Clause Union and Clausal Position

Josef Bayer, Tanja Schmid and Markus Bader

University of Konstanz, Germany

1 Introduction

Inspired by Evers (1975), the classical generative view of German control infinitives with *zu* has been since the eighties that (i) the infinitive is a CP which, in the context of a certain class of coherence-inducing verbs, can be pruned so that reanalysis can take place, and that (ii) the infinitival CP can stay in situ, i.e. to the immediate left of the matrix verb, or undergo extraposition. Thus, the string in (1) should in principle be ambiguous between (2a) and (2b).

(1) *daß Max mir das Lexikon zu kaufen empfohlen hat.*
     that Max me the lexicon to buy   recommended has
     'that Max has recommended to me to buy the lexicon.'

(2) a. $[\dots [_{CP} \dots V_2] V_1]$ bi-clausal
     b. $[\dots V_2 V_1]$ mono-clausal

(2a) would essentially be equivalent to the extraposition option in (3).

(3) *daß Max mir empfohlen hat [das Lexikon zu kaufen].*
     that Max me recommended has   the lexicon to buy
     'that Max has recommended to me to buy the lexicon.'

However, for Bech (1955/57) (1) would have been only a coherent, i.e. mono-clausal construction. We shall see below that with respect to language use, (2b) is preferred over (2a), and (3) is preferred over (2b). In other words, why should (2a,b) exist, if (3) exists? And in particular, why should (2a) as an intraposed version of the preferred extraposed version in (3) exist, if it is the least preferred construction anyway? In this article we will provide evidence that in German the structure in (2a) does not arise as such. There are two reasons for this: (i) If restructuring is an option, this option *must* be taken. In this case, (1) will always be structured as in (2b). (ii) If restructuring is impossible, the clausal infinitive is either not positioned as in (2a) or it is not a CP. If [… *zu*+V] is a CP with a null-complementizer, it may be scrambled or topicalized to SpecCP. If it is a bare clause without a complementizer, it can arise to the immediate left of the verb. In that case it is subject to the same licensing conditions as V_2 in the mono-clausal structure seen in (2b).

We will approach the question of clause union and clausal position from both an empirical and a comparative-theoretical perspective. The article is structured as follows: In section 2, *Infinitival Complementation in Language Processing*, we will present the results of an empirical investigation of intra- and extraposed coherent and non-coherent constructions including frequency in a corpus of written language, graded grammaticality judgments, and processing load during reading. Section 3, *Syntactic and Comparative Investigations*, describes the syntactic structure of intraposed coherent and non-coherent infinitives, and the licensing of

pre-verbal infinitives in immediately pre-verbal ('intraposed') position. This part contains among other things information about complementation in Bangla, a South-Asian SOV-language which resembles German in allowing both pre- and post-verbal complements, and seems to shed new light on the apparently free variation between (2a) and (3). Section 4 contains *Conclusions* in which we will try to integrate the findings presented. An appendix on *Re-positioned 'zu'* contains material that gives additional support to the account presented in section 3.

2 Infinitival Complementation in Language Processing

In this section, we will summarize several results from an ongoing series of experimental investigations on infinitival complementation in German. In particular, we will cite results from a corpus study, a questionnaire study, and a self-paced reading study that have investigated intraposition and other clause union properties. For reasons of space, we will discuss these experiments only insofar as they address the question as to which structure is preferably assigned to intraposed infinitival complements in German.

With regard to the Human Sentence Processing Mechanism (HSPM), we will make three assumptions.[1] First, the HSPM is a serial device, which means that it always computes a single, fully specified structure. Second, the HSPM works in an incremental manner, as captured in the *Left-to-Right Constraint* of Frazier and Rayner (1988) (cf. (4)). Third, when faced with a syntactic ambiguity, the HSPM is guided by economy principles like the *Simplicity* principle of Gorrell (1995) (cf. (5).

(4)   *Left-to-Right Constraint* (Frazier and Rayner, 1988)
      Each item is incorporated into a constituent structure representation of a sentence (essentially) as the item is encountered.

(5)   *Simplicity* (Gorrell, 1995)
      No vacuous structure building.

What do these assumptions predict when applied to a sentence with an intraposed infinitival complement? First, when parsing such a sentence, each word will be integrated immediately into an ongoing syntactic structure. For sentences with an intraposed infinitive, there will always be a local or global ambiguity as to whether the sentence ultimately receives a mono-clausal or bi-clausal analysis. Given *Simplicity,* the HSPM should impose the simplest structural analysis that is possible, which, for the constructions at hand, clearly is the coherent, mono-clausal structure. Thus, when considering data from language comprehension, we should find evidence for a preference of coherent infinitivals over non-coherent infinitivals.

2.1 Extraposition and Intraposition: A Corpus Study

A corpus study was conducted in order to obtain quantitative information about the 56 control verbs used in a following questionnaire study (cf. 2.2. below). The corpus search was conducted on the Mannheim COSMAS corpus consisting of more than 1.500 million word-forms of newspaper texts and literature (http://www.ids-mannheim.de). A main objective of the corpus search was to extract the frequency of preverbal, i.e. intraposed, as opposed to post-verbal, i.e. extraposed, infinitival complements. For each verb, sixty randomly selected occurrences were analyzed.

2

With respect to the position of infinitival complements, we found a strikingly low rate of intraposed constructions in the corpus. The main results are the following:

1. Less than 4% of infinitival complements occurred in intraposed position.
2. Less than 20% of all verbs in our sample occurred with an intraposed verbal element.
3. These were the following nine verbs: *beabsichtigen* ('to intend'), *versuchen* ('to try'), *drohen* ('to threaten'), *aufhören* ('to stop'), *angeben* ('to indicate'), *hoffen* ('to hope'), *raten* ('to recommend'), *anfangen* ('to start'), *beginnen* ('to begin").

With respect to the structure of the intraposed infinitival complements, all occurrences in our sample were such that matrix verb and embedded verb were adjacent. That is, we did not find sentences like the one in (6) in which the embedded infinitive is separated from the matrix verb by an intervening element, which automatically forces a non-coherent construction.

(6)   *..., daß sie das Buch zu lesen **mehrfach**   versucht hat.*
         that she the book to read repeatedly tried   has
        'that she repeatedly has tried to read the book.'

To sum up, according to the corpus study, intraposition is clearly disfavored for infinitival complements: Intraposed infinitival complements occurred at a low rate and only with a few verbs. This shows that they - although not ungrammatical - represent a marked option. That embedded verb and matrix verb are never separated in the intraposed constructions of our sample hints at a coherent, i.e., mono-clausal structure in cases where an intraposed infinitive is used after all.

2.2 Coherent and Non-Coherent Constructions: A Questionnaire Study

A questionnaire study was carried out in order to investigate native speakers' intuitions with respect to various (non-)coherence-sensitive tests as found in the literature (cf. Bech (1955/57), Grewendorf (1988), Stechow and Sternefeld (1988) among others). Only control matrix verbs were used in the bulk of this study. The main objectives were to investigate (i) for which verbs restructuring is an option at all, and (ii) whether the preverbal, i.e., intraposed construction patterns with coherent or non-coherent structures. The study and the results are documented in full detail in Schmid, Bader and Bayer (forthcoming). Here, we will focus on the structural status of intraposed constructions: Are they preferably analyzed as coherent, non-coherent, or are both structures equally available?

Altogether, seven different constructions involving infinitival complements were tested. Four constructions required a clause union structure (coherent constructions), two constructions disallowed clause union (non-coherent constructions), and one construction was ambiguous with respect to clause union according to standard generative references on this topic (e.g., Grewendorf, 1988; Stechow and Sternefeld, 1988). One sample sentence from the questionnaire study is shown exemplarily for two coherent constructions, the two non-coherent constructions, and the ambiguous construction in

3

(7) through (9).

(7) **Coherence Tests**

    a. **Long scrambling of a pronoun**
        Was den Aufsatz betrifft,   so ist klar, warum ihn Karla
        what the article   concerns so is clear why   it   Karla
        zu schreiben beschlossen hat.
        to write   decided   has
        'As for the article, it is clear why Karla decided to write it.'

    b. **Wide scope of negation**
        Karla hat keinen Aufsatz zu schreiben beschlossen und Friederike auch nicht.
        Karla has no   article   to write   decided   and Friederike also not
        'Karla hasn't decided to write an article, and neither has Friederike.'

(8) **Non-Coherence Tests**

    a. **Extraposition of the infinitival complement**
        Karla hat beschlossen, den Aufsatz zu schreiben.
        Karla has decided   the article to write
        'Karla decided to write the article.'

    b. **Narrow scope of negation**
        Karla hat keinen Aufsatz zu schreiben beschlossen   und Friederike auch.
        Karla has no   article   to write   decided   and Friederike too
        'Karla decided not to write an article, and so did Friederike.'

(9) **Ambiguity with respect to Coherence? – Intraposition**

        Karla hat den Aufsatz zu schreiben beschlossen.
        Karla   has the article to write   decided
        'Karla has decided to write the article.'

Coherent constructions have in common that they are transparent for otherwise clause-bounded processes. In the questionnaire study four constructions were tested for this property in total, two of which are given in

5

(7):

(7a) illustrates 'long' scrambling of a pronoun. When scrambling of a pronoun from a position in the embedded infinitive over the matrix subject leads to a grammatical result, this suggests a coherent construction. No clause boundary is present between the embedded verb and the matrix verb.

The same argumentation holds when a negated element in the embedded infinitive takes scope over the matrix verb, as must be the case in sentence

(7b). The wide scope reading of the negation that inheres the negative indefinite keinen Aufsatz ('no article') is forced by a negated conjunct (and … also not) following the main predicate.

The two constructions that only allow a non-coherent, bi-clausal structure, are shown in (8). (8a) illustrates the extraposition of the infinitival complement. When extraposition leads to a grammatical sentence, the extraposed phrase must have been a constituent of its own. When the infinitival construction is opaque for trans-clause processes, the construction in question is non-coherent. This is the case in (8b). In this sentence, the negated element inside the infinitival complement may not take scope over the matrix verb but only over the embedded verb. In the questionnaire, the narrow scope negation was forced by a positive conjunct (and … too).[2]

The last construction tested in the questionnaire study is intraposition of the embedded infinitive with no additional material intervening between the embedded and the matrix verb. This is illustrated by the example in (9). Sentences with intraposed infinitives were included in the questionnaire study to see how acceptable they are, and especially to see whether they are perceived as ambiguous between a coherent and a non-coherent construction as suggested by, e.g. Grewendorf (1988) and von Stechow and Sternefeld (1988).

In total, 56 control verbs were tested as matrix verbs. We concentrated on control verbs as matrix verbs in this study because they appear both in coherent and non-coherent constructions, and because there is still a debate in the literature about the range of control verbs allowing clause union. The 56 control verbs were classified according to three criteria:

1. The control property of the verb: Subject control versus object control.
2. The syntactic function of the embedded infinitive: The embedded infinitive either replaced a direct object or a prepositional object.
3. Additional objects: In addition to the infinitival complement, the verbs either had no further object, or an additional dative object, or an additional accusative object.

The resulting five subclasses according to these criteria are shown in Table 1:

Table 1: Materials – subclasses of control verbs

Control	Infinitive in function of	Additional objects	Examples
Subject	Accusative	∅	versuchen ('to try"), beschließen ('to decide')
Subject	Accusative	Dative	drohen ('to threaten"), versprechen ('to promise")
Subject	PP	∅	aufhören ('to stop"), klagen ('to complain")
Accusative Object	PP	Accusative	auffordern ('to ask"), ermahnen ('to urge")
Dative Object	Accusative	Dative	erlauben ('to allow"), verbieten ('to forbid")

For the questionnaire study, 70 test sentences (14 sentences per verb class), and 60 filler sentences were randomly ordered. 35 participants were asked to rate the acceptability of each sentence on a scale from 1 ('very plausible, easy to understand") to 5 ('very complicated, difficult to understand").

The main results of this study are shown in Figure 1 (for more details cf. Schmid et al., forth-coming). The y-axis shows the rating-scale from 1 to 5, and the x-axis shows the different constructions. Whereas the results for the ambiguously coherent intraposed construction and the two non-coherent constructions (extraposed and narrow scope) are shown separately, the results for the four coherent constructions have been collapsed and are shown under the label 'mean coherence". The mean rating that has been assigned to each of the four constructions is given separately for the five verb classes.
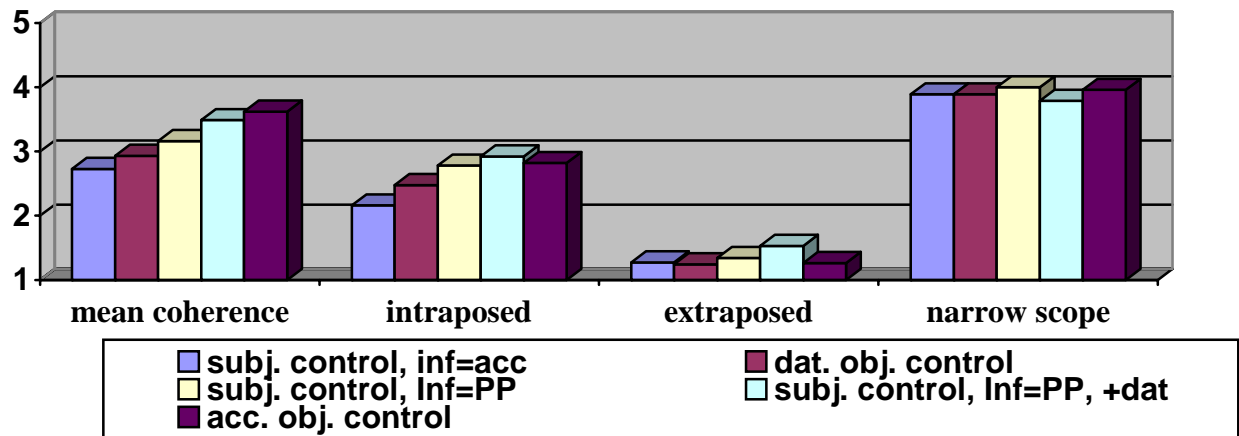

Figure 1: Main Results of Questionnaire Study on Coherent and Non-Coherent
          Infinitival Constructions

The main results of the questionnaire study concerning the status of intraposed infinitival complements can be summarized as follows:

1. The coherent constructions show a rather different picture when compared to the two non-coherent constructions (extraposed and narrow scope): The rating of coherent constructions showed a clear verb-class effect which was absent for non-coherent constructions.
2. Sentences with intraposed infinitival complements – which are assumed to be ambiguous between a coherent and a non-coherent structure – clearly patterned with the coherence and not with the non-coherence tests in showing clear verb-class effects. This is further support for the hypothesis that intraposed infinitivals are treated by the HSPM as coherent constructions.
3. The verb class that was judged best for the coherent constructions was the subclass of subject-control verbs in which the infinitival complement replaces the direct object. Next came the subclass of object control verbs in which the controller is a dative object, and the infinitival also replaces the direct object. That these two verb classes fared best on the coherence tests fits well with the hypothesis that coherence is only allowed when the infinitival complement is a direct object (cf. Bayer and Kornfilt (1990); Haider (1993: 251)).
4. For each single verb type, extraposition was rated better than intraposition (mean 1.3 vs. 2.6).
5. Intraposed sentences with narrow scope of negation were rated as being rather deviant (mean 3.9). Because these sentences require a non-coherent structure, this deviance would at least partially follow if the HSPM indeed has a preference for interpreting intraposed in-finitives as coherent. While the sentence-final conjunct *und ... auch* ('and ... also') which

9

enforces the narrow scope reading might have contributed to the enhanced complexity of these sentences, conjuncts of this sort pose no problems as such. The questionnaire also contained different types of control sentences, among which sentences like

(10). These received rather good acceptability ratings.

(10) *Karla hat einen Aufsatz schreiben müssen, und Werner auch.*
       Karla had an   article write   must   and Werner too
       'Karla had to write an article, and Werner had too.'

2.3 Scope of Negation: A Reading Time Experiment

To test whether the linguistic intuitions that were obtained in the rating study are reflected in online language processing of (non-)coherent infinitives, we conducted a self-paced reading experiment. The main objective of this experiment was to test the hypothesis that intraposed infinitives are preferentially assigned a coherent, mono-clausal structure during on-line language comprehension. As described above, the results on intraposed infinitivals obtained in the prior questionnaire study provide some initial evidence for this hypothesis, insofar as intraposed infinitivals clearly patterned with the unambiguous coherent constructions in being dependent on verb-class.

To test this hypothesis of a coherence-preference for sentences with intraposition, subjects had to read sentences in which an intraposed infinitive contained a negative quantifier such as *keines von den Büchern* ('none of the books"). The test sentences always ended with the negated conjunct *und … auch nicht* ('and … also not"):

(11) *Der Opa   hat keines von den Büchern   zu lesen versucht*
       the grandpa   has none of   the books   to read tried
       und der Onkel auch nicht.
       and the uncle also not
       'Grandpa didn't try to read any of the books, and the uncle didn't either.'

A negated conjunct is only compatible with the wide scope reading of negation, i.e., with a reading like *it is not the case that grandpa tried to read any of the books*. This reading in turn is only possible when the embedded infinitival complement is constructed coherently.

30 sentences of the type shown in (11) were constructed, with each sentence appearing in 6 conditions according to six verb classes. Five of these verb classes were the verb classes of control verbs already used in the questionnaire study. The sixth verb class were modal verbs. These are always coherent. The experimental method was the self-paced reading method with word-by-word non-cumulative presentation. This method works as follows. Participants are seated in front of a computer screen. They read sentences by pressing a response button for each word. Upon each button press, the next word appears, and the last word disappears. Thus, always only a single word is visible.

What do our hypotheses about the HSPM predict for reading such sentences? As already explained above, due to the *Left-to-Right Constraint* and *Simplicity* (cf. (4) and (5)), the HSPM will compute a coherent structure for a sentence with intraposed infinitival. This means that the clause-final matrix verb (*versucht* in (11)) will be attached into a mono-clausal structure, irrespective of its lexical specification. As soon as it has been attached, the HSPM will engage in processes of checking and linking (cf. Mitchell, 1994) with regard to the verb's lexical properties. As a by-product, the HSPM will determine whether the verb is compatible with a coherent structure or not. In case the verb cannot engage in the formation of a coherent construction, some kind of reanalysis is called for, which should be reflected in enhanced reading times on the matrix verb. No such reanalysis is necessary for verbs which allow for the coherent construction. Thus, when taking the mean coherence rating that was obtained in the questionnaire study described in the last section as a measure of how well a verb is compatible

with a mono-clausal structure, we expect that the better a verb was judged as coherent in the questionnaire study, the faster it should be read in the reading experiment. In sum, there should be a positive correlation between mean-coherence score in the questionnaire and reading times on the matrix verb.

What should we expect if the HSPM does not prefer to compute a coherent structure, but instead has a preference for the non-coherent structure? Such a preference might result, for example, if the HSPM works according to a head-driven schema, that is, if the parser can insert a phrase into the ongoing syntactic representation only after its head has been encountered in the input. Such a parser, which would not work strictly incremental, might first come up with a bi-clausal structure. When processing the matrix verb, there are then two possibilities. Either the parser simply sticks to the bi-clausal representation, or it reanalyzes it in case the verb has a strong tendency for coherence. Thus, we would expect either no correlation between reading times on the matrix verb and mean-coherence scores, or even a negative correlation because reading times should be slower for more coherent verbs which, under the head-driven schema, would induce reanalysis as an extra step.

Turning to the clause-final negation, recall that this negation is only compatible with the wide scope construal of the negative quantifier in the embedded infinitival (*keines von den Büchern* ('none of the books" in (11)). This wide scope construal in turn is only compatible with a coherent construction. Whenever a matrix verb that is not allowed in a coherent construction induces reanalysis toward a non-coherent structure, reading times should increase on the clause-final negation. No such effect is expected when the construction stays coherent throughout the parsing process.

Reading times of special interest were the reading times on the matrix verb, the 'spill over" region (*und …*), and the last two words (*auch nicht*). Since our main interest was in the relationship between reading times and coherence properties of verbs, we show in Table 2 the correlations between reading times and mean coherence scores as obtained in the questionnaire study presented in the preceding section.

Table 2: Correlations between reading times and questionnaire results
(starred correlations are significant)

	infinite verb	matrix verb	spill-over region	*auch* 'also"	*nicht* 'not"
clause union compatibility	-0,035	0,479*	0,410*	0,292	0,231

Reading times on the matrix verb as well as on the immediately following spill-over region showed a significant positive correlation with the mean coherence score: The better a verb scored on the coherence tests the faster it was read. This is exactly what was predicted under the hypothesis that a coherent structure is computed for these sentences. If a coherent structure is computed while reading the sentences, the matrix verb will fit into the ongoing structure according to its coherence properties. A coherence-inducing verb fits in well, allowing the sentence processor to proceed with the next words quickly. A verb disallowing coherence, in contrast, does not fit in well, and thus reading times will be slowed down. That this effect was not only visible on the matrix verb but also in the spill-over region is not unusual for self-paced reading experiments and probably reflects just some kind of inertia in button-pressing.

13

If readers switched to a non-coherent structure when encountering a verb not allowing coherence, we would expect a correlation with mean coherence also on the final negation *nicht* because the negative conjunct acts as a trigger of coherence. Contrary to expectation, reading times on the two final words (*auch nicht*) requiring wide scope did not correlate significantly with the mean coherence scores obtained in the questionnaire study. This suggest that readers do not reanalyze the initial coherent structure toward a non-coherent structure on encountering a verb which resists clause union. That is, the mismatch between structural analysis and lexical requirements that occurs if a verb is not a clause union verb does not seem to be strong enough to trigger any kind of structural revision.

One possible alternative to assuming a general parsing preference for coherence might be that there is a semantic preference for the wide-scope reading (corresponding to the mono-clausal structure) as compared to the narrow-scope reading (bi-clausal structure). However, in a further selfpaced-reading study which investigated sentences like (11) but without any negative element (e.g. *Der Opa hat ein Buch zu lesen versucht, und der Onkel auch*), a similar result was found, that is, a positive correlation between reading times and mean-coherence scores. Thus, while it is certainly true that in many instances a narrow-scope reading is less expected than a wide-scope reading (cf. note 2), these further results show that a semantic bias alone cannot explain the whole pattern of results described above. On the other hand, this bias might be a further reason why we did not find a correlation at the clause-final negation. It might simply be that the syntactic parsing-preference for coherence and the semantic bias towards wide-scope were strong enough for subjects to hold on to the structure computed from the outset, even if this structure was not fully compatible with certain requirements of the verb.

2.4 Summary of Empirical Investigations

All three of the empirical studies summarized above lend support to the hypothesis that intraposed infinitives are preferentially assigned a coherent, mono-clausal structure. Whereas extraposed infinitival complements are always clause-like and involve therefore a bi-clausal analysis, this option is strongly marked for intraposed infinitival complements being preferably analyzed as non-clause like by the parser. In the next section we will turn to syntactic observations which pertain to the question of bi-versus mono-clausality.

3 Syntactic and Comparative Investigations

Quantitative and behavioral data as shown in the previous section are important contributions to linguistic theorizing. The strong preference in parsing pre-verbal infinitival complements as coherent constructions may invite the conclusion that Bech was right after all in not considering any clausal status for them. Syntactic structures such as (2a) would then simply not exist, and the quantitative and behavioral data would map onto the admissible structural assignments of competence grammar more or less directly. In the present section we will take a closer look at infinitival complementation in order to avoid premature conclusions. In doing so, we will also draw on data from Bangla (also known as Bengali), an Indo-Aryan South Asian language in which finite complements may appear both in pre- and in post-verbal position. Our focus will be on the formal properties of complements as expressed by complementizers or by the infinitival marker *zu* in German.

3.1 The Distribution of *daß*-CPs in German

It is a well-known though still badly understood fact that German (as well as Dutch) resists the appearance of finite CPs in direct object (DO) position.

(12) a. *Hans hätte vermutlich bezweifelt* [*daß Maria kommt*].
    Hans had-subj presumably doubted that Maria comes
    'Hans would have presumably doubted that Maria will show up.'
  b. *?*Hans hätte vermutlich* [*daß Maria kommt*] *bezweifelt*.
  c. *Hans hätte* [*daß Maria kommt*]₁ *vermutlich* t₁ *bezweifelt*.
  d. [*Daß Maria kommt*]₁ *hätte Hans vermutlich* t₁ *bezweifelt*.

Within GB-theory, Stowell (1981: 146) proposed a principle of *Case Resistance* (CRP) according to which 'Case may not be assigned to a category bearing a Case-assigning feature". Finite clauses such as those in (12) (and also *to*-infinitives) have a +Tense feature which is a Case-assigning feature. So CPs would be banned from the DO-position (as well as from the subject-position, P-object position etc.).[3] (12a) would then be the result of extraposition as a strategy to evade this principle, similarly the topicalization seen in (12d). The scrambling operation in (12c) would be seen as adjunction outside the site of Case assignment. In more recent theorizing the question is more whether the CP would ever be merged as in (12b) or raised into the pre-verbal position. If no merging in or movement into this position is possible, structures like (12b) would never arise.

3.2 Distribution of Infinitives in German and Adjacency

Infinitives are even more challenging because they do appear in pre-verbal position. If Stowell (1981) was right in subsuming *to-/zu*-infinitives under the categories blocked by *Case Resistance* they should not be allowed there. German *zu*-infinitives would be subject to the same fate as finite CPs. Given that there is no string-wise discrimination between coherence and non-coherence for verbs which allow both construals, it seems difficult to approach the question. Nevertheless, there is an interesting test that can shed light on the issue. This test rests on the fact that German (like many other SOV-languages) shows the phenomenon of rightward movement (or stranding, depending on one's analysis) of relative clauses or PPs of different

phonological sizes. Instead of *Ich habe mich [dafür] entschieden* ('I have decided on it") one can also say *Ich habe mich entschieden [dafür]*.[4] As Haider (1993; 2003), van Riemsdijk (1994) and Bayer (1995) observe, such movement / stranding is not allowed when the infinitive occupies the left-adjacent position of the matrix verb.[5]

(13) **Coherent construction:** *versuchen*

    a. *Ich habe [mich dafür   zu entscheiden] versucht.*   **pre-V**
       I   have REF there-for to decide   tried
       'I have tried to decide on it.'
    b. **Ich habe [[mich ~~dafür~~ zu entscheiden] dafür] versucht.*   **pre-V + ex**
    c. *Ich habe versucht [[mich ~~dafür~~ zu entscheiden] dafür].*   **extraposed+ex**
    d. *Ich habe [[mich ~~dafür~~   zu entscheiden] dafür]$_1$*   **scramb+ex**
       I   have   REF there-for to decide   there-for
       schon   mehrmals   erfolglos   t$_1$ versucht.
       already more-than-once successlessly   tried
       'I have already more than once without success tried to decide on it.'

Extraposition leads to ungrammaticality only if the infinitive is in immediately pre-verbal position. Is this so because coherence is mandatory for *versuchen* as suggested by Haider (2003)? The intervention of *dafür* would block the process of head-to-head movement or base-generation of the V-cluster. Indeed, material can extrapose in such cases but only to the very end of the clause, i.e. to the right of *versucht*: *Ich habe mich ~~dafür~~ zu entscheiden versucht dafür*. The fact that this is impossible with verbs that disallow coherence is one of the classical tests for clause union in cases such as *versuchen*. Notice that the non-coherent verb *auffordern* ('to ask") does not allow this: **Ich habe ihn [sich ~~dafür~~ zu entscheiden] aufgefordert dafür* ('I asked him to decide on it'). The less widely known surprise is that even in the context of a non-coherence verb extraposition from the infinitival complement is blocked as long as the clause is in V-adjacent position:

(14) **Non-coherent construction:** *auffordern*

    a. *Ich habe ihn [sich dafür   zu entscheiden] aufgefordert.*   **pre-V**
       I   have him REF there-for to decide   asked
       'I have asked him to decide on it'
    b. **Ich habe ihn [[sich ~~dafür~~ zu entscheiden] dafür] aufgefordert.* **pre-V + ex**
    c. *Ich habe ihn aufgefordert [[sich ~~dafür~~ zu entscheiden] dafür].* **extraposed+ex**
    d. *Ich habe ihn [[sich ~~dafür~~ zu entscheiden] dafür]$_1$ schon*   **scramb+ex**
       mehrmals t$_1$ *aufgefordert.*
       'I have already more than once asked him to decide on it.'

What is so special about the left V-adjacent position that any sort of manipulation of the right edge of the infinitival complement will lead to a strong effect of ungrammaticality? The parser's strong preference for coherent constructions which has been discussed in section 2 could indicate that non-coherent intraposed infinitives such as (14a) are simply not admitted by the competence grammar. Given an input like (14a), the parser would then try to cope with it 'as if it were' a coherent construction. According to this reasoning, the deviance of (14b)

could still be explained along the lines of obligatory V-cluster formation as predicted by head movement or base-generation of a coherent construction. Thus, the adjacency requirement either holds independently of coherence, or this particular positioning of the infinitive triggers coherence no matter what, i.e. even in the presence of a verb like *auffordern* which, according to general wisdom, does not trigger coherence.[6] In order to approach this alternative we will next take a look at an SOV-language which shows no overt signs of clause union.

3.3 An Answer from Bangla

Bangla (also known as Bengali) is an Eastern Indo-Aryan SOV-language with a mixed system of pre- and post-verbal finite complements. Like various other head-final languages, Bangla also shows rightward movement of backgrounded light elements that we have used as a test of adjacency or adjacency disruption in 3.2. Unlike German, Bangla allows also personal pronouns to the right of the verb.[7] Why this is so is so far unclear, but does not matter here because we are simply using the construction as a test. (15b) shows that rightward pronoun movement is fine in post-verbal complements. (16b) shows that, just as in German, such movement leads to ungrammaticality in immediately pre-verbal position.[8] The adjacency requirement is also relaxed when the complement has undergone leftward movement as seen in (16c). Between the topicalized complement and the main clause there is an intonation break.

(15) **Post-verbal complement**

    a. *ami ʃune -cʰilam [paʃ-er baɽi -r kukur-ʈa toma-ke kamre-cʰe].*
       I   heard-have   next-poss house-poss dog -cl you -obj bitten -has
       'I heard the next door's dog has bitten you.'
    b. *ami ʃune-cʰilam [[paʃ-er baɽi-r kukur-ʈa ~~toma-ke~~ kamre- cʰe] toma-ke].*

(16) **Pre-verbal complement**

    a. *ami [paʃ-er baɽi -r kukur-ʈa toma-ke kamre-cʰe] ʃune -cʰilam.*
       I   next-poss house-poss dog -cl you -obj bitten-has heard-have
       'I heard the next door's dog has bitten you.'
    b. **ami [[paʃ-er baɽi-r kukur-ʈa ~~toma-ke~~ kamre-cʰe] toma-ke]] ʃune-cʰilam.*
    c. *[[paʃ-er baɽi-r kukur-ʈa ~~toma-ke~~ kamre-cʰe] toma-ke]]₁ ami t₁ ʃune-cʰilam.*

How should (16b) be ruled out in Bangla? Coherence cannot be responsible for the following reasons: The verb of the dependent clause is finite and would have to move out of a finite clause in order to trigger clause union by virtue of verb raising. But even independently of finiteness, Bangla shows no signs of verb movement (head movement) or rearrangement of verb order as in West-Germanic. Although the strictly head-final nature of the language creates strings of adjacent verbs, there are no independent signs of a verbal complex. For example, none of the coherence tests that play a role in the syntax of German would yield a result in favor of coherence. There is also no indication that the two finite verbs in (16a), *kamrecʰe* and *ʃunecʰilam* would form a prosodic constituent. In fact, there appears to be sustained a high tone on $[cʰe]_\sigma$ of the first and a low tone on $[ne]_\sigma$ of the second.

If the similarity between Bangla and German with respect to clausal licensing is not accidental, we may conclude from this that the restriction seen in (14b) of section 3.2 is not the result of forced clause union but rather the result of a more general constraint that monitors the shape of a clausal complement in immediately pre-verbal position.[9] This leads us to a question about the nature of the pre- and post-verbal complements which appear in Bangla finite complementation as well as in German infinitival complementation. Notice that the complements considered so far offer little to distinguish them. It looks as if the same type of complement can be placed either to the right or to the left of the matrix verb. From a theoretical point of view, this would be a disturbing situation because, given the same numeration, it would be unclear why the complement can in one case stay in post-verbal position whereas it has to raise into pre-verbal position on other occasions. The issue can be addressed empirically because next to the bare finite clauses shown in (15) and (16), Bangla also has clauses with overt complementizers. Post-verbal clauses show the clause-initial complementizer *je* which is homophonous with the relative clause particle *je* (/ʤe/ 'which'). *Je*-clauses can be linked to a pronominal or NP-like correlate in the matrix clause as seen in (17a). They are strictly impossible in pre-verbal position as shown in (17b) and (17c).[10]

(17) *je*-clauses

    a. *cʰele-ʈa (e kɔtʰa) ʃune -cʰe* [*ʤe or baba aʃ -be*].
       boy -cl this news heard-has that his father come-will
       'The boy heard (it) that his father will come'.
    b. **cʰele-ʈa* [*ʤe or baba aʃ -be*] *ʃune -cʰe.*
       boy -cl that his father come-will heard-has
    c. **[ʤe or baba aʃ -be] cʰele-ʈa ʃune -cʰe*
       that his father come-will boy -cl heard -has

<div align="right">Bayer (1996: 255)</div>

Pre-verbal complements do exist, but they never have an initial complementizer. They rather show the post-verbal complementizer-like element *bole*; *bole* is a verbal form which consists of the stem *bol* ('to say') and the perfective participle morpheme *-e*. *Bole* is a quotative suffix as found in the Dravidian languages but also elsewhere in genetically and geographically unrelated languages. (18a) shows a *bole*-clause in immediately left-peripheral position of the matrix verb. (18b) shows that such a clause can move to the left, but cannot be related to a correlate. The question marks in (18c) should indicate that a *bole*-clause in post-verbal position is not fully normal but may occur.

(18) *bole*-clauses

    a. *cʰele-ʈa* [[*or baba aʃ -be*] *bɔl-e*] *ʃune -cʰe*
       boy -cl his father come-fut say-prt heard-has
       'The boy heard that his father will come'
    b. [[*or baba aʃ -be*] *bɔl-e*] *cʰele-ʈa* (**e kɔtʰa*) *ʃune -cʰe*
       his father come-fut say-prt boy -cl this news heard-has
       '[That his father will come] the boy has heard'

c. ??*cʰele-ʈa ʃune -cʰe* [[*or baba aʃ   -be*] *bɔl-e*]
    boy   -cl heard-has   his father come-will say-prt

 Bayer (1996: 255)

Abstracting away from the somewhat unclear status of (18c), these clauses are in complementary distribution as summarized in Table 3.[11]

Table 3: Distribution of *je-* and *bole*-clauses

	in pre-verbal position	*in post-verbal position*
[… V *bole*]	✓	*/??/marked
[*je …* V]	*	✓

Although *je* and *bole* seem to both function as complementizers or subordinating morphemes, they have very different properties which go beyond lexical and etymological differences and also beyond the positional differences summarized in Table 3. *Je*-clauses allow extraposition in the same way as the C-less example in (15b), whereas *bole*-clauses insist on strict adjacency between the finite verb, the complementizer *bole* and the matrix verb according to which material from … cannot intervene in the positions indicated by ▲ in the configuration …V_{fin}] *▲ *bole*] *▲ V_{matrix}].

The distribution of complements seen in Bangla conforms to a widely attested pattern of word order typology.[12] Assuming that complements without an overt complementizer do not deviate from this pattern, Bayer (1996: 266f.) suggested that post-verbal complements may have a clause-initial zero complementizer, while pre-verbal complements may have a clause-final zero complementizer or simply the finite verb functioning as the complement's head in place of *bole*. The adjacency effect observed in (16b) but not in (15b) can now be tied to the formal nature of the complement: Post-verbal complements are CPs which are licensed by an initial complementizer. The complement of C must be a finite clause, but there is no constraint on its right edge. Pre-verbal complements have no initial complementizer. They rely either on the verbal element *bole* or on the finite verb itself. Both appear invariably in final position. Material that moves to the right of this position seems to obliterate the complement's formal features and turn it into an unusable structure.

As a result of this digression to Bangla we can maintain that the adjacency effect cannot universally be reduced to restructuring by head-to-head movement (clause union). First of all, Bangla does not show the typical signs of restructuring; secondly, even if it would, the two relevant verbs in (16b) – *kamrecʰe* and *ʃunecʰilam* – could not engage in such a process because they are both finite and clearly serve as heads of distinct syntactic domains. If we are allowed to use this result as a suggestion for the analysis of German, the no-intervention effect observed in non-coherent constructions such as (14b) should not prematurely be reduced to the parser's preference for coherent constructions and resulting ignorance of non-coherent construal. With this result in mind, we will next explore the structural nature of pre- and post-verbal infinitival complements in German.

 19

3.4 Status checking

Bech (1955/57) has suggested a formal theory of German infinitive constructions in which he draws a parallel between verb-form licensing and Case licensing. Infinitival complements are governed in analogy to nominal complements. Bech calls the relevant verb form *status*. Table 4 shows the three Statuses for the verb *lieben* ('to love') that Bech distinguished for the verb-governed form, in his system the 'supine':

Table 4: Status (Bech, 1955/57: ch.1)

		s u p i n e
1. status	*-en*	lieb-en
2. status	*zu -en*	zu lieb-en
3. status	*ge- -t*	ge-lieb-t

A verb which merges with an infinitive must check its Status just like a verb that merges with an NP must check its Case. Modals, the future auxiliary *werden* and certain light verbs check 1st Status (*-en*), control and raising verbs check 2nd Status (*zu ...-en*), the auxiliaries *haben, sein* and the passive auxiliary *werden* check 3rd Status (*ge-...-t*). Since verbs in 1st and 3rd Status are always coherent, we do not worry about them. The interesting cases are those which appear in 2nd Status. Here we find both coherent and non-coherent constructions. Thus, although Bech considered the string … $V_2 V_1$] as generally indicative of a coherent construction, a bi-clausal analysis cannot be excluded. The fact that Bangla shows precisely the same linear constellation and the same adjacency constraint but lacks clause union should be seen as support for the availability of a bi-clausal analysis. But then the important question arises how the Status of the complement clause can be checked. According to a widespread view, *zu*-infinitival clauses are CPs with a zero complementizer.[13] While this makes sense for post-verbal infinitives and conforms to the typological pattern mentioned in section 3.3, it is less clear that it does for pre-verbal (non-coherent) infinitives. In German one cannot find overt complementizers on argument CPs which would provide the right kind of evidence. However, Dutch has an optional complementizer *om* after the predicates *proberen* ('to try') and *weigeren* ('to refuse"). And there, an *om*-complement is only possible in post-verbal (or perhaps in scrambled) position, but not in immediately pre-verbal position:[14]
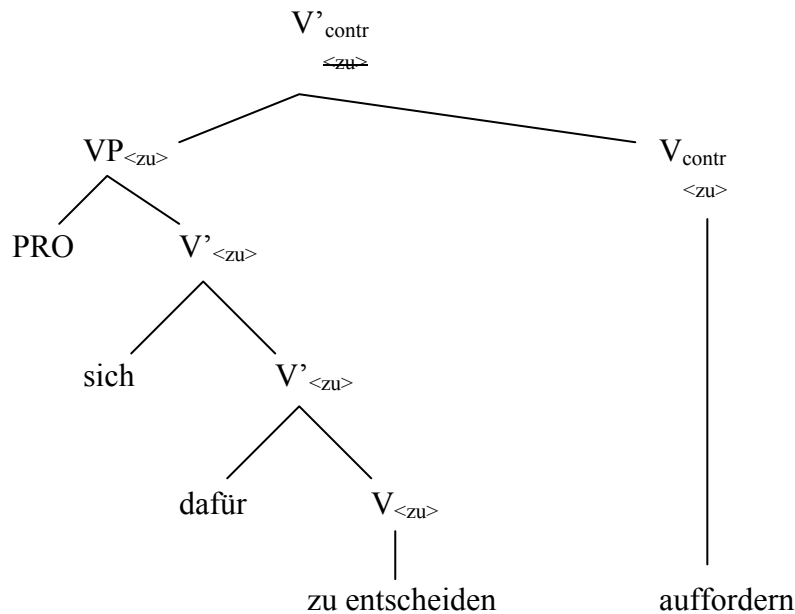
(19) a. *Jan heeft geprobeerd* [*om weg te lopen*].
     Jan has tried   C   away to run
     'Jan tried to run away.'
  b. **Jan heeft* [*om weg te lopen*] *geprobeerd.*

Let us therefore consider the idea that in pre-verbal *zu*-infinitives with clausal status there is no complementizer. What would be the head of such a clause? There seems to be only one reasonable possibility: *zu*. Notice, however, that there have been convincing arguments against *zu* as a VP-final functional head that triggers rightward movement of the verb (cf. Haider (1993) and Vikner (2001)). Following essentially Haider's theory, we assume here that functional heads have distinct properties as a function of their positioning: Initial functional heads such as *v*, T, C etc. attract the verb and normally project a specifier with an EPP-feature. It is far from obvious that final functional heads behave alike. If they are treated on a par with initial functional heads, one has to accept the full consequences of a Kayne-style
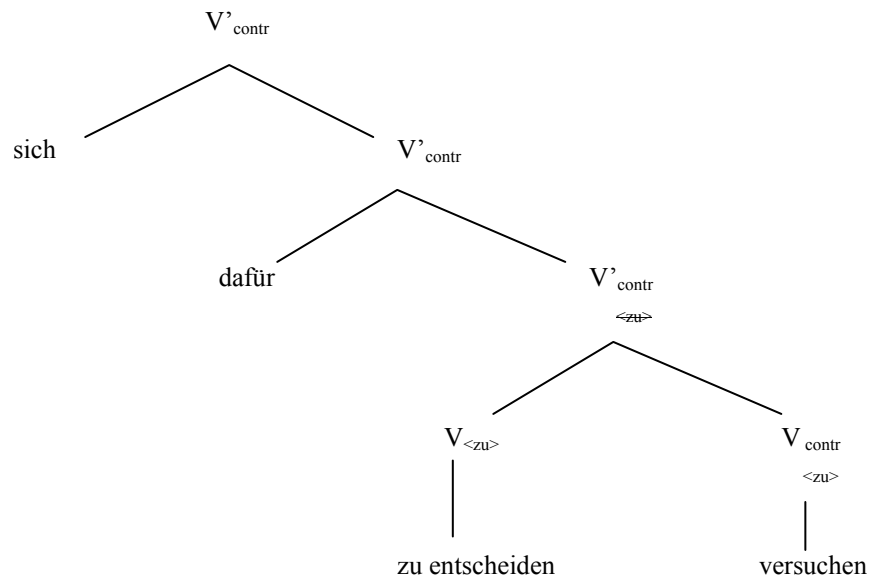
analysis. Alternatively, they could enter the derivation by virtue of morphological composition alone. In that case all there is is morphological merger between a functional element F and a lexical element X. When X projects in syntax, F will be carried along as a feature on X.[15] F is not responsible for displacement; it does not give rise to a functionally motivated specifier etc. It is available for checking anywhere in the domain of X.[16] Let us adopt such a system for the case at hand and assume that *zu* is a prefix which endows the verb with 2nd Status, and that the verb thus enriched projects to maximality. The Status of the clause will be straightforwardly determined by the head which is lexically V and functionally *zu*, i.e. <2nd status>.[17] Instead of attributing the determination of Status to an empty C-head we can now say that Status is signaled directly at the right edge of the complement. Thus, the right edge of a clausal *zu*-infinitive may be all there is of functional information. Assume now that extraposed or stranded material somehow obliterates the right edge of the clause.[18] If the right edge is the only formal signature of the clause this will lead to a licensing problem: the Status feature of the clause ceases to be visible for the selecting verb. This enables us to explain why pre-verbal clauses must be head-final in the strictest sense. Scrambled and extraposed/post-verbal clauses lack this constraint. In German they could be headed by a zero complementizer which is in an agreement relation with the feature <2nd status> provided by the final *zu*+V. There is much evidence that displaced material (even if it undergoes reconstruction) does not need to echo the exact formal structure of its in-situ counterpart.[19] The distribution of finite *daß*-CPs which has been considered in section 3.4 shows exactly this. Thus, although *daß*-clauses are generally ungrammatical in immediately pre-verbal position they are grammatical in scrambled or post-verbal position.

If we are on the right track, there are (i) special licensing requirements for complements which appear in left-adjacent position of the selecting verb, and (ii) these requirements hold independently of clause union, i.e. *zu* can be simply a prefix to a verb (in which case we have a mono-clausal construction) or prefix to a verb that heads a clause (in which case we have a bi-clausal construction). (iii) The adjacency requirement indicates that an agreement relation is at stake which is checked at PF, i.e. essentially a relation of concord as Juan Uriagereka (p.c.) suggested. This is compatible with a movement analysis in the sense of Kayne (1994) according to which the *zu*-infinitive raises to a checking position associated with the matrix verb. It is, however, equally compatible with a base-generation analysis according to which the *zu*-infinitive is directly merged into the domain in which the Status feature can be checked.[20] The bi-clausal and the mono-clausal structures are shown in (20) and (21) respectively. The Status feature <zu> should indicate that the control verb (V_{contr}) searches for a complement that checks off this feature:

(20) (=14a) **bi-clausal / non-coherent**

V'_{contr}
$<zu>$ (struck through)

$VP_{<zu>}$ V_{contr}
           $<zu>$

PRO $V'_{<zu>}$

sich $V'_{<zu>}$

dafür $V_{<zu>}$

zu entscheiden   auffordern

(21) (=13a) **mono-clausal / coherent**

V'_{contr}

sich V'_{contr}

dafür V'_{contr}
       $<zu>$ (struck through)

$V_{<zu>}$ V_{contr}
          $<zu>$

zu entscheiden   versuchen

(20) roughly corresponds to a CP-analysis, and (21) corresponds to a VP-analysis with remnant movement in Wurmbrand (2003). While we will abstain from both a CP-analysis and remnant movement, the basic distinction remains the same.[21] (21) presupposes the formation of verbal complexes without running through phrasal complementation and subsequent pruning operations. If computational steps in the grammar map onto steps of the parser, it would be wrong to assume that the unionized structure in (21) is derived from the bi-clausal structure in (20). The corpus, judgment and behavioral data we have seen in section 2 convey consistently that the mono-clausal analysis is simpler and more readily available. It does not fol-

low from this, however, that the less available parse in terms of a clausal analysis would be prevented by the competence grammar.

To sum up, three things have emerged from the preceding considerations: (i) Pre-verbal clausal infinitives are in all likelihood complementizer-less. (ii) They are licensed via a checking process which relies on their Status feature, in our case <zu>, i.e. Bech's 2nd status. (iii) The Status feature must be visible to the head with which the infinitive is merged; extraposed material destroys the required visibility. In the next section, we will try to make this last point more precise.

3.5 Status Checking and Agreement

In the previous section it was assumed that material which has been extraposed from the infinitival clause or has otherwise been stranded in post-verbal position obliterates the Status feature and therefore prevents feature checking. The situation is familiar from the syntax of pre-nominal modification and has been accounted for in earlier work with a filter called the *Head Final Filter* (HFF). The formulation of the HFF in (22) follows Escribano (2004: 5)

(22) HFF: Base-generated pre-modifiers must be head-final

In Dutch and German, the pre-nominal inflected adjective must not be separated from the nominal by any non-agreeing interveners. The Dutch examples in (22) are from van Riemsdijk (1994). Witness that null-inflection is only possible in indefinite neuter singulars. Others show the schwa-ending -*e*.

(23) a. *een snel genoeg vliegtuig*   (neuter)
           a   fast enough plane
       b. **een snel genoeg auto*   (non-neuter)
             a   fast   enough car
       c. **een snell-**e** genoeg auto*
       d. *%een snel genoeg-**e** auto*

(23d) shows a frequently chosen repair strategy by which the left-adjacent inflection is retained at the cost of inflecting the wrong constituent, a deviance which is indicated by %. *Genoeg* is an adverb and would normally not be inflected. If it follows the actual adjective (*snel*) it seems to be treated like an adjective. German retains more morphology than Dutch. The following case of a masculine indefinite shows the same effects:

(24) a. *ein genug groß-**er** Teller*
           a   enough big -agr plate
       b. **ein groß-**er** genug Teller*
       c. %ein groß genug-**er** Teller

Although what van Riemsdijk refers to as a 'frequent mistake' violates a norm, the examples (23d) and (24c) are clearly distinct from the seriously ungrammatical examples in which the inflected adjective is separated from the agreeing noun by some non-agreeing intervener. The repair seen in (23d) and (24c) is an interesting proof of the obligatoriness of the adjacency relation between A+agr and N. In English, the adjectival inflection has been lost. As a result, N

may be pre-modified by non-adjectival categories, and A and N may in principle be separated by non-adjectival interveners.

(25) a. *the **only** book* (*that John has ever read*)
     b. *a big **enough** plate*
     c. *a **tongue-in-cheek** remark*[22]
     d. *a wrong **in my opinion** view*[23]

Agreement has to be checked in (23) and (24) by virtue of the categories D, A and N of which D and A show overt agreement morphology in modern German. We assume that uninterpretable (and interpretable but redundant) features have to be removed in core syntax, and that this can be done by the agree relation suggested in Chomsky (2001). In (24a) the AP agrees with N by virtue of the head A which bears the features agreeing with N, and D agrees with [AP N] by virtue of the features it encodes. Assume now that due to a PF-operation the head A is affected. How this can happen is not totally clear. Let us for concreteness follow the proposal of Truckenbrodt (1995) and assume that an extraposable element (in German usually an adverb or a PP) attaches to the right of its minimal containing XP for a prosodic reason, e.g.[24]

(26) a. (*Die Leute*)$_\phi$ (*waren drauf   stolz*)$_\phi$ $\Rightarrow$ extraposition $\Rightarrow$
        the   people   were   there-of proud
     b. (*Die Leute*)$_\phi$ (*waren* ~~*drauf*~~ (*stolz*)$_\omega$ (*drauf*)$_\omega$)$_\phi$

A similar process applying to (24a) would lead to the ill-formed output in (24b). The PF-relevant agreement morphology of A (*-er*) would be inaccessible because [*…großer*] has now ceased to exist in favor of [*… großer genug*].

At this point we can return to the adjacency problem of intraposed infinitival clauses. Considering the most challenging case, namely the transition from (14a) to (14b), one can take extraposition as a prosodic realignment of the phonological word dafür. The relevant transition is as in (27).

(27) a. *... dafür* (*zu entscheiden*)$_\phi$ $\Rightarrow$ extraposition $\Rightarrow$
     b. *...* ~~*dafür*~~ (*zu entscheiden*)$_\omega$ (*dafür*)$_\omega$)$_\phi$

Although the verb form marked with the prefix zu remains the head of the infinitival clause, this head has now become inaccessible for PF. The string (zu entscheiden dafür) is not a proper exponent of the formal features of the PF-representation that is required to check off the feature <zu> of the verb auffordern. The uninterpretable feature <zu> will stay in the derivation and lead to a crash at the LF-interface.

Why is it that extrapositions of this sort do not harm displaced (i.e. extraposed or scrambled) clauses? Our conjecture was that extraposed infinitives are headed by a null complementizer. This complementizer must be in a selection relation with the infinitival V-projection, but selection being distinct from agreement, we do not expect adjacency/head-final effects. Thus, one might want to correlate the absence of adjacency/head-final effects with selection of a CP. Involvement of a CP-shell would also be predicted for infinitives in scrambled or topicalized positions. While this might be appropriate for German, we have to acknowledge that it is cer-

tainly not appropriate for Bangla. Recall that in this language C-initial complements are strictly confined to the post-verbal position. Pre-verbal complements are either bare clauses or head-final bole-clauses. So at least for Bangla one may ask why scrambled infinitives should be exempted from PF-checking? One hypothesis could be that the infinitival complement is merged in or raised into the checking domain of the matrix verb, at which point it would be vulnerable by the process referred to in (26). Under this hypothesis the question is how these restrictions could be suspended after scrambling. One solution to this problem provides additional support for the PF-nature of rightward movement that has been argued for above. This solution says that PF-based processes of rightward movement are confined to the concrete occurrence in the flow of speech and are not relevant for the computation of filler-gap dependencies. Thus, the fact that some prosodic realignment in constituent X has taken place at a certain point in the derivation and has been spelled out at this point, cannot be held responsible for the representation that counts for the interpretation of X in its underlying (trace) position.[25]

We have now arrived at a sub-theory of infinitival licensing which can account for the non-intervention effect without abandoning the possibility that intraposed infinitives may retain their clausal Status. The trees in (20) and (21) suggest that Status-checking is independent of clause union. Infinitives which are merged at the immediate left side of the verb (or have been raised into this position) and are spelled out there are subject to PF-checking at this place. As we could show, this holds for coherent infinitives (cf. 21) in the same way as for non-coherent ones (cf. 20). If we are right, the licensing of pre-verbal finite complements in Bangla should follow the same mechanism. Extraposed material which is prosodically aligned at the right edge of the Status checker will invariably destroy the constituent's Status checking ability. We assume that phonological phrasing respects the clause boundary of non-coherent constructions. Thus, extraposition cannot skip over a potential prosodic attachment site.[26]

To sum up at this point, we have made a proposal as to why extraposition from an infinitival clause may render it useless as a Status-checker, and why such a process would only hold for clauses in immediately pre-verbal position.

4 Conclusions

We have started the present investigation with a question about intraposed (pre-verbal) infinitives in German: Does the grammar license infinitives with CP-status (cf. (2a)) in this position simultaneously with the mono-clausal construction by which matrix verb and dependent infinitive form a V-cluster (cf. (2b))?

(12) a. $[\dots [_{CP} \dots V_2] V_1]$ bi-clausal
      b. $[\dots V_2 V_1]$ mono-clausal

This question was approached from two sides,

1. an empirical investigation which involved a corpus study, a questionnaire study and a processing experiment;
2. a syntactic investigation that concentrated on the positioning of infinitival complements and their formal properties. This part involved a comparative part in which we considered complementation in a more distant non-Germanic V-final language.

The empirical studies referred to in 1. have consistently shown that post-verbal complements are the preferred option, and that sentences with intraposed infinitives are preferentially analyzed as coherent, mono-clausal constructions. Why should this be so? If UG allows clausal complementations as in (2a), why should they be dispreferred by the parser? The answer follows from properties of the HSPM. The study of head-final languages and constructions has revealed that the HSPM builds structure also in the absence of the verbal head, and that it does so under minimal assumptions about the input (cf. (4) and (5)).[27] Considering an incomplete parse of (1) (*daß Max mir nur das Lexikon zu kaufen empfohlen hat*) as in (28),

(28) *daß Max mir nur das Lexikon ...*
       that Max me only the lexicon ...

the HSPM has assigned a structure according to which *Max* is [DP-nom], *mir* is [DP-dat], and *nur das Lexikon* is [DP-acc]. Under the most parsimonious analysis, these three constituents are part of the same clause whose lexical head is a di-transitive verb.[28] There is no function word which would signal to the HSPM a new clause boundary before the on-going clausal analysis is completed. Given that the grammar allows coherent constructions, there is no reason for the parser to first build a CP which is pruned down to a coherent construction once the matrix verb has been received and has been identified as a coherence-inducing verb. The expectation is rather that there is only one parse, and that this parse is consistent with a mono-clausal, coherent construction. Since back-tracking and reanalysis are highly unfavorable options, it can be explained why the listener or reader would try to retain a coherent construction as long as possible. The fact that such an analysis tends to be retained even in the face of verbs which resist a coherent analysis is consistent with this explanation.

One conclusion could then be that intraposed infinitives as in (2a) are not available for purely syntactic reasons, and that this is the reason why they would be forced into a mono-clausal analysis during parsing. Our syntactic investigation referred to under ii. has shown that this conclusion would be premature. To demonstrate this we concentrated on an adjacency (non-intervention) effect that has been observed in German intraposed infinitives, and can be described in the style of a filter as in (29).

(29) *$[... [[...t_1 ...V] XP_1] V]$

The existence of such a constraint could mean that (2b) is in fact the only structural option available. XP would prevent this by blocking restructuring. We could show, however, that (29) has a wider coverage than expected in such a scenario. Bangla, the language we used for comparison, has intraposed finite complements, i.e. complements which cannot be affected by clause union for principled reasons. Nevertheless, (29) holds in Bangla as well. Since (29) does not hold for clauses which are arguably CPs, this result led to the conclusion that (2a) is available but that the complement may be inappropriately characterized by the label CP. Evidence was adduced that intraposed complements are formally typed via their right edge, that this typing has a function in checking features of the selecting verb, and that this typing is affected by an intervener as in (29). In this case it would be wrong to ascribe the typing of the intraposed infinitives to a C-head. In fact, typological research revealed a strong correlation of post-verbal clauses with initial C and pre-verbal (intraposed) clauses with final C.[29] German

does not have final C, but the fact that intraposed clauses and sub-clausal units insist on the peripherality of the verb (here *zu*+V) can be seen as a reflex of this typological generalization.

Various questions remain which we can not yet answer in a satisfactory way. Why would a language afford the luxury of intraposed clausal complements, which are known to cause processing problems, if extraposition or base-generation in post-verbal position is a much preferred option? How do languages with intraposed finite complements resolve the tension between structure and processing? Notice that these languages have a final C like Bangla *bole*, if they use a special functional element at all. Given that new clauses are normally signaled where they begin, the clause boundary would in these languages be signaled 'far too late'.[30] Under the assumption that the HSPM is universal, the same dispreference for intraposed clauses should be expected as the one found in German. The examples from Bangla which have been considered here involved finite complements with or without overt signs of subordination. Here the overtly expressed subject of the dependent clause could be an advantage for the parser. Thus, even if no functionally marked left clause boundary is present, the appearance of a (nominative) subject could be used as a clue for inserting a clause boundary. This clue cannot be considered very reliable though. In Bangla, subject and direct object NPs can be morphologically distinguished only if the object NP is animate and then suffixed with -*ke*. But even then Case disambiguation can be temporally postponed by pre-nominal material like stacked possessive NPs which are frequently used, and which do not give a clue to the ultimately correct analysis. Bangla has, in addition, null subjects as well as non-nominative (genitive) subjects. In combination with the fact that there is no V-clustering or any other form of coherence, these possibilities could turn the processing of pre-verbal clauses into a risky enterprise. We would therefore predict that intraposed clausal complements are approximately as dispreferred in Bangla as they are in German. We do not have quantitative data, but intuitive judgments by native speakers point in the expected direction. Since clause union is not an option, intraposed bare clauses and *bole*-clauses are for many speakers best when they are topicalized, i.e. moved over the matrix subject, and prosodically disconnected.[31] This is in full agreement with Hawkins' (1990: 248ff) generalizations about the placement of head-final CPs in SOV-languages.

Even if various questions cannot be answered conclusively, we hope to have shown here that quantitative and behavioral data about syntactic structures and processes as well as cross-linguistic comparison can shed new light on issues of Germanic syntax that have been studied quite extensively over the past 30 years, and that theoretical and empirical work can fertilize each other.

Appendix : **Re-positioned** *zu*

(23d) and (24c) in the text above show a repair strategy by which it is guaranteed that the pre-nominal AP's inflection appears at the right edge and thus adjacent to the following noun. Given the adjacency requirement that has been observed in intraposed *zu*-infinitives, a similar repair strategy could be expected. There is indeed the phenomenon of 're-positioned (or misplaced) *zu*". Bech (1963) was one of the first to provide relevant data and give an account of this phenomenon. The head-final verb order in German leads to the expectation that the placement of *zu* in a cluster of three verbs is as in (ia). However, German has a rule by which

in a cluster V+V+*haben* the auxiliary *haben* is preposed, the so-called '*haben* inversion", as seen in *als er ihn **hat** kommen sehen* instead of *als er ihn kommen (ge)sehen **hat*** ('when he saw him coming").[32] If (ia) is the input to this rule we expect (ib), but (ib) is entirely unacceptable. What happens instead is a re-positioning of *zu* as seen in (ic).

(i) a. *Ich glaubte es tun können **zu** haben.*   ⇒ *haben inversion* ⇒
         I   thought it do can   to have
         'I thought I would be able to do it.'
     b. **Ich glaubte es **zu** haben tun können.*   ⇒ *zu re-positioning* ⇒
     c. *Ich glaubte es haben tun **zu** können.*

Bech (1963: 291f.)

This re-positioning always targets the last verb.[33] In the context of our considerations, this phenomenon can be seen as reflecting the strength of the requirement to mark the right edge of the infinitive construction with the relevant Status feature, similarly to the requirement of marking the right edge of an AP with the relevant features for gender, number and Case agreement.

The infinitival complement with re-positioned *zu* in (ic) appears in post-verbal position. Under the assumption that it is headed by a zero complementizer (or is at least not in a PF-relevant checking relation with the matrix verb) its right edge may not be an issue. However given that in German *zu*-infinitives have the option of appearing as complementizer-less clauses in immediately pre-verbal position, the right-peripheral visibility of their Status *is* an issue. A comparison with Dutch may be interesting at this point. With the exception of scrambled clauses headed by *om*, the grammar of Dutch precludes *te*-infinitives from pre-verbal position. Within our account, this could be related to the fact that *te* does not undergo re-positioning. Consider the following run-of-the-mill examples from the press.

(ii) a. *Aan het wetsvoorstel wordt een nieuwe regeling   toegevoegd*
          to   the law-proposal was   a   new   regulation added
          *om   sneller **te** kunnen bouwen.*
          in-order faster to can   build
          'A new regulation was added to the bill in order to speed up
          permissions for construction.'

      b. *Het Amerikaanse leger zegt in het zuiden van de opstandige stad*
          the American   army says in the south   of the rebellious   town
          *Fallujah een huis   **te** hebben ontdekt dat ...*
          Fallujah a   house to have   found   which
          'The American army claims to have found a house in the south
          of the rebellious town of Fallujah which ...'

      c. *pil zonder   naar de dokter **te** moeten gaan*
          pill without to   the doctor to must   go
          'contraceptive for which one does not need to see a doctor'

Acknowledgements

References

Bader, M. 2000. "On reanalysis: Evidence from German". In *German Sentence Processing*, B. Hemforth and L. Konieczny (eds.), 187-246. Dordrecht: Kluwer.

Bader, M., and Lasser, I. 1994. "German verb-final clauses and sentence processing: Evidence for immediate attachment". In *Perspectives on Sentence Processing*, C. Clifton, Jr., L. Frazier and K. Rayner (eds.), 225-242. Hillsdale, NJ: Lawrence Erlbaum.

Bayer, J. 1995. "On the origin of sentential arguments in German and Bengali". In *Studies in Comparative Germanic Syntax*, H. Haider, S. Olsen and S. Vikner (eds.), 47-75. Dordrecht: Kluwer.

Bayer, J. 1996. *Directionality and Logical Form: On the Scope of Focusing Particles and wh-in-situ*. Dordrecht: Kluwer.

Bayer, J. 1999. "Final complementizers in hybrid languages". *Journal of Linguistics* 35: 233-271.

Bayer, J. 2001. "Two grammars in one: sentential complements and complementizers in Bengali and other South Asian languages". In *The Yearbook of South Asian Languages and Linguistics 2001*, P. Bhaskararao and K. V. Subbarao (eds.), 11-36. New Delhi: Sage Publications.

Bayer, J, and Kornfilt, J. 1990. "Restructuring effects in German". In *Parametric Variation in Germanic and Romance. Edinburgh Working Papers in Cognitive Science* 6, E. Engdahl, M. Reape, M. Mellor and R.P. Cooper (eds.), 21-42. University of Edinburgh.

Bayer, J. and Kornfilt, J. 1994. "Against scrambling as an instance of Move-α". In *Studies on Scrambling*, N. Corver and H. van Riemsdijk (eds.), 1-15. Berlin: Mouton de Gruyter.

Bech, G. 1955/57. *Studien über das deutsche verbum infinitum*. (Det Kongeliege Danske Videnskabers Selskab; Dan. Hist. Filol. Medd. Bind 35, no.2 (1955) & Bind 36, no.6 (1957)). New edition 1983. Tübingen: Niemeyer.

Bech, G. 1963. "Grammatische Gesetze im Widerspruch". *Lingua* 12: 291-299.

Chomsky, N. 2001. "Derivation by phase". In *Ken Hale: A Life in Language*, M. Kenstowicz (ed.), 1-52. Cambridge, MA: MIT Press.

Dasgupta, P. 1980. *Questions and Relative and Complement Clauses in a Bangla Grammar*. Ph.D. dissertation, New York University.

Dryer, M. S. 1980. "The positional tendencies of sentential noun phrases in universal grammar". *The Canadian Journal of Linguistics* 25: 123-195.

Escribano, J. L. G. 2004. "Head-final effects and the nature of modification". *Journal of Linguistics* 40: 1-43.

Evers, A. 1975. *The Transformational Cycle in Dutch and German*. Doctoral dissertation, Rijksuniversiteit Utrecht. Distributed by Indiana University Linguistics Club.

Frazier, L. and Keith R. 1982. "Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences". *Cognitive Psychology* 14: 178-210.

Gorrell, P. 1995. *Syntax and Parsing*. Cambridge: Cambridge University Press.

Grewendorf, G. 1988. *Aspekte der deutschen Syntax*. Tübingen: Narr.

Grosu, A. and Thompson, S. 1977. "Constraints on the distribution of NP clauses". *Language* 53: 104-151.

Haider, H. 1993. *Deutsche Syntax - generativ. Vorstudien zu einer projektiven Theorie der Grammatik*. Tübingen: Narr.

Haider, H. 2003. "V-clustering and clause union. Causes and effects". In *Verb Constructions in German and Dutch*, P. Seuren and G. Kempen (eds.), 91-126. Amsterdam: John Benjamins.

Hawkins, J. A. 1990. "A parsing theory of word order universals". *Linguistic Inquiry* 21: 223-261.

Hoekstra, T. 1984. *Transitivity. Grammatical Relations in Government-Binding Theory*. Dordrecht: Foris.

Inoue, A. and Fodor, J. D. 1995. "Information-Paced Parsing of Japanese". In *Japanese Sentence Processing,* R. Mazuka and N. Nagai (eds.), 9-63. Hillsdale, NJ: Lawrence Erlbaum.

Kayne, R. S. 1994. *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press.

Kimball, J. 1973. "Seven principles of surface structure parsing in natural language". *Cognition* 2: 15-47.

Koster, J. 1986. *Domains and Dynasties*. Dordrecht: Foris.

Mitchell, D.C. 1994. "Sentence parsing". In *Handbook of Psycholinguistics*, M. A. Gernsbacher (ed.), 375-409. San Diego: Academic Press.

Reuland, E. and Kosmeijer, W. 1988. "Projecting inflected verbs". *Groninger Arbeiten zur Germanistischen Linguistik* 29: 88-113.

van Riemsdijk, H. 1988. "The representation of syntactic categories". *Proceedings of the Conference on the Basque Language*. Basque World Congress. vol 1. 104-116.

van Riemsdijk, Henk. 1989. "Movement and regeneration". In *Dialect Variation and the Theory of Grammar*, P. Benincá (ed.), 105-136. Dordrecht: Foris.

van Riemsdijk, H. 1994. Adjunction and adjaceny. ms. Tilburg University.

Ross, J. R. 1967. *Constraints on Variables in Syntax*. Ph.D. dissertation. MIT.

Schmid, T., Bader, M. and Bayer, J. forthcoming. "Coherence - an experimental approach". ms. University of Konstanz.

Singh, U. N. 1980. "*Bole* an unresolved problem in Bengali syntax". *Indian Linguistics* 41: 188-195.

von Stechow, A. and Sternefeld, W. 1988. *Bausteine syntaktischen Wissens*. Opladen: Westdeutscher Verlag.

Stowell, T. 1981. *The Origins of Phrase Structure*. Ph.D. dissertation. MIT.

Truckenbrodt, H. 1995. "Extraposition from NP and prosodic structure". *NELS* 25: 503-517.

Vasishth, S. 2003. *Working Memory in Sentence Comprehension. Processing Hindi Center Embeddings*. London: Routledge.

Vikner, S. 2001. *Verb Movement in Germanic and Optimality Theory*. Habilitation thesis. University of Tübingen.

Wurmbrand, S. 2001. *Infinitives. Restructuring and Clause Structure*. Berlin: de Gruyter.

¹ We do not have the space here to justify these assumptions in detail. For general reviews of human parsing, cf. Mitchell (1994); for a defense of these assumption for parsing verb-final languages, cf. e.g. Bader and Lasser (1994); Inoue and Fodor (1995).

² One can, of course, not a priori exclude the possibility of a positive conjunct attaching to a negative sentence, but the result is often semantically awkward as seen in *John ate nothing, and Bill didn't either/* (??) *and so did Bill* or paraphrases like (??) *What John did was eat nothing*.

³ A generalization of the CRP followed with Hoekstra's (1984) *Unlike Category Condition* (UCC) and various developments stemming from later work such as van Riemsdijk's (1988) *Unlike Feature Constraint* (UFC).

⁴ For PPs of this kind it is important that they are backgrounded / phonologically light and can integrate into the prosodic phrase of the preceding focused verb, here the phonological word *entSCHIEden*. Without intending to suggest a particular analysis, we will refer to this phenomenon as **ex(traposition)**.

⁵ When we speak of 'adjacency' we refer only to the case in which the verbal complex of the infinitive is linearly adjacent to the matrix verb by virtue of the absence of rightward movement / stranding related to the infinitival complement. We will not speak of the independent case in which the linear adjacency of the verbs is broken by scrambling of the clause or by verb-projection raising.

⁶ One reviewer suggested that there may be prosodic reasons for the non-intervention effect. As shown by Truckenbrodt (1995), PP-extraposition and relative clause extraposition must obey a prosodic constraint which requires that a prosodic constituent π attaches to the immediate right of another prosodic constituent of the same type. If the light PP *dafür* in (13) and (14) is a phonological word ω, it could attach to *zu*+V which is the next ω with the effect of blocking clause union. Truckenbrodt's constraint is observed if the PP can attach at the end as in *Ich habe mich* ~~dafür~~ *zu entscheiden versucht dafür*. Assuming that non-coherence must be expressed prosodically by two phonological phrases (φ), this would perhaps explain why **Ich habe ihn [*sich* ~~dafür~~ *zu entscheiden*] *aufgefordert dafür* ('I asked him to decide on it') or **Ich habe ihn [*sich* ~~dafür~~ *zu entscheiden*] *gezwungen dafür* ('I forced him to decide on it') are out; *dafür* would illegally have skipped a potential attachment site. This would not explain, however, why the non-intervention effect still holds in non-coherent constructions such as (14b).

⁷ The transcription of the examples follows IPA. The glosses should be read as follows: -cl (classifier-style determiner), -poss (possessor/'genitive' Case), -obj (objective/'accusative' Case)

⁸ The original observation of this is due to Probal Dasgupta (p.c.).

⁹ In fact, we are quite sure that the observed similarity between Bangla and German is not accidental. Informal testing with speakers of the SOV-languages Japanese and Turkish, both of which allow light pronoun extraposition, revealed the same adjacency constraints as those observed in Bangla and German. Thanks to Baris Kabak for his help in clarifying this issue.

¹⁰ The same holds for *ki*-clauses in Hindi and all other Indo-Aryan languages we could obtain information about. It also holds for the SOV-languages Persian and Turkish both of which have post-verbal clauses with an initial complementizer *ke* or *ki*, cf. Dryer (1980).

¹¹ For detailed discussion see Bayer (1995; 1996; 1999; 2001); Dasgupta (1980); Singh (1980).

¹² According to Grosu and Thompson (1977), initial complementizers in pre-verbal complements are almost universally unavailable.

¹³ Cf. Wurmbrand (2001) to mention a more recent analysis.

¹⁴ Cf. Koster (1986: 123)

¹⁵ Arguments in favor of co-projection of this kind can be found in Reuland and Kosmeijer (1988), Haider (1993), Bayer and Kornfilt (1990; 1994) and elsewhere.

¹⁶ One immediate advantage of such a system is that Case checking is independent of positions. Given the possible base order non-nominative < nominative of German and other SOV-languages, there is a clear advantage over the standard position-based accounts of Case checking. Bayer (2001) points out other advantages in accounting for negation in South Asian languages.

¹⁷ For a caveat cf. the appendix.

¹⁸ We will try to make this more precise in section 3.5.

¹⁹ Widely known examples are so-called 'quantifier-floating' and other split-NP-constructions. Cf. note 25.

²⁰ In this situation, complement and specifier are collapsed, a technical state of affairs which seems to characterize head-final constructions and their apparent non-distinctness of specifiers and complements quite correctly.

²¹ For arguments against clause union in terms of VP-remnants cf. Haider (2003).

²² From Escribano (2004: 26)

[23] This is an observed example which may be considered ungrammatical. The fact that such constructions are produced by speakers of English shows, however, that the HFF cannot hold in full generality in this language. For a recent account of the HFF in purely phrase structural terms and without any reference to agreement cf. Escribano (2004). One of the problems of Escribano's theory is that it cannot account for Dutch and German. Another problem is that it considers only modifiers but no complements or specifiers. It would thus be impossible to apply it to pre-verbal infinitival complements.

[24] φ and ω are symbols for the phonological phrase and for the phonological word respectively.

[25] For supporting demonstrations of this cf. Truckenbrodt (1995). According to the idea that the trace is a copy of the displaced element it should be clear that this copy is a syntactic and not a phonological entity. NP-split in German provides convincing empirical support of this claim. Thus, although there is reliable evidence for movement (cf. van Riemsdijk, 1989) in (i), the verbatim reconstruction of *ein Auto* leads to a seriously ill-formed construction as seen in (ii). The only permissible form would be the morpho-syntactically different form seen in (iii).

(i)   [**Ein** Auto]$_1$ *kann er sich leider*   [*kein**es** t$_1$*] *leisten*
            a car   can he REF unfortunately none   afford
        'As for a car, he can unfortunately not afford any'

(ii)   *(*daß*) *er sich leider* [*kein**es*** [***ein** Auto*]] *leisten kann*

(iii)   (*daß*) *er sich leider* [*kein Auto*] *leisten kann*

[26] Notice that this amounts to the upward boundedness which Ross (1967) captured with the *Right Roof Constraint*.

[27] For the processing of German, cf. Bader and Lasser (1994) and Bader (2000).

[28] Sentence completion tasks from our experimental work in progress never yield constructions in which the provided constituents would distribute over more than a single clause. As realistic completions of (28) one can, thus, expect di-transitive verbs like *zeigte* ('showed'), *gab* ('gave'), *verkaufte* ('sold') etc. but not *zu kaufen empfohlen hat* ('has recommended to buy') or *zu schenken dem Onkel empfohlen hat* ('has recommended to the uncle to buy') etc.

[29] Cf. Table 3, which describes Bangla. For reasons of space we cannot support this here with data from other languages.

[30] Cf. Kimball (1973).

[31] Hindi is closely related to Bangla but lacks pre-verbal finite complements altogether. Experimental results about the processing of center embedded non-finite clauses in Hindi (cf. Vasishth, 2003) reflect the expected complexities.

[32] This rule also extends to *werden* and sometimes even to modals, but less frequently so.

[33] The purely linear nature of *zu* re-positioning should throw doubts on the representation in (20). In (20) the Status feature <zu> is projected with the verbal head of the clause. Re-positioning of *zu* shows, however, that the appearance of *zu* as a prefix to the verbal head may only be an artifact of non-inversion. We have to leave this important issue for future research.