

Word order in German verb final sentences: A linguistic and psycholinguistic problem

Thomas Pechmann¹, Hans Uszkoreit², Johannes Engelkamp¹ & Dieter Zerbst¹

University of the Saarland
Department of Psychology
and
Department of Computational Linguistics

This paper summarizes the first year of research on the order of major constituents in German sentences, and has been supported by the Deutsche Forschungsgemeinschaft (Pe 357-3). We are grateful to Matthias Nickola and Peter Heinrich for their substantial contribution to this project. We would also like to thank Karel Oliva for valuable discussions and helpful comments.

Abstract

This paper deals with linearization of complements of German verbs. In German all permutations of the subject, the indirect, and the direct object do occur. Yet, they are supposed to differ regarding their degree of acceptability. Uszkoreit (1987) proposed a set of rules which aimed at representing such preferences as the product of different factors. This theoretical account leads to a predicted ranking of the possible syntactic forms. In a set of experiments we tested some of these predictions by application of different methods for tapping into the actual processing of the sentences. In particular, the predictions were (a) that sentences are more acceptable if subjects precede objects than vice versa and (b) that sentences are more acceptable if indirect objects precede direct objects than vice versa. Both comprehension and production experiments were carried out. The methods we used included a ranking task, delayed sentence matching, delayed articulation, rapid serial visual presentation and a sentence generation task. The findings yielded a very consistent picture concerning the position of the subject. Sentences were particularly easy to process if the subject was in initial position and particularly hard to process in subject-final constructions. Furthermore, there is somewhat weaker evidence for the assumption that sentences are easier to process if direct objects are preceded by indirect objects. Since these results were obtained by rather different methods they can be regarded as particularly reliable. Moreover, the data did provide evidence for a gradual increase or decrease of acceptability and no evidence for a jump function, sharply separating grammatical from ungrammatical forms. One of the principal aims of this first phase of our investigations which is reported in the present paper was to find experimental methods which consistently differentiate between the various permutations of verb complements as predicted by theoretical assumptions. This aim could be achieved. The next step will be to include pragmatic factors which are supposed to play a significant role in determining the acceptability of the sentences we are studying.

Word order in modern linguistics

Natural languages differ significantly with regard to the degree of variation in their word order. On one end of the scale, we find languages such as English that exhibit a rather strict order, at least among heads and complements. On the opposite end of the scale there are languages such as the Australian language Warlpiri for which only very few ordering constraints can be observed. For some time it was claimed that this language allows for any permutation of the words of a sentence (cf. .i.Hale, 1983);. Although this extreme claim turned out to be untenable, Warlpiri remains to be a language which allows variation of word order to a very large extent. The majority of languages, including German, mix fixed and free word order in numerous exciting ways.

Though it might be regarded a core concept of syntax, word order variation has not played the role it deserves in modern linguistics. This might in parts be due to the anglocentric view of the disciplin. However, it also reflects the lack of a good understanding of the interface between syntax and pragmatics. In grammars based on the notion of constituency, the prevalent descriptive means has been the phrase-structure tree. In Chomsky's Generative Transformational Grammar (TG; cf. .i.Chomsky, 1957, 1965); mappings from one phrase structure tree to another were introduced, but each derivation started from an ordered tree. The implication from transformational grammar was that of a default or basic word order for a given language. Obligatory deviations could be achieved by obligatory transformations, stylistic variations by stylistic transformations.

In the theory of Government and Binding (GB; .i.Chomsky, 1981); which is the historical child of TG, word order is still explained by instances of alpha-movement applied to a basic word order. However, in European linguistics grammars based on the TG model never succeeded in completely reigning the field. Particularly for Slavic languages (cf. .i.Melcuk, 1974;; .i.Sgall, Hajicov & Panevov , 1986); alternative models always kept the floor, in particular Dependency Grammar (.i.Tesnišre, 1959);. The same holds for German (.i.Kunze, 1975); and Romanic languages (.i.Tesnišre, 1959);. Dependency Grammar permits a more intuitive treatment of word order variation, for it provides a level of description at which heads and their arguments and adjuncts are not ordered with respect to each other. From this level one can define rules for the linearization at the surface.

In feature-oriented declarative grammar formalisms, commonly referred to as unification grammars, word order is often described by a method which was first proposed by Gazdar and Pullum (.i.Gazdar & Pullum, 1982); for the framework of Generalized Phrase Structure Grammar (GPSG). The descriptive tool is called the ID/LP format. It is based on the separation of the immediate dominance (ID) concept and the linear precedence (LP) concept in the phrase structure component of a grammar. ID-rules determine the hierarchical structure of a sentence. LP-rules restrict the sequence of sister-nodes in a tree. A grammar without LP-rules generates a language in which all sister-nodes are free to permute.; ID-rules in GPSG take the form $X \rightarrow Y_1, Y_2, \dots, Y_3$ where X is an element of VN and Y_i is an element of the union of VN and VT. The commas between the symbols on the right-hand-side of the rule indicate that the rule does not specify the sequence of these symbols. LP-rules take the form $Y_1 < Y_2$ saying that each constituent Y_1 must precede each sister constituent Y_2 if constituents of both categories are present in a local tree.

By applying the ID/LP format it became possible to chunk the regularities of word order for one grammar which had previously required a large number of phrase structure rules. Sequential variants could simply be described by the absence of certain LP-rules. By using a feature-oriented representation of categories various word order regularities could be described by, for instance, morphological features, the syntactic category of constituents, or thematic

roles (cf. .i.Gazdar, Klein, Pullum & Sag, 1985;; .i.Uszkoreit, 1987);. The following three rules may serve as examples :

```
[cat:np] < [cat:pp]
[case:dat] < [case:acc]
[th-role:goal] < [th-role:theme]
```

The mixture of fixed and free word order in German could be handled by the application of LP-rules to certain phenomena, but not to others. LP-rules could be applied, for instance, to the linear sequence of articles and nouns or to the position of the verb, but not to the position of adverbials. The positions of finite verb forms could be represented as equivalent variants of linearization without the assumption of a basic word order (cf. .i.Uszkoreit, 1987);. This method was adopted in various forms by other declarative grammar formalisms. The variants differ with respect to the expressive power of the LP-rules (cf. .i.Pollard & Sag, 1987,; concerning HPSG) or with respect to the domain of their application (e.g. .i.Reape, 1990);. However, the problem of partially free word order cannot be solved by simple LP-rules, since these do not express preferences, but strict regularities. If the sequence of two constituents A and B is not restricted by an LP-rule, both orderings, A<B and B<A, are equally grammatical. Yet, if an LP-rule exists such that A<B, then only this one sequence is grammatical.

The 'Mittelfeld' (middle field) in German verb final sentences as a critical case

A revealing case regarding partially free word order and one to which the framework just mentioned has already been applied concerns the variation of verb complements in the so-called 'Mittelfeld' (middle field) in German verb final sentences (cf. .i.Thiersch, 1982;; .i.Koster, 1987;; .i.Grewendorf, 1988;; .i.Fanselow, 1990;; .i.Sternefeld, 1990);. The middle field may consist, for instance, of the subject and two objects:

- (1) Dann wird der Richter dem Anwalt den Beweis liefern.
Then will the judge the lawyer the proof present.

In German, any permutation of the subject, the indirect object, and the direct object is considered to be grammatical and acceptable. However, this is not to say that the sequence of the verb complements is not rule-governed. There are clear preferences which depend upon the thematic roles, the discourse functions, and the pronominalization of subject and objects (cf. .i.Lenerz, 1977);. Sometimes the factors mentioned above strongly exclude some sequences. In order to formally represent such preferences as the product of different factors, .i.Uszkoreit (1986); has proposed complex LP-rules consisting of a number of sequential rules which must not all be fulfilled in a particular case. Different weights can be attached to single principles. A decrease in grammaticality then results from the sum of violations of single principles. For the German middle field .i.Uszkoreit (1987); has proposed the following complex LP-rule which is based on morphological marking (case; the dichotomy of noun/pronoun) as well as on discursive roles (focus):

[case:nom] < [case:acc]
[case:nom] < [case:dat]
[case:dat] < [case:acc]
[cat:pron] < [cat:noun]
[focus:-] < [focus:+]

These principles are not supposed to have the same weight. In particular, the tendency to place the subject (nominative) before the indirect (dative) or direct (accusative) object, for instance, is assumed to be considerably stronger than the tendency to have the accusative object preceded by the dative object. Also, the inclination to put personal pronouns in front of other elements is rather strong in German. Another ordering principle that has been observed, concerns the length or syntactic heaviness of the permuting elements. Shorter and less complex elements tend to precede longer and more complex ones. Following .i.Behaghel (1932),; this strong tendency has become known as the Gesetz der wachsenden Glieder (law of increasing elements). To our knowledge there have been no attempts to formalize this performance principle in any modern competence-based generative grammar model.

Although the observed principles belong to different domains of linguistic description, it seems justified to bundle them into a single complex rule since they all affect the syntactic sequence. Whereas each of the individual preferences had been observed before, Uszkoreit (1987) first provided a formal framework for describing the interaction of partially conflicting principles, for determining varying degrees of grammaticality, and for explaining preferences for certain readings in cases of ambiguity.

His approach is strongly connected with the following three hypotheses:

1. There is a gradual scale in acceptability between clearly grammatical and clearly ungrammatical linearizations.
2. There is no break point or jump function separating grammatical and ungrammatical variants.
3. Violations of ordering principles are normally caused by some overriding conflicting principle.

The necessity of complex weighted LP-rules has since been accepted by other linguists (cf. .i.Pollard & Sag, 1987;; .i.Reape, 1990);. The new framework preserves the descriptive tool of the original LP-rules as a special case. Simple LP-rules which specify, for example, the ordering of articles and nouns can be regarded as special cases of complex LP-rules in which there exists only one principle of linearization.

The account of complex LP-rules is not necessarily at odds with the word order theory of the Prague School (cf. .i.Sgall et al., 1986);. In this theory it is assumed that word ordering is affected by different factors from different domains (syntax, morphology). The major role, however, is assigned to the theme/rheme ('topic/focus') dichotomy and related phenomena. The basic claim is that in the case of unmarked ('objective') word order all elements of the theme precede all elements of the rheme. Theme-internally the sequence of the elements is organized according to their salience in the unfolding discourse (salient elements precede less salient elements). The rheme internal sequence is not dependent upon the discourse, but upon the so-called systemic organization. This is a language-dependent hierarchy which is based on the thematic roles of the corresponding elements of a sentence. Deviations from the systemic organization are possible. They can be described as syntactically based deviations ('shallow word order rules'; for instance, the position of the verb in German or the almost obligatory initial position of the subject in English) or they can be explained by the markedness of the utterance (so-called 'subjective' word order). In this case, the rheme, or a part of it, can be inserted in front of the theme, becoming at the same time strongly accentuated.

The problem of empirical validation

As sketched above, several (as yet underspecified) theories have been proposed concerning the interactive influence of factors which restrict the flexibility of word order. Furthermore, we dispose of descriptive tools to formalize language-specific analyses in accordance with particular theories. Computational linguistics even provides first attempts to apply formal grammars to language processing by machines (cf. .i.Engelkamp, Erbach & Uszkoreit, 1992;; .i.Erbach, 1992;; .i.Oliva, 1992;; .i.Reape, 1990);. What is urgently needed are methods which allow us to test and evaluate competing theories, formalisms, and analyses.

A decision between different word order theories often depends upon the correct evaluation of the relative grammaticality of particular sentences. In linguistics grammaticality judgements are typically based upon individual intuitions or, in rare cases, upon acceptability ratings obtained from a sample of native speakers. This last method has proven useful as long as it is applied to clear yes/no decisions. Yet, in many cases theories have become so fine-grained that decisions between pairs of sentences both containing doubtful cases of grammaticality, are required. Such decisions are hard to obtain from metalinguistic judgements of native speakers. Furthermore, the value of acceptability ratings is often clearly restricted. First, metalinguistic judgements are notoriously inconsistent (cf. .i.Levelt, 1972);. Second, the sentences to be rated are often embedded in question-answer sequences (cf. .i.Lenerz, 1977;; .i.Pfeiffer, Pucek & Sgall, 1994);. By providing the subject

both with the lexical material and with the sentence structure of the question, it is very likely that effects of structural parallelism in the sentences to be judged can be obtained (cf. .i.Levelt & Kelter, 1982);. Third, linguistic theories usually arise from the metalinguistic judgement of individual scientists. All that can be achieved by obtaining the metalinguistic judgements of independent subjects in a rating experiment is to increase the statistical reliability and validity of these intuitions. But by using the very same kind of data the danger of methodological circularity is evident. Therefore, it seems highly desirable to design additional experimental methods which allow us to assess the acceptability of various syntactic forms independent from subjects' metalinguistic intuitions. This seems possible if the assumption that grammatical acceptability is closely related to ease of processing is correct, and this can be tested by appropriate procedures. We take this as our working hypothesis. In the following, we will report upon a series of experiments which can be regarded as a first step in this direction with regard to the linearization of verb complements in the middle field of German. As outlined above German allows any permutation of grammatical subjects, indirect objects and direct objects. Take as an example the sentence

- (2) Dann wird der Dirigent dem Geiger den Taktstock geben
(Then will the conductor the violinist the baton give)

All other permutations are possible as well:

DANN WIRD DER DIRIGENT DEM GEIGER DEN TAKTSTOCK GEBEN
DANN WIRD DER DIRIGENT DEN TAKTSTOCK DEM GEIGER GEBEN
DANN WIRD DEM GEIGER DER DIRIGENT DEN TAKTSTOCK GEBEN
DANN WIRD DEM GEIGER DEN TAKTSTOCK DER DIRIGENT GEBEN
DANN WIRD DEN TAKTSTOCK DER DIRIGENT DEM GEIGER GEBEN
DANN WIRD DEN TAKTSTOCK DEM GEIGER DER DIRIGENT GEBEN

Though all these sentences are perfectly grammatical, they are considered to differ regarding their degree of acceptability. The question facing us is, which principle(s) these differences depend upon. As mentioned above, .i.Uszkoreit (1986); proposes five such principles:

- (OP1) The agent precedes the theme.
(OP2) The agent precedes the goal.
(OP3) The goal precedes the theme.
(OP4) Focused constituents follow other constituents.
(OP5) Personal pronouns precede nonpronominal constituents.

The acceptability of a particular sentence is assumed to be a function of the combined weights of the observed or violated rules. Uszkoreit, however, could not provide evidence for numerical weights, since empirical methods for determining such weights have not yet been found.

The five principles refer to three different levels: thematic roles (related, but not identical, of course, to grammatical functions), focusing, and pronominalization. Thus, the theory strongly emphasizes the impact of thematic and pragmatic factors.

Evidently, the theoretical assumptions proposed by Uszkoreit can only be ultimately tested empirically by taking into account pragmatic contexts. Nevertheless, in our first set of experiments we decided to start with the investigation of isolated sentences excluding focus and pronominalization as factors. The reasons are the following. First, we wanted to keep our experimental material at the beginning as simple as possible, restricting the number of critical variables to a minimum. Second, no established experimental

methods exist which can deal with the problems in question. Therefore, our first aim was to find techniques which consistently differentiate between sentences varying in the position of verb complements and which do not rely on subjects' metalinguistic judgements. Again, we did not want to further complicate this initial step by including linguistic contexts from the very start. Yet, it stands to reason that such a strategy must necessarily be supplemented in subsequent experiments by including such contexts.

In general, the theory claims that sentences become less acceptable the more operating principles are violated. This would yield the following picture for the principles governing the middle field in German whilst disregarding focus and pronominalization in isolated sentences:

	OP1	OP2	OP3	sum of violations
SDA	+	+	+	0
SAD	+	+	-	1
DSA	+	-	+	1
DAS	-	-	+	2
ASD	-	+	-	2
ADS	-	-	-	3

(S=subject; D=dative object; A=accusative object; for the sake of simplicity we identify agent with subject here, theme with dative object, and goal with accusative object)

If the three operating principles are equally weighted, SDA-sentences should be the standard form. SAD- and DSA-sentences, including one violation, should be somewhat less acceptable. DAS- and ASD-sentences should be even less acceptable, because they include two violations. The least acceptable form should be ADS-sentences with three violations. Thus the rank order of acceptability would be:

SDA > SAD, DSA > DAS, ASD > ADS

Yet, as we have repeatedly mentioned, it is unlikely that the principles are equally weighted. In particular, we assume that the two principles which state that objects are preceded by subjects (OP1 and OP2) are particularly strong. There is a large number of empirical investigations dealing with the special status of 'subjecthood' (cf. .i.Bock, Loebell & Morey, 1992;; .i.Engelkamp & Zimmer, 1983;; .i.Ertel, 1977);. However, these studies discuss at length which factors language users base their decision when assigning subjecthood to a particular concept but not to another. In contrast, they are usually not concerned with the linearization of subject and several objects, as they might constitute the middle field in German sentences. Again, this might be due to the anglocentric bias of linguistic and psycholinguistic research, since this issue does not often arise in a fixed word order language, like English. Yet, if the problem is noted, all agree that subjects are by default put into the initial slot provided for verb complements. Consequently, we assume that the first two ordering principles will have a stronger effect than the third one. On the basis of this claim, the predicted rank order of sentence formats is somewhat different. Giving priority to the subject position and assuming that the sequence of indirect and direct objects is simply nested under this first factor, the following rank order of acceptability arises:

SDA > SAD > DSA > ASD > DAS > ADS.

Experimental studies

A set of experiments was designed to test this hypothesis by investigating the actual processing of corresponding sentences. However, it should be emphasized that our principal concern in the initial phase of experimental studies, was to find methods which consistently differentiate between different word order variants, so that the effect of pragmatic contexts could be studied in later experiments.

As we mentioned earlier, we started from the working hypothesis that ease of processing is closely related to acceptability. Since we could not rely on established methods, we decided to start by testing a number of paradigms which had earlier been applied in investigations concerned with differences in syntactic form, but in different respects. These paradigms were the following: delayed sentence matching (.i.Forster, 1987;; .i.Forster & Stevenson, 1987;; .i.Freedman & Forster, 1985);, delayed articulation (cf. .i.Ferreira, 1991);, rapid serial visual presentation (.i.Forster, 1970);, a sentence generation method (cf. .i.Pechmann, 1994);, and, in addition, a traditional rating procedure. A study of event-related brain potentials has not yet been carried out, but is in preparation at present.

Experiment I: Rating

The classical method for studying degrees of grammaticality or acceptability usually applied by linguists is the rating procedure. Subjects are provided with a set of sentences which they are asked to rank on a particular scale. We already discussed some crucial problems with this method and argued for the necessity that metalinguistic judgements should be cross-validated by the application of independent methods which tap into the on-line processing of language.

However, in order to be able to compare metalinguistic judgements with data from experiments studying the actual processing, in the first experiment to be reported here we asked subjects to rate the different sentence formats we are interested in with respect to their acceptability.

Method

Subjects

Twenty-four subjects volunteered in this rating study. They were all students of the Philipps-University at Marburg.

Material and procedure

The stimuli were six proto-test sentences in six sentence formats: SDA, SAD, DSA, DAS, ASD, and AVDS. One of the formats, AVDS which was included instead of ADS, is an ungrammatical form, since the verb is moved from the end of the sentence into a sentence-internal position, e.g.:

DANN WIRD DEN ERFOLG GÖNNEN DEM SIEGER DER SCHWIMMER.

This deviant format was included to test whether the method would at least be sensitive to an unacceptable word order in our sentences.

The six proto-test sentences were:

DANN WIRD DER SCHWIMMER DEM SIEGER DEN ERFOLG GÖNNEN.
 BALD WIRD DER JUNGE DEM VATER DEN HAMMER HOLEN.
 DORT WIRD DER KUNDE DEM H?NDLER DEN AUFTRAG GEBEN.
 DANN WIRD DER LEHRER DEM SCHŞLER DEN FEHLER SAGEN.
 BALD WIRD DER MALER DEM NACHBARN DEN SCHUPPEN STREICHEN.
 DORT WIRD DER MIETER DEM ANWALT DEN VERTRAG ZEIGEN.

All sentences were structured as follows: adverb-copula-verb complements-verb. The sentences were matched according to the number of syllables. All nouns and verbs were bi-syllabic, all other words were mono-syllabic. All nouns were masculine. Note, that in German, case is marked by the article. For masculine gender the nominative, dative and accusative are unequivocally marked by the corresponding articles. This does not hold for the feminine and neuter gender. Therefore, only masculine gender was used.

A random sequence of the 36 sentences was prepared such that two sentences of the same format or two variants of one proto-sentence did not directly succeed each other.

Subjects received the sentences printed on paper. They were asked to judge each sentence on a 5-point-scale ranging from "completely unacceptable" (1) to "completely acceptable" (5).

Results

Table 1 shows the mean rating values for the six sentence formats:

SDA	SAD	DSA	ASD	DAS	AVDS
4.71	3.64	2.85	2.26	1.81	1.66

A one-way analysis of variance was calculated, yielding a highly significant effect ($F(5,115)=142.13$; $p<.001$). According to a post-hoc Newman-Keuls test, the differences between all formats are significant, with the exception that DAS does not differ from AVDS (in this and all following Newman-Keuls test we used the significance level of $p=.05$).

Discussion

The rating data strongly support the acceptability ranking of formats predicted by the assumed word order theory, as far as isolated sentences are concerned. Thus, the linguistic intuitions on which the theory is based could be confirmed by studying an independent sample of subjects and by the application of appropriate statistics. Yet, we have already mentioned some reservations against metalinguistic judgements as sole data for testing linguistic theories. Therefore, we proceeded to apply further experimental methods to the issue in question. The following experiment used the method of delayed sentence matching.

Experiment II: Delayed sentence matching

The delayed sentence matching paradigm was developed by Ken Forster in order to study the issue of overgeneration (.i.Freedman & Forster, 1985;; .i.Forster, 1987;; .i.Forster & Stevenson, 1987);. In this paradigm the subject is presented with a first sentence which is followed by a second sentence displayed below the first one two seconds later. Subjects are asked to decide as quickly as possible whether both sentences are identical or not. Later, only identical sentences are analysed.

Of course, half of the sentences in the experiment are identical and half are different. If the sentences do not match, one of the words is replaced by another one:

JOHN KNEW THAT MARY WAS COMING.
JOHN FELT THAT MARY WAS COMING.

Using this method, Forster studied syntactically correct sentences and sentences with different violations of syntactic structure. The findings showed that subjects' reaction time was clearly dependent upon the degree of grammaticality. As compared to normal sentences, reaction times were delayed for sentences including violations of number agreement or regarding the position of quantifiers:

- (3) Where does bears usually hibernate?
- (4) The maid all put the rubbish out.

Forster argues that the comparison process can be carried out at different levels simultaneously. These levels include, for instance, graphemic features, letters, words, syntax, or meaning. This claim is substantiated by corresponding experimental findings. To give an example, semantically plausible sentences like

- (5) The dentist extracted the tooth.

are judged significantly faster than sentences without a valid semantic interpretation like

- (6) The clergy advertised the bones.

However, it seems not to be true that all levels are involved in all cases. This can be concluded from the following observation. The comparison of words (HOUSE/HOUSE) consumes less time than the comparison of letter sequences (HSEUO/HSEUO). The most plausible explanation is that in the first case the comparison is carried out at the word level whereas in the second case it is carried out at the grapheme level. The time differences can then be explained by

the number of elements which have to be taken into account in both cases (one vs. five, respectively).

That comparisons are carried out at the syntactic level is demonstrated by experiments testing 'syntactic prose' vs. random word sequences:

- (7) The jolder gints to his gloob.
- (8) The to gloob jolder gints his.

Here, subjects prove to be faster when presented with syntactic prose as compared to random word order. Extensive experimentation, however, revealed that this method is not sensitive to all kinds of syntactic violation. No effect was demonstrated, for example, for violations of the Specified Subject Constraint or of the Subjacency Constraint. It is still a matter of debate why reaction times are systematically delayed by certain syntactic violations, but not by others (cf. .i.Crain & Fodor, 1985;; .i.Forster, 1987;; .i.Forster & Stevenson, 1987);. Under discussion are the correctability and the semantic plausibility of sentences.

Method

Subjects

Subjects in our experiment were 30 students of the University of the Saarland at Saarbrücken. They were paid for participation.

Material

Six proto-sentences formed the basic material. Again, they all consisted of an adverb, a copula, a subject, an indirect object, a direct object and a verb:

DANN WIRD DER LEHRER DEM SCHÜLER DEN ATLAS LEIHEN.

BALD WIRD DER JUNGE DEM ONKEL DEN SCHLÜSSEL HOLEN.

DORT WIRD DER PFLEGER DEM KRANKEN DEN VERBAND WECHSELN.

DANN WIRD DER VATER DEM SÜGLING DEN SCHNULLER GEBEN.

BALD WIRD DER OBER DEM RENTNER DEN KUCHEN REICHEN.

DORT WIRD DER TANKWART DEM KUNDEN DEN WAGEN WASCHEN.

From all six proto-sentences all six possible word order variants (SDA, SAD, DSA, ASD, DAS, ADS) were realized, yielding 36 test sentences. Another 36 distractor sentences were needed because in the delayed sentence matching paradigm subjects are asked to judge whether two sentences are identical or not. They were constructed by changing either one of the nouns or one of the articles per test sentence. Across the six variants of each proto-sentence, all nouns and all articles were changed once. Nouns were changed such that a minimal number of letters was replaced by others in order to form another German noun. The changing of articles always included the last letter only. Thus, words were modified minimally in order to guarantee subjects' allocation of attention to the material. An example:

DANN WIRD DER LEHRER DEM SCHÜLER DEN ATLAS LEIHEN.

DANN WIRD DER LEHRER DEM SCHÜFER DEN ATLAS LEIHEN.

In addition to these 72 sentences, 36 filler-sentences were used. They only consisted of two verb complements and were not restricted according to the same principles as the test- and distractor-sentences. The 36 filler sentences were also derived from 6 proto-sentences by variation of word order. The six proto-sentences were:

DANN WIRD DIE MUTTER DAS MÄSLI ESSEN.

GESTERN HAT DAS HÄNDCHEN DEN HAUSHERRN GEBISSEN.

MORGENS MUSS DIE MAGD DIE KUH MELKEN.

EBEN HAT DER PATIENT DIE SPRITZE ERHALTEN.

MORGEN WIRD ERNA DIE TELLER KAUFEN.

ZULETZT WIRD DER BAUER DIE GERSTE MÜHEN.

The variants of the proto-filler-sentences were constructed by not only changing the order of the verb complements, but by including the verbs, copulas and adverbs as moving elements as well. For the filler sentences, an equal number of non-identical distractor pairs had to be constructed also. For these sentences, any word in a sentence could be altered, the only restriction being that the number of letters of the corresponding word was maintained and that the new letter string was a regular German word.

Eight different lists of the 144 sentences were prepared. First of all, 12 short lists were set up, each including the 12 proto-sentences (12 proto-test-sentences and 12 proto-filler-sentences). In each of these short lists, each sentence format (SDA to ADS) appeared once and half of the sentence pairs were identical. Identical and non-identical pairs of the very same sentence were separated by six short lists. The items in the short lists were randomized and

the 12 short lists were combined to yield a complete list of 144 items. By varying the combination of short lists and by varying the sequence of items (forwards vs. backwards), eight complete lists were constructed which were distributed to an equal number of subjects, as best as possible.

Procedure

The two sentences of each item were presented one above the other in the middle of a monochrome CRT-screen of a Personal Computer. Subjects were asked to decide as fast as possible whether two sentences were identical or not by pressing one of the two control-keys of the computer.

Before the experiment proper started six practice items were presented which did not appear again later. Subjects received feedback about their reactions. If they erroneously judged a non-identical pair to be identical, the position of the altered word was marked by a small star.

The 144 experimental items were presented in continuous succession. Each trial began with a short attention-signal which was a 500 Hz tone lasting 100 ms. After 200 ms the first sentence was presented for 4000 ms, followed by the second sentence presented below the first one. Both sentences were visible on the screen for again maximally 4000 ms. Reaction time measurement started as the second sentence was presented. When a subject pressed a key in time, the stimuli disappeared from the screen. Acoustic feedback was given in the case of a false reaction. If subjects did not react during the presentation of the second sentence, both sentences disappeared and a phrase was flashed on the screen asking the subject to respond faster. The inter-trial interval was set at 8500 ms which was independent from subjects' response times.

Results

In total, the percentage of errors was 11.3%. Subjects made more errors when the sentences were non-identical (19.8%) as compared to identical (2.7%). Filler-sentences yielded a smaller number of errors (6.3%) than test-sentences (16.3%). With regard to the test sentences only, the percentage of errors was 4.2% for identical items and 28.4% for non-identical items.

The analysis of reaction times was restricted to correct responses to identical items. Data more than three standard deviations from the subject's mean were eliminated from further analyses (< 0.5% of the data). A one-way analysis of variance was calculated for the six word order formats. This analysis yielded a significant effect ($F(5,145)=4.93$; $p<.001$). Table 2 presents the mean reaction times for the six sentence formats.

Table 2. Mean reaction times for the six sentence formats in Experiment II (in ms):

SDA	SAD	DSA	ASD	DAS	ADS
2574	2833	2618	2662	2689	2717

In a Newman-Keuls test SAD significantly differed from all other formats.

Discussion

The results obtained in the delayed sentence matching paradigm are not very revealing. The statistical analyses yielded a significant effect for the sentence format, but post-hoc tests showed that this is only due to particularly long reaction times for the format SAD, which is an unexpected finding. Regarding the non-significance of the remaining five formats, two alternative interpretations are conceivable. First, we must, of course, take into account

the possibility that, in contrast to our theoretical assumptions, the sentence formats are not really processed differently. Second, it might also be true that the delayed sentence matching paradigm turns out to be insensitive for revealing differences which actually exist.

The second hypothesis is substantiated by the findings which have already been mentioned above, showing that the method does not yield significant effects for violations of the Specified Subject Constraint or the Subjacency Constraint (.i.Forster, 1987),; for example. It might be the case that the delayed sentence matching paradigm is only sensitive to relatively gross violations of syntactic structure, such as violations of gender agreement, etc.

In order to provide evidence favoring one of the two hypotheses, we decided to repeat the experiment using exactly the same procedure, but including the ungrammatical word order variant used in the ranking study. If the method does not prove to be sensitive to such a deviation concerning the middle field in German, it is obviously inappropriate for studying the issue in question.

Experiment III: A modified replication of Experiment II

Method

Subjects

Subjects were 24 students of the Philipps-University at Marburg.

Material and procedure

The method and procedure were exactly the same as in the second experiment with two exceptions. First, we used different test sentences, namely those which had been prepared for the ranking procedure. Second, one of the formats, ADS, was replaced by an ungrammatical sentence structure, AVDS, as mentioned above.

Results

Table 3 shows the mean reaction times for the six sentence formats (in ms):

SDA	SAD	DSA	ASD	DAS	AVDS
2471	2519	2500	2566	2621	2526

The data were subjected to a one-way analysis of variance. No significant effect for sentence format ($F(5,115)=2.07$; $p=0.074$) was obtained.

Discussion

The replication of the delayed sentence matching paradigm yielded no difference between the six word order formats. In contrast to the second experiment, reaction times for the SAD-format did not differ from the other formats. Since even the ungrammatical form did not differ significantly from the others, we conclude that the delayed sentence matching paradigm is inappropriate for differentiating between word order variants in the German middle field.

Experiment IV: Delayed articulation

A second experimental paradigm which has been applied to the study of syntax is the method of delayed articulation (cf. .i.Ferreira, 1991);. In this paradigm, the subject is presented with a sentence and is asked to process the sentence

for as long as necessary to be able to reproduce it. When the subject presses a key, the sentence disappears from the screen. After a variable interval, the subject is prompted to begin to reproduce the sentence as fast as possible. The application of this method to the study of syntactic processing originates from the study of motor behavior. It was demonstrated that the time necessary to initiate a complex movement is dependent upon the degree of hierarchical structure underlying this movement (cf. .i.Rosenbaum, 1985);. Since speech production is a motor process as well, it seemed possible to study the impact of syntactic structures of different hierarchical complexity on articulation latencies.

Ferreira tested three levels of syntactic complexity. All sentences included the same number of words, but differed with respect to the complexity of the grammatical subject. Each sentence consisted either of an article, one or two adjectives and a noun, or of an article, a noun and a prepositional phrase, or of a noun and an embedded relative sentence:

- (9) The large and raging river empties into the bay
that borders the little town.
- (10) The river near their city empties into the bay
that borders the little town.
- (11) The river that stopped flooding empties into the bay
that borders the little town.

The analysis of encoding times and articulation latencies yielded significant effects. The more complex the sentences, the longer the initiation times for articulation were. Furthermore, only the complexity of grammatical subjects had any effect, whereas the complexity of grammatical objects turned out to be ineffective. Also, a variation of semantic plausibility failed to yield any effect.

With regard to articulation latencies, Ferreira explained her findings by arguing that subjects are confronted with sentences which are specified semantically and syntactically, but not phonologically. The transformation of a semantic-syntactic representation into a phonological representation takes place immediately before articulation. The time consumed by this transformation depends upon the number of nodes a sentence contains. This claim refers to the X-max-algorithm proposed by .i.Selkirk (1984);. This algorithm converts the syntactic representation of a sentence into a phonological one based on the maximal projection of a lexical category. Thus, Ferreira hypothesizes that the syntactic complexity of a sentence mirrors the number of its syntactic phrases which must be transformed into phonological phrases. The greater the number of phonological phrases which have to be constructed, the longer the initiation time for articulation will be.

Method

Subjects

Thirty students at the University of the Saarland participated in our experiment as subjects. They were paid for their services.

Material and procedure

The material consisted of the 36 test-sentences which had been used in Experiment II (i.e. the six proto-sentences in the six formats SDA, SAD, DSA, ASD, DAS, ADS). Each subject received a different sequence of these sentences. The 36 sentences were subdivided into 6 blocks of 6 sentences. Each block included all six proto-sentences and all six different formats. For each subject a random sequence of items in each block was generated. Furthermore, the sequence of blocks was varied across subjects.

The sentences were presented in the centre of a CRT computer screen. Each trial began with an attention signal (a tone of 100 ms duration). After another 100 ms, the test sentence was flashed on the monitor. Subjects were instructed to carefully read and encode the sentence, so that they were able to reproduce it directly afterwards. The encoding time was restricted to a maximum of 20 seconds. When the subject was ready, (s)he pressed a button and the sentence immediately disappeared. Shortly afterwards the subject was visually prompted to begin to reproduce the sentence as quickly as possible. The delay between the offset of the sentence and the articulation prompt varied in steps of 250 ms from 500 to 1000 ms. The prompt remained visible for 500 ms. Then the monitor was cleared for five seconds before the next trial began.

When the prompt for reproducing the sentence appeared, a pulse was sent to a DAT-recorder (Digital Audio Tape recorder) which in turn started a millisecond clock. When articulation began, a voice-key triggered the clock to stop. Because of the unreliability of voice-key measurements (cf. .i.Pechmann, Reetz & Zerbst, 1988),; the latencies of all utterances were later checked off-line by visualizing the acoustic signal.

Before the experiment proper started, subjects received six practice items.

Results

Both the encoding times of the sentences and their articulation latencies were subjected to statistical analyses.

(a) Encoding times

Table 4 displays the mean encoding times for the six sentence formats (in ms):

SDA	SAD	DSA	ASD	DAS	ADS
5094	5802	6262	6763	6647	7525

A one-way analysis of variance yielded a highly significant effect ($F(5,145)=16.69$; $p<.001$). A Newman-Keuls test showed that SDA yielded the shortest encoding times, whereas ADS yielded the longest encoding times of any format. In addition, SAD differs from ASD and DAS.

(b) Errors

All sentences reproduced in the experiment were checked to see if they were identical to the original sentences. Because of technical problems during recording, the data from one subject had to be excluded from further analyses. The data of the remaining 29 subjects were inspected for errors in wording. The number of errors and repairs in using the different sentence formats is shown in Table 5:

	errors	repairs	sum
SDA	15	3	18
SAD	41	5	46
DSA	34	4	38
ASD	33	6	39
DAS	48	6	54
ADS	35	4	39

(number of utterances per sentence format: 174)

Since the number of errors in the experiment is too small, no inferential statistics were calculated. However, it is obvious that less errors were made in the canonical format SDA than in the other formats. The most revealing error concerning us is the exchange of verb complements (other errors being the omission of words, the production of wrong words etc.). The Table 6 below depicts the number of verb complement exchanges for the different sentence formats:

	S <-> D	S <-> A	A <-> D
SDA	0	0	7
SAD	0	0	29
DSA	2	12	1
ASD	9	1	1
DAS	1	15	0
ADS	9	0	0

The absolute number of exchanges is small. Therefore, an interpretation of these data is rather speculative. However, two points are striking. On the one hand, the first complement in a sentence very rarely takes part in an exchange. On the other hand, the highest number of errors occurs in the format SAD.

(c) Articulation latencies

All utterances which were non-fluent or not reliably measurable were excluded from the analysis. Also, extreme reaction times (more than three standard deviations from the subject's mean) were discarded. Using this procedure, it turned out that for eight subjects, less than two-thirds of all data points remained. The data of these eight subjects were not considered in the statistical analyses, leaving 21 subjects in the sample. A one-way analysis of variance was calculated on the basis of the data of these 21 subjects. The mean articulation latencies for the six sentence formats (in ms) are depicted in the following Table 7:

SDA	SAD	DSA	ASD	DAS	ADS
776	786	761	768	834	774

The analysis of variance yielded a significant effect ($F(5,100)=3.11$; $p=.01$). According to a Newman-Keuls test, DAS differs from all other formats.

Discussion

In contrast to the method of delayed sentence matching, delayed articulation proved to be sensitive to differences between the six sentence formats. Yet, this claim must be restricted to the analysis of encoding times. The number of errors was too small to allow statistical analyses. However, the pattern observed (in particular the high number of exchanges between direct and indirect objects in the format SAD) might suggest that if a sentence begins with the grammatical subject, the tendency to continue with the indirect object is particularly strong, thus confirming principle OP3 (indirect objects precede direct objects). The articulation latencies were not particularly revealing. In general, the finding that they did not differ for the six sentence forms (with the exception of DAS) was not unexpected. It supports Ferreira's claim that differences in delayed articulation depend upon the number of phonological phrases. In all of our sentences, however, the number of phrases was identical. The analysis of encoding times yielded a very close correspondence between the data and the theoretical prediction. In statistical terms, we found that SDA, which was predicted to be the easiest format, had shorter encoding times than all other formats. The encoding times of ADS, which should be the most difficult format, were significantly longer than those of all others. Finally, the times required to memorize the SAD format, which was supposed to rank at the second position, were consistently shorter than those of ASD and DAS, which rank at positions four and five. Thus, the method differentiates between at least the two ends of the range of predicted acceptability and it confirms the special status of subject position.

Experiment V: Rapid serial visual presentation

Experiments dealing with the processing of very rapidly presented non-linguistic visual stimuli showed a close correspondence between the degree to which the material was structured and subjects' performance of recognition. The more the material was structured the better the recognition was (cf. .i.Mayzner, Tresselt & Cohen, 1966);. This method was applied by Ken Forster to the study of linguistic stimuli (.i.Forster, 1970);. He reasoned that syntactic structure should affect subjects' performance likewise.

He studied simple and complex sentences as well as random word sequences. The individual words were presented very briefly (for approximately 70 ms) in direct succession. Subjects were asked to repeat the sentences they had seen as exactly as possible immediately after presentation. Examples for simple sentences are the following:

Predicates: The young boy was very hungry.
Transitives: Alan has broken my mother's vase.
Adverbials: The kitten climbed over the fence.

Examples for complex sentences:

Reduced relatives: The truck Susan was driving crashed.
Unreduced relatives: The clothes that Mary wore vanished.
Complements: They persuaded him to work harder.
Time adverbials: Bob celebrated after his friends arrived.
Manner adverbials: Jim escaped by unlocking the door.

Note that these experiments were not designed to test any particular linguistic theory, but to develop an instrument for the assessment of linguistic complexity (as defined by a traditional transformational grammar). However, although simple and complex sentences consisted of the same number of words, a deep structure analysis reveals that one sentence was perceived in the case of simple sentences, and two were perceived in the case of complex sentences. The results yielded significant differences. On the one hand, sentences were better recognized than random word orders. On the other hand, subjects' performance with simple sentences was superior to their performance with complex sentences. Thus, the method proved to be sensitive to syntactic complexity. This claim was confirmed in follow-up experiments in which Forster was able to rule out the hypothesis that the differences are simply due to visual processing. Our next experiment made use of this technique.

Method

Subjects

Twenty-six students at the University of the Saarland volunteered as subjects. They were paid for participating.

Material

The critical test-sentences used were 36 different sentences which were prepared following the restrictions described in the method section of Experiment II. All six sentence formats occurred with the same frequency. These items were ordered using the same procedure as in Experiment IV. Altogether, twelve different lists of the 36 sentences were prepared and distributed to an equal number of subjects as best as possible.

Procedure

Each sentence was partitioned into six fragments: the adverb, the copula, the three verb complements, and the verb, e.g.

BALD / WIRD / DER OBER / DEM RENTNER / DEN KUCHEN / REICHEN

These fragments were presented on a computer screen in direct succession, each fragment for 150 ms. The subjects' task was to reproduce the sentence immediately afterwards as correctly as possible. They were explicitly instructed not to change the order of words and not to produce possible completions of sentences they had only partially recognized, but only to articulate what they

had seen. The instruction was made even stronger by announcing incomplete sentences (which did not actually appear).

The sentences were displayed as soon as the subject pressed a key on the computer's keyboard. Thus, the pace of the experiment was determined by the individual participant. Subjects' utterances were recorded on-line and additionally also recorded onto cassette. Later, the on-line recordings were compared with the tape-recordings and corrected if necessary.

Results

A number of different analyses were carried out in order to account for different measures of reproduction. Here, we restrict our report to the number of words, nouns, and complements correctly reproduced.

Table 8 shows the mean number of words for the six sentence formats which were correctly reproduced (independent from their position):

SDA	SAD	DSA	ASD	DAS	ADS
6.58	6.60	6.53	6.24	6.47	6.14

According to an analysis of variance the six formats differ significantly ($F(5,125)=3.63$; $p<.01$). A post-hoc Newman-Keuls test showed that in the formats SDA, SAD and DSA more words were reproduced than in the format ADS. All other differences failed to be significant.

Table 9 presents the mean number of nouns which were correctly reproduced (independent from their position):

SDA	SAD	DSA	ASD	DAS	ADS
1.99	2.01	1.90	1.81	1.88	1.75

An analysis of variance yielded a significant effect ($F(5,125)=3.07$; $p=.012$). A post-hoc Newman-Keuls test showed that in the sentence formats SDA and SAD more nouns could be reproduced than in the format ADS.

If we also take the position of the nouns (or their grammatical function) into account, the following table emerges showing how often subjects, indirect or direct objects were reproduced in the six sentence formats. Table 10:

	1st noun	2nd noun	3rd noun	sum
SDA	.86	.51	.62	1.99
SAD	.88	.57	.56	2.01
DSA	.84	.46	.60	1.90
ASD	.84	.51	.46	1.81
DAS	.84	.59	.45	1.88
ADS	.83	.55	.37	1.75

According to analyses of variance, the six sentence formats do not differ regarding the first two nouns ($p>.16$). With respect to the third noun, however, the analysis was highly significant ($F(5,125)=7.68$; $p<.001$). Post-hoc tests

(Newman-Keuls) revealed that a significantly larger number of nouns was reproduced in the formats SDA, SAD and DSA compared to ASD and ADS. Furthermore, SDA and DSA significantly differ from DAS.

Table 11 depicts the mean number of complements (independent from their position):

SDA	SAD	DSA	ASD	DAS	ADS
1.38	1.35	0.65	0.50	0.59	0.47

It is obvious from Table 11 that more complements (correct article + noun) were reproduced correctly if the first complement was the subject. This observation is strongly supported by an analysis of variance ($F(5,125)=28.9$; $p<.001$). A post-hoc Newman-Keuls test confirmed that SDA and SAD differ significantly from all other formats. No further differences are significant.

If one additionally takes into account the correct position of the complements, the following table emerges. Table 12:

	1st compl.	2nd compl.	3rd compl.
SDA	.72	.25	.40
SAD	.75	.39	.21
DSA	.12	.12	.42
ASD	.20	.18	.12
DAS	.10	.42	.07
ADS	.21	.13	.13

Three one-way analyses of variance were calculated. They yielded significant effects for all three complements (in all cases $p < .001$).

After post-hoc Newman Keuls tests, the following picture emerges. With regard to the first complements, SDA and SAD differ significantly from all other formats. Second complements were most often correctly reproduced in the formats DAS and SAD, which both differ significantly from all other formats. In addition, SDA differs from DSA and ADS. With respect to the third complements DSA and SDA both result in significantly more reproductions than all other formats. In addition, SAD and DAS differ from each other.

This picture can be basically reduced to the following observation. First complements are easier to reproduce if they are the subject of the sentence. Second and third complements are easier to reproduce if they are the direct objects of the sentence.

Discussion

These data support the claim that language users have a strong preference for sentences beginning with the grammatical subject. We have already provided evidence for this 'subject bias'. Taking this preference for granted, it is not surprising to find that first complements are better reproduced as subjects. But what about a noun of a direct or indirect object? If such a noun is erroneously reproduced in initial position although it occurred in second or third position, is it still marked as direct or indirect object? Mostly, this is not the case. When the indirect object was the first complement, it was reproduced as nominative in 86% of all cases by assigning the corresponding article ("der"). The corresponding percentage for direct objects was 70%. The tendency to utter the grammatical subject as first complement is proven when you consider that in 78% of all cases the nominative was assigned to the first complement. Only rarely did subjects move an actual nominative complement from second or third position into the initial position. Instead, they tended to assign the nominative to initial complements which had actually been direct or indirect objects.

A related phenomenon is that in second and third position direct objects were mostly reproduced correctly. This was partially due to an erroneous transformation of subjects and indirect objects into direct objects.

The data of this experiment strongly support the predictions derived from the assumed ordering principles regarding the easiest and the most difficult sentence formats. This holds for a number of dependent variables, such as the number of correctly reproduced words independent from their position, the number of correctly reproduced nouns independent from their position and, taking position into account, the number of correctly reproduced complements irrespective of position and the number of correctly reproduced first complements. All these measures show that SDA and SAD are the easiest to process and ADS the hardest.

Another very striking finding of this experiment is the impact of the tendency to assign the nominative to the first complement (the 'subject bias'). This is

substantiated by a great number of transformations of direct and indirect objects into grammatical subjects. Furthermore, we observed a significant tendency to assign accusative to the second complement, erroneously leading to transformations of grammatical subjects and indirect objects into direct objects. One possible explanation for this unexpected finding might be that if sentences in everyday language only include two verb complements, the second complement is more often a direct than an indirect object. This assumption is substantiated by the fact that if indirect objects and subjects were transformed into direct objects, this primarily occurred when only two complements were reproduced (in 77% of all cases). Yet, if direct objects or subjects were transformed into indirect objects the percentage of cases when only two complements were remembered was much smaller (22%).

Experiment VI: Sentence generation

In a series of production experiments word order in complex noun phrases has been studied experimentally by Pechmann and co-workers (cf. .i.Pechmann, 1994;; .i.Pechmann & Zerbst, 1992, 1993);. We investigated phrases which consisted of two adjectives denoting size and color and a noun. One of the methods we designed was the following. Subjects were presented with the three words of a phrase vertically aligned in morphologically unspecified form, e.g.

klein		blau	(blue)
blau	or	klein	(small)
Hut		Hut	(hat)

The subjects' task was to generate as quickly as possible a regular phrase following the order top-down. To do this, subjects had to morphologically specify the adjectives: "kleiner blauer Hut" or "blauer kleiner Hut" (note that in German adjectives are marked for case and gender). The analysis of articulation latencies showed that this method is sensitive to the order of prenominal adjectives. Canonical phrases in which the adjective denoting size preceded the adjective denoting color systematically had shorter latencies. The same holds for simple sentences of the form:

Dagmar		Perlen
kaufen	or	kaufen
Perlen		Dagmar

Here, the verbs were presented in the infinitive and had to be inflected by the subjects. Subject-initial sentences ("Dagmar kauft Perlen"; Dagmar buys pearls) had significantly shorter articulation latencies than object-initial sentences ("Perlen kauft Dagmar"; .i.Pechmann & Zerbst, submitted);. In our sixth experiment we applied a modification of this method to the study of word order in the middle field of German sentences. We prepared two versions of the experiment. In the first version (the so-called structure section), subjects were informed about the abstract structure of the sentences (e.g. subject-indirect object-direct object) and only then got the words to be inserted into the sentence slots. In the second version (the so-called content-section), subjects first received the words and then they were informed about the sentence structure. In both cases, subjects were asked to begin to generate the corresponding sentences as fast as possible. In the following, the two subtasks will be reported successively.

Experiment VIa: Application of a given syntactic form to different sentence contents

Method

Subjects

Twenty-six subjects participated in the experiment. They were paid for participating.

Material and procedure

The sentences tested in this experiment were the same 36 sentences we used in Experiments II and IV. These 36 test sentences were subdivided into six blocks. Each block contained all six proto-sentences, but in a different order. Before the items of a particular block were presented, subjects were informed about the syntactic structure they were supposed to assign to this block. This instruction was given by supplying the subjects with the abstract order of verb complements. Subjects were presented with five lines vertically aligned on a computer screen. The second to fourth line informed about the order of complements. If subjects saw, for example:

sentence beginning
indirect object
direct object
subject
predicate

they knew the abstract syntactic structure they were supposed to produce. In addition they were given a sample sentence, e.g.

sentence beginning:	Bald wird
indirect object:	Hauswirt
direct object	Mangel
subject	Mieter
predicate	zeigen

They were told that the expected sentence would be "Bald wird dem Hauswirt den Mangel der Mieter zeigen". The critical task of the subjects was to choose the correct articles. They were instructed to take their time until they were confident that they were able to assign the sentence format to different sentence contents. If a subject felt ready, (s)he pressed a key. After a brief pause of 200 ms an acoustic attention signal sounded for 100 ms, followed by a pause of 150 ms. Then, the five lines which specified the order of sentence elements were displayed until the subject initiated the presentation of the following item by pressing the key again. Each format-block started with three dummy items which were not considered for later analyses. The order of sentence formats was systematically varied across subjects. Subjects' utterances were recorded on DAT. The measurement of articulation latencies followed the same procedure which was described for Experiment IV.

Results

All utterances were checked for errors and repairs. All errors and repairs were excluded from the data pool. The data of six subjects were completely excluded from further analyses because these subjects had a very high percentage of errors and repairs. The data of another subject were excluded because this subject delivered less than four valid data points in more than two sentence formats. Thus, the statistical analyses were based on the results of 16 subjects.

Extreme values were excluded from the remaining data pool by following the usual criterium of plus/minus three standard deviations from the subject's mean. For each utterance, four time points were determined by visual inspection of the digitalized speech signal: articulation onset and the onset of the three articles. The procedure for excluding extreme data points was done for each dependent measure separately.

Table 13 depicts the mean latencies for the onset of articulation for the six sentence formats. The latencies in this and all other tables of Experiments VI and VII refer to milliseconds.

SDA	SAD	DSA	ASD	DAS	ADS
815	766	800	820	849	1034

A one-way analysis of variance yielded a highly significant effect ($F(5,75)=5.74; p<.001$). A post-hoc Newman-Keuls test showed that the latencies of ADS exceeded those of all other formats which do not differ statistically. The onset times of the first articles are depicted in Table 14.

SDA	SAD	DSA	ASD	DAS	ADS
1320	1244	1367	1402	1428	1743

The analysis of variance was significant ($F(5,75)=11.92$; $p<.001$). Again, a Newman-Keuls test showed that ADS differs from all other formats.

Table 15 displays the articulation latencies of the second articles

SDA	SAD	DSA	ASD	DAS	ADS
2035	1954	2127	2153	2349	2635

The analysis of variance was significant again ($F(5,75)=10.02$; $p<.001$). According to a Newman-Keuls test, the latencies under ADS were longer than those under all other formats. In addition, DAS yielded longer latencies than SDA and SAD.

The onset times of the third articles are provided by Table 16.

SDA	SAD	DSA	ASD	DAS	ADS
2729	2655	2761	2838	3091	3434

The analysis of variance turned out to be significant ($F(5,75)=11.11$; $p<.001$). A Newman-Keuls test proved that ADS differs from all other formats. Furthermore, DAS yielded longer latencies than SDA, SAD, DSA, and ASD.

Summarizing the statistical analyses regarding the speech onset latencies of the utterance as such, and of the three complements, a consistent picture emerges. In all cases the format which should be the least acceptable regarding the theoretical assumptions (ADS) has significantly longer latencies than all other formats. The remaining five formats do not differ statistically with regard to articulation onset and first article. With regard to the second article, DAS yielded longer articulation latencies than the two forms with the subject in initial position. In the case of the third article, DAS yielded longer articulation latencies than all other forms except ADS. Thus, we find strong evidence that if the subject is in third position, utterances become particularly hard to process.

Of course, the onset times just discussed are cumulative data. Latencies of the third article include the latencies of articulation onset and of the onsets of the first and second article. Thus, one might hypothesize that differences at later points in time are just a function of differences in speech onset. In order to check this argument the differences were calculated between ADS and all other formats for each time point. They are presented in Table 17:

	SDA	SAD	DSA	ASD	DAS	ADS
speech onset			219	268	234	214
1st article	423		499	376	341	315

2nd article	600	681	508	482	286
3rd article	705	779	673	596	343

This table shows that the differences between ADS and the other formats cannot be reduced to differences in speech onset. Rather, the differences increase the later the point of measurement in the sentence.

In further analyses, we inspected the durations of three different sentence segments: the durations from sentence beginning until the onset of the first article, from the first to the second article, and from the second to the third article.

Table 18 depicts the mean durations from articulation onset to the onset of the first article split up separately for the six sentence formats.

SDA	SAD	DSA	ASD	DAS	ADS
506	478	567	582	579	709

The formats differ significantly ($F(5,75)=10.26$; $p<.001$). The duration of the first segment was longest under ADS. Additionally, the first segment was shorter under SAD than under DSA, DAS, and ASD.

Table 19 below shows the mean durations from the first to the second articles (second segment).

SDA	SAD	DSA	ASD	DAS	ADS
715	710	760	751	921	892

Once more, the analysis of variance yielded a significant effect ($F(5,75)=6.21$; $p<.001$). According to a Newman-Keuls test, both ADS and DAS significantly differ from all other formats.

Finally, we calculated the durations from the second to the third articles (third segment) which are depicted in Table 20.

SDA	SAD	DSA	ASD	DAS	ADS
694	701	634	686	742	799

The analysis of variance proved that significant differences exist ($F(5,75)=5.21$; $p<.001$). Descriptively, the two formats with the subject in second position, DSA and ASD, yielded the shortest durations. According to a Newman-Keuls test, the duration of the third segment under ADS exceeded all other durations, except under DAS. In addition, third segments are longer under DAS than under DSA.

It is striking that the two formats with the subject in second position yielded the shortest durations, i.e. subjects needed less time than objects to be produced as second complements. Again, under DAS and ADS the longest durations

are obtained, i.e. under those formats in which a second object is being produced as second complement, and where the subject comes in third position.

In a final set of analyses, we looked at the durations of complements which are dependent upon the position of the subject. The first analysis showed that the duration from sentence onset to the first article increased with a later position of the subject (Table 21):

SXX	492
XSX	574
XXS	644

(S=Subject, X=other complements)

The difference between the three durations is highly significant ($F(2,30)=12.25$; $p<.001$). A Newman-Keuls test confirmed that durations under SXX differ from durations under XSX, which also differ from durations under XXS. Note, however, that the long durations under XXS are due to the ADS format only.

Finally, the factors 'position of subject' and 'order of complements' (first vs. second verb complement) were analyzed in a two-factorial statistical design. Table 22 depicts the corresponding means of duration.

	SXX	XSX	XXS	
C1	713	756	907	792
C2	698	660	770	709
	705	708	838	751

(C1 = duration of first complement (onset of first article to onset of second article), C2 = duration of second complement (onset of second article to onset of third article))

The analysis of variance yielded significant main effects for subject position ($F(2,30)=12.89$; $p<.001$) and order of complements ($F(1,15)=15.92$; $p<.01$) and a significant interaction ($F(2,30)=4.89$; $p=.015$).

The main effect subject position was further tested in a post-hoc Newman-Keuls test which revealed that first and second complements in subject-final sentences (XXS) had significantly longer durations than in sentences in which the subject occupied the first or second position (SXX, XSX).

The main effect order of complements demonstrates that second complements were shorter in duration than first complements.

The interaction was tested by application of a Newman-Keuls test which showed that the longest durations are to be observed for first complements with subject in third position (Xxs; critical complements are marked by capital letters). They were longer than all other complements. In addition, subjects in second position (xSx) had shorter durations than objects in second position in subject-final sentences (xXs) and objects as first complements followed by a subject (Xsx).

The interaction can be summarized as follows. For both first and second complements, no difference can be observed between subject-initial and subject-middle sentences. In contrast, significantly longer durations are to be observed in subject-final sentences. Furthermore, first complements yielded longer

durations than second complements in subject-middle and subject-final, but not in subject-initial sentences.

Experiment VIb: Application of a given sentence content to different syntactic forms

Method

Subjects

The same 26 students of the University of the Saarland participated in this part of the experiment, as participated in Experiment VIa.

Material and procedure

The same 36 sentences were used (six proto-sentences and six formats) as in the other part of Experiment VI. Here, subjects were informed about the words which constituted the verb complements of a particular proto-sentence, the adverb and the verb. Then, they received information about the sentence form (the order of complements) and were asked to start articulation of a sentence as fast as possible.

First, subjects were informed about the words of a particular sentence and about which word went into which complement, e.g.

opening: Bald wird
indirect object: Hauswirt
direct object: Mangel
subject: Mieter
predicate: zeigen

Given this information, subjects learned that they were supposed to produce the sentence "Bald wird dem Hauswirt den Mangel der Mieter zeigen". As soon as they were confident of the sentence content, they pressed a key and presentation of the first item started. Seven blocks of six items each were prepared. The first block was a practice block which was not analyzed later. The proto-sentence of this practice block was 'Dann wird der Anwalt dem Richter den Beweis liefern'. During the presentation of experimental items, subjects only received the words of the sentence without articles, but no information about the grammatical functions was given, e.g.

Bald wird
Hauswirt
Mieter
Mangel
zeigen

Thus, the sentence format was initiated by the vertical order of the words which constituted the different complements. The order of item blocks (i.e. sentence formats) was systematically varied across subjects. Block internally, two sentences with the same first verb complement did not directly succeed each other. For each subject, each of the six formats was used once as first item. The sentence content task was always the second part of Experiment VI.

Results

In this part of the experiment two subjects were removed from the subject pool because their percentage of errors and repairs was too high. Another seven subjects were excluded because they provided less than four data points in more than two sentence formats. Thus, the sample which was used for statistical analysis included 17 subjects. In all other respects, the data analysis of this second part of Experiment VI proceeded along the same lines as the analysis of the first part.

Table 23 depicts the mean speech onset times for the six formats.

SDA	SAD	DSA	ASD	DAS	ADS
801	838	865	852	824	830

An analysis of variance turned out to be insignificant ($F(5,90)=1.35$; $p=.25$).

The next Table 24 presents the latencies from stimulus onset to the onset of the first article.

SDA	SAD	DSA	ASD	DAS	ADS
1309	1344	1575	1564	1593	1476

The analysis of variance yielded a highly significant effect ($F(5,90)=10.69$; $p<.001$). According to a Newman-Keuls test, SDA and SAD yielded shorter latencies than all other formats.

Table 25 depicts the onset times of the second articles:

SDA	SAD	DSA	ASD	DAS	ADS
2018	2074	2353	2339	2519	2396

The analysis of variance was significant ($F(5,90)=16.43$; $p<.001$). A Newman-Keuls test revealed that SDA and SAD yielded shorter latencies than all other formats. In addition, ASD and DSA yielded shorter latencies than DAS.

The onset times of the third articles are displayed by Table 26:

SDA	SAD	DSA	ASD	DAS	ADS
2674	2749	2975	2959	3199	3101

These onset times are also statistically different ($F(5,90)=15.65$; $p<.001$). A Newman-Keuls test revealed the same picture as with the second articles: SDA and SAD yielded shorter latencies than all other formats, ASD and DSA yielded shorter latencies than DAS.

Overall, the picture which emerges from the statistical analysis of articulation latencies of the sentence beginnings and of the different articles is very consistent. No differences with respect to sentence formats can be found regarding the sentence beginnings. Yet, with regard to the articles, those formats which have the subject as first complement yielded the shortest latencies, and the format DAS always yielded the longest latencies.

In the next stage we analyzed the durations of different sentence segments. Again, the first segment was determined as the delay from sentence beginning to the onset of the first article. The mean durations are presented in Table 27:

SDA	SAD	DSA	ASD	DAS	ADS
507	506	710	712	769	646

The analysis of variance turned out to be significant ($F(5,90)=13.15$; $p<.001$). The Newman-Keuls test showed that durations of first segments of SDA and SAD significantly differ from all those of all other formats. Additionally, durations under DAS significantly differ from durations under ADS.

Next, we examined the mean durations from onset of the first articles to onset of the second articles (Table 28):

SDA	SAD	DSA	ASD	DAS	ADS
709	730	778	775	926	920

Once more, the analysis of variance was significant ($F(5,90)=12.59$; $p<.001$). According to a Newman-Keuls test, DAS and ADS yielded significantly longer durations than all other formats.

The durations from onset of the second to the onset of the third articles are depicted in Table 29:

SDA	SAD	DSA	ASD	DAS	ADS
656	675	622	620	680	705

A significant difference between the sentence formats was obtained by an analysis of variance ($F(5,90)=12.78$; $p<.001$). This finding was qualified by a Newman-Keuls test, which showed that the durations of these segments are shorter of the formats ASD and DSA than of all other formats. Additionally, durations under SDA differ from durations under ADS.

Finally, we determined the durations of the different sentence segments dependent upon the position of the subject. Table 30 depicts the corresponding durations from sentence onset to onset of the first article.

SXX	509.8
XSX	710.5
XXS	699.0

The analysis of variance was significant ($F(2,36)=25.66$; $p<.001$). A Newman-Keuls test showed that the durations of SXX are significantly shorter than the durations of the other two formats neither of which differs from the other.

As a final step, a two-factorial analysis of variance was calculated again, with 'position of subject' and 'order of complements' as factors. The mean durations are presented in Table 31:

	SXX	XSX	XXS	
C1	721	779	933	811
C2	667	622	692	660
	694	701	813	736

(C1 = duration of first complement (onset of first article to onset of second article), C2 = duration of second complement (onset of second article to onset of third article))

Both main effects and the interaction are highly significant (position of subject: $F(2,36)=40.64$; $p<.001$; order of complements: $F(1,18)=50.71$; $p<.001$; interaction: $F(2,36)=24.99$; $p<.001$).

Inspection of the corresponding means showed that second segments were significantly shorter than first segments. Regarding the factor subject position, a Newman-Keuls test proved that the durations of first and second complements in subject-final sentences were significantly longer than in sentences in which the subject occupied the first or second slot for verb complements. For the interaction, the following picture emerges. The shortest segments were subjects as second complements (xSx). They were shorter than any other segment. In contrast, the durations of objects in sentence-initial position followed by another object (Xxs) were longer than all other segments. This also held for objects in first position followed by the subject (Xsx), except that these segments were shorter than the segments mentioned before (Xxs). Finally, the durations of subjects in first position (Sxx) were longer than the durations of objects in second position followed by another object (sXx). One might summarize that in all three cases (i.e. subject positions), first complements yielded longer durations than second complements. With regard to first complements, durations increased the later the subject appeared in the sentence. Regarding second complements, this is also true for subject-middle and subject-final sentences, whereas for subject-first and subject-middle sentences the pattern is reversed.

Discussion

In both sections of Experiment VI a rather similar pattern of findings was obtained, which is quite consistent with our findings from the previously reported experiments. Regarding speech onset latencies, data from the structure section provide strong evidence for the sentence form ADS being the most difficult to process. ADS yielded longer latencies for all of the four time measurement points than all other sentence forms. Inspection of the differences from measure point to measure point revealed that these differences cannot be reduced to a difference occurring at the beginning of the sentences. In addition, particularly long latencies were obtained in DAS-sentences. Both findings support the claim that the position of the subject is crucial for the ease of syntactic planning processes. Whereas the speech onset data from the structure section mainly differentiate between the most difficult sentence form and all others, the speech onset data from the content section primarily reflect the other end of the range of predicted acceptability. No differences at articulation onset could be observed. This is probably due to the fact that the first two words were identical for each set of six items which only differed in syntactic form. This did not hold for the structure section of the experiment. However, with regard to the onsets of all three articles, it is true that the two sentence forms with subject as the first complement, SDA and SAD, had shorter latencies than all other forms. In addition, we found for the second and third articles that DAS had significantly longer latencies than DSA and ASD, i.e. the subject in final position again led to an increase in articulation latencies. A very similar picture also emerged with respect to the durations of different sentence segments. We analyzed the durations from speech onset to the onset of the first article, from onset of the first article to onset of the second article, and from onset of the second article to onset of the third article. Regarding the first segments, ADS turned out to yield longer durations than all other forms in the structure section, whereas SAD yielded particularly short ones. Both findings again strongly underline the impact of the subject's position. In the content section SDA and SAD yielded shorter durations than all other forms. In addition, initial segments in DAS-sentences were particularly long.

Regarding the second segments (onset of first article to onset of second article, which roughly corresponds to the first complement), both sections of the experiment yielded exactly the same outcome saying that durations were significantly longer in ADS- and DAS-sentences than in all other sentence forms. This again supports the impact of the subject in final position.

Evidence for the strong influence of the subject role was also provided by the durations of the third segments. In the content section DSA- and ASD-sentences yielded significantly shorter durations than all other sentence forms, i.e. when subjects were produced as second complements, they had shorter durations than when objects were produced as second complement. In addition, the indirect objects (second complements) in ADS-sentences yielded longer durations than in SDA-sentences, again demonstrating the impact of the subject position.

Particularly long durations of the second complements in subject-final sentences (ADS and DAS) could also be observed in the structure section of the experiment. In a final series of analyses, we inspected the durations of segments dependent upon the position of the subject. Again, a very similar picture emerged in the structure section and in the content section. The impact of the subject in final position was strongly supported by an analysis of the sentences' beginnings. In both sections of the experiment the durations of the first segments (from speech onset to the onset of the first article) were consistently shortest if the subject was in first position (SXX-sentences). In the structure section, we also found that the durations were shorter if the subject was the second complement as compared to the third one.

Inspection of the durations of the first and second complements revealed that in both sections of the experiment, second complements were significantly shorter than first complements. This is mainly due to the particularly long durations of first complements in subject-final sentences (see below).

If the data are collapsed across the durations of first and second complements, we found that in both sections of the experiment XXS-sentences yielded longer durations than both SXX- and XSX-sentences. This again is mainly a result of the long durations of first objects in subject-final sentences.

Finally, we analyzed the interaction of subject position and order of complements. The patterns obtained in both sections of the experiment are very similar. Once again, they emphasized the impact of the subject's position. Particularly long durations were yielded by first objects in sentence-final sentences, and particularly short durations were yielded by subjects as second verb complement.

General Discussion

The starting point of the empirical part of the project reported in this paper was a theory concerning the order of verb complements in the German Mittelfeld. The theory is based on certain ordering principles and on Uszkoreit's approach to describing their interaction. This theory strongly emphasizes the influence of pragmatic factors. However, since we lack established experimental methods to assess the degree of acceptability concerning the sentence types in question, we decided to study isolated sentences in the first stage, thus initially excluding the impact of pragmatic contexts. One of our principal aims was to assess different degrees of acceptability using data obtained from language processing. Therefore, our experiments tapped into various processing aspects by examining, for instance, reaction times, articulation latencies, or the number of reproduced words.

Concerning isolated sentences, only three of the five operating principles proposed by Uszkoreit are relevant. They regulate the relative order of verb complements depending on their thematic role. For the verbs we selected for the experiments, they predict that grammatical subjects precede objects and indirect objects precede direct objects. It is assumed that the acceptability (reflected in the ease of processing) of different word order variants corresponds to the number of principles which are violated. We further assumed that priority is given to the position of the subject as compared to the position of the two objects relative to each other. Thus, the following rank order for the six sentence formats which we studied was predicted:

SDA > SAD > DSA > ASD > DAS > ADS

A series of six experiments was carried out to put the predicted ranking to test. The methods we used were: a metalinguistic ranking procedure, the delayed sentence matching paradigm (with replication), delayed articulation, rapid serial visual presentation, and a sentence generation experiment. Before we draw our conclusions from the data we collected, we will summarize the important findings of all experiments (except the delayed sentence matching experiments which turned out to be unrevealing) in the following table.

Table 32. Summary of statistically significant differences between the six sentence formats obtained in all experiments reported above (without delayed sentence matching)

- (1) Rating SDA > SAD > DSA > ASD > DAS

- (2) Del. articulation
 - a) encoding times SDA < all others
 - SAD < ASD, DAS
 - ADS > all others

 - b) speech onset DAS > all others

- (3) RSVP
 - a) number of words ADS < SDA, SAD, DSA
 - b) number of nouns ADS < SDA, SAD
 - c) number of compl. SDA, SAD > all others

- (4) Sentence generation
 - Structure section
 - a) onset latencies
 - speech onset: ADS > all others
 - 1st article: ADS > all others
 - 2nd article: ADS > all others;
 - DAS > SDA, SAD
 - 3rd article: ADS > all others
 - DAS > SDA, SAD, DSA, ASD

 - b) durations
 - 1st segment: ADS > all others
 - SAD < DSA, ASD, DAS
 - 2nd segment: DAS, ADS > all others
 - 3rd segment: ADS > all others except DAS
 - DAS > DSA
 - c) first complements yielded longer durations than second complements
 - d) durations of first and second complements: XXS > SXX, XSX
 - e) Xxs > all others
 - xSx < xXs, Xsx

Content section

- a) onset latencies
 - speech onset: no differences
 - 1st article: SDA, SAD < all others
 - 2nd article: SDA, SAD < all others
 - DAS > DSA, ASD
 - 3rd article: SDA, SAD < all others
 - DAS > DSA, ASD
- b) durations
 - 1st segment: SDA, SAD < all others
 - DAS < ADS
 - 2nd segment: DAS, ADS > all others
 - 3rd segment: ADS > SDA
 - DSA, ASD < all others
- c) first complements yielded longer durations than second complements
- d) durations of first and second complements: XXS > SXX, XSX
 - e) Xxs > all others
 - xSx < all others
 - Xsx > all others except Xxs
 - Sxx > sXx

In the experiments, a number of rather different dependent variables were tested: metalinguistic judgements, memory encoding times, number of reproduced words and complements, articulation latencies and durations during generation. Yet, the data provide strong evidence for the two principles which we regarded as most influential, the ones that state that subjects tend to precede objects. However, one restriction has to be made concerning this point. In all of our experiments, the grammatical subjects were the agents, in terms of thematic roles. Thus, our results cannot simply be generalized to non-agentive subjects. In order to study whether non-agentive subjects would yield a different pattern of findings, further experiments are needed.

The third operating principle which applies to isolated sentences suggests that indirect objects precede direct objects. We assumed that this principle is less important than the principle concerning the subject position. Although the evidence supporting this third principle is less convincing, it is nevertheless substantial. In the delayed articulation experiment, for instance, only SDA-sentences needed significantly less time to be encoded than DSA-, ASD-, DAS- and ADS-sentences. This was not true for SAD-sentences. Also, subjects needed less time to encode DSA-sentences than ADS-sentences. This did not hold for ASD-sentences. The same picture emerged with rapid serial visual presentation, again supporting the claim that DSA-sentences were easier than ASD-sentences. In addition, only ADS-sentences, but not DAS-sentences, achieved the worst rate of reproduction in this experiment. Finally, in the sentence generation experiment we often found that ADS-sentences, but not DAS-sentences yielded the longest articulation latencies or durations.

Taken together, we conclude from our findings that (1) we succeeded in setting up experiments which consistently differentiated between the word order variants we were investigating. (2) The ordering principles, which predict a rank order of these variants with regard to their acceptability or ease of processing, could be confirmed. There is very strong evidence in favor of the subject-first-principle, but also substantial evidence for the third principle, which claims that indirect objects precede direct objects. The pattern of results is consistent with the claim that the last principle is nested under the subject-first-principle. We would like to stress the fact that rather different methods yielded a very consistent picture. Thus, the data have been successfully cross-validated and can be regarded as rather reliable.

Our data also support two of the three hypotheses underlying Uszkoreit's theory on the interaction of ordering principles. We observed a gradual increase in processing effort, spanning the scale between grammatical and ungrammatical sentences. We did not see any evidence for a jump function separating grammatical linearizations from the ones that are considered unacceptable. Of course, it would be desirable to further explain our findings by a detailed analysis of the fine-grained processes which unfolded in our experiments. Regrettably, we feel that psycholinguistics is still far from providing the knowledge necessary to do this adequately. We do not dispose of generally accepted processing models for the production or understanding of natural language. One of the most promising models, at least from our point of view, is the model of simulated annealing proposed by Gerard Kempen and his colleagues (.i.de Smedt & Kempen, 1987;; .i.Kempen & Vosse, 1989);. Using a metaphor derived from the biosynthesis of proteins, they claim that fragments of linguistic structures are relatively free to combine. Restrictions necessary to avoid random results are set by corresponding feature lists. Of course, one might assume that the linearization of verb complements is controlled by such lists as well. This should be investigated theoretically and, possibly, by performing appropriate computer simulations. However, this is clearly beyond the scope of the present paper.

We set out to cross-validate metalinguistic judgements about the acceptability of different syntactic structures by performing experiments tapping into the actual processing of corresponding sentences. One might wonder why we did not additionally make use of existing corpora. It seems probable that the acceptability of linguistic structures is reflected in the frequency of their usage. There is a simple answer to this question. Although work on such corpora is now increasing rapidly, we still do not dispose of a German corpus suited for our purposes. The crucial problem is that the unit of existing corpora is the word. What would be needed are annotated corpora which allow us to search for abstract syntactic structures of the kind we are interested in. They simply do not exist yet.

Two further points will be raised. The first point concerns the relationship between metalinguistic judgements and processing data. Of course, linguistic intuitions are data. The aim of our project has not been to invalidate linguistic judgements as empirical data. Rather, its aim was, first, to empirically test the reliability of such judgements and, second, to test whether the metalinguistic intuitions can be backed up by studying the processing of corresponding sentences. However, the relationship between metalinguistic judgements and processing aspects must not at all be simple. This is to say that we do not claim, of course, that linguistic preferences based on metalinguistic judgements will always be explainable in terms of actual processing data. In contrast, linguistic preferences might be based on stylistic factors and the like. However, if degrees of acceptability of different syntactic structures are found to closely correspond to differences in the processing of such structures, we regard this as a plausible explanation for the differences in acceptability. The second, and final point, concerns our future prospects. So far, we have restricted ourselves to the analysis of isolated sentences. Yet, linguistic theory strongly emphasizes the impact of the pragmatic context for the acceptability of sentence forms. We started with the investigation of isolated sentences in order to obtain some kind of baseline data. These data are now available and our plan is to proceed studying the impact of pragmatic factors by embedding the sentences into appropriate contexts. Thus, we will try to test to what extent pragmatic factors overrule the impact of grammatical roles given appropriate thematic contexts.

References

- Behaghel, O. (1932). *Deutsche Syntax*. Heidelberg: Winter.
- Bock, J.K.; Loebell, H. & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, 99 (1), 150-171.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Crain, S. & Fodor, J. (1985). How can grammars help parsers? In D. Dowty, L. Karttunen & A. Zwicky (eds.), *Natural language parsing: Psycholinguistic, computational, and theoretical perspectives*. Cambridge: Cambridge University Press.
- Smedt, K. & Kempen, G. (1987). Incremental sentence production, self-correction and coordination. In G. Kempen (ed.), *Natural language generation: New results in Artificial Intelligence, psychology, and linguistics*. Dordrecht: Kluwer.
- Engelkamp, Johannes & Zimmer, H.D. (1983). *Dynamic aspects of language processing*. Berlin: Springer.
- Engelkamp, Judith; Erbach, G. & Uszkoreit, H. (1992). Handling linear precedence constraints by unification. In: *Proceedings of the 30th Meeting of the ACL*, Berkeley.
- Erbach, G. (1992). Ambiguity and linguistic preferences. In: *Proceedings of ECAI 1992*, Trento.
- Ertel, S. (1977). Where do the subjects of sentences come from? In S. Rosenberg (ed.), *Sentence production*. Hillsdale: Lawrence Erlbaum.
- Fanselow, G. (1990). Scrambling as NP-Movement. In G. Grewendorf & W. Sternefeld (eds.), *Scrambling and barriers*. Amsterdam: Benjamins.
- Ferreira F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, 30, 210-233.
- Forster, K.I. (1970). Visual perception of rapidly presented word sequences of varying complexity. *Perception & Psychophysics*, 8, 215-221.
- Forster, K.I. (1987). Binding, plausibility, and modularity. In J.L. Garfield (ed.), *Modularity in knowledge representation and natural-language understanding*. Cambridge, MA: MIT Press.
- Forster, K.I. & Stevenson, B.J. (1987). Sentence matching and well-formedness. *Cognition*, 26, 171-186.
- Freedman, S.E. & Forster, K.I. (1985). The psychological status of overgenerated sentences. *Cognition*, 19, 101-131.
- Gazdar, G. & Pullum, G.K. (1982). *Generalized phrase structure grammar: A theoretical synopsis*. Bloomington, Indiana: Indiana University Linguistics Club.

Gazdar, G.; Klein, E.; Pullum, G.K. & Sag, I.A. (1985). Generalized phrase structure grammar. Cambridge: Blackwell.

Greenberg, J.H. (1966). Some universals of grammar with particular reference to the order of meaningful elements. In J.H. Greenberg (ed.), *Universals of language*. Cambridge, MA: MIT Press.

Grewendorf, B. (1988). *Aspekte der deutschen Syntax: Eine Rektions-Bindungs-Analyse*. Tübingen: Narr.

Hale, K. (1983). Warlpiri and the grammar of nonconfigurational languages. *Natural Language and Linguistic Theory*, 5-47.

Kempen, G. & Vosse, Th. (1989). Incremental syntactic tree formation in human sentence processing: A cognitive architecture based on activation decay and simulated annealing. *Connection Science*, 1, 275-292.

Koster, J. (1987). *iDomains and dynasties: The radical autonomy of syntax*. Dordrecht: Foris.

Kunze, J. (1975): *Abhängigkeitsgrammatik*. *Studia grammatica XII*. Berlin.

Lenerz, J. (1977). Zur Abfolge nominaler Satzglieder im Deutschen. *Studien zur deutschen Grammatik 5*. Tübingen: Narr.

Levelt, W.J.M. (1972). Some psychological aspects of linguistic data. *Linguistische Berichte*, 17, 18-30.

Levelt, W.J.M. & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, 14, 78-106.

Mayzner, M.S.; Tresselt, M.E. & Cohen, A. (1966). Preliminary findings on some effects of very fast sequential input rates on perception. *Psychonomic Science*, 6, 513-514.

Melcuk, I.A. (1974). *The experience of the linguistic models "Meaning-Text"* (auf Russisch). Moskau: Nauka.

Oliva, K. (1992). The proper treatment of word order in HPSG. In: *Proceedings of COLING 1992*, Nantes.

Pechmann, Th. (1994). *Sprachproduktion*. Wiesbaden: Westdeutscher Verlag.

Pechmann, Th.; Reetz, H. & Zerbst, D. (1989). Kritik einer Meámethode: Zur Ungenauigkeit von voice-key Messungen. *Sprache & Kognition*, 8, 65-71.

Pechmann, Th. & Zerbst, D. (1992). Zum Einfluá syntaktischer Faktoren auf verschiedenen Stufen des Sprachproduktionsprozesses. In G. Antos & H. P. Krings (Hg.), *Textproduktion. Neue Wege der Forschung*. Trier: Wissenschaftlicher Verlag Trier.

Pechmann, Th. & Zerbst, D. (1993). Zur allmählichen Verfertigung komplexer Nominalphrasen beim Reden. In S. Felix, Ch. Habel & G. Rickheit (Hg.), *Kognitive Linguistik: Repräsentation und Prozesse*. Opladen: Westdeutscher Verlag.

Pechmann, Th. & Zerbst, D. (submitted). Zum Einfluß der syntaktischen Struktur auf den Zeitverlauf von Prozessen der grammatischen Kodierung bei der Sprachproduktion.

Pfeiffer, O.E.; Pucek, M. & Sgall, P. (1994). Die Thema-Rhema-Gliederung im Deutschen und ihre automatische Analyse. In U. Klenk (Hg.), *Computatio Linguae* II (ZDL-Beiheft 83). Stuttgart: Franz Steiner Verlag.

Pollard, C. & Sag, I.A. (1987). *Information-based syntax and semantics*, Vol. 1: Fundamentals. CSLI Lecture Notes No. 13. Stanford: Stanford University.

Reape, M. (1990). A theory of word order and discontinuous constituency in West Continental Germanic. In E. Engdahl & M. Reape (eds.), *Parametric variation in Germanic and Romance: Preliminary investigations*. ESPRIT Basic Research Action 3175 DYANA, deliverable R1.1.A.

Rosenbaum, D.A. (1985). Motor programming. In H. Heuer, U. Kleinbeck & K.-M. Schmidt (eds.), *Motor behavior: Programming, control, and acquisition*. Berlin: Springer.

Selkirk, E.O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Sgall, P.; Hajicova, E. & Panevova, J. (1986). The meaning of the sentence in its semantic and pragmatic aspects. Dordrecht: Reidel.

Sternefeld, W. (1990). Scrambling and minimality. In G. Grewendorf & W. Sternefeld (eds.), *Scrambling and barriers*. Amsterdam: Benjamins.

Tesnišre, L. (1959). *Elements de syntaxe structurale*. Paris.

Thiersch, C. (1982). A note on "Scrambling" and the existence of VP. *Wiener Linguistische Gazette*, 27/28.

Uszkoreit, H. (1986). Constraints on order. *Linguistics*, 24, 883-906.

Uszkoreit, H. (1987). Word order and constituent structure in German. CSLI Lecture Notes No. 8. Stanford: Stanford University.

.Anfang Index.
Behaghel (1932), 6
Bock, Loebell & Morey, 1992 10
Chomsky, 1957, 1965) 3
Chomsky, 1981) 3
Crain & Fodor, 1985 14
De Smedt & Kempen, 1987 47
Engelkamp & Zimmer, 1983 10
Engelkamp, Erbach & Uszkoreit, 1992 7
Erbach, 1992 7
Ertel, 1977) 10
Fanselow, 1990 5
Ferreira, 1991) 11, 19
Forster & Stevenson, 1987 11
Forster & Stevenson, 1987) 13, 14
Forster, 1970) 11, 24
Forster, 1987 11, 13, 14
Forster, 1987) 17
Freedman & Forster, 1985 13
Freedman & Forster, 1985) 11
Gazdar & Pullum, 1982) 3
Gazdar, Klein, Pullum & Sag, 1985 4
Grewendorf, 1988 5
Hale, 1983) 3
Kempen & Vosse, 1989) 47
Koster, 1987 5
Kunze, 1975) 3
Lenerz, 1977 7
Lenerz, 1977) 5
Levelt & Kelter, 1982) 7
Levelt, 1972) 7
Mayzner, Tresselt & Cohen, 1966) 23
Melcuk, 1974 3
Oliva, 1992 7
Pechmann & Zerbst, 1992, 1993) 29
Pechmann & Zerbst, submitted) 30
Pechmann, 1994 29
Pechmann, 1994) 11
Pechmann, Reetz & Zerbst, 1988), 20
Pfeiffer, Pucek & Sgall, 1994) 7
Pollard & Sag, 1987 6
Pollard & Sag, 1987, 4
Reape, 1990) 4, 6, 7
Rosenbaum, 1985) 19
Selkirk (1984) 20
Sgall et al., 1986) 6
Sgall, Hajicov & Panevov , 1986) 3
Sternefeld, 1990) 5
Tesnišre, 1959) 3
Thiersch, 1982 5
Uszkoreit (1986) 5, 8
Uszkoreit (1987) 5
Uszkoreit, 1987) 4
.Ende Index.