Local syntactic coherence interpretation. Evidence from a visual world study.

Lars Konieczny (lars@cognition.uni-freiburg.de)
Daniel Müller (daniel@cognition.uni-freiburg.de)
Wibke Hachmann (wibke@cognition.uni-freiburg.de)
Sarah Schwarzkopf (sarah.schwarzkopf@cognition.uni-freiburg.de)
Sascha Wolfer (sascha@cognition.uni-freiburg.de)

Center for Cognitive Science, University of Freiburg, Friedrichstr. 50
D-79098 Freiburg i. Br., Germany

Abstract

This paper addresses the question whether locally coherent word sequences in spoken sentences interfere with the global interpretation. We present a visual world experiment where three depicted events were presented simultaneously while a sentence was spoken. The results suggest that local syntactic coherences are not only processed at a syntactic level, but they are interpreted on-line.

Keywords: local coherence; sentence processing; parsing complexity; visual world, dynamical systems; simple recurrent networks.

Introduction

Sentences sometimes contain sequences that form partial or even whole sentences on their own when viewed in isolation. For example, the sentence

i. *The coach smiled at the player tossed a Frisbee by the opposing team.*

contains the grammatically correct main clause *the player tossed a Frisbee.* Similar constructions exist in German:

ii. *Die Tatsache, dass die Astronautin überrascht den Außerirdischen entdeckte, erregte Aufsehen*.
The fact, that the astronaut$_{[fem]}$ surprisedly/surprises the alien discovered, caused a sensation.
"The fact that the astronaut suprisedly discovered the alien, caused a sensation."

Sentence (ii) contains the well-formed main clause *die Astronautin überrascht den Außerirdischen.* (*The astronaut surprises the alien.*). In both cases the meaning of the embedded clause diverges from the global content of the sentence and should not be considered a possible interpretation because it is ruled out by the context of the preceding words. Do locally well-formed sequences, or *local syntactic coherences* (LSCs) nevertheless affect human sentence processing? This question has attracted a growing amount of attention recently because a measurable influence of LSCs would challenge the *global consistency assumption* (GCA) underlying most current theories of language processing: syntactic integrations of incoming material are always consistent with the preceding input.

In recent years, several researchers reported evidence that LSCs indeed influence sentence processing and have to be considered a potential source of linguistic complexity. Tabor, Galantucci and Richardson (2004) found higher reading times in sentences like (i) on the morphologically ambiguous (past participle or a past tensed main verb) *tossed* than in the control conditions where the LSC was destroyed either by replacing *tossed* with the unambiguous *thrown* and/or by using the unreduced relative clause (*…the player (who was) tossed/thrown…)*.

Further evidence was provided by Konieczny (2005) from an online anomaly detection experiment, showing that the detection of errors took longer when they were locally coherent. Blicknell, Demberg and Levy (2008) report *local coherence effects* (local coherence effects) in an analysis of the Dundee Corpus of eyetracking data from reading of newspaper texts, and Cai, Pickering and Sturt (2008) report similar results from a self-paced-reading study.

The human sentence processing mechanism (HSPM) is generally conceived of as an incremental parser whose strategy combines top-down strategies with a varying degree of bottom-up processing, depending on the particular model. Common to virtually all models, and naturally emerging from this view, is that sentence processing is maximally integrated at any given point in time, i.e. that each new word is integrated into a partial parse representation that includes all words of the sentence processed so far. Such an algorithm is parsimonious because only potentially viable structures are computed so that no memory and processing energy is wasted for unnecessary computations. Various complexity metrics based on this idea have since been proposed with considerable success in accounting for processing difficulties (e.g. Gibson 1998; Hawkins 1994). However, local coherence effects pose problems for these accounts because for LSCs to have an effect, the parser would have to compute a virtually infinite amount of analyses which could have been ruled out if the sentence context had been considered right away.

Attempts to explain local coherence effects

Several attempts have been made to explain local coherence effects. Tabor, Galantucci and Richardson (2004) interpret the effects as an interference of local and global analysis, based on a self-organized parser (SOPARSE, Tabor & Hutchins, 2004; see also Vosse & Kempen, 2000): Local lexical fragments – sometimes spanning several words –

compete with one another for being linked to form a coherent parse tree. Gibson (2006), on the other hand, explained Tabor, Galantucci and Richardson's (2004) effects by a combination of lexical unigram bottom-up statistics and global syntactic top-down predictions, eliminating the need for local coherence processing of multiple adjacent words. Finally, the cue-based parsing approach (Lewis & Vasishth, 2005; van Dyke, 2007; Gordon, Hendrick & Johnson, 2001) tries to relate sentence processing to general properties of human memory and claims that local coherence effects may stem from interference of multiple similar NPs during retrieval (van Dyke, 2007). While both bottom-up statistics and memory retrieval models operate merely on a single-word, lexical basis, only Tabor's account attributes local coherence effects to sequences of several adjacent words.

All models have in common that they operate on discrete symbolic representations of syntactic elements. A radically different view stems from eliminative connectionist models, such as Simple Recurrent Networks (SRNs, Elman 1990). SRNs implicitly acquire grammatical knowledge by developing expectations about, and being corrected by each next incoming item. The only source of information is hence contained in the sequence of items the network is trained with. In this framework, the crucial factors of processing complexity appear to be word-order regularity and frequency of exposure (MacDonald & Christiansen, 2002). Although SRNs overcome the limitations of pure n-gram predictions (Christiansen & Chater, 1999), their performance has been demonstrated to be affected by LSCs (Konieczny, Müller and Ruh, submitted). Since local sequential information is an important source of information in the process of grammar acquisition (Elman, 1993) influencing the way language is processed later, local coherence effects can be viewed as an epiphenomenon of the underlying mechanisms of acquisition and processing which SRNs are based on.

Local coherence effects are hence an important window to the fundamental properties of the HSPM and its underlying representations, and a cornerstone in the discussion and evaluation of the theories of human language processing.

However, empirical evidence for LSC processing thus far is exclusively based on reading data, and is mainly based on evidence from English – with the exception of Konieczny (2005). As a consequence of the latter, reduced relative clauses have almost exclusively been tested. Besides the questionable implications of being restricted to just one class of constructions in only one language, a more direct consequence is that only LSCs crossing clause boundaries have been looked at, which might have obscured the real impact LSCs may have on processing. Furthermore, and most importantly, current data provide no conclusion about how LSCs influence processing beyond the syntactic level. To address these questions, we have conducted a visual world experiment with German complement clauses.

Visual world experiment

With our experiment, we addressed three issues of local coherence processing:

1. Do local coherence effects also show up in spoken language processing?
2. Can local coherence effects be demonstrated within clause-boundaries of globally correct sentences?
3. Do local coherences affect language processing beyond the syntactic level? Are locally coherent structures being interpreted to some degree?

Spoken target sentences with (1) and without (2) an embedded local coherence were played while participants were presented with three depicted events on a computer screen. Only one of these depicted events corresponded to the global meaning, while a second event corresponded to the local meaning. The third corresponded to neither the global or local meaning.

If the existence of a local coherence in the spoken target sentence increases the proportion of fixations on pictures depicting their content, the LSC must have been computed and interpreted to some degree.

Materials and Design

We constructed forty-eight German complement clauses embedded in simple matrix clauses. Each complement clause contained a local syntactic coherence that resulted from the use of ambiguous participial adverbs (*überrascht* in (1)) in between the subject and the object of the complement clause. In the control condition, the participial adverb is replaced by a synonymous unambiguous form (*ungläubig* in (2)), with the effect of destroying the local syntactic coherence.

1) Die Tatsache, dass ***die Astronautin überrascht den Außerirdischen vom Mars*** entdeckte, erregte Aufsehen.
The fact, that the astronaut$_{[fem]}$ astonished the alien from Mars discovered, caused quite a stir.
"The fact that the astronaut suprisedly discovered the alien from mars, caused quite a stir."

2) Die Tatsache, dass die Astronautin ungläubig den Außerirdischen vom Mars entdeckte, erregte Aufsehen.
The fact, that the astronaut$_{[fem]}$ surprisedly the alien from Mars discovered, caused quite a stir.
"The fact that the astronaut surprisedly discovered the alien from mars, caused quite a stir."

The word *überrascht* (*astonished*) is lexically ambiguous, as it can be a participle or a full verb. Globally, *überrascht* in (1) must be analyzed as a participle-adverb (*surprisedly*) of the clause-final verb *entdeckte* (*discovered*). Locally however, it can also form a local syntactically coherent clause *die Astronautin überrascht den Außerirdischen vom Mars* (*the astronaut surprises the alien from Mars*) as a full verb. In the control condition without a local coherence, the ambiguous adverb (*überrascht, surprisedly*) was replaced with an unambiguous one (*ungläubig*), as in (2)).

We crossed this factor with the introduction of another potentially LSC-destroying item, a temporal adverb (*gerade*, *just*) placed between the subject and the ambiguous/unambiguous participle (see 3,4). While this eliminates the possibility to form a complete locally coherent main clause within the complement clause, it introduces another locally coherent construction starting with a topicalized temporal adverb (*Gerade überrascht den Außerirdischen* ...). Note, however, that this LSC lacks a Subject-NP and is never actually completed because right after the Object-NP (the alien) the finite verb (*discovers*) finishes the complement clause.

3) Die Tatsache, dass die Astronautin *gerade* überrascht den Außerirdischen vom Mars entdeckte, erregte Aufsehen.
   The fact, that the astronaut$_{[fem]}$ *just* surprisedly /surprises the alien from Mars discovered, caused quite a stir.
   "The fact that the astronaut just suprisedly discovered the alien from mars, caused quite a stir."

4) Die Tatsache, dass die Astronautin *gerade* ungläubig den Außerirdischen vom Mars entdeckte, erregte Aufsehen.
   The fact, that the astronaut$_{[fem]}$ *just* perplexedly the alien from Mars discovered, caused quite a stir.
   "The fact that the astronaut just perplexedly discovered the alien from mars, caused quite a stir."

We added this factor to address two potential problems: First, differential fixation proportions on the depicted events corresponding to the global and the local event might simply reflect the quality of the match between the visual properties of the depicted events and the two adverbs. If the additions of *gerade* (*just*) alters or even destroys the effect, this possibility can be ruled out. Secondly, a similar logic holds for the evaluation of the unigram statistics model (Gibson, 2006), where any effect would have to be attributed to the lexical properties of the adverb, regardless of the existence or absence of a local coherence.

Visual materials. Each visual stimulus was composed of three different depicted events with the same pair of actors. One depicted event represented the globally correct content of the subordinate clause (here: *the astronout surprisedly discovers the alien*, Figure 1a), the second picture depicted the content of the local coherence (here: *the astronout surprises the alien*, Figure 1b), and the third picture showed the two actors in a unrelated interaction (such as both not taking notice of each other) neither matching the global nor the local content (Figure 1c).

The placement of the events on the screen was controlled, such that each event occurred equally often in each position.

Auditory materials. In aural perception we have to deal with the problem that prosodic cues might weaken local coherence processing. To minimize the probability, locally coherent sequences were recorded separately – as main clauses – and spliced into the spoken matrix sentences. The result still sounded very natural, partly because LSCs did not cross clause-boundaries. Naturalness was established in a pre-test with 5 naïve native speakers who did not notice any splicing and claimed that the sentences sounded well-formed.

To minimize prosodic differences between conditions, control conditions were produced by splicing the critical words (unambiguous adverb, "gerade") into the target sentence (1):

Die Tatsache, dass die {✂ Astronautin (✂ gerade ✂) überrascht (✂ ungläubig ✂) den Außerirdischen ✂} vom Mars entdeckte, erregte Aufsehen.

Materials were constructed according to a 2x2 design, comprising the factors *adverb-type* (*ambiguous* vs. *unambiguous*) and *coherency-interruption* (*with "gerade"* vs. *without "gerade"*).

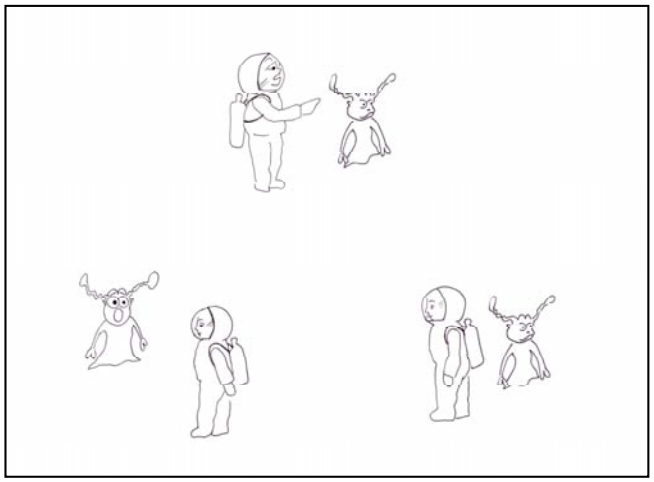

Figure 1: A visual stimulus with three depicted events: a. global event (top), b. local event (left), c. unrelated event (right).

Procedure

Before the experiment started, participants were read a description of the following procedure and were told that they had to judge the overall match of the pictures and the text used in the study in a questionnaire after the experiment.

Each trial started with the presentation of the visual stimulus, followed by a short description of the depicted agents:

5) Hier sieht man eine Astronautin und einen Außerirdischen in drei verschiedenen Szenen.
   Here you see an astronaut and an alien in three different scenes.

After a pause of eight seconds the auditory stimulus was presented in its natural speech rate. During the trials, participants had no explicit task but to listen to the sentences and look at the pictures.

24 participants were presented 48 stimuli each (twelve in each condition), as well as 24 filler items.

Fixations on each of the three depicted events were recorded while the spoken sentences were played to the participants. Data was collected with an SR Research Eyelink II head-mounted eyetracker sampling pupil position at 500 Hz.

Hypotheses

If locally coherent sequences are mentally processed and interpreted, fixation proportions should be higher on the local event in (1), where sentences actually contain a local coherence, than in (2), where they don't. This difference was expected to show up shortly after the participle/adverb, which is required for the LSC interpretation.

In (3-4), where the subject-NP was additionally separated from the adverb by an inserted temporal adverb, we expected to see fewer fixations on the local event at, or shortly after, the ambiguous adverb, as a potential locally coherent event schema still lacks an actor (*Gerade überrascht …*). So if local coherence interpretation can be established between (3) and (4), we expected to see the effect only later in the sentence.

Results

To account for length differences of each segment in the auditory stimuli, trial data were synchronized at the offset of the PP/adverb. Each fixation was assigned to time bins each 50 ms in size, ranging from 1500 ms before, and 4000 ms after the synchronization point. Fixation proportions on the three scenes were calculated for each time bin. Aggregated data are plotted in Figure 2 and 3.

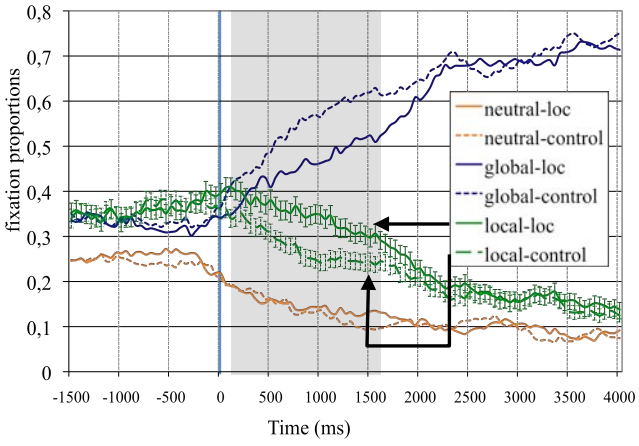

Figure 2: time course of fixation proportions on local, global and neutral scenes, condition 1 (ambiguous adverb, without *gerade*) vs. condition 2 (unambiguous adverb, without *gerade*). Whiskers represent standard errors.

Figure 2 illustrates that when the sentence unfolds and the meaning becomes clearer, fixations on the global event increase to about 80%, indicating that participants grasp the meaning of the sentence and its corresponding depicted event even in the absence of a trial by trial task. With respect to the hypotheses, we found a reliable difference in fixation proportions on the LSC-scene between (1) and (2), starting at about 100 ms and lasting until 1700 ms after the offset of PP/adverb (marked in grey, figure 2). There were more looks at the local event when participants listened to the sentence containing a local coherence than when there was no local coherence in the auditory stimulus (Figure 2).

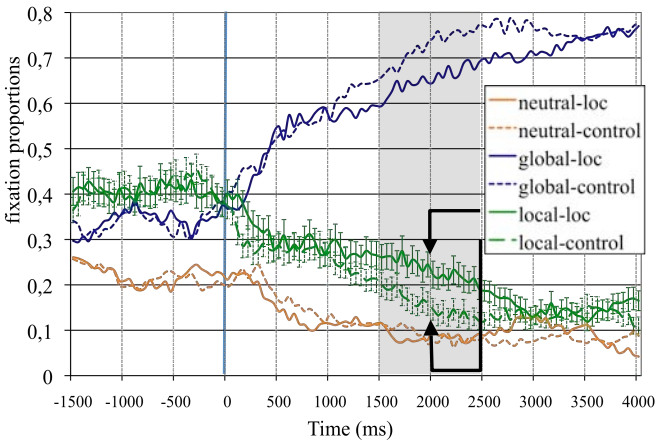

Figure 3: time course of fixation proportions on local, global and neutral scenes, condition 3 (ambiguous adverb, with *gerade*) vs. condition 4 (unambiguous adverb, with *gerade*). Whiskers represent standard errors.

Between (3) and (4) we also found higher fixation probabilites on the LSC-scene in the ambiguous condition, however starting and ending substantially later in the sentence (between 1500ms and 2500ms after offset PP/adverb, see Figure 3, marked in grey), while in the time frame of the local coherence effect between (1) and (2), the fixation proportions on the local event were decreased (Figure 3).

Discussion

The results provide further evidence for the psychological reality of local syntactic coherence processing in sentence processing. Furthermore, they suggest that LSCs not only affect processing at the syntactic level, but that sequences are interpreted to some degree even if they are incompatible with the global content of the sentence they are embedded in. We take the results as strongly supporting language processing accounts that are not based on the *global-consistency assumption*.

Since most of the theories that account for local coherence effects don't make explicit predictions about interpretational processes, our experiment cannot readily be taken as distinguishing between the particular theories. However, the results allow at least some claims about the theories in question.

Concerning the cue-based-retrieval/similarity-based-interference account, it is unclear how the effects could be

explained. Firstly, these models are to the best of our knowledge underspecified with respect to whether or not the verb-form of the ambiguous participle is considered at all during parsing. Interference stems from multiple and potentially ungrammatical retrieval candidates (mostly NPs), and not from ungrammatical cues. Hence, if only grammatical retrieval cues are generated throughout processing, then whether or not the verb interpretation of the participle is retrieved depends on the representations build so far and on the grammatical constraints used for generating cues. If the finite verb interpretation would be considered though, it might have an effect on the looks on the scenes.

Regarding the bottom-up/top-down statistics account (Gibson 2006), it may well be that the difference in type/token probability between the ambiguous and unambiguous adverb elicits processing difficulty in the former case. But in this case the difference should elicit comparable effects in the *gerade* conditions (3-4). However, the local coherence effect shows up only late, about one and a half second after the offset of the ambiguous participle. This renders a lexical unigram effect highly unlikely.

In our view, the data are best accounted for by dynamical self-organizing systems, like Tabor and Hutchins SOPARSE (2004), as well as SRN-based connectionist models, because local coherence effects are a natural by-product of these models. Furthermore, even if no explicit claims are made about the linkage between syntactic and interpretational processes, these frameworks assume no strictly separated syntactic and semantic mechanisms or processes, but rather interactive constraints guiding comprehension. SOPARSE however, due to its hybrid/symbolic nature, must assume a competition between discrete analyses, such that a local coherence should generally increase processing load. In SRNs, on the other hand, LSCs should elicit both competition and facilitation, depending upon the position. At the beginning and throughout the LSC, word predictions of the global and local analysis are in line, so that its quality should actually be increased. At the end of the LSC however, predictions begin to diverge, so that competition should show up in an increased prediction error.

While Tabor, Galantucci and Richardson's (2004) data are compatible with the competition assumption, the elevated reading times might as well have been produced by the reanalysis from the incorrect local to the correct global parse, so that no parallelism is required.

It remains to be seen whether competition effects do also show up in much more straight forward sentences like those used in the present study, where the correct global interpretation does not require a rare construction. Such reading studies are currently under way in our lab, and preliminary results suggest that there is no competition effect before the end of a LSC.

Rationality

In theories dicussed thus far local coherence effects appear more or less unfortunate, regardless of whether they naturally emerge from the system, as for SRNs, or not, because they indicate distraction of attention and thus point to a deficiency of the system. In other words, local coherence effects are seen as reflecting "irrational" properties of the HSPM that allows or even forces it to process irrelevant information while at the same time available contextual information that could have been used for constraining the search must be ignored.

The *Good Enough approach* (Ferreira & Patson, 2007) provides a potential loophole. In everyday communicational contexts, reaching complete accuracy and consistency is not always required and often too wasteful or demanding. Comprehenders therefore use rather fast and frugal heuristics often leading to incomplete or multiple representations. However, these "errors" do not outweigh processing speed, as long as representations are computed that are "good enough" to follow the conversation. So, local coherence effects may reflect the interference of the heuristics-generated meaning of the LSC eventually clashing with the global meaning (reminiscent of the *sausage machine*, Frazier & Fodor, 1978).

In contrast to this bounded rationality approach, Levy (2008) accounts for local coherence effects without abandoning rationality. While he also takes into account the circumstances of realworld communication, he focuses on the nature of linguistic input as noisy and potentially erroneous.

In Levy's account, what appears to be a local coherence might be a valid analysis within a "syntactic neighbour" arising from the repair of previously read words to form an analysis consistent with the "local" coherence. For example, if the word *at* had been read as *as*, the "locally" coherent main-verb reading of *tossed* in "The coach smiled at/as the player tossed the Frisbee …" makes perfect sense. In other words, in the attempt to arrive at a consistent analysis the parser not only has to integrate the incoming material in the actual interpretation consistent with the already processed input, but also considers potential errors made in the past input. Very regular and frequent local sequences may thus lead the parser to reconsider the accuracy of the already processed input – were there any words missing, misunderstood, misspelled or left out? While the idea seems generally very tempting, it remains to be seen whether the assumptions can be supported empirically, for instance by showing that the likelihood of errors in the previously processed material has an effect on LSC processing.

Rationality and SRNs

SRNs, as an instance for distributed processing accounts, are generally quite robust when confronted with erroneous and noisy input. It is possible that SRNs implement a mechanism for dealing with errors in previously processed material inherently. Further research along these lines is necessary to answer this question.

More importantly however, SRNs are not only models for adult sentence processing but first and foremost for language acquisition. As Elman (1993) has argued, in an early stage of language acquisition it is useful to focus on smaller constructions, so that fundamental dependencies can be acquired at all. So the property of SRNs that local syntactic coherences are processed in a particular way could be viewed as relict of the underlying functional language acquisition mechanism.

At the same time, the computation of multiple predictions in SRNs does not require additional resources for multiple computations and storage, as in most symbolic/algorithmic models. Making local predictions is therefore not necessarily costly, while it provides easily accessible and, in the majority of cases, valid information about the remainder of the sentence. SRNs are thus a good implementation of a "good enough" approach (Ferreira & Patson, 2007).

Conclusion

The local coherence effect is an important phenomenon providing insights into the fundamental mechanisms and representations in human language processing. Our results demonstrate the psychological reality of local coherence processing, and show that local coherences can even trigger interpretational processes incompatible with the global context. We take the results as supporting self-organizing dynamical systems, and in particular SRNs, where local coherence effects emerge as a natural consequence of central principles of language acquisition and processing.

Acknowledgements

This work was granted by the German Research Foundation (DFG, grant KO 1932/3-1).

References

Blicknell, K., Demberg, V. & Levy, R. (2008). Local coherences in the wild: An eye-tracking corpus study. *Poster presented at the 21th CUNY Conference On Human Sentence Processing.* Chapel Hill, North Carolina, March 13-15, 2008.

Cai, Z., Pickering, M. J., & Sturt, P. (2008). The filler may be a gentleman: Local coherence effect in long-distance dependency processing. *Poster presented at AMLaP 2008*, Cambridge, UK, 4-6 September 2008.

Christiansen, M. H. & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, *23 (2)*, 157-205.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179-211.

Elman, J. L. (1993). Learning and development in neural networks: the importance of starting small. *Cognition, 48 (1),* 71-99.

Ferreira, F. & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass* 1/1–2, 71–83.

Frazier, L. & Fodor, J.D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6 (4),* 291-325.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, *68 (1)*, 1-76.

Gibson, E. (2006). The interaction of top-down and bottom-up statistics in the resolution of syntactic category ambiguity. *Journal of Memory and Language*, *54*, 363-388.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1411–1423.

Hawkins, J. A. (1994). *A performance theory of order and constituency.* Cambridge/New York: Cambridge University Press.

Konieczny, L. & Müller, D. (2004). Does word order account for the advantage of subject-extracted over object-extracted relative clauses? *Poster presented at AMLaP 2004,* Aix en Provence, France, September 16-18 2004.

Konieczny, L. (2005). The psychological reality of local coherences in sentence processing. *Proceedings of the 27th Annual Conference of the Cognitive Science Society.* Stresa, Italy, August 2005.

Konieczny, L., Müller, D. & Ruh, N. (submitted). Prozesse des sprachlichen Arbeitsgedächtnisses.

Lewis, R. L. & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*, 375-419.

Levy, R. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Waikiki, Honululu, Hawai, October 25-27, 2008.

MacDonald, M. C. & Christiansen, M. H. (2002). Reassessing working memory: A comment on Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review*, *109*, 35-54.

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of Merely Local Syntactic Coherence on Sentence Processing. *Journal of Memory and Language*, *50 (4)*, 355-370.

Tabor, W. & Hutchins, S. (2004): Evidence for self-organized sentence processing: digging-In effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30 (2)*, 431– 450.

Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33 (2)*, 407– 430.

Vosse, T. & Kempen, G. (2000). Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and a lexicalist grammar. *Cognition, 75* (2), 105-43.