Unifying syntactic theory and sentence processing difficulty through a connectionist minimalist parser

Sabrina Gerth (gerth@ling.uni-potsdam.de)
Department of Linguistics, Karl-Liebknecht-Str. 24-25
14476 Potsdam, Germany

Peter beim Graben (p.r.beimgraben@reading.ac.uk)
School of Psychology and Clinical Language Sciences, Whiteknights, PO Box 217,
Reading, RG6 6AH, United Kingdom

Keywords: Computational psycholinguistics; Modeling human sentence processing; Minimalist Grammars; Connectionist architecture;

Introduction

Syntactic theory provides a rich array of representational assumptions about linguistic knowledge and processes. Such detailed and independently motivated constraints on grammatical knowledge ought to play a role in sentence comprehension. However most grammar-based explanations of processing difficulty in the literature have attempted to use grammatical representations and processes per se to explain processing difficulty. They did not take into account that the description of higher cognition in the mind encompasses two levels: On the one hand, at the macrolevel, symbolic computation is performed, and on the other hand, at the microlevel, mathematical computation is achieved through processes within a dynamical system. One critical question is therefore how linguistic theory and dynamical systems can be unified to provide an explanation for processing effects. Here, we present such a unification for a particular account to syntactic theory: namely a parser for Stabler's Minimalist Grammars, in the framework of Smolensky's Integrated Connectionist/Symbolic architectures. In simulations we demonstrate that the connectionist minimalist parser produces predictions which mirror empirical findings from psycholinguistic research.

Method

Materials In contrast to English, the word order in German is relatively free, which offers the opportunity to vary syntactic processing difficulties for the same lexical items by changing their morphological case. For this study mild garden-path sentences in German (subject-object vs. object-subject) sentences were used which are known for eliciting a P600 in an event-related brain potential (ERP) experiment (Frisch, Schlesewsky, Saddy, & Alpermann, 2002). Consider the following example sentences in German:

(1) *Der* *Detektiv* *hat* ***die*** ***Kommissarin*** *gesehen.*
The $\text{detective}_{\text{MASC|NOM}}$ has the $\text{investigator}_{\text{FEM|ACC}}$ seen.
'The detective has seen the investigator.'

(2) *Die* *Detektivin* *hat* ***den*** ***Kommissar*** *gesehen.*
The $\text{detective}_{\text{FEM|AMBIG}}$ has the $\text{investigator}_{\text{MASC|ACC}}$ seen.
'The detective has seen the investigator.'

(3) *Den* *Detektiv* *hat* ***die*** ***Kommissarin*** *gesehen.*
The $\text{detective}_{\text{MASC|ACC}}$ has the $\text{investigator}_{\text{FEM|NOM}}$ seen.
'The investigator has seen the detective.'

(4) *Die* *Detektivin* *hat* ***der*** ***Kommissar*** *gesehen.*
The $\text{detective}_{\text{FEM|AMBIG}}$ has the $\text{investigator}_{\text{MASC|NOM}}$ seen.
'The investigator has seen the detective.'

The sentences (1)-(2) have subject-object order whereas (3)-(4) have object-subject order. Previous work (Weyerts, Penke, Münte, Heinze, & Clahsen, 2002) has shown, that sentence (3) is harder to process than sentence (1) due to the scrambling operation which has to be applied to the object of sentence (3) and leads to higher processing load. A second effect for these syntactic constructions in German is that (2) and (4) contain a case ambiguous nominal phrase (NP). Bader and Meng (1999) found that readers assume that the first NP is a subject when it is case-ambiguous; Frisch et al. (2002) showed in an event-related brain potentials study that sentences like (4) lead to a mild garden-path effect. This work is able to model both effects - the scrambling operation as well as the disambiguation effect.

Symbolic Representation The symbolic representations of human sentence processing are well-established in the linguistic literature covering a wide range of grammatical formalisms e.g. lexical-functional grammars (LFG), head-driven phrase grammars (HPSG), tree-adjoining grammars (TAG), Minimalist Grammars (MG) and so on. Until now, the present work is the first study which uses the Minimalist Grammars formalism for German, so far it has been only applied to English (Stabler, 1997; Harkema, 2001; Hale, 2003). In order to use MG for a language with relatively free word order, a new pair of features was introduced into the formalism. These scrambling feature expands the movement operation, thereby accounting for the possibility to rearrange arguments of the sentence signaled by morphological case.

Mathematical Representation The second part of this study deals with the encoding of the particular parse steps carried out by the grammar formalism. The minimalist tree of each parse step is mapped onto the fractal tensor product encoding as follows: role vectors represent the positions in the binary minimalist tree (root, left child, right child), while fillers account for the symbols of the tree and the minimalist features of the lexicon entries (e.g. $>$, $<$, +acc, -acc, d, =d etc.). The tensor product (Smolensky & Legendre, 2006) is calculated by the binding of role and filler which results in a tensor product representation of each parse step. In other words each symbolic representation will be presented as a numerical value in an activation space and can be visualized in a coordination system by trajectories. These trajectories visualize the sentence processing difficulties by exploring different areas in the vector space.
Finally the numerical values of the encoding are used as input to a neural network. This study will use Tikhonov-Hebbian learning to simulate the underlying language processes with the help of autoassociators.

Results

Figure 1 shows the trajectories of sentence (1) and (3) which only differ in the scrambling operation for (3). Both graphs

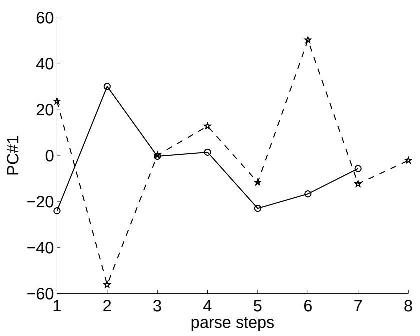

Figure 1: Time series for the scrambling operation.

start with different initial conditions and converge until parse step 6. At this point the second NP is moved (scrambling) at which the trajectories diverge significantly reflecting the disambiguation process and a high syntactic processing difficulty.

Figure 2 shows the trajectories for the sentences (2) and (4). The trajectories start with the same initial conditions and proceed equally because both sentences are parsed equally (following the subject preference strategy) until parse step 5. At that point the graphs diverge significantly which can be interpreted as processing difficulties as encountering the second NP (disambiguation). The scramble operation becomes

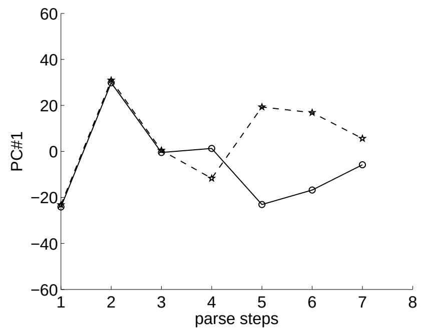

Figure 2: Time series for the garden-path effect.

inevitable for sentence (4) and requires a reanalysis of the built syntactic structure. Further the trajectory for sentence (4) breaks down at parse step 7 simulating the garden-path effect.

By modeling these kinds of processing difficulties (Gerth & beim Graben, submitted; beim Graben, Gerth, & Vasishth, 2008) on both levels–macrolevel and microlevel–this approach bridges the gap between the symbolic computation and the mathematical representation and combines the functionalities of established linguistic theories.

References

Bader, M., & Meng, M. (1999). Subjekt-object ambiguities in german embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research*, *28*(2), 121-143.

beim Graben, P., Gerth, S., & Vasishth, S. (2008). Towards dynamical system models of language-related brain potentials. *Cognitive Neurodynamics*, *2*(3), 229 - 255.

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, *85*, B83-B92.

Gerth, S., & beim Graben, P. (submitted). Unifying syntactic theory and sentence processing difficulty through a connectionist minimalist parser. *Cognitive Neurodynamics*.

Hale, J. T. (2003). *Grammar, uncertainty and sentence processing*. The Johns Hopkins University.

Harkema, H. (2001). *A recognizer for minimalist grammars*. University of California, Los Angeles.

Smolensky, P., & Legendre, G. (2006). *The Harmonic Mind: From neural computation to optimality-theoretic grammar vol. 1: Cognitive architecture* (Vol. 1). MIT Press.

Stabler, E. P. (1997). Derivational Minimalism. *Lecture Notes in Computer Science*, *1328*, 68-95.

Weyerts, H., Penke, M., Münte, T. F., Heinze, H.-J., & Clahsen, H. (2002). Word order in sentence processing: An experimental study of verb placement in german. *Journal of Psycholinguistic Research*, *31*(3), 211-268.