium linguistic complexity as this version may be influenced by both, court decisions and newspaper reports: On the one hand, as they were written by legal experts, the press releases may contain specific features of German legalese. On the other hand, as they are considered to be rephrased variants of the court decisions for specialised journalists they may show general language features of newspaper texts.

Research design

Complex grammatical structures are major indicators of the linguistic complexity of texts. Therefore, we focus on the empirical analysis of the syntactic properties of the above mentioned intralingual versions of German legal texts. We apply a range of corpus-linguistic methods to verify and intensify the hypothesis about properties of legal and/or administrative texts stated in example-based studies (cf. Wagner 1981 for administrative texts and Altehenger 1983 for court decisions).

We intend to gain quantitative results about the syntactic properties responsible for the varying degrees of complexity in the three intralingual versions. In addition, the annotation results are supposed to reflect which features on which syntactic level are unique to the court decisions and which features are modified or deleted in the press releases and newspaper reports respectively. We interpret syntactic complexity by investigating the following features:

- Sentence length
- Embedding in sentences
- Length of noun phrases and prepositional phrases
- Embedding of noun phrases and prepositional phrases
- Nominalisations

For the purpose of obtaining more information about the syntactic complexity on sentence and phrase level, we use the notion of field topology which was especially developed for investigating the structure of the German sentence. It constitutes a relatively theory-neutral description

of German syntax and considers its flexible constituent structure. The German sentence is split up into smaller parts with regard to the distributional properties of the verb complex – so called topological fields: ‘Vorfeld’, ‘linke Satzklammer’ (finite part of the verb complex), ‘Mittelfeld’, ‘rechte Satzklammer’ (non- finite part of the verb complex) and ‘Nachfeld’.

We combine automatic parsing and fine-grained manual annotation to get detailed information about the linguistic complexity on the syntactic levels of the three versions. We analyse sentences from the three corpora using a topological parser (Braun 1999). It is based on the notion of field topology for German syntax as mentioned above. The parser structures the sentence into a series of neighbouring and embedded topological fields. The noun phrases and prepositional phrases are analysed manually using an XML editor. For practical reasons, only the ‘Vorfeld’ is annotated: The ‘Vorfeld’ constitutes the leftmost part of the German sentence, in front of the finite verb, and is expected to contain mainly noun and prepositional phrases. Furthermore, we annotate only phrases which have a minimum length of seven tokens (newspaper reports) or ten tokens (decisions and press releases) as we expect that phrases of this length already have a certain complexity. Both the automatic output of the parser and the output of the manual annotation are double-checked by the annotators.

We query nominalisations with the help of a concordance tool which displays all matching constructions. We count all tokens which are classified as nouns by a part-of-speech tagger and which contain the suffixes "-ung", "-ion", "-ismus", "-heit", "-keit", "-ität", "-schaft" as well as the respective plural forms.

Findings

In the following, we present some of the results of our corpus-linguistic analysis before we move on to a general interpretation. We start by looking into the syntactic complexity on sentence level.

	Decisions	Press releases	Newspaper reports
Av. sentence length in token	24.32	20.31	14.89
No. of subordinate clauses per sentence	0.57	0.50	0.34
Embedding			
Level 0	56.77 %	62.78 %	68.33 %
Level 1	34.11 %	29.42 %	28.47 %
Level 2	7.90 %	6.58 %	3.20 %
Level 3	1.13 %	1.22 %	0.00 %
Level 4	0.08 %	0.00 %	0.00 %

Table 1: Complexity on sentence level

Table 1 displays the annotation results concerning the complexity of the sentence in the court decisions, press releases and newspaper reports, respectively. Considering the average number of tokens per sentence in the three subcorpora we find longer sentences in the court decisions than in the press releases. The newspaper reports by far have the shortest sentences. In addition, more than every second sentence in the decisions also contains a subordinate clause, whereas in the press releases every second, and in the newspaper reports every third sentence contains a subordinate clause. The higher number of subordinate clauses in the court decisions indicates that the sentences contain more embedding than the press releases, and even far more embedded clauses than the newspaper reports. This tendency can be confirmed when looking at the percentage of embedded clauses on the respective levels of embedding. The lower percentage of clauses on Level 0 in the court decisions tells us that there are more clauses to be found on deeper levels of embedding. In most instances, the percentage of embedded subordinate clauses on deeper levels is higher in the court decisions than in the press releases, and the newspaper reports clearly contain less embedded clauses on the deeper levels. Generally speaking, the findings from our analysis of the sentences reflect an extreme syntactic complexity to be found in the court decisions. The newspaper reports show a tendency towards much simpler sentence constructions, whereas the press releases are of medium complexity on sentence levels. We can conclude that extreme syntactic complexity on sentence level is one prominent feature of court decisions, especially compared to newspaper reports.

We now consider the syntactic complexity on phrase level. Table 2 displays the annotation results for the complexity of the phrases in the court decisions, press releases and newspaper reports, respectively.

	Decisions	Press releases	Newspaper reports
Max. phrase length in token	62	52	27
Av. phrase length in token	5.80	4.69	3.71
Embedding			
Level 0	29.70 %	24.92 %	48.51 %
Level 1	36.11 %	42.25 %	39.55 %
Level 2	25.21 %	22.49 %	10.45 %
Level 3	6.62 %	7.60 %	1.49 %
Level 4	2.14 %	2.74 %	0.00 %
Level 5	0.21 %	0.00 %	0.00 %

Table 2: Complexity on phrase level

Comparing the length of phrases in the three intralingual versions as shown in Table 2, we see that the court decisions have the longest phrases and that the newspaper reports have the shortest phrases. The average phrase length in the press releases lies between the other two versions.

In order to obtain the depth of embedding, we count all head nouns on the respective levels of embedding. The results for the newspaper reports correspond to our expectations: Most of the noun phrases and prepositional phrases appear on level 0. On level 1 the percentage of embedded phrases is much higher than on level 2 and level 3. Apart from that, they contain only four levels of embedding, whereas the press releases have five levels and the court decisions even six levels of embedding. Comparing the press releases and the court decisions, two peculiarities have to be noted: First, the percentage of phrases on level 0 in the press releases is lower than in the court decisions. Second, on level 1, the percentage of embedded phrases is higher in the press releases than in the decisions and the difference from level 0 to 1 is even bigger than in the decisions. This may be due to terminology and paragraphs in the court decisions being explained in the press releases by using embedded phrases. Apart from that, the higher percentage of phrases on level two in the decisions and press releases may be caused

by coordinations within phrases. This tendency towards organising phrases in extremely compact structures may result from the need to present the information in a concise way: Explanations and further information on the court decisions may be packed into phrases rather than being realised by verbal constructions which would require more space. It thus leads to a high nominal density.

Finally, the amount of nominalisations as deverbal derivations to be found in each intralingual version is also an indicator for more or less syntactic complexity.

	Decisions	Press releases	Newspaper reports
All nominalisations	7.15 %	7.30 %	4.54 %
„ung“-nominalisations	5.31 %	5.56 %	3.30 %

Table 3: amount of nominalisations

Table 3 shows that 7.15 % of all nouns in the court decisions are nominalisations. The percentage of nominalisations in the press releases is even higher (7.30 %). This higher proportion of nominalisations may also be an indicator of the above described tendency to compress information: press releases have to inform the recipient in a very condensed way. The newspaper reports exhibit a much lower amount of nominalisations (4.54 %) than the two other intralingual versions.

Our data clearly show that derivations on "-ung" are the most frequent German nominalisations in all three intralingual versions. For this reason, we concentrate on these forms when elaborating rephrased versions of nominalisations as will be explained in section 2.2.

Summarising our findings, we can say that the annotation results mostly confirm our hypotheses concerning the syntactic complexity of German legal texts. Comparing the court decisions, press releases and newspaper reports the following important tendencies have to be outlined:

- The court decisions display more complexity on most syntactic levels than the press releases and the reports: They have the longest sentences, they contain the highest number of subordinate clauses and the highest depth of embedding. They also contain the longest noun phrases and prepositional phrases. Furthermore, on most lev-

els, the decisions show more embedded phrases than the press releases and the newspaper reports and they contain a high number of nominalisations. We assume that this high complexity on all syntactic levels has a negative effect on the comprehensibility of the court decisions. It complicates processing and risks destructing the recipient's capacity of storing information in the short term memory.

- The press releases' complexity lies between that of the court decisions and the newspaper reports. In some cases, they approximate the court decisions or are even more complex: They contain more embedded noun phrases on level two and have slightly more nominalisations than the court decisions. These factors indicate a tendency to informational density similar to that in the court decisions: Information is packed into heavy noun phrases and nominalisations rather than being distributed onto larger grammatical units like sentences. Therefore, we assume that the syntactic structures in the press releases, on most levels, are still too complex for lay persons. And this complexity still affects the reception process. For this reason, the press releases cannot be used as a starting point to resolve the complexity of the court decisions.
- The newspaper reports are less complex than the court decisions, and are still clearly less complex than the press releases on all investigated syntactic levels. As stated in our assumptions above, they mostly contain general language features with which lay persons as readers are confronted every day. We assume that the newspaper reports are much easier to process than the two other intralingual versions. Therefore, they can serve as a yardstick for elaborating methods to rephrase the court decisions, as will be explained in the following section.

2.2. Rephrasing register-specific syntactic constructions

For the rephrases, it is necessary to restrict the number of syntactic features as well as the number of versions to be elaborated. We thus concentrate on the three syntactic features of deeply embedded sentences (abbreviated „S”), deeply embedded phrases („P”) and „-ung”-nominalisations („N”). With a view to the feasibility of the psycholinguistic experiment, we restrict ourselves to two rephrases (versions B

and C) of the original version A retrieved from the corpus of court decisions. Our goal is to work out three degrees of complexity: a highly complex version A, a medium complex version B and a simple version C. All versions have to have the same legal contents. This is guaranteed by the interdisciplinary team consisting of both jurists and linguists.

Facing the highly complex structures of the court decisions, the question arises how to delimit the rephrased versions. The medium complex version follows the idea of an optimal approach to comprehensibility as proposed by Groeben & Christmann (1989). This approach claims that texts which conform to the reader's expectations – for instance by means of maximally simple structures (see below) – do not offer any cognitive stimulus to the reader. Extremely simplified texts are said to destruct the recipient's motivation to keep reading (cf. Groeben & Christmann 1989:175). We apply this optimal approach to comprehensibility to the medium version by taking the corpus results of the newspaper reports as a benchmark. As previously mentioned, we assume that newspapers use a range of language well adapted to the reading habits of the lay reader⁴ – or, to put it the other way round – newspapers form their reading habits. Thus, structures typical for newspaper reports should be familiar to lay readers of court decisions and still challenge them enough to pay attention to the unfolding text.

The simple version C realises a maximum strategy: the complex sentence, for instance, is broken down to one clause per sentence. This corresponds to the maximum approach to comprehensibility as advocated by Langer et al. (1974). They report on a study in which experts score texts in view of the four dimensions 'linguistic simplicity', 'structure-organisation', 'brevity-shortness' and 'interest-liveliness'. They argue that the higher a text is scored with respect to these dimensions the better the text will be memorised. We apply their dimension of simplicity to the three syntactic features under investigation in our study.

The rephrased syntactic dimensions in detail

As shown in the corpus analysis, the authors of court decisions make use of more complex and embedded sentence structures than the au-

⁴ Particularly those readers participating in our psycholinguistic experiment, see section 3

thors of the other text types in the corpus. Initially, those structures were rephrased which displayed the highest level of embedding on the sentence level. They turned out to be rather exceptional examples with non-representative peculiarities, so we broadened the scope for extracting sentences from the corpus to slightly less striking examples. Example 1 shows the rephrasing process for the syntactic dimension S, with 1a showing the original sentence (version A), 1b the medium version B and 1c the maximum version C:

- 1a *Paragraph 1626 BGB ist mit Artikel 6 Grundgesetz insoweit nicht vereinbar, als eine Übergangsregelung fehlt, die eine gerichtliche Einzelfallprüfung, ob das Wohl des Kindes einer gemeinsamen elterlichen Sorge der nicht miteinander verheirateten Eltern entgegensteht, für die Fälle vorsieht, in denen die Eltern mit dem Kind zusammengelebt, sich aber noch vor Inkraft-Treten des Kindschaftsrechtsreformgesetzes am 1. Juli 1998 getrennt haben.*
- 1b *Paragraph 1626 BGB ist mit Artikel 6 Grundgesetz insoweit nicht vereinbar, als eine Übergangsregelung fehlt. Diese müsste eine gerichtliche Einzelfallprüfung für die Fälle vorsehen, in denen die Eltern mit dem Kind zusammengelebt haben, sich aber vor dem Inkrafttreten des Kindschaftsrechtsreformgesetzes am 1. Juli 1998 getrennt haben. In diesem Fall wäre zu prüfen, ob das Wohl des Kindes einer gemeinsamen elterlichen Sorge der nicht miteinander verheirateten Eltern entgegensteht.*
- 1c *Paragraph 1626 BGB ist mit Artikel 6 Grundgesetz in einem Punkt nicht vereinbar: Eine Übergangsregelung fehlt. Diese müsste eine gerichtliche Einzelfallprüfung unter zwei Bedingungen vorsehen. Erstens müssten die Eltern mit dem Kind zusammengelebt haben. Zweitens müssten diese sich vor dem Inkrafttreten des Kindschaftsrechtsreformgesetzes am 1. Juli 1998 getrennt haben. In diesem Fall könnte das Wohl des Kindes einer gemeinsamen elterlichen Sorge der nicht miteinander verheirateten Eltern entgegenstehen.*

As this example already shows, the logical relation within the sentence poses a problem for the rephrasing process. Transferring the intrasentential logical relation to a sequence of logically related sentences may involve major restructuring, because sentence splitting can lead to ambiguous intersentential reference and to a loss of textual coherence. In Example 1 restructuring is realised by postponing the clause containing the postmodification of „Einzelfallprüfung“ to the end of the newly created sequence and adding an introductory anaphoric prepositional phrase.

Embedding on the phrase level is resolved first by transferring a nominal structure into a verbal structure. This is typically realised by expressing the nominal meaning in a subordinate clause with the formerly only nominal meaning distributed on nominal and verbal elements. For our rephrasing process this means that version B is more complex on the sentence level than version A. Therefore, the second rephrase leading to version C consists in a break up of the complex sentence structure as for the embedded sentences (see above). Example 2 below illustrates syntactic dimension P.

- 2a *Im vorliegenden Fall braucht auf die Erfolgsaussichten einer noch einzulegenden Verfassungsbeschwerde gegen das angegriffene Gesetz nicht eingegangen zu werden, weil jedenfalls die Folgenabwägung zu Lasten der Antragsteller ausfällt.*
- 2b *Im vorliegenden Fall braucht nicht darauf eingegangen zu werden, welche Erfolgsaussichten eine Verfassungsbeschwerde gegen das angegriffene Gesetz haben würde, die noch einzulegen wäre, weil jedenfalls die Folgenabwägung zu Lasten der Antragsteller ausfällt.*
- 2c *Im vorliegenden Fall braucht nicht darauf eingegangen zu werden, welche Erfolgsaussichten eine Verfassungsbeschwerde gegen das angegriffene Gesetz haben würde. Diese wäre noch einzulegen. Jedenfalls fällt die Folgenabwägung zu Lasten der Antragsteller aus.*

With respect to nominalisations, we concentrate on deverbal derivations with the suffix „-ung”, as they have the highest frequency of all nominalisations in the three intralingual versions (see section 2.1). The high frequency of this feature, however, raises the question, which instances we want to rephrase. If we rephrased all instances, the sentence structure would become overly complex. Furthermore, this would not match the frequency in the newspaper reports. We therefore have to make a selection. One should think that nominalisations created on an ad-hoc basis to condense clausal meaning into a nominal structure qualify for rephrases. However, most instances of nominalisations are lexicalised through frequent use. Often both, the nominal and the verbal form of the lemma, are equally frequent. The most plausible candidates therefore are accumulations of „-ung”-nominalisations within the same noun phrase construction. The first rephrasing step consists in using the verbal form of the lemma, thus creating a clausal structure. In those cases where the rephrasing results in a non-finite subordinate

clause, this clause is again rephrased into a finite clause. Example 3 is a case in point.

- 3a *Bei der Abwägung zwischen den Belangen der Beschwerdeführerin einerseits und dem Persönlichkeitsrecht der Kläger andererseits gebühre dem Interesse der Kläger an einer Veröffentlichung der Richtigstellung auf der Titelseite der Vorrang.*
- 3b *Bei der Abwägung zwischen den Belangen der Beschwerdeführerin einerseits und dem Persönlichkeitsrecht der Kläger andererseits gebühre dem Interesse der Kläger daran, eine Richtigstellung auf der Titelseite zu veröffentlichen, der Vorrang.*
- 3c *Bei der Abwägung zwischen den Belangen der Beschwerdeführerin einerseits und dem Persönlichkeitsrecht der Kläger andererseits gebühre dem Interesse der Kläger daran, dass eine Richtigstellung auf der Titelseite veröffentlicht wird, der Vorrang.*

Limitations of the rephrasing process

Originally, we had expected to be able to develop rules for rephrasing each syntactic dimension. However, the instances proved too heterogeneous to maintain the same strategy for each sentence in one syntactic dimension. In order to be able to systematise the procedure, it may help to categorise the structures from a functional perspective and then develop rules for each function identified. Within the given limitations of our study this has to be left for future research.

As previously mentioned, the rephrasing process has consequences for the coherence structure of the text. When we break up complex sentences into simple sentences without or almost without embedding – as we do in the maximum version C –, we risk causing a loss of textual coherence. As it is no longer possible to insert local intrasentential cohesive ties, we can only create coherence by establishing cohesive ties on intersentential level. However, adding cohesive elements to every simple sentence often produces an awkward style inappropriate for written language. Therefore, in this study we try to minimalise the elements added in the rephrases, thus inserting cohesive elements only where absolutely necessary. A future study focussed on textual characteristics of legal texts may provide further insight into the requirements for improved comprehensibility regarding cohesion and coherence.

Preparation of the experiment

We restrict our experimental data to those sentences which clearly display one of the three syntactic dimensions mentioned above. This is done in order to avoid phenomena assumed to impede the understanding of the sentence in question, which could distort the results of the test. Where necessary, we therefore slightly change the original sentences (version A) in order to discard any possible interfering elements. Example 4 below shows a sentence from the corpus used for testing the syntactic dimension P. The target phrase is underlined in the test sentence in 4b. Compared to the original sentence in 4a, it becomes clear that the changes affect the length of other phrases which are not in the focus of the present stimulus.

4a Original sentence from the corpus

Allerdings stehe der familienrechtlichen Lösung in § 1626 a BGB im Falle einer Trennung der Eltern eines nichtehelichen Kindes nach längerem Zusammenleben mit diesem die verfassungsrechtliche Wertung entgegen, dass weder dem Elternrecht der Mutter noch dem des Vaters ein Vorrang eingeräumt werden könne.

4b Test sentence (target phrase underlined)

Allerdings stehe der familienrechtlichen Lösung im Falle einer Trennung der Eltern eines nichtehelichen Kindes nach längerem Zusammenleben mit diesem entgegen, dass weder der Mutter noch dem Vaters ein Vorrang eingeräumt werden könne.

The third part of the study consists in testing the rephrases in a psycholinguistic experiment in order to determine whether rephrasing the complex structures of the court decisions improves the comprehensibility for lay persons. As this is the main focus of the current paper, it will be described in more detail in the following section.

3. Processing complex and rephrased versions: the psycholinguistic test

3.1. Method

The underlying assumption of our psycholinguistic experiment is that longer reading times equal either deeper processing or more complex texts. Reading times can be measured in a self paced reading experiment (Mitchell 1987) which is chosen as the optimal testing method given

the limited resources of the study. Measuring the reading times for the different versions gives an indication of the processing effort the participants need for these versions. Furthermore, self paced reading represents a good indicator of comprehensibility when combined with comprehension questions and measuring the response latencies, i.e. the time, participants need to answer the comprehension questions.

The 45 participants consist of 36 lay persons, mainly students of Saarland University, and 9 experts, i.e. legal experts and advanced law students. As we intend to test these subjects on the three versions A, B and C, they are split up into three groups consisting of 15 participants (12 lay persons and 3 experts), respectively (see below). We exclude students or experts of linguistics and law from the group of lay persons. The subjects were paid 4 Euros for participating in the experiment.

In the experiment, the texts are presented on two portable computers using the software DMDX⁵. The participants' task is to read the texts on the screen. The words of a sentence appear on request by mouse click. The programme logs all mouse clicks and computes the time from one mouse click to the next, thus recording the time the participant takes to process one word.

We use 30 sentences from the rephrasing process as stimuli. For each syntactic dimension (S, P, N) 10 sentences are chosen. Each of the sentences is realised in the versions A, B and C described above in section 2.2. In order to give the sentences some context, we include filler sentences where appropriate. These sentences serve to introduce the circumstances and remain unchanged in all three versions. They are partly taken from the corpus and partly individually written adapted to the stimuli. Example 5 shows that the filler sentences are simpler in their structure and introduce some content which is then elaborated on in the stimulus.

- 5 Filler 1:
Die Gesetzgebungszuständigkeit der Länder ist nicht durch Bundesrecht ausgeschlossen.
- Filler 2:
Auch die Finanzverfassung des Grundgesetzes steht der Abgabenerhebung nicht entgegen.

⁵ <http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm>

Stimulus dimension S, version A:

Der Zweite Senat hat geklärt, dass es für die kompetenzrechtliche Zulässigkeit einer nichtsteuerlichen Abgabe nicht darauf ankommt, ob sie den Anforderungen standhält, die sich aus der Begrenzungs- und Schutzfunktion der bundesstaatlichen Finanzverfassung ergeben.

The stimuli (and filler) appear in randomised order and are each followed by a comprehension question, which is proofread by a jurist checking the legal content of the question. The question corresponding to the stimulus in example 5 is given here as example 6.

6 *Wurde geklärt, worauf es für die kompetenzrechtliche Zulässigkeit ankommt?*

Like the fillers these questions also remain unchanged in all versions and have to be answered with yes or no. A short break is included before the question to avoid inadvertent clicking instead of answering the question.

We thus obtain the following design. Three complexity conditions (versions A, B, C) and two expertise conditions (lay persons, experts) result in a 3x2 factorial design. 15 participants – 12 lay persons and 3 experts – are assigned to each of the three complexity conditions at random.

We do not follow the moving windows paradigm where a word disappears as soon as the next is requested. We expect that the participants might lose track of the highly complex sentences from the court decisions if they cannot move back within one sentence. However, we analyse the reading times with regard to so called windows, i.e. recurring areas in a sentence which indicate additional cognitive load. The analysis does not suggest any noticeable windows. Probably, the participants request the words until they see a complete sentence on the screen and then process the whole sentence. Therefore we only consider aggregated reading times for complete sentences. The reading times for all individual words of a sentence as logged by DMDX are summed and divided by the number of tokens of each complexity condition as the versions differ in length. The time span from appearance of the comprehension question to its answering, the response latency, is interpreted as logged by the programme. The number of correct responses is summed. Any outliers are eliminated.

3.2. Hypotheses

The psycholinguistic test aims at identifying the differences in the cognitive processing of the three rephrased versions of legal texts described above. We review processing by measuring three dependent variables: 1) reading times, 2) response latencies and 3) correctness of responses given by the subjects. Independent variables are 1) degree of complexity and 2) expertise. Combining these two types of variables in hypotheses allows us to draw conclusions about the comprehensibility of the three rephrased versions A, B and C.

The group of experts only serves as an explorative group, since it is too small to yield any significant results. In order to permit a comparison between lay persons and experts both groups should be of comparable size. Therefore, we do not formulate any hypotheses in connection with the comparison. However, we will briefly discuss the results for this group in section 3.3. We concentrate on detecting differences in the comprehensibility of the three rephrased versions A, B and C by varying the degree of complexity along three syntactic dimensions S, P and N, as has been explained above. However, we do not intend to compare the difference in comprehensibility between the three grammatical dimensions as this comparison is problematic from a linguistic point of view. While comparing the effect of varying different syntactic structures on the comprehensibility constitutes an interesting research question, we cannot analyse it with our research design. This question would require a design focussed on the different structures and leaving aside the study of legal language.

In consideration of the above mentioned conditions we can formulate the general assumption that rephrasing version A, which displays a very high complexity on all syntactic levels, will increase comprehensibility. The expected effect of the two rephrases on the comprehensibility for lay readers can be described as follows. Maximally simplifying the syntactic structures should result in the shortest reading times in version C. Simple sentences do not require tracing back complicated structures and should therefore be the version requiring less reading time than the other two versions. However, cognitive processing is not stimulated by these simple structures. According to Groeben & Christmann (1989) this should destruct the reader's motivation to read carefully and then memorise what he/she read. We thus expect that participants reading

version C take longer to answer the comprehension questions and answer them less well than those reading version B which have a cognitive stimulus in the medium complex sentences. Version A is expected to perform worst because the complex structures ask too much from the lay persons not familiar with these structures. More precisely, we can establish the following hypotheses for the interpretation of the findings about the three above mentioned dependent variables:

H1 Reading times: $A > B > C$

The reading times for version A are longer than the reading times for version B, which is of medium complexity. The reading times for version B are longer than the reading times for version C, which constitutes a maximum simplification.

H2 Response latencies: $A > C > B$

The response latencies for version A are longer than the response latencies for version C. The response latencies for version C are longer than the response latencies for version B.

H3 Correctness of responses: $A < C < B$

In version A less correct responses are given than in version C. In version C less correct responses are given than in version B.

3.3. Findings

Our main focus in connection with the findings is the performance of the lay participants. The results for this group's reading times in the three versions A, B and C as displayed in Figure 1 are partially significant, i.e. they partially confirm our hypothesis H1. In more detail, the difference between version A and B is not significant. The difference between version B and version C is significant. This means that the test persons read version C much more quickly than the other two versions. However, reading times for version B which is of medium syntactic complexity almost equal those for version A as the most complex version. While being much less complex, version B still seems to cost much processing effort.

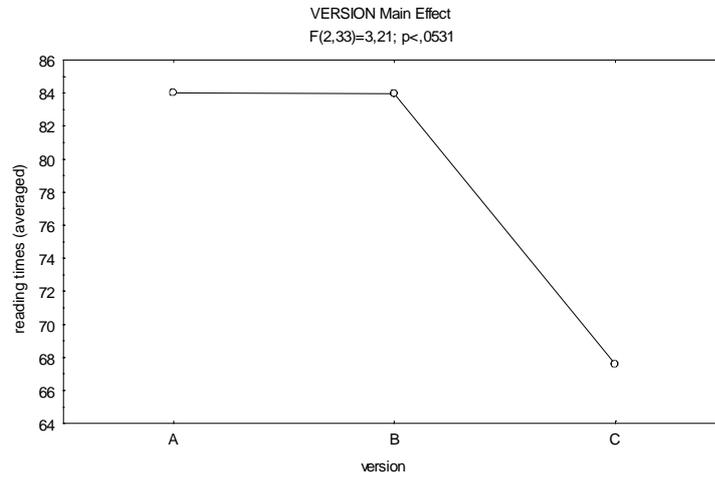


Figure 1: Lay persons' reading times for the three versions

Note here, that reading times do not allow inferences about the comprehension itself. They measure the duration of processing but are not a direct indicator of the comprehensibility of a text. This can only be determined on the basis of the results for all three dependent variables.

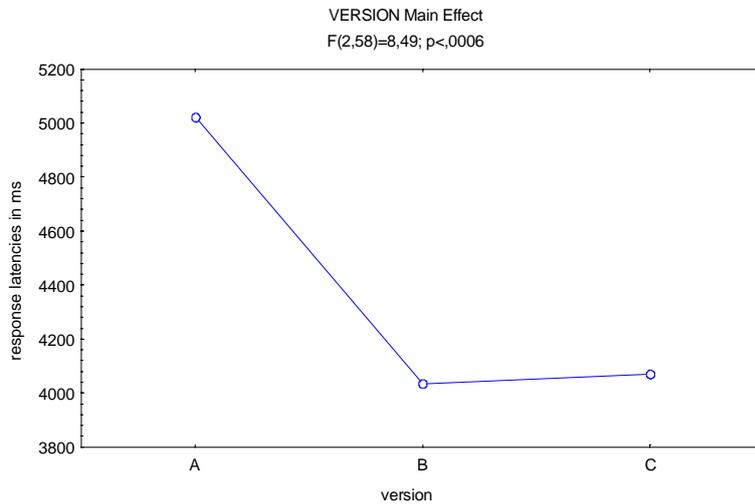


Figure 2: Lay persons' response latencies for the three versions

Our hypothesis H2 regarding the response latencies is partly confirmed as depicted in Figure 2. While the mean latency for lay readers of version A is significantly higher than for readers of version B, the difference between version B and C is not significant. Participants thus have most difficulties in processing version A, therefore they need much more time to answer the comprehension questions than in the other two versions. In the rephrased versions, the processing effort seems clearly reduced. However participants do not need significantly more time to process the question in version C than in version B. This result in combination with the result for the reading times suggests that the medium complex version B is not processed clearly better than version C. This picture changes when looking at the findings for correctness of responses for the three rephrased versions A, B and C.

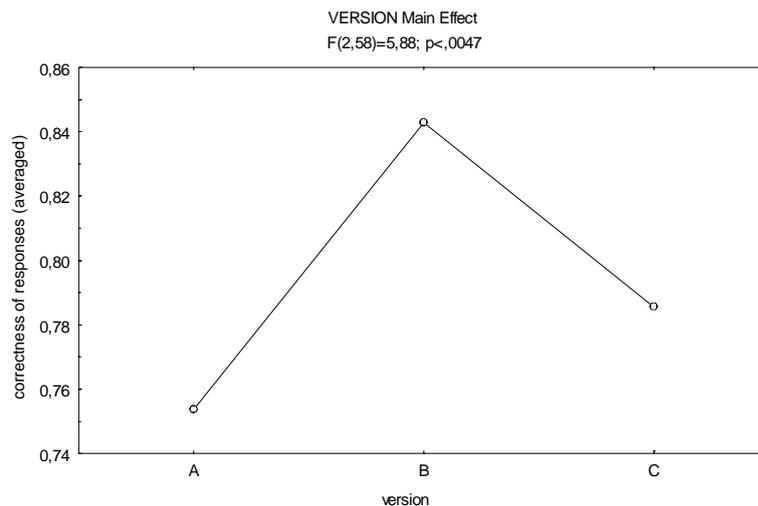


Figure 3: Correctness of lay persons' responses in the three versions

As with the other two hypotheses, H3 again is partially confirmed. Figure 3 displays the amount of correct responses given for versions A, B and C.

The difference between versions A and B is significant as is the difference between versions B and C. The participants reading version B are able to answer much more often correctly than are those reading ver-

sion A or C. The difference between A and C is not significant, meaning that both groups of participants have similar difficulties responding to the comprehension questions, albeit for different reasons. The difficulties of participants assigned to version A most likely begin with processing the complex sentences, a problem manifested in the long reading times and continue with considering a response to the comprehension question. Participants assigned to version C seem to simply run through the sentences without really memorising what they are reading and feel capable of answering the questions quickly, a fact reflected in the short response latencies. Because of its medium complexity version B provides enough incentive for the subjects to read the legal text and work towards understanding it correctly.

The results compared by degree of expertise show that the jurists participating in our experiment have longer reading times and response latencies in all three versions. This is probably due to the fact that versions B and C contradict the reading habits of legal experts: the expert recognises the sentences as originating from a legal text and has to incorporate the differences to what he/she expects. Jurists give more correct responses than the lay participants. This does not come as a surprise as the experts comprehend the legal content regardless of the syntactic form. A cautionary note is in place, since these results only constitute tendencies not based on statistical tests. They show the difference between the two recipient groups of court decisions and have to be kept in mind when varying the syntactic structure of court decisions.

3.4. Interpretation

We are able to show that, broadly speaking, syntactic rephrases improve processing with lay readers, thus confirming our general assumption. The experiment leaves no doubt that version A, i.e. sentences from court decisions, massively impedes the comprehensibility for lay readers. The lay persons assigned to this version read longer, consider their responses longer and still answer the comprehension questions less well than those assigned to the rephrased versions. We can say that the syntactic structures of version A are too complex for the participants to understand them correctly.

While reading times do not support our hypothesis that the B-version is processed faster by the lay readers, the combined interpretation of all

three dependent variables – reading times, response latencies and correctness – show that version B leads to an optimal improvement of the comprehensibility. Compared to version A, the participants assigned to version B do not read faster in this version but take significantly faster to answer the comprehension questions and – what is most important – are in the position to answer the questions more often correctly. This shows that, in comparison to version A, the comprehensibility has improved.

The C-version does not perform better in this overall interpretation. Although the reading times are significantly shorter in this version than in the other two versions and the participants answer the questions as quickly as the readers of the B-version, they are not able to give as many correct responses as in the B-version. The simple syntactic structures seem to lead to a loss of cognitive motivation for the subjects to read the texts properly in order to understand them correctly. There are two possible explanations for this: first, the overly simple structures of this version may induce the readers to not start cognitive processing of what they are reading. This explanation is in line with Groeben & Christmann's (1989) motivation- and cognition-based approach. The second explanation is that the structures are resolved beyond what the lay reader expects. The logical relation between clauses is shifted to the sentence level, leading to a text structure which is less cohesive. Version B creates enough textual coherence to linguistically link parts of the text in a meaningful way. This is not the case in the maximal simplification of version C. In this version information is spread so sparsely in short and simple sentences that it partially lacks cohesive ties to establish meaningful relations in the text as discussed in section 2.2. This lack of textual coherence may complicate processing and decrease comprehensibility and, finally, result in a low amount of correct responses to comprehension questions. Thus, both the A- and the C-version do not conform to the reading habits of the lay reader, leaving version B which follows the degree of syntactic complexity found in the newspaper reports as the version understood best by the lay readers.

4. Conclusions and outlook

The study presented here gives us an in-depth look at the workings of three syntactic features of court decisions. Each of the three methodol-

ological steps contributes its part to the overall picture. First, we built a corpus annotated with syntactic information on German court decisions and related text types. The interpretation of this corpus in itself may yield valuable results for the analysis of German legal language. We used the corpus to investigate how syntactic specificities of court decisions are varied in the related text types of press releases and newspaper reports on the decisions. The corpus analysis also enabled us to retrieve the most distinctive instances from the subcorpus of court decisions, which were then used for the rephrasing process.

The second step showed the possibilities and limitations of the rephrasing process. On the basis of the instances retrieved from the corpus, we elaborated two rephrases: a medium complex one with syntactic structures demanding some cognitive effort from the recipient and a simple one which builds on the assumption that maximally simple structures are memorised best. We identified limitations with each of the three syntactic features varied in this process. The specificities of each instance rephrased made it obvious that there is a need for more fine-grained (functional) analyses of complex structures in a future study. This may ultimately lead to rules for automatically rephrasing complex syntactic structures.

Finally, the results of the psycholinguistic experiment help us understand what a comprehensible court decision could look like. The study showed that rephrases with a medium degree of syntactic complexity similar to that of newspaper reports score better when tested with lay persons. This clearly shows that the language of court decisions is not adapted to the needs of lay citizens as one of two recipient groups of court decisions. However, the comparison of the two groups of expertise, i.e. the two recipient groups, indicated that legal experts will not easily accept rephrases from the type scoring best with lay persons as they contradict the jurists' expectations. This may change in the future if research findings such as those presented here are used for teaching legal writing to law students. If they learn a plain writing style at an early stage in their formation it will become natural to them to put complex facts in a simple way.

In a broader view, a cautionary note is in place. Syntactic changes cannot remedy the inherent incomprehensibility of the law itself. They only operate on the surface level. Furthermore, on this surface level, the

changes should not be limited to the level of syntax on which we have concentrated in this study but have to be combined with changes in the lexis used as well as in the overall cohesive structure of the text. For this purpose, continuative studies on these levels have to be conducted. It remains to be seen, however, what these combined linguistic efforts can achieve towards the goal of improving the comprehensibility of legal language.

References

- Altehenger, Bernhard 1983: Die richterliche Entscheidung als Texttyp. In Petöfi, János 1983 (ed): *Texte und Sachverhalte*. Hamburg: Buske. 185-227.
- Basedow, Jürgen 1999: Transparenz als Prinzip des (Versicherungs-)Vertragsrechts. In *VersR* 50, 1045ff.
- Braun, Christian 1999: *Flaches und robustes Parsen deutscher Satzgefüge*. Diploma thesis, Universität des Saarlandes, Saarbrücken.
- Engberg, Jan 1997: *Konventionen von Fachtextsorten. Kontrastive Analysen zu deutschen und dänischen Gerichtsurteilen*. Tübingen: Narr.
- Flesch, Rudolph F. 1948: A new readability yardstick. In *Journal of Applied Psychology* 32, 221-233.
- Groeben, Norbert/Christmann, Ursula 1989: Textoptimierung unter Verständlichkeitsperspektive. In Antos, Gerd/Krings, Hans Peter 1989 (eds): *Textproduktion. Ein interdisziplinärer Forschungsüberblick*. Tübingen: Niemeyer. 165-196.
- Grönert, Kerstin 2004: *Verständigung und Akzeptanz in der Kommunikation zwischen Bürger und Verwaltung*. PhD Dissertation, Universität Bielefeld.
- Hansen-Schirra, Silvia/Neumann, Stella 2004: Linguistische Verständlichmachung in der juristischen Realität. In Lerch, Kent D. 2004 (ed): *Recht verstehen. Verständlichkeit, Missverständlichkeit und Unverständlichkeit von Recht*. Volume 1. Series „Die Sprache des Rechts der Berlin-Brandenburgischen Akademie der Wissenschaften“. Berlin, New York: de Gruyter. 167-184.
- Jaspersen, Andrea 1998: *Über die mangelnde Verständlichkeit des Rechts für den Laien*. PhD Dissertation, Rheinische Friedrich-Wilhelms-Universität.
- Mitchell, D.C. 1987: Reading and syntactic analysis. In Beech, J.R./Colley, A.M. 1987 (eds): *Cognitive approaches to reading*. Chichester, UK: John Wiley & Sons, 87-112.
- Langer, Inghard/Schulz von Thun, Friedemann/Tausch, Reinhard 1974: *Verständlichkeit in Schule, Verwaltung, Politik und Wissenschaft*. München u.a.: Reinhardt.
- Lerch, Kent D. 2004a (ed): *Recht verstehen. Verständlichkeit, Missverständlichkeit und Unverständlichkeit von Recht*. Volume 1. Series „Die Sprache des Rechts der

- Berlin-Brandenburgischen Akademie der Wissenschaften". Berlin, New York: de Gruyter.
- Lerch, Kent D. 2004b: Verständlichkeit als Pflicht? Zur Intransparenz des Transparenzgebots. In Lerch, Kent D. 2004 (ed): *Recht verstehen. Verständlichkeit, Missverständlichkeit und Unverständlichkeit von Recht*. Volume 1. Series „Die Sprache des Rechts der Berlin-Brandenburgischen Akademie der Wissenschaften". Berlin, New York: de Gruyter. 239-283.
- Neumann, Stella/Hansen-Schirra, Silvia 2004: Der Konjunktiv als Verständnisproblem in Rechtstexten. In *Zeitschrift für Angewandte Linguistik (ZfAL)* 41/2004, 67-87.
- Oksaar, Els 1988: *Fachsprachliche Dimensionen*. Tübingen: Narr.
- Rave, Dieter/Brinckmann, Hans/Grimmer, Klaus (eds) 1971: *Paraphrasen juristischer Texte*. Darmstadt: Deutsches Rechenzentrum.
- Rickheit, Gerd 1995: Verstehen und Verständlichkeit von Sprache. In: *Sprache: Verstehen und Verständlichkeit*. Kongressbeiträge zur 25. Jahrestagung der Gesellschaft für Angewandte Linguistik GAL e.V., ed. by Bernd Spillner. Frankfurt/Main: Lang. 13-30.
- Wagner, Hildegard 1981: *Die deutsche Verwaltungssprache der Gegenwart*. 3rd print. Düsseldorf: Schwann.