SOUL: A Cognitive Parser

Lars Konieczny
Center for Cognitive Science
Institute of Computer Science and Social Research
University of Freiburg, Germany
lars@cognition.iig.uni-freiburg.de

Gerhard Strube
Center for Cognitive Science
Institute of Computer Science and Social Research
University of Freiburg, Germany
strube@cognition.iig.uni-freiburg.de

Abstract

In this paper, we introduce a new model of human sentence processing. The psychological issues addressed include the use of lexical information, specifically subcategorization information, during the initial stage of syntactic structure assembly, the issue of *linear* parsing, i.e. immediate attachment of words to the sentence structure, within a *head-driven* grammar framework, and the resolution of attachment ambiguities. We will demonstrate that the variety of psycholinguistic phenomena can be accounted for by the assumption of principled behavior of linguistic signs, which are implemented in an object-oriented fashion. The model provides a serial implementation of *Parametrized Head Attachment*.

Parsers as adequate cognitive models

Cognitive parsers are computational models of human cognitive processes with respect to the syntactic analysis of natural language. To qualify for that, those models first have to be computationally sufficient, which means that they must be able to parse sentences of some natural language (or at least a non-trivial subset of those). The second criterion is that cognitive parsers must be shown to proceed in a way that leads to the same behavior that has been found in studies of human parsing, i.e., the same kind of errors, and the same preferences in analysis. A third, ambitious criterion, labeled '*type transparency*', has been specified by Berwick and Weinberg (1985), who suggest that 'grammatical representations are embedded directly into parsers, without intervening derived predicates or multiplied-out rule systems'. The SOUL (*Semantics-Oriented Unification-based Language*) system has been built with this requirement in mind. If *transparency* holds, then cognitive parsers should not only model linguistic performance, but the linguistic competence of the idealized 'native speaker' as well (Konieczny & Hemforth, 1994). Our approach is therefore based on a principled theory of grammar (namely, the *Head-driven Phrase Structure Grammar*, HPSG, Pollard & Sag, 1994).

The status of lexical information in parsing

In a *'head-driven'* grammar framework, such as HPSG, *transparently* assembling syntactic structure depends directly upon the availability of lexical information. Although there is a broad consensus among psycholinguists that lexical information influences the parsing process very shortly after a word is processed, it is still an open question as to whether this information is used to guide the initial structure-building process (*lexical guidance, lexically directed assembly)*, or whether it is only used to monitor or evaluate the independently built structures (*lexical filter, structure checking models*), and to rule out inconsistent ones. In our approach, we take the former view.

The most prominent proponent of the latter view is the *garden-path model* by Frazier and her colleagues, which still constitutes a substantial part of its successor, the *construal theory* (Frazier & Clifton, in press). The *garden-path model* focuses on structural economy in parsing (*minimal attachment, late closure*). In addition to the 'constituent structure subsystem', Frazier (1987, p. 582) distinguishes another subsystem, the *thematic processor*, which serves the purpose of checking or monitoring the initially built structure based on lexical and higher level information. In terms of processing, this amounts to the structure being built from syntactic category information alone, and a later 'filtering' stage, where structures proposed by the constituent structure subsystem can be rejected by means of additional lexical information, for example, semantic features.

Structure checking models, such as the garden-path model, therefore presume that: 1. there is an independent grammar, such as CFG-rules, outside the lexicon. This grammar is constructive in the sense that syntactic structure can be built from it without further lexical information. 2. only the category information of lexical items is used initially, while subcategorization information is ignored at the first stage. The model presented in this paper depends on neither of these assumptions. It will nevertheless be shown that the data can be accounted for with a transparent one-stage parser built on a state-of-the-art lexicalized grammar.

Psycholinguistic evidence. Although the issue of lexical guidance vs. filter is one of the most extensively investigated in psycholinguistics, a final decision is yet to come. Among the many researchers reporting evidence in favor of the *guidance* assumption are Boland and Tanenhaus (1991) and others. In a series of self-paced reading and eye-movement experiments, reported in (Hemforth et al., 1991; Strube et al., 1990; Konieczny et al., 1994ab), we investigated PP-attachment preferences in German sentences, such as (1) and (2).

(1) Marion *beobachtete* das Pferd {a. mit dem neuen Fernglas, b. mit dem weißen Fleck}.

Marion *watched* the horse {a. with the new binoculars, b. with the white fleck}.

(2) Marion *erblickte* das Pferd {a. mit dem neuen Fernglas, b. mit dem weißen Fleck}.

Marion *caught sight of* the horse {a. with the new binoculars, b. with the white fleck}.

The verb was either biased to "expect" an instrument, as in the case of *beobachten (to watch)*, or not to do so, as in the case of *erblicken (to catch sight of)*. Subjects spent longer reading the PP, if the verb's bias contradicted the plausibility of the attachment, as in (1b) and (2a), compared to (1a) and (2b), where the verb-bias is confirmed by plausibility. These data are thus in line with findings in support of the assumption that the verb-bias influences initial parsing decisions.

On the contrary, Mitchell (1987, 1989), and recently Adams, Clifton, and Mitchell (1991, in press) have reported studies that seem to provide evidence against the lexical guidance hypothesis. Subjects read sentences like (3ab).

(3) Although the audience {a. yawned, b. booed} (,) the comedian continued telling very bad jokes.

Readers spent longer reading the disambiguating second verb (*continued*) in a sentence with a transitive first verb (*booed*, 3b), than in any other condition, including control conditions with a comma. However, they spent reliably longer reading the NP (*the comedian*) following an intransitive verb, such as *yawned* (3a), than in any of the other conditions. Their results indicate that in the absence of a comma readers preferred to interpret the NP as a direct object of the first verb. In case of a transitive first verb, the initial choice was only corrected at the disambiguating second verb. However, the longer reading times on the NP in sentences with an intransitive verb (3a) indicate that the NP has initially been interpreted as a direct object despite the verb´s intransitivity. The results thus seem to support Mitchell´s interpretation that only information at the major category level is used during the initial stage of structure assembly, thus allowing the "minimal" attachment of the NP to the current VP. Only during the second stage can the verb´s subcategory rule out the initial attachment and initiate a reanalysis, resulting in increased processing load at the NP. Hence, one-stage approaches like ours appear to be challenged.

If lexical information were essential for the assembly of structure, and hence for interpretation, head-final structures, which are very common structures in some languages, such as German or Dutch, must obviously be processed differently from head-initial structures. This aspect is often pointed out by proponents of two-stage structure checking models, such as Frazier (1989) and Mitchell (1989). Even if lexical information can only help the parser to build the "best" structure more often, head-final structures, such as German subclauses, will hence be disadvantageous, since lexical information enters the process too late to prevent the parser from building the wrong structure. Frazier (1989) claims that head-final structures had disappeared over time if the initial assembly of structure were dependent upon lexical information. We will argue against this position later.

Immediate vs. delayed attachment

There is common agreement that human parsing proceeds incrementally, essentially in a word-by-word fashion. From a technical point of view, parsing with lexicalized head-driven grammars bears the problem that the assembly of a certain piece of structure without its head does not seem to be possible at all. Lexicalized head-driven grammars seem to imply *head-driven parsing* accounts, which have been proposed in psycholinguistics (Abney, 1989; Pritchett, 1992), as well as in computational linguistics (e.g. Kay, 1989). Contrary to "*linear*" parsing approaches (Konieczny, in prep.), in which 'perceivers incorporate each word of an input into a constituent structure representation of the sentence, roughly as each item is encountered' (Frazier, 1987, p. 561), *head-driven* parsing accounts assume that structure is not built before its head is processed and, as a consequence of this, the attachment of complements to the structure is delayed until it is licensed by the head.

However, there is an increasing amount of evidence against *delayed attachment* (e.g. Inoue & Fodor, 1994; Bader and Lasser, in press). Hemforth, Konieczny and Strube (1993) found evidence for attachment to a phrase marker even before the head of the phrase was read. Overall, there is good reason to believe that structure is built even before the head is processed, and that complements can be attached into the structure before the attachment is licensed by the head. However, there is no need to reject *head-driven grammars* as a competence base for strictly *linear* parsers because of these findings, as will be demonstrated later on in this paper.

The SOUL system

A short note on the competence base

HPSG is a system of linguistic *signs* (in roughly the Saussurean sense). Signs can be either phrases or words, and consist of a set of attributes, such as the phonological form, syntactic and semantic features, etc. Constituent structure is accounted for by an attribute as well, which is only permitted for phrasal signs, however. Whenever signs are combined, parts of their feature structures are *unified*. After successful unification, the feature structure consists of all the information that constituted the formerly separated feature structures.

An important extension is that feature structures are labeled. A label represents a specific *sort* of feature structure, and only certain attributes are permitted by a particular sort. *Sorted feature structures*, then, represent a partial description of a linguistically relevant generalized piece of information, and can be organized in hierarchical systems of linguistic sorts. Thus, sorts can have subsorts, which are *subsumed* by their supersorts, i.e. they inherit all the information of their supersorts but are more specific to a certain extent.

Sorts are a powerful device that distinguish HPSG from many other grammar frameworks. Valence-classes, for example, can be accounted for by a hierarchical system of sorts, which represent the respective feature structure of subcategorization requirements of *intransitive, transitive, strictly-transitive, ditransitive* verbs, etc. These general sorts can later be referenced in the concrete lexical entries.

In our model, sorts play an important role in solving the problem of linear incremental parsing with a head-driven grammar, as described later in this paper.

Apart from this, there are only a few very general constraints on the well-formedness of signs left outside the lexicon, namely general rule-schemas (about five in number) without any category information, which serve as the basic building-blocks of constituent structure, and a few principles of well-formedness, which further constrain the way in which signs can be put together (see Pollard & Sag, 1994, for a detailed introduction to HPSG).

The processing mechanism

The parser is implemented in an object-oriented manner. Linguistic *sign*-tokens, such as words or phrases, are implemented as objects and come equipped with methods for combining themselves with other *sign*-objects[1]. Feature structures, i.e. the internal representations of signs, are implemented as objects with methods suited for *unification, consistency-checks, copying, etc.* The object-class *sign* inherits these methods from the *feature structure* class, and adds a certain amount of parsing "behavior".

Parsing starts out from the lexicon: as soon as a word is encountered, a lexical sign is "activated", i. e. after it is retrieved from the lexicon, an "act"-message is sent to it (*active sign*). When activated, a lexical sign behaves in a very principled way.

The *active sign* and the *waiting sign*. Besides the *active sign*, there is a *waiting sign*, which represents the entire sentence-structure that was parsed before a new word was processed. During processing, the active sign passes through several stages of development, which correspond to a level in the constituent structure. At each stage, the *active sign* attempts to successfully invoke one of the following three methods in the order of appearance:

1. *attach* self to the waiting sign (if there is one). If it is impossible to do so at the current state,

2. *project* self up to the next phrasal level, as a "daughter" in one of the basic HPSG-rule-schemas, but never higher than sentence level. As a part of projection, unify the new phrasal unit with the principles of well-formedness in order to rule out illegal projections. Then start again with 1.

3. If either the attachment (1.) worked, or the projection (2.) reached the sentence level, *predict* the subsequent constituents at each level top-down, in order to provide attachment sites for subsequent items. This is done by collecting the yet unfilled daughters, which stem from the subcategorization frame of the *preferred* lexical forms of each item read so far. The daughters are put onto a stack, with the predicted adjacent sisters (*predicted signs*) on top.

Two types of attachments have to be distinguished: 1. *predicted* attachment, i.e. the unification with one of the *predicted signs*, and 2. the *active* attachment search, which starts at the left adjacent item and proceeds upwards to each immediate dominator of the current constituent, if the attachment is prohibited at the current level. The *active* attachment search incorporates the search for alternative lexical forms of the items read so far, which might provide a matching complement, and the attempt to attach the sign to the current node as an adjunct.

The basic *attach-project-predict* behavior of signs results in an "eager left-corner"-like process. The first item in a sentence is projected upwards until the sentence level is reached. Each following item is integrated immediately, either as a complement, or as an adjunct. In head-final structures, however, complements precede their heads. Thus, complements have to predict their heads in some way.

Strictly *linear* parsing of head-final constructions. As demonstrated in Konieczny and Hemforth (1994), HPSG has some specific devices that can either be used directly, or easily extended to avoid the procedural shortcomings of a head-driven account.

Firstly, some non-heads, such as determiners, modifiers and complementizers, are equipped with information about their heads, which allows the bottom-up projection of the structure beyond the level of the 'maximal projection' of the particular item itself, and hence the (top-down) prediction of the head it specifies, or modifies, respectively. As a modest deviation from transparency, Konieczny and Hemforth (1994), secondly, propose the use of general sentence schemes for those cases, where this kind of information is not available from the lexicon. These schemas can be triggered by anything that can start a new sentence, including non-heads. Importantly, sentence schemas are initially underspecified at the level of verbal complements, i.e. they are simply restricted to the most general sort in the lexical sort-hierarchy for subcategorization frames, namely *verbal-complements*. The restriction

[1] A sign-oriented mechanism is thus inserted before the undirected type-deduction mechanism of unification-based grammars.

to *verbal-complements*, thirdly, serves two purposes during the parsing of head-final structures: 1. it allows only those complements to be attached that are consistent with other complements within a legal subsort of the verbal subcategory, and hence 2. it successively constrains the subcategory of the subsequent verb. Thus, with each attached complement of a verb, the predicted verb-sign is restricted to a more specific subcategory.

The sortal hierarchy is intended to express linguistic generalizations, such as the valence-classes of verbs, which would otherwise have to be encoded within each lexical entry. Sorts can further be used in parsing, as described above. This kind of "type-inference" thus serves a purpose comparable to CFG-rules in a standard parsing paradigm.

Equipped with these devices, parsing with HPSG avoids most of the "shortcomings" of other principled accounts. However, the processing of verb-final and verb-initial structures is carried out differently: whereas heads in the initial position project their subcategorization requirements directly into the structure, the successive attachment of complements in head-final structures results in a sortal inference on the head's subcategory. However, since constraining lexical information is available in head-final constructions to permit the attachment of the complements of subsequent heads, processing these structures is not significantly disadvantageous compared to head-initial structures. Thus, Frazier's and Mitchell's argument does not apply to the approach presented here.

Contrary to our approach, in an account based on an extra-lexical grammar, such as CFG-rules, the parsing of head-final and head-initial structures results in the same kind of process, namely the selection and application of a particular rule. It is important to point this out in the light of the data presented in the next sections.

Results

PP-attachment revisited

We have presented data on PP-attachment sentences such as (1) and (2), which strongly support the lexical guidance hypothesis. In Konieczny et al. (1994ab) we also investigated the processing of similar German subclauses, such as (4), in which the verb was placed at the end of the subclause.

(4) Ich habe gehört, daß Marion das Pferd {a. mit dem neuen Fernglas, b. mit dem weißen Fleck} beobachtete.

I have heard, that Marion the horse {a. with the new binoculars, b. with the white fleck} watched.

"I have heard, that Marion watched the horse {a. with the new binoculars, b. with the white fleck}."

The material was used in two experiments, in one of which we measured word-by-word self-paced reading times, whereas in the other one eye-movements were recorded. Subjects spent longer processing the PP, if it represented an object which could not plausibly be attributed to the preceding noun (*horse with binoculars*), as in (4a). As we have demonstrated in Konieczny et al. (1994ab), these findings strongly contradict the predictions of the *garden path theory* and its current successor *construal theory* (Frazier and Clifton, forthcoming), since the subjects did not seem to choose the structurally more parsimonious attachment to the VP initially, but preferred to attach the PP to the preceding direct object in the initial stage, before world knowledge rendered this analysis implausible.

In an eye-movement study based on very similar materials, Konieczny (in prep.) found an interaction of verb-placement and the referential ambiguity of the direct object NP in sentences like (4). Findings like these were further supported by results from a self-paced reading study and an eye-movement experiment reported in (Scheepers, Hemforth, & Konieczny, 1994) on sentences with NP-attachment ambiguities. Taken together, the data provide overwhelming evidence in support of the *head attachment principle* (Konieczny, Hemforth, & Strube, 1991) as expressed in (5).

(5) *head attachment*

Prefer to attach an item to a phrasal unit whose lexical head has already been read.

The *head attachment principle* applies to all those cases, where an attachment ambiguity can be resolved by either attaching to a preceding head, or to a head yet to come.

In many other cases, such as the verb-second sentences in (1), there are two or more heads that are potential attachment sites for an ambiguous item. As we have seen above, the decision now depends on the *lexical preferences* of either of the heads, as expressed in (6).

(6) *preferred-role attachment*

Prefer to attach an item to a phrasal unit whose head *preferentially* subategorizes for it.

This is to an extent the reincarnation of Ford et al.'s (1982) principle of *lexical preference*, which was built into their LFG-based parsing model. In the current version of our model, the preference to expect either the occurrence, or the non-occurrence of an optional complement is implemented by the assignment of a *strength*-value to distinct subcategorization frames. After the lexicon was accessed, only one frame, and most probably the strongest one, will be used throughout the initial analysis[2].

Of course, there are cases where two (or more) preceding heads do not differ in their preference to bind a constituent. In these cases, a decision is supposed to be based

[2]The use of a strength value does not necessarily imply the assumption of an "exposure-based" aspect in our model. On the contrary, we are currently approaching a theory of lexical preference that tries to relate subcategorization preferences to the argument structure of verbs (Scheepers, in prep.), to the issue of ontological necessity of thematic roles, and to properties of the situation model (Konieczny, in prep).

upon *recency* (7).

(7) *most-recent head attachment*

Prefer to attach an item to the head that was read most recently.

The predictions of *most-recent head attachment* compare to those of *late closure* in the garden-path model. However, it is only applied if the other principles fail to provide a decision. This has been expressed in the unified *parametrized-head attachment principle, PHA* (8), which furthermore serves the purpose of emphasizing the fact that attachment ambiguities are resolved on the basis of certain parameters of lexical heads, such as *relative position* and *lexical preferences*.

(8) *parametrized-head attachment, PHA* (Konieczny, Hemforth, Scheepers, & Strube, 1994)

(Attempt to) apply *head attachment* (5) before *preferred-role attachment* (6) before *most-recent head attachment* (7).

According to PHA, and in particular due to *head attachment*, the parser initially builds a syntactic structure, which can be evaluated semantically as soon as possible (*"semantics-oriented" processing*). Note, however, that semantics is not supposed to *guide* the parser, as in a strongly interactive *"semantics-driven"* account.

The implementation of *PHA*. Fortunately, *PHA* can be accounted for without major extensions to the current model. This is most obvious in the case of *head attachment*, which can be derived directly from the fact that an active sign first attempts to attach itself to the waiting structure before it continues to do something else.

The observed attachment phenomena in verb-final constructions, such as (4), can easily be explained: firstly, let us assume that the waiting sign is the structure built before the PP was processed. The list of predicted signs now contains the predicted verb-fragment, but also a pointer to the list of the complements of the verb. Since the direct object NP has been processed before, the list is restricted to the sort *transitive*. However, due to the lack of the verb in this position, no PP has been predicted yet. Thus, no *predicted* attachment of the PP in (4) can take place. However, the *active* attachment search starting at the left-adjacent NP succeeds in attaching the PP as an adjunct. Only then can the attachment be evaluated with respect to plausibility. If the "plausibility check" fails, the *active* attachment search is continued at the next level, namely the VP. Since no verb was read so far, the search for an alternative lexical form does not make much sense. Instead, the sort-hierarchy is accessed in order to search for a subsort of the proposed verb's subcat-sort *transitive* that permits the attachment of a PP-object. It will finally attach itself as a PP-complement into the list of complements, which is now restricted to a subsort of *ditransitive*.

Hence, the process of *active* attachment is refined to 1. the search of an alternative lexical form, in the case that a lexical item is already available, and 2. *permitted* attachment, which requires a sortal inference in the hierarchy of valence-sorts, otherwise, followed by *adjunct* attachment.

The situation in verb-second structures is different. By the time the PP is about to be processed, the verb has already predicted its preferred complements. If the verb is biased towards the expectation of an instrument, a PP is among the predicted signs and thus succeeds in attaching itself to the VP. If not, the PP has to initiate the active attachment search, starting at the most recent head, the direct object NP, to which it can easily be adjoined. *Preferred role attachment* therefore directly results from the distinction between *predicted* and *active* attachment[3].

If the *active* attachment to the NP fails for reasons of plausibility, the attachment to the VP is pursued.

Finally, *most-recent head attachment* is a direct result of 1. the prediction stack with the most recent predictions always on the top, and 2. the *active* attachment search starting from the left-adjacent item upwards. The recency effect is thus accounted for within a serial search model, as opposed to some current accounts, which provide an "activation" based explanation. E.g. Gibson et al. (in prep.) implement recency as a *cost*-function, which assumes that all potential attachment sites have to be computed in parallel before the cost for the attachment to each of them can be calculated in the second run. We regard the serial-search account as more elegant and parsimonious.

The immediate use of lexical information revisited. We still have to provide an explanation for Mitchell's (1987, 1989) and Adams et al.'s (in prep.) challenging results on sentences like (3). A closer look to their material uncovers that many of their verbs, such as *to talk* and *to yawn*, have rare but permissible transitive forms[4]. Although this does not weaken their argument against lexical guidance models that only pursue the lexically preferred path, e.g. the model of Ford et al. (1982), the studies provide no evidence against the immediate use of lexical information in a model, that gives highest priority to *head attachment* (5): If no complement is preferentially predicted at the point when the NP is to be processed, the *active* attachment search provides another lexical form that permits its attachment as a complement, at least in case of verbs like *pray, talk, yawn, cough, doze, smile,* etc.. However, the unification with the now predicted complement fails very early, e.g. due to a thematic role restriction mismatch. Only then can the active sign continue to project itself further bottom-up and finally produce the correct structure.

The increased reading times at the NP following an "intransitive" verb, compared to the transitive verb condi-

[3]Note that there are no grades of preference during parsing, since *lexical strength* only affects the initial choice of a subcategorization frame.

[4]according to Webster's Ninth New Collegiate Dictionary.

tion, are thus firstly due to the search for another verb-form that permits the attachment and secondly, and more importantly, due to the additional processing load of several steps of bottom-up-projection *after* the attempted attachment / unification failed. Compared to the comma conditions, however, the processing load is increased because with a comma, the search for another verb-form and the unification with the predicted complement is not even attempted. Although *attachment* and unifying in lexical information takes place at one and the same stage, because it is actually one and the same process, Mitchell's results are easily accounted for.

Conclusions

We introduced the SOUL system as a computationally sufficient, implemented[5] model of human language competence and performance. HPSG sign tokens are implemented as objects, whose basic *attach-project-predict* behavior results in attachment preferences predicted by the *Parametrized Head Attachment* (PHA) principle. Since the principles of the grammar are put to use directly, parsing is supposed to be 'transparent'. It could be demonstrated that this behavior fits the available psycholinguistic data well, i.e., it is psychologically adequate.

Acknowledgments

This research was supported by the German National Research Foundation (Deutsche Forschungsgemeinschaft, DFG, Str 301/4-3). We want to thank Barbara Hemforth and Don Mitchell for many fruitful discussions and comments, and Kim O'Brien and an anonymous reviewer for their helpful comments on an earlier draft of this paper.

References

Abney, S. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research*, 18.1, 129-144.

Adams, B. C., Clifton, C., Jr., & Mitchell, D. C. (in prep). Lexical guidance in sentence processing: further support for a filtering account. *Manuscript submitted.*

Berwick, R. C. & Weinberg, A. S. (1985). Deterministic parsing and linguistic explanation. *Language and Cognitive Processes*, 1, 109-134.

Boland, J. E., & Tanenhaus, M. K. (1991). The role of lexical representation in sentence processing. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 331-366). North-Holland: Elsevier.

Ford, M., Bresnan, J. & Kaplan, R.M. (1982). A comptence-based theory of syntactic closure. In J. Bresnan (Ed.) *The mental representation of grammatical relations*. Cambridge, Mass.: MIT Press

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *The psychology of reading* (pp. 559-586). Hove/ London/Hillsdale: Lawrence Erlbaum.

Frazier, L. (1989). Against lexical generation of syntax. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 506-528). Cambridge, MA: MIT Press (Bradford).

Frazier, L., & Clifton, C. (in press). *Construal theory.*

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickock, G. (in prep.). Recency Preference in the Human Sentence Processing Mechanism. *Submitted manuscript.*

Hemforth, B., Konieczny, L., Strube, G., & Wrobel, H. (1990). Kognitive Modellierung und empirische Analyse von Prozessen der Satzverarbeitung. *Technical report.*

Hemforth, B., Konieczny, L., & Strube, G. (1993). Incremental syntax processing and parsing strategies. *Proceedings of the 15th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum (pp. 539-545).

Inoue, A., & Fodor, J. D. (1994). Information-paced parsing of Japanese. In R. Mazuka & N. Nagai (Eds.), *Japanese syntactic processing*. Hillsdale, NJ: Erlbaum.

Kay, M. (1989). Head-driven parsing. *Proceedings of the international workshop on parsing technology.* Carnegie-Mellon University, Pittsburgh, PA.

Konieczny, L. (in prep.). *Sentence processing*. Doctoral dissertation at the University of Freiburg, Germany

Konieczny, L. & Hemforth, B. (1994). Incremental parsing with lexicalized grammars. *Submitted to Linguistics.*

Konieczny, L., Hemforth, B., Scheepers, C., & Strube, G. (1995). PP-attachment in German: results from eye movement studies. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.) *Eye movement research. Mechanisms, processes and applications.* Amsterdam: North Holland, pp. 405-420.

Konieczny, L., Hemforth, B., & Strube, G. (1991). Psychologisch fundierte Prinzipien der Satzverarbeitung jenseits von Minimal Attachment. *Kognitionswissenschaft*, 2.

Konieczny, L., Scheepers, C., Hemforth, B. & Strube, G. (1994). Semantikorientierte Syntaxverarbeitung. In S. Felix, C. Habel, & G. Rickheit (Eds.), *Kognitive Linguistik: Repräsentationen und Prozesse*. Opladen: Westdeutscher Verlag.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading*. London: Lawrence Erlbaum Associates Ltd.

Mitchell, D. C. (1989). Verb-guidance and other lexical effects in parsing. *Language and Cognitive Processes*, 4, SI 123-155.

Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar.* Chicago, London: University of Chicago Press and CSLI Publications.

Pritchett, B. (1992). *Grammatical competence and parsing performance*. Chicago: University of Chicago Press.

Scheepers, C. (in prep.) *Zur Verarbeitung von Argumentrelationen.* Doctoral dissertation at the University of Freiburg.

Scheepers, C., Hemforth, B. & Konieczny, L. (1994). Resolving NP-attachment ambiguities in German verb-final constructions. In B. Hemforth, L. Konieczny, C. Scheepers & G. Strube (Eds.), *First Analysis, reanalysis, and repair* (IIG-Berichte 8/94, pp. 51-76). Freiburg: IIG.

Strube, G., Hemforth, B., & Wrobel, H. (1990). Resolution of Structural Ambiguities in Sentence Comprehension: On-line Analysis of Syntactic, Lexical, and Semantic Effects. *The 12th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum, pp. 558-565.

[5]The SOUL system is implemented in the Objective-C programming language and runs on computers with NEXTSTEP 3.2 or higher.