International Journal of Bifurcation and Chaos, Vol. 14, No. 2 (2004) 531–549

PARALLELIZING GRAMMATICAL FUNCTIONS: P600 AND P345 REFLECT DIFFERENT COST OF REANALYSIS

STEFAN FRISCH[*,†], PETER BEIM GRABEN[‡] and MATTHIAS SCHLESEWSKY
Institute of Linguistics, University of Potsdam, Germany
[*]*Max-Planck-Institute of Cognitive Neuroscience, Leipzig, Germany*
[‡]*Institute of Physics, Nonlinear Dynamics Group, University of Potsdam, Germany*
[†]*frisch@rz.uni-potsdam.de*

Received May 1, 2002; Revised September 30, 2002

It is well-known from psycholinguistic literature that the human language processing system exhibits preferences when sentence constituents are ambiguous with respect to their grammatical function. Generally, many theories assume that an interpretation towards the subject is preferred in such cases. Later disambiguations which contradict such a preference induce enhanced processing difficulty (i.e. reanalysis) which reflects itself in late positive deflections (P345/P600) in event-related brain potentials (ERPs). In the case of phoric elements such as pronouns, a second strategy is known according to which an ambiguous pronoun preferentially receives the grammatical function that its antecedent has (*parallel function strategy*). In an ERP study, we show that this strategy can in principle override the general *subject preference strategy* (known for both pronominal and nonpronominal constituents) and induce an object preference, in case that the pronoun's antecedent is itself an object. Interestingly, the revision of a subject preference leads to a P600 component, whereas the revision of an object preference induces an earlier positivity (P345). In order to show that the latter component is indeed a positivity and not an N400-like negativity in the same time range, we apply an additional analysis based on symbolic dynamics which allows to determine the polarity of an ERP effect on purely methodological grounds. With respect to the two positivities, we argue that the latency differences reflect qualitative differences in the reanalysis processes.

Keywords: Event-related brain potentials (ERP); P600; P345; subject preference; parallel function; symbolic dynamics.

1. Introduction

A major part of the empirical research on human language processing addresses the question on how our language processing system deals with locally ambiguous sentences ([Mitchell, 1994]; for a formal treatment see [beim Graben *et al.*, 2004]). The sentence fragment in (1), e.g. is syntactically ambiguous seeing that the relative pronoun *who* can either be the subject of the relative clause (i.e. as in (2)) or the object (i.e. as in (3)), depending on the subsequent input.

It is Peter **who**, ... (1)

It is Peter who likes the girl. (2)

It is Peter who the girl likes. (3)

Whether all disambiguations (such as (2) and (3)) are equally easy to process or whether some of them are more easy than others can tell us something

[†]Address for correspondence: Institute of Linguistics, University of Potsdam, P.O. Box 601553, 14415 Potsdam, Germany.

about the human language processing architecture. If, e.g. one continuation is more difficult than another, this shows (i) that processing is not suspended at the point of ambiguity and (ii) that the processor tries to resolve the ambiguity on the basis of a *preference* for one analysis over others. Numerous factors have been proposed to account for such preferences, i.e. structural simplicity, frequency, working memory limitations etc. (see [Mitchell, 1994] for a review). Experimentally, preferences are measurable in that a disconfirmation of a preferred analysis by following input materials induces enhanced processing cost, because the human sentence processor, the *parser*, has to abandon the preferred choice and to find the appropriate alternative (*reanalysis*). This difficulty can then be determined by different experimental measures, such as self-paced reading times[1] or event-related brain potentials (ERP) [Kutas & Petten, 1994]. With respect to constituents which are ambiguous between subject and object, it has been found in numerous studies across languages that a subject interpretation is preferred and that a later disambiguation towards an object interpretation is more difficult to process [Schlesewsky *et al.*, 2000]. In German, e.g. Schriefers *et al.* [1995] have shown this in sentences such as (4) and (5).

Das	*ist*	*die*	*Tante*,	*die*	*die*	*Nichten*	*gesehen*	**hat**.	
this	is	the	aunt	who[sg]	the	nieces[pl]	seen	has[sg]	
"This is the aunt who has seen the nieces."									(4)
Das	*ist*	*die*	*Tante*,	*die*	*die*	*Nichten*	*gesehen*	**haben**.	
this	is	the	aunt	who[sg]	the	nieces[pl]	seen	have[pl]	
"This is the aunt who the nieces have seen."									(5)

Up to the final word of the sentence, in both (4) and (5), the relative pronoun *die* (*who*) as well as the subsequent noun phrase *die Nichten* (*the nieces*) are ambiguous with respect to their grammatical functions. i.e. *die* can either be the subject and *die Nichten* the object (such as in (4)) or vice versa (such as in (5)). The fragment is disambiguated by the sentence final auxiliary *hat* or *haben*, respectively, which has to agree in number with the subject of the relative clause. In a self-paced reading paradigm, Schriefer *et al.* [1995] found that subjects took longer to read the auxiliary in (5) than in (4), indicating that such a continuation was dispreferred. Testing analogous sentences with event-related brain potentials, Mecklinger *et al.* [1995] found a positivity peaking at around 345 ms on the auxiliary in the dispreferred condition (5) compared to (4).

In German (as in many other languages), a preference to interpret an ambiguous noun phrase (NP) argument as the subject of a clause, rather than the object, has been replicated with different types of experimental measures (cf. [Schriefer *et al.*, 1995; Bader & Meng 1999; Schlesewsky *et al.*, 2000; Mecklinger *et al.*, 1995; Friederici *et al.*, 1998; Frisch *et al.*, 2002]). With respect to language processing architecture, these results demonstrate that, in case of a local ambiguity, not all continuations are predicted to be equally probable, but that some are more expected than others. Seeing that the *subject preference* has shown to be rather immune to nonsyntactic information such as plausibility [Mecklinger *et al.*, 1995] and context [Bader *et al.*, 2000], it is assumed that the preference is based on syntactic factors alone, i.e. lies in the simpler sentence structure of subject-first compared to object-first sentences [Gibson, 1998].

Although nonsyntactic information does not seem to play a role, it is not yet clear whether there are other types of nonstructural, but nevertheless syntactic information on which a preference might be based. One potential candidate in this regard is the so-called *parallel function strategy* [Grober *et al.*, 1978]. This mechanism assumes that the parser endeavors to unify the grammatical functions of elements such as pronouns which need another element (*antecedent*) in order to be interpreted.

[1] In self-paced reading, a sentence is presented in segments, i.e. single words or phrases, on a computer monitor and subjects successively read through the sentence by repeatedly pressing a button while the time for reading each segment is registered [Haberlandt, 1994].

Grober *et al.* [1978] demonstrated this preference by making subjects complete sentence fragments such as (6).

John may scold **Bill** because **he** .... (6)

Grober *et al.* [1978] found a general preference to interpret the pronoun *he*, which is subject of the second clause, as to be coreferent with the subject of the first clause (*John*) rather than the object (*Bill*). The authors present *parallel function* as a heuristic strategy for ambiguity resolution. The question arises, however, whether this mechanism is not merely one of selecting one antecedent out of several possible ones (as is the case in (6)), or whether the parallelity of grammatical functions is also used to resolve the syntactic ambiguity of a constituent. Such an influence was demonstrated by Schlesewsky [1997] (see also [Fanselow *et al.*, 1999]) who found the reading time advantage normally found in sentences such as (4) to disappear, when the antecedent of the relative pronoun *die* was the object of its own clause as in (7).

Der	*Onkel*	*besuchte*	*die Tante,*	*die*	*die Nichten*	*gesehen*	*hat.*	
the	uncle	visited	the aunt[obj]	who[sg]	the nieces[pl]	seen	has[sg]	
"The uncle visited the aunt who has seen the nieces."								(7)

This finding can be explained under the assumption that when the processor encounters the ambiguous relative pronoun (*die*), it attempts to resolve the ambiguity by transferring the grammatical function (i.e. *direct object*) of the antecedent (*die Tante*) to the pronoun, which is exactly what *parallel function strategy* assumes.

1.1. *The present study*

As we have outlined above, psycholinguistic theory has identified two different syntactic mechanisms which trigger preferred readings of ambiguous pronouns in German: A general *subject preference* (also applied to arguments other than pronouns) and a (pronoun-specific) *parallel function strategy* according to which pronouns are preferentially assigned the same grammatical function as their antecedents. In the present study, we present an ERP experiment in which we test the effects of both mechanisms in the case when the two strategies make opposing predictions. More precisely, we want to test whether the positive deflection elicited by a dispreferred disambiguation of a pronoun towards the object [Mecklinger *et al.*, 1995] can be eliminated or even reversed (and an object preference induced) when the antecedent of this pronoun has itself the grammatical function of object.

1.2. *Data analysis*

1.2.1. *Voltage average analysis*

Event-related brain potentials are usually assumed to be small-amplitude, invariant and time-locked signals, the responses of the brain to perceptive or cognitive processes, veiled by the ongoing spontaneous brain activity as it is measured by the electroencephalogram (EEG). In order to recover these signals one has to collect an ensemble of N EEG trials (the so-called *epochs*) all containing the ERPs responding to similar stimuli that are presented at time $t = 0$ with respect to each trial. These trials can then be regarded as realizations of a stochastic process mixed up by the ERP signal and the background EEG, that is modeled by stationary and ergodic noise [Regan, 1972, 1989; Niedermeyer & da Silva, 1999; beim Graben *et al.*, 2000].

Let us briefly review the voltage averaging approach of ERP analysis. We start with the model of a measured ERP trial $x_i(t)$ as a mixture of the invariant ERP signal $s(t)$ and a realization of the noisy background EEG, $\eta_i(t)$,

$$x_i(t) = s(t) + \eta_i(t)\,. \tag{8}$$

Additionally, one assumes that $s(t)$ is uncorrelated with the background EEG and it obeys

$$s(t) = 0\,, \quad \text{for } t < 0\,, \tag{9}$$

i.e. there is no event-related potential before presenting the stimulus. For each epoch, $\eta_i(t)$ is a stationary and ergodic stochastic process mimicking the spontaneous EEG activity and observational noise so that the different $\eta_i(t)$, $\eta_j(t)$ must be stochastically independent for $i \neq j$ and all $\eta_i(t)$ will be drawn from the same probability distribution.

Because the ERP epochs are cut from the measured EEG data stream in a noncontinuous way and because slow waves and sensor artifacts (e.g. a sweating scalp) cause drifts and trends in the EEG, one has to supply a well-defined baseline in the prestimulus time interval $[-T, 0]$ to compare different experimental conditions subsequently. Slow waves can be diminished by filtering the data in a high frequency band. However, it cannot be assumed that the time averages of the processes $\eta_i(t)$ vanish generally. One therefore computes the time averages in the prestimulus intervals

$$\beta_i = \lim_{T\to\infty} \frac{1}{T} \int_{-T}^{0} x_i(t)\mathrm{d}t\,. \tag{10}$$

This leads according to Eq. (9) to the random numbers

$$\beta_i = \lim_{T\to\infty} \frac{1}{T} \int_{-T}^{0} \eta_i(t)\mathrm{d}t\,. \tag{11}$$

Next, one subtracts these baseline values from the corresponding ERP epochs

$$\zeta_i(t) = x_i(t) - \beta_i = s(t) + \eta_i(t) - \beta_i\,.$$

From these processes the (empirical) means

$$\overline{\zeta(t)} = \frac{1}{N}\sum_{i=1}^{N}\zeta_i(t) = s(t) + \frac{1}{N}\sum_{i=1}^{N}(\eta_i(t) - \beta_i)$$

are computed, thus obtaining a further stochastic process. These mean values are estimators of the expectation values

$$E(\overline{\zeta(t)}) = E\left(s(t) + \frac{1}{N}\sum_{i=1}^{N}(\eta_i(t) - \beta_i)\right) = s(t)\,,$$

i.e. the ERP content of the signal [beim Graben *et al.*, 2000].

The discussion above essentially relies on the assumption that the ERP signal is invariant across trials. This is certainly not the case since there are changes in amplitude as well as in latency time (e.g. the signal onset time) from trial to trial e.g. caused by habituation, learning or changes in attention [Callaway & Halliday, 1973; Coppola *et al.*, 1978; Gasser *et al.*, 1983; Möcks *et al.*, 1984]. Thus, it is more realistic to describe the signal content of the EEG at least by the ansatz of Puce *et al.* [1994] in order to tackle this latency jitter problem:

$$s_i(t) = s(t + \tau_i)\,. \tag{12}$$

In Eq. (12), $s(t)$ is the invariant signal as before but it is randomly shifted in time by some discrete stochastic process τ_i with mean $E(\tau_i) = 0$ and variance $D^2(\tau_i) = \sigma_\tau^2$. Beim Graben [2001] has shown how the latency jitter smears the ERP signal out in the course of time (see also [Callaway & Halliday, 1973; Niedermeyer & da Silva, 1999; Krieger *et al.*, 1995]).

Summarizing the prerequisites of the voltage ERP averaging technique, we note that the ERP epochs are supposed to be superpositions of an event-related signal $s(t)$ that should be invariant from trial to trial, and realizations of independent and identically distributed stationary and ergodic stochastic processes $\eta_i(t)$. These conditions are usually not met in real ERP data. The signal $s(t)$ at least jitters in time and correlations between the signal and the background EEG and within the background EEG across trials provide a deterioration of the signal-to-noise ratio. Thus, averaging of signals that are not time locked according to the stimulus yields a damping of waveforms. Furthermore, Eq. (8) states that there is no impact of the noise on the dynamics; it is assumed to be purely observational noise.

1.2.2. *Symbolic dynamics*

In order to present a complementary approach of data analysis that does not rely on requirements that are not fulfilled in measured EEG data, beim Graben *et al.* [2000] have developed an alternative approach to analyze ERP data based on *symbolic dynamics* of time series. According to this approach, an ERP measurement trial $x_i(t)$ is mapped onto a sequence of very few symbols $s = a_1, \ldots, a_L$ obtained by a coarse-graining of the state space of the underlying dynamical system [Hao, 1991; Wackerbauer *et al.*, 1994]. This technique neglects some detailed information about the dynamics. But on the other hand it leaves "robust" properties invariant [Hao, 1991; Tang & Tracy, 1998; Kurths *et al.*, 1995] by "ignoring information about the details of the trajectory in phase space" [Buchner & Zebrowski, 1999]. When the details are contributed by noise, symbolic encoding acts as a highly effective filtering technique. Moreover, symbolic dynamics is able to cope with nonstationarity as has been shown by Schwarz *et al.* [1993], Buchner and Zebrowski [1999], and beim Graben [2001]. In the latter study, beim Graben [2001] has proven that symbolic dynamics also enhances the signal-to-noise ratio of time series disturbed by additive noise. As a tool for analyzing natural data, symbolic dynamics has been successfully applied by Schwarz *et al.* [1993], Buchner and Zebrowski [1999], Flepp *et al.* [1991], Kurths *et al.* [1995], Schiek *et al.* [1998], Engbert *et al.* [1997], Rapp *et al.* [1994], Steuer *et al.*

[2001, 2004], Keller and Wittfeld [2004], Schack [2004], and Mischaikow *et al.* [1999].

Consider an ensemble of ERP epochs $x_i(t)$, where i ($1 \le i \le N$) is the ensemble index and t is the (discrete) time index ranging from 1 to L. N is the cardinality of the ensemble again. The simplest way to get a binary symbolic encoding is to use a certain threshold θ_i depending on each ERP trial. Then, the encoding rule

$$s_{i;t} = \begin{cases} 0: & x_i(t) < \theta_i \\ 1: & x_i(t) \ge \theta_i \end{cases} \tag{13}$$

assigns the symbol "0" to the value $x_i(t)$ of the ith time series at time t if $x_i(t)$ is below the threshold θ_i, and to the letter "1" otherwise. This procedure is called *static encoding* of the time series [Wackerbauer *et al.*, 1994]. θ_i can be chosen by any statistics of the time series, e.g. as the time average of the epoch $x_i(t)$ [Tang & Tracy, 1998] or as the median of $x_i(t)$ [beim Graben *et al.*, 2000; Rapp *et al.*, 1994]. A *dynamic encoding* looks at the changes of the signals $x_i(t)$ in the course of time. Using a *lookahead lag* l the rule

$$s_{i;t} = \begin{cases} 0: & x_i(t+l) < x_i(t) \\ 1: & x_i(t+l) \ge x_i(t) \end{cases} \tag{14}$$

defines a binary dynamic encoding of the ERPs [Wackerbauer *et al.*, 1994].

In the present study we shall apply another, newly developed, encoding algorithm to ERP data for the first time that detects up- and down-going half waves of the signals [beim Graben, 2001]. This *half wave encoding* can be regarded as a mixture of the static and the dynamic encoding strategy. As the static encoding, it makes use of a trial dependent threshold, but now, this threshold becomes also time dependent. First, the algorithm detects the signal's inflection points by estimating the extrema of the first derivatives of the time series. This captures the dynamics of the system. Second, a "jumping baseline" is introduced by the secants between succeeding inflection points. By applying the static encoding rule (13) with respect to the linear threshold functions

$$\theta_{i;j}(t) = a_{i;j}t + b_{i;j}\,, \tag{15}$$

where j is the number of the time interval between inflection points j and $j+1$ of ERP epoch i, one obtains sequences consisting only of consecutive "0"s for the negative half waves and sequences of consecutive "1"s for the positive half waves. For a pure sinusoid this encoding is equivalent to the static encoding (13). However, we have shown that the algorithm is very robust against additive as well as latency noise and against slow nonstationarities, e.g. drifts and trends in the EEG, because its acts on the signal like computing the second derivative [beim Graben, 2001].

By using one of the three encoding strategies, we obtain a matrix $(s_{i;t})_{i\le N;t\le L}$ of symbols "0" and "1". This matrix is considered as a set of rows

$$E = \{s_i | s_i \in \{0,1\}^L, 1 \le i \le N\}\,, \tag{16}$$

where $\{0,1\}^L$ denotes the Lth Cartesian power of the set of symbols called *alphabet* $\{0,1\}$. A subset $[a_{i_1},\ldots,a_{i_n}]_t$ of the ensemble E is called *n-cylinder* at time t, if there are n letters $a_{i_1},\ldots,a_{i_n} \in \{0,1\}$ and a time point t such that all sequences in the subset $[a_{i_1},\ldots,a_{i_n}]_t$ match in the subsequence $a_{i_1},\ldots,a_{i_n} \in \{0,1\}$ beginning at time t [beim Graben *et al.*, 2000]. The symbol sequence $w = (a_{i_1},\ldots,a_{i_n}) \in \{0,1\}^n$ is called *n-word*. This definition was introduced into information theory by McMillan [1953]. For an instructive example consider the set of all books ever printed. Then the cylinder $[\text{book}]_{42}$ is the subset of all books having the word "book" beginning with the 42nd letter. However, this cylinder is different from the set $[\text{book}]_{24}$ containing all books that have the word "book" beginning with the 24th letter.

A probability measure for cylinder sets of experimental data is provided by the set theoretic cardinality function "$\#(\cdot)$" counting the members of the cylinder set

$$p(a_{k_1},\ldots,a_{k_n}|t) = \frac{\#([a_{k_1}\ \ldots,a_{k_n}]_t)}{N}\,. \tag{17}$$

Considering all cylinders of given length n at given time t together, the resulting distribution $\{([a_{k_1},\ldots,a_{k_n}]_t,\ p(a_{k_1},\ldots,a_{k_n}|t))|t,n \text{ fixed}\}$ is called *word statistics* of order n. The word statistics can be further characterized by measures of complexity, e.g. the Shannon and the Rényi entropy [Shannon & Weaver, 1949; Rényi, 1970]. The Shannon cylinder entropy of order n at time t of the ensemble E is given by

$$H_n(t) = -\sum_{(a_{k_1},\ldots,a_{k_n})} \times p(a_{k_1},\ldots,a_{k_n}|t)\,\log\,p(a_{k_1},\ldots,a_{k_n}|t)\,. \tag{18}$$

The *relative cylinder entropy*

$$H(t) = \frac{H_n(t)}{n} \tag{19}$$

measures the information per letter. A generalization leads to the quantity

$$I_{n;q}(t) = \frac{1}{1-q} \log \sum_{(a_{k_1},\ldots,a_{k_n})} p(a_{k_1},\ldots,a_{k_n}|t)^q \tag{20}$$

that is called n-order *Rényi cylinder entropy* depending on the parameter q, which plays a role as an *inverse temperature* in statistical mechanics [Rényi, 1970]. The base of the logarithm in the formulas above is arbitrary, but we recommend using the $\log_I$, where I is the cardinality of the letter alphabet, because relative entropy will then be normalized to the range $[0, 1]$. Entropy is a measure of uncertainty of a given probability distribution. It reaches its maximum value +1 for uniformly distributed events. It assumes its minimum 0 if there is only one certain event with probability 1. For uniform distributions all q-Rényi entropies have the same value +1. But for nonuniform distributions the q-Rényi entropies differ significantly. For $q > 1$ high word probabilities are enhanced, whereas small probabilities are suppressed. Entropies of symbol distributions (1-word entropies) measure the amount of order in the system at one instance of time and are therefore comparable with instantaneous ensemble averages of voltage ERPs. Thus, we shall refer to them as *instantaneous entropies*. However, the so-called *block entropies*, i.e. cylinder entropies for longer words ($n > 1$), additionally reflect dynamical properties of the system. Event-related entropies are able to detect disorder–order phase transitions in the brain [Kelso *et al.*, 1992; Başar, 1980, 1983]. Similar measures are the wavelet entropy of Quiroga *et al.* [2001] and the linear intertrial coherence of Makeig *et al.* [2002].

2. Methods

2.1. *Experimental design and hypotheses*

We tested German sentences (see (21) to (24) below) consisting of a main clause which introduced two noun phrases of different gender, and a subsequent embedded clause containing an ambiguous personal pronoun *sie* (*she*) and a noun phrase which was always clearly marked for either subject or object. The pronoun was disambiguated by a subsequent noun phrase either towards the subject ((21) and (23)) or the object ((22) and (24)). Furthermore, we varied the grammatical function of the antecedent of the pronoun in a preceding clause between subject (as in (21) and (22)) and object ((23) and (24)). The determiner of the disambiguating second NP in the second clause was the critical (=disambiguating) word for the ERP analysis and is therefore underlined, while the pronoun and its gender-matching antecedent are marked in bold font.

Nachdem **die Kommissarin** *den Detektiv getroffen hatte, sah*
after **the cop**[S] the detective met had saw
sie *den Schmuggler.*
she[amb] the[O] smuggler
"After the cop had met the detective, she saw the smuggler." (21)

Nachdem **die Kommissarin** *den Detektiv getroffen hatte, sah*
after **the cop**[S] the detective met had saw
sie *der Schmuggler.*
she[amb] the[S] smuggler
"After the cop had met the detective, the smuggler saw her." (22)

Nachdem der Detektiv **die Kommissarin** *getroffen hatte, sah*
after the detective **the cop**[O] met had saw
sie *den Schmuggler.*
she[amb] the[O] smuggler
"After the detective had met the cop, she saw the smuggler." (23)

Nachdem	*der*	*Detektiv*	**die Kommissarin**	*getroffen*	*hatte,*	*sah*
after	the	detective	**the cop**[O]	met	had	saw
sie	*der*	*Schmuggler.*				
she[amb]	the[S]	smuggler				

"After the detective had met the cop, the smuggler saw her." (24)

With respect to the different strategies which could be used in order to resolve the ambiguity of the pronoun *sie* (*she*), two different comparisons between conditions are interesting, namely (22) versus (21) and (24) versus (23).

A general subject preference based on structural simplicity of the second clause (cf. [Gibson, 1998]) would predict that *sie* is preferentially analyzed as the subject of the second clause in general. Therefore, this strategy would predict that those conditions induce additional processing cost in which the continuation does not confirm the initial choice of subject, irrespective of the first clause. Accordingly, conditions (22) and (24) in both of which the pronoun is disambiguated as the *object* should be more difficult than their compare conditions (21) and (23) in which the subject choice is confirmed. On the basis of previous studies ([Mecklinger *et al.*, 1995; Friederici *et al.*, 1998]) we would expect to find late positive ERP deflections in (22) compared to (21) and in (24) compared to (23).

If, however, our processing systems attempted to match the grammatical functions of pronoun and antecedent in order to resolve the ambiguity (what *parallel function strategy* would assume, cf. [Grober *et al.*, 1978]), then we would have different predictions with respect to the second comparison. In conditions (21) and (22) in which the antecedent is subject of its own clause, *parallel function* would lead to a subject preference for the pronoun. Thus, (22) in which this choice turns out to be false should be more difficult than (21). Therefore, both a *subject preference strategy* and *parallel function strategy* would, though on different grounds, predict the same ERP behavior, namely, a late positivity in (22) compared to (21). This, however, is different in the second comparison (23 versus 24) in which the antecedent is object and in which *parallel function strategy* should therefore induce an object preference for the pronoun. Thus, since *parallel function* would predict (23) to be the unexpected continuation, a processing difficulty in form of a late positive ERP deflection would be expected in (23) compared to (24) and not vice versa (as would follow from a general *subject preference strategy*).

In sum, whereas for the first comparison, both strategies would predict (22) to be more difficult than (21), they would make exactly opposing predictions for the second comparison, (24) versus (23).

2.2. *Participants*

Sixteen students (eight female) from the University of Leipzig took part in the experiment after giving informed consent. They were paid for their participation. Mean age was 24 years (range 19 to 27). All of them were right handers, monolingual native speakers of German and had normal or corrected-to-normal vision.

2.3. *Materials*

All sentences consisted of a first clause which introduced two nominal phrases (NPs) of different gender (masculine versus feminine). These were the potential antecedents of the pronoun. This subclause was followed by a second clause with a syntactically ambiguous, feminine pronoun which was coreferent with the feminine antecedent in the first clause. The pronoun was followed by a non-pronominal NP which was always masculine (and thereby unambiguous with respect to its grammatical function) and which disambiguated the preceding pronoun. Note that a disambiguation towards object-before-subject does not imply a non-canonical word order, seeing that in German, the unmarked order is that pronouns precede non-pronominal NPs, irrespective of their grammatical function (see [Schlesewsky *et al.*, 2003] for a discussion). Forty sentences in each of the four critical conditions (21) to (24) were constructed out of 40 NP-NP-NP-triplets. In addition, 160 filler sentences analogous to those in the critical conditions but with unambiguous masculine pronouns were used as distractors.

2.4. *Procedure*

All 320 sentences were presented visually in the center of a computer screen and in a pseudorandomized order which varied between subjects. The NPs in the first clause were presented as a whole for 700 ms, all other words for 600 ms. 1000 ms after the final word of each sentence, a comprehension question had to be answered within a maximal interval of 5000 ms. Four types of questions were used, referring to either the first or the second clause with equal probability. Half of the questions were critical since they concerned the grammatical function of the pronoun in the second clause (i.e. "did the cop see someone or did someone see the cop?"). Behavioral performance only served as a criterion that the subjects had read the sentences attentively, but seeing that it is an offline measure, it was not seen as critical for the hypotheses (which were intended to address online processing).[2]

2.5. *Recording of the ERP data*

The EEG was recorded by means of Ag/AgCl electrodes placed on the basis of the 10–20-system [Sharbrough *et al.*, 1995] with C2 as ground electrode and the left mastoid as recording reference. EEGs were re-referenced to linked mastoids offline. In order to control for eye movement artifacts, the electro-oculogram (EOG) was monitored from two electrodes at the outer canthus of each eye and from two electrodes located above and below the subject's right eye. Electrode impedances were kept below 5 kΩ. EEG and EOG channels were recorded continuously with a band pass from DC to 30 Hz and a digitization rate of 250 Hz. ERPs were filtered with a 0.4 Hz high pass filter in order to compensate for drifts.

2.6. *Data analysis*

2.6.1. *Voltage average analysis*

Voltage average analysis ERPs were computed for each subject per condition per electrode. Averages were aligned to a 200 ms baseline (see Sec. 1.2.1) before the onset of the critical word (determiner following the pronoun). Trials with incorrect responses in the comprehension task and/or with ocular or amplifier saturation artifacts were excluded from the averages. On average, 35.9% of the trials were rejected across the four critical conditions. The statistical analysis on the voltage average means was done in an ANOVA with two condition factors, namely, grammatical function of the ANTECEDENT in the first clause (with the levels *subject* versus *object*) and a second factor grammatical function of the PRONOUN in the second clause (again with the levels *subject* versus *object*). ANOVAs were computed for midline and lateral electrodes separately, including a topographical factor ELECTRODE with the three levels *FZ*, *CZ* and *PZ* for the midline and REGION OF INTEREST (ROI) with the four levels *left-anterior* (electrodes F7, F5, F3, FT7, FC5), *right-anterior* (electrodes F8, F6, F4, FT8, FC6), *left-posterior* (electrodes P7, P5, P3, PO7, PO3), *right-posterior* (electrodes P8, P6, P4, PO8, PO4) for the lateral sites. In the report of the statistical results, we follow a strictly hierarchical procedure, going from global interactions to single comparisons at single electrodes or ROIs, respectively [Frisch, 2000]. Interaction between the two condition factors were always resolved by the factor ANTECEDENT and the reanalysis effect (disambiguation of the pronoun towards object versus subject) was computed separately for subject and object antecedent conditions. This procedure corresponds to the above hypotheses and is methodologically "clean" insofar as the respective conditions in each of the comparisons are identical up to the critical word (determiner). Seeing that the topographical factor ROI alone is irrelevant to the hypotheses, it is only reported when it interacts with a condition factor. On the basis of visual inspection, the ANOVA was performed by averaging

[2]For reasons of completeness, we will briefly report the behavioral performance for the questions concerning the interpretation of the pronoun in the second clause (half of the questions). See Sec. 2.6.1 for the statistical design of this analysis. For the error percentages, we found a main effect of ANTECEDENT ($F(1,15) = 9.52$, $p < 0.05$), due to more errors when the antecedent was object (10.8%) compared to when it was subject (7.5%). Furthermore, more errors in the conditions with a disambiguation of the pronoun towards the object (13.2%) versus the subject (5.2%) lead to a main effect of PRONOUN ($F(1,15) = 10.31$, $p < 0.01$). The interaction between both factors was not significant ($F < 1$). For the response latencies, we found a main effect of ANTECEDENT ($F(1,15) = 7.15$, $p < 0.05$) for shorter latencies in the subject- (1941 ms) compared to the object-antecedent conditions (2059 ms). Longer latencies for an object (2098 ms) compared to a subject disambiguation (1902 ms) of the pronoun were reflected in a main effect of PRONOUN ($F(1,15) = 34.15$, $p < 0.001$). Furthermore, there was an interaction between both factors ($F(1,15) = 5.17$, $p < 0.05$) due to the fact that the disadvantage for an object disambiguation of the pronoun was larger if the antecedent was subject (274 ms; $F(1,15) = 30.27$, $p < 0.001$) compared to when it was object (117 ms; ($F(1,15) = 6.45$, $p < 0.05$).

the single subject averages over time in two different time windows, namely, one going from 300 to 500 ms and a second one from 500 to 800 ms.

2.6.2. *Symbolic dynamics*

The half wave encoding algorithm qualitatively described in Sec. 1.2.2 works as follows [beim Graben, 2001]: From a given time series $x(t)$ we determine the ratio of averaged secant slopes over their variances within a sliding time window of length T_1 yielding the *secant slope function*

$$u(t) = \frac{1}{T_1} \sum_{k=1}^{T_1/2} \frac{x(t+k) - x(t-k)}{k} \tag{25}$$

$$v(t) = \frac{1}{T_1 - 1} \sum_{k=1}^{T_1/2} \left[\frac{x(t) - x(t-k)}{k} - u(t) \right]^2 + \left[\frac{x(t+k) - x(t)}{k} - u(t) \right]^2 \tag{26}$$

$$w(t) = \frac{u(t)}{v(t)}\,. \tag{27}$$

Afterwards, the resulting $w(t)$ is smoothed by a moving rectangular low pass filter of length T_2

$$\overline{w}(t) = -\frac{1}{T_2} \sum_{k=-(T_2-1)/2}^{(T_2-1)/2} w(t+k)\,. \tag{28}$$

Finally, the inflection points of the original signal are provided by the extrema of $\overline{w}(t)$. The monotonic branches of $\overline{w}(t)$ correspond to the half waves of the time series. Monotonicity is tested using the dynamic encoding rule (14) with a lookahead window of length l.

The crucial parameter of the encoding is the width T_1 of the window where the secant slopes are computed to estimate the first derivative of the signal. This parameter defines a very narrow frequency band such that the half wave encoding can be considered as symbolic dynamics applied to band pass filtered data (like Başar's *combined analysis procedure* [Başar, 1980; Başar *et al.*, 2004]). To choose the parameter T_1 appropriately, we take the voltage ERPs obtained from the averaging procedure as a heuristics and estimate the durations of the respective components and their first harmonics. The other parameters T_2 and l are then chosen by visual inspection of the cylinder entropies within the time windows of the ERP components.

When using the half wave encoding, the word statistics provides a measure of coherence of the ERP across all measured trials [Saddy & beim Graben, 2002] in a similar fashion as the linear intertrial coherence used by Makeig *et al.* [2002]. A positive half wave, e.g. occurring in almost all epochs at the same time leads to an overlap of the corresponding sequences of consecutive "1"s. This overlap yields high values of the probability estimates [Eq. (17)], $p(a_{k_1}, \ldots, a_{k_n}|t)$, for words containing many consecutive "1"s at a certain instance of time. A qualitative model for the word statistics in the presence of additive and latency noise is discussed by beim Graben and Frisch [to appear].

3. Results

3.1. *Voltage average results*

3.1.1. *Descriptive analysis*

Figures 1 and 2 display the voltage grand averages for the five left-posterior electrode sites.[3] The first comparison between the conditions in which the antecedent is the subject is displayed in Fig. 1. As can be seen from the figure, condition (22) (in which the pronoun is disambiguated towards the object) displays a positivity at around 600 ms compared to condition (21), which is to be expected on the basis of previous studies on subject preferences (see Sec. 1). In the second comparison between the two conditions (23) and (24) with the antecedent being object, as shown in Fig. 2, the two conditions differ from one another between approx 300 and 500 ms. With respect to the latency and topography, this effect could either be an N400-like negativity for a dispreferred disambiguation of the pronoun towards the object in (24) compared to (23) or a P345 positivity in (23) compared to (24) reflecting the revision of an object preference.

3.1.2. *Statistical analysis*

The ANOVA on the voltage average data in the time window between 300 to 500 ms at the midline revealed an interaction PRONOUN by ELECTRODE ($F(2,30) = 5.18$, $p < 0.05$) and a (marginal) three-way interaction PRONOUN by ANTECEDENT

[3]Note that positivity is plotted downwards due to an electrophysiological convention.

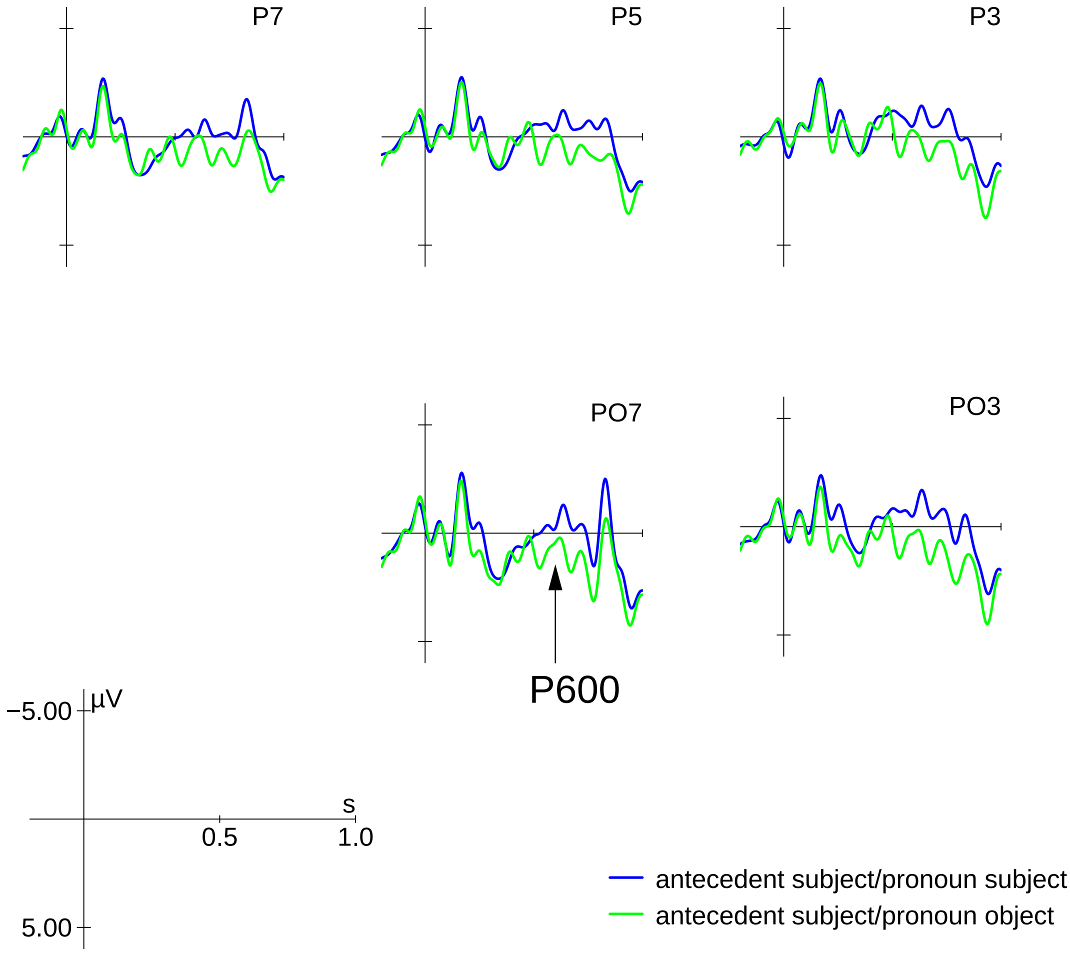


Fig. 1. Average ERPs on the critical disambiguating determiner (onset at vertical bar) for the conditions in which the antecedent of the pronoun is subject. Blue stands for condition (21) in which the pronoun is disambiguated towards the subject, green for condition (22) in which the pronoun is disambiguated towards the object.

by ELECTRODE ($F(2,30) = 3.70$, $p = 0.06$). There was no main effect of PRONOUN at any of the three electrodes, but there was a significant interaction PRONOUN by ANTECEDENT at electrode PZ ($F(1,15) = 5.00$, $p < 0.05$) which was due to a positivity in condition (23) compared to (24) ($F(1,15) = 7.92$, $p < 0.05$) but not in the comparison between (22) and (21) ($F < 1$). Over the lateral sites, we found an interaction PRONOUN by ROI ($F(3,45) = 4.04$, $p < 0.05$), the resolution of which did not reveal significant main effects of PRONOUN in any of the ROIs. Furthermore, we found a three-way interaction ANTECEDENT by PRONOUN by ROI ($F(3,45) = 4.16$, $p < 0.05$). An interaction ANTECEDENT by PRONOUN was found for the left-posterior ROI only ($F(1,15) = 8.61$, $p < 0.05$) and was due to a positivity in (23) compared to (24) ($F(1,15) = 8.35$, $p < 0.05$), whereas there was no difference between (21) and (22) ($F < 1$).

In the statistical analysis in the second time window (500 to 800 ms) for the midline sites, we found a main effect of PRONOUN ($F(1,15) = 10.31$, $p < 0.01$) due to more positive going waveforms in the conditions in which the pronoun was disambiguated towards the object. An interaction PRONOUN by ANTECEDENT by ELECTRODE ($F(2,30) = 3.57$, $p < 0.05$) was also found, but no significant interaction PRONOUN by ANTECEDENT at any of the three electrodes. Over lateral sites, there was a main effect of PRONOUN ($F(1,15) = 11.24$, $p < 0.01$) due to more positive going waveforms in the conditions with an object-before-subject disambiguation at the critical determiner.

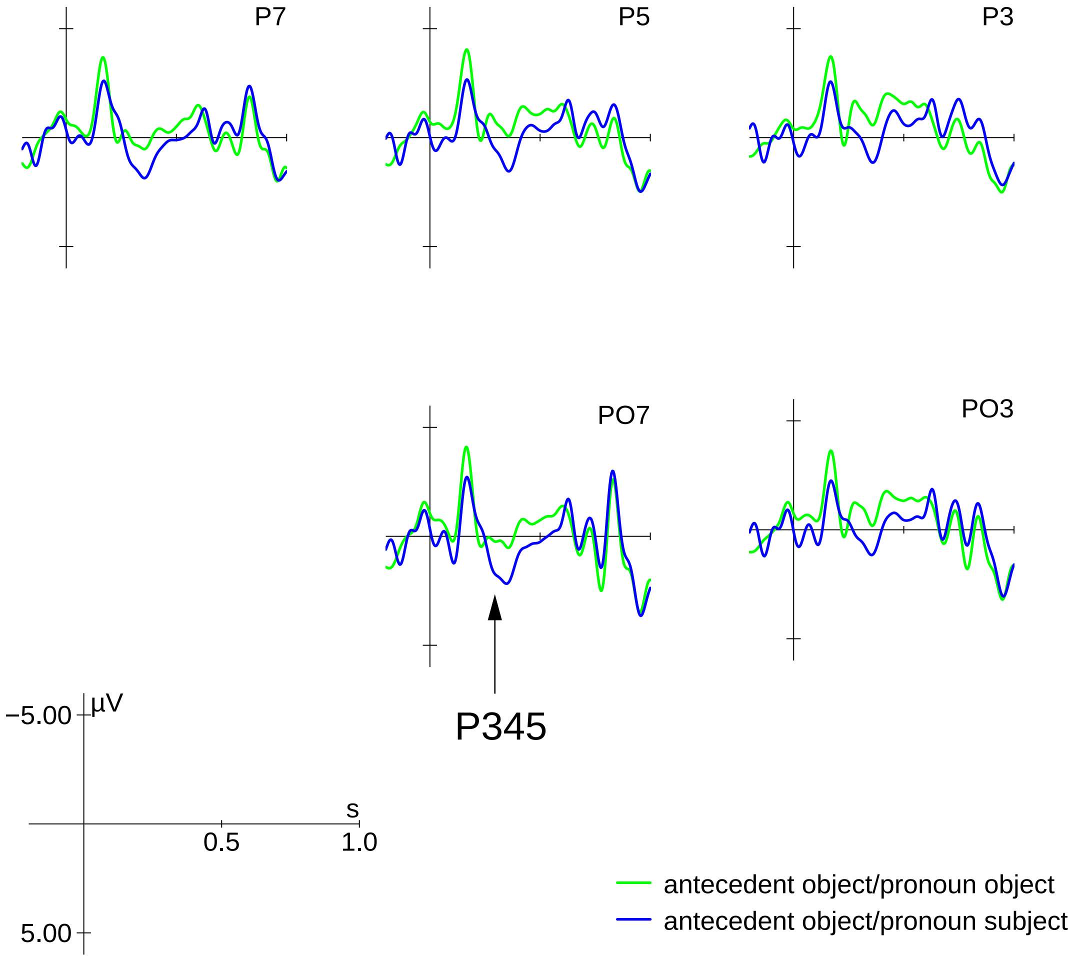


Fig. 2. Average ERPs on the critical disambiguating determiner (onset at vertical bar) for the conditions in which the antecedent of the pronoun is object. Blue stands for condition (23) in which the pronoun is disambiguated towards the subject, green for condition (24) in which the pronoun is disambiguated towards the object.

Furthermore, we found an interaction ANTECEDENT by PRONOUN by ROI ($F(3, 45) = 6.21$, $p < 0.01$). An interaction between the two condition factors was found in both the left-posterior ($F(1, 15) = 11.59$, $p < 0.01$) and the right-posterior ROI ($F(1, 15) = 4.76$, $p < 0.05$). In both ROIs, this interaction was due to a significant positivity in (22) compared to (21) (left-posterior: $F(1, 15) = 13.92$, $p < 0.01$; right-posterior: $F(1, 15) = 6.39$, $p < 0.05$), whereas the conditions (23) and (24) did not differ in any ROI (both left-posterior and right-posterior: $F < 1$).

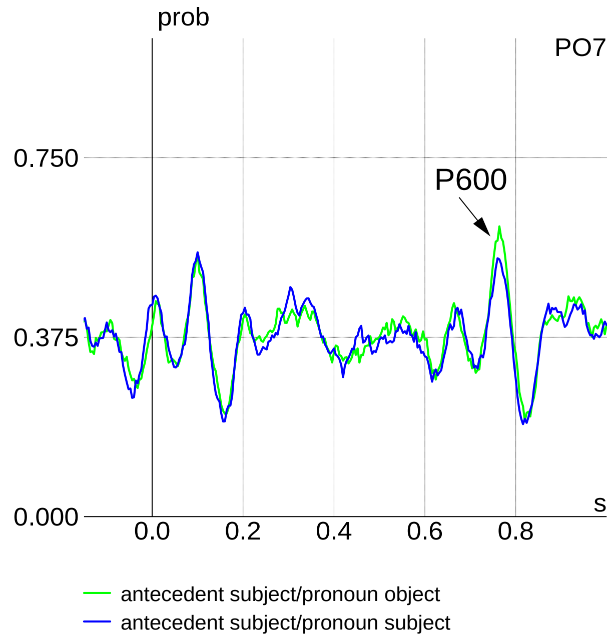
3.2. *Symbolic dynamics results*

3.2.1. *Descriptive analysis*

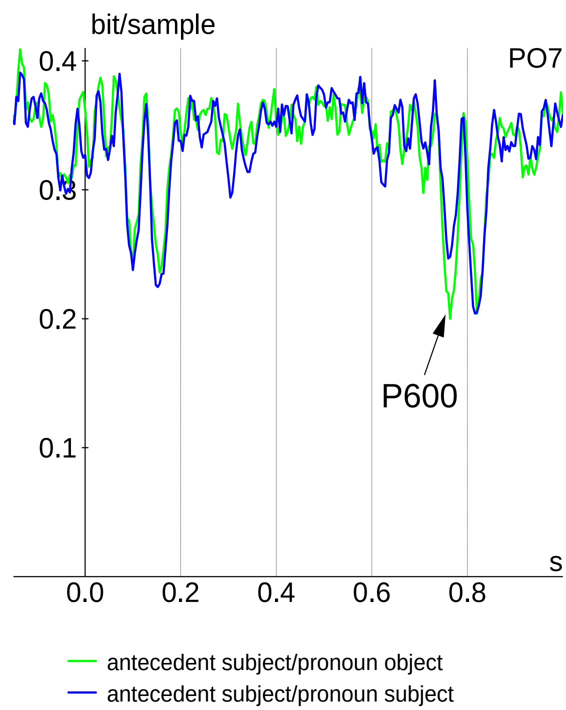
Although there is a clear difference between the two conditions in which the antecedent of the pronoun is object, namely between (23) and (24), we cannot definitely decide whether we deal with a positivity (P345) in (23) compared to (24) or with a (N400-like) negativity in (24) compared to (23). In almost all other ERP studies, theoretical assumptions normally give us a clear definition of what the experimental condition and what the compare condition is [Coles & Rugg, 1995]. In language experiments, e.g. the experimental condition is, on theoretical grounds, anomalous in some respect (i.e. either ungrammatical or dispreferred) and is then compared to a condition which is unmarked. As an example, an N400 in response to a semantic anomaly is a more negative going waveform in a semantically anomalous condition *relative to* a condition which is semantically coherent. In the present study, however, such a determination is only clear in the first comparison between (22) versus (21), seeing that all existing psycholinguistic theories would predict that (22) is the more marked condition and (21) the compare condition (see Secs. 1 and 2.1). In (23) versus (24), by contrast, this question is unclear seeing that different psycholinguistic assumptions allow opposing predictions, both on the basis of previous experimental data: One strategy (*subject preference*) predicts that (24) is dispreferred (and therefore the experimental condition) over (23) (which is then the control condition), whereas another strategy (*parallel function*) predicts the opposite. There is also little evidence as far as component parameters, i.e. morphology, latency, topography, or results from previous studies are concerned. As the visual inspection of many ERP results reveals, component morphology is a very vague way to decide whether one deals with a negativity or a positivity [Coles & Rugg, 1995]. Neither latency nor topography are helpful, seeing that e.g. P345 and N400 occur in very similar time windows and over the same areas of the scalp. And although previous studies have found positivities for a dispreferred disambiguation rather than negativities (see Sec. 1), Friederici *et al.* [1998] reported a positivity only when the disambiguation took place via number agreement, but a negativity when the disambiguation was induced by the morphological case marking of the following argument (as in the present study).

Thus, since we do not have a clear idea on theoretical grounds which of the two conditions (23) and (24) is the experimental and which is the compare condition and since this is crucial for the definition of an ERP component [Coles & Rugg, 1995], beim Graben and Frisch [to appear] suggested a way to determine a questionable effect on purely methodological grounds. Following their approach, we used the half wave encoding technique based on symbolic dynamics as described in Sec. 2.6.2. We performed a half wave encoding of the ERP epochs using two different parameter settings: $T_1 = 24$, $T_2 = 17$, $l = 2$ samples for the subject antecedence conditions (21, 22) and $T_1 = 52$, $T_2 = 17$, $l = 4$ samples for the object antecedence conditions (23, 24). As mentioned in Sec. 2.6.2, the crucial parameter T_1 was adjusted by using the durations of the averaged voltage ERPs as a heuristic. In the former case, the period of the first harmonic of the P600 component in the condition (22) from Fig. 1, whereas in the latter, twice the duration of the whole P345 in the condition (23) from Fig. 2 was used.

After symbolic encoding all trials of all subjects were collected into the grand ensemble Eq. (16).



(a)



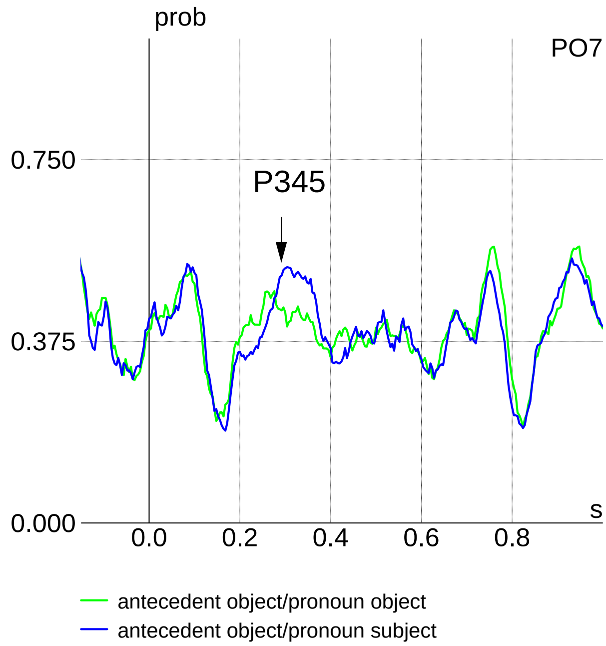
(b)

Fig. 3. (a) Relative frequency of the word "1111" of the half wave encoded ERPs [Eqs. (25)–(28) and (14) with parameters $T_1 = 12$, $T_2 = 17$, $l = 2$ samples] on the critical disambiguating determiner (onset at vertical bar) for the conditions in which the antecedent of the pronoun is subject. Blue stands for condition (21) in which the pronoun is disambiguated towards the subject, green for condition (22) in which the pronoun is disambiguated towards the object. (b) Running cylinder entropy [Eqs. (20) and (19)]. Plotted are 4-word Rényi entropies with parameter $q = 10$.

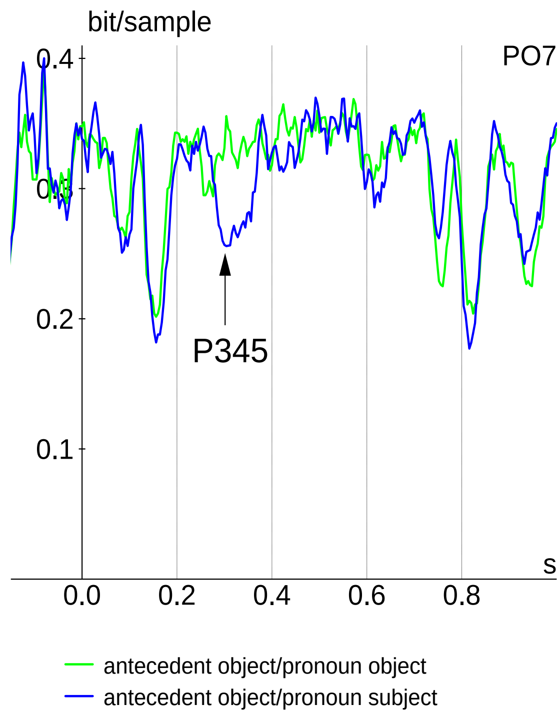
From this, the relative frequencies of all 16 4-words ("0000", "0001", "0010", "0011", "0100", "0101", "0110", "0111", "1000", "1001", "1010", "1011", "1100", "1101", "1110", "1111") were determined according to Eq. (17). Figure 3(a) displays the relative frequency of the word "1111" for one of the selected five left-posterior electrode sites, namely for electrode PO7, for the first comparison of condition (22) against (21) where the antecedent of the pronoun is subject. The P600 ERP component is reflected by an increase of the probability of "1111" thus showing coherence of positive half waves [Saddy & beim Graben, 2002]. As a measure of coherence Fig. 3(b) presents the normalized Rényi cylinder entropy for the 4-word statistics with $q = 10$ [Eqs. (20) and (19)]. As beim Graben [2001] showed, the coherence of half waves due to overlapping words increases with the word length even in the presence of latency jitter Eq. (12). The P600 ERP causes an entropy drop around 760 ms indicating a disorder–order phase transition of the system [Başar, 1980; Kelso *et al.*, 1992].

Figure 3 displays the action of the half wave encoding as a narrow pass band filter. While the voltage averages for the conditions (22) versus (21) diverge in the whole P600 time window [Fig. 1], the corresponding word statistics [Fig. 3(a)] and cylinder entropies [Fig. 3(b)] exhibit such behavior only for a half cycle around 760 ms while the long-lasting divergence of the voltage averages is filtered out. The half wave encoding in the given frequency band therefore enhances *local* timing differences while slow changes are suppressed. Similarly, the spatial topography of ERPs becomes more pronounced at certain electrode sites when using symbolic dynamics. This was the reason for plotting only one electrode (PO7), where the effects were most illustrative. Symbolic dynamics therefore allows to determine the spatiotemporal foci of ERPs with more precision. Furthermore, also the peak latencies of ERP components may differ slightly, because voltage averaging is sensitive against outliers. This problem is not present in symbolic dynamics.

As for the second comparison between the two conditions (23) and (24) where the antecedent of the pronoun is object, Fig. 4(a) reveals a similar increase of the word probability $p(1111|t)$ in (24) compared to (23). Noteworthy, there is an increase in positive half wave coherence indicated by a maximum of this probability around 350 ms where $p(1111|t) > 0.375$ in the opposite direction



(a)



(b)

Fig. 4. (a) Relative frequency of the word "1111" of the half wave encoded ERPs [Eqs. (25)–(28) and (14) with parameters $T_1 = 52$, $T_2 = 17$, $l = 4$ samples] on the critical disambiguating determiner (onset at vertical bar) for the conditions in which the antecedent of the pronoun is object. Blue stands for condition (23) in which the pronoun is disambiguated towards the subject, green for condition (24) in which the pronoun is disambiguated towards the object. (b) Running cylinder entropy [Eqs. (20) and (19)]. Plotted are 4-word Rényi entropies with parameter $q = 10$.

where the pronoun is disambiguated as a subject (23). This finding uniquely shows that the relative positivity in condition (23) with respect to (24) is indeed a positivity (i.e. P345) rather than a negativity [beim Graben & Frisch, to appear].

Figure 4(b) displays the cylinder entropy of the 4-word statistics (Rényi entropy, $q = 10$). Both the P600 in condition (24) and the P345 in condition (23) give rise to decreasing entropy. Note that the durations of both events are different indicating that the components belong to different frequency bands. This finding was supported by the choice of the time window T_1 of the half wave algorithm. $T_1 < 50$ samples enhanced the P600 while $T_1 > 50$ samples boosted the P345.

3.2.2. *Statistical analysis*

The relative frequencies of the word "1111" denoting a positive half wave sustaining for four consecutive samples (16 ms) were subjected to an ANOVA with the same factorial design as for the voltage averages (see Sec. 3.1.2). In the first time window (300 to 400 ms) there was an interaction ANTECEDENT by PRONOUN $F(1, 15) = 4.42$, $p = 0.05$) over the midline, due to the fact that there was no reanalysis effect between the subject antecedent conditions ($F < 1$), but that in the comparison between the two object-antecedent conditions, there was an increase of coherence in the condition in which the pronoun was disambiguated towards the subject compared to a disambiguation towards object ($F(1, 15) = 5.22$, $p < 0.05$). Over lateral sites, we also found an interaction ANTECEDENT by PRONOUN ($F(1, 15) = 6.53$, $p < 0.05$). As in the midline analysis, the two conditions with a subject antecedent did not differ in this time window ($F < 1$), but there was again a stronger increase of coherence in the object antecedent condition in which the pronoun was disambiguated towards the subject compared to when it turned out to be the object ($F(1, 15) = 7.36$, $p < 0.05$).

In the second time window (700 to 800 ms), there was an interaction PRONOUN by ELECTRODE ($F(1, 15) = 6.86$, $p < 0.01$). Resolving it did not reveal a significant main effect of pronoun at any of the three midline electrodes. Over lateral sites, we found an interaction PRONOUN by ROI ($F(1, 15) = 6.18$, $p < 0.01$). Its resolution revealed a significant main effect of pronoun over the left-posterior ROI only ($F(1, 15) = 9.49$, $p < 0.01$) which was due to an increase of coherence in both conditions in which the pronoun was disambiguated towards the object compared to when it turned out to be the subject.

4. Discussion

In a language processing study in German using event-related brain potentials, we opposed two different ambiguity resolution strategies, namely, *subject preference* and *parallel function* in order to investigate the reanalysis processes induced when both of them stand in conflict. We addressed this question by presenting sentences with pronouns which were ambiguous with respect to their grammatical function and were disambiguated by the following word either towards the subject or the object. What we manipulated in addition was whether the antecedent of the pronoun was the subject or the object of its own clause. When the antecedent of the pronoun was itself subject, we observed a preference for the pronoun towards the subject seeing that a disambiguation towards the object induced a positivity in the ERP between 500 and 800 ms (*P600 component*). By contrast, when the antecedent was object, we found a disambiguation of the pronoun towards the *subject* to elicit a positive deflection between 300 and 500 ms (*P345 component*). Although the P600 finding can be explained by both *subject preference* and *parallel function* (seeing that both make the same prediction here, though on different grounds), the P345 effect clearly shows that the resolution of an ambiguous pronoun is influenced by the grammatical function of the pronoun's antecedent.

4.1. *Parallel function as an ambiguity resolution strategy*

That a subject preference can be turned into an object preference is an interesting finding seeing that so far, it has been widely held that a general subject preference is applied independently of other types of information, solely on the basis of some conception of syntactic simplicity (see Sec. 1). Although *parallel function strategy* is based on the syntactic notion of grammatical function, it is independent of structural simplicity, but relies upon the relationship between coreferential elements, that is, on a type of syntactic context. In processing, the preference to match the grammatical functions of coreferent elements seems strong enough not only to eliminate, but even to reverse the subject preference as normally found for ambiguous arguments in noncoreferential contexts.[4]

4.2. *P345 and P600 as indicators of two types of reanalysis*

What remains, however, is the question of the functional difference between the P345 and the P600 component. To say it in other words, why does the revision of a subject preference ((22) versus (21)) elicit a P600 and the revision of an object preference ((23) versus (24)) lead to an earlier P345?

Friederici and Mecklinger [1996] proposed that the difference in latency between the positivities is dependent on different types of reanalysis, the reanalysis reflected in a P345 being easier to perform for the human parser than the reanalysis reflected in a P600. They illustrate this by contrasting the P345 that they found in relative clauses such as (5) compared to (4) with the finding of a P600 in complement clauses such as (30) compared to (29).

Er	*wusste*,	*dass*	*die*	*Tante*	*die Nichten*	*gesehen*	**hat**.	
he	knew	that	the	aunt[sg]	the nieces[pl]	seen	has[sg]	
"He knew that the aunt has seen the nieces."								(29)
Er	*wusste*,	*dass*	*die*	*Tante*	*die Nichten*	*gesehen*	**haben**.	
he	knew	that	the	aunt[sg]	the nieces[pl]	seen	have[pl]	
"He knew that the nieces have seen the aunt."								(30)

As Friederici and Mecklinger [1996] explained with reference to linguistic theory, reanalysis in the first case is easier seeing that a revision from (5) to (4) does not imply changes in the syntactic tree structure of the sentence (i.e. no additional nodes have to be created), but it does when (30) is revised to (29). With respect to our results, this explanation suggests that the revision of a subject preference ((22) versus (21)) makes necessary a "breaking-up" of the tree structure and the insertion of additional nodes, this being reflected in a positivity with a longer latency (P600). By contrast, the revision of an object preference ((23) versus (24)) takes place without adding nodes to the tree, but just by reassigning the lexical elements to the structural positions, thereby eliciting an earlier positivity (P345). Figures 5 and 6 illustrate the tree structures[5] of the second clause that the processor is supposed to adopt when encountering the ambiguous pronoun on the basis of either a subject preference (Fig. 5) or an object preference (Fig. 6).

As can be seen from Fig. 5, when our processor predicts a subject interpretation for the pronoun, there is no reason to assume further arguments (such as a direct object). In other words, projecting an intransitive structure (i.e. a structure with only a verb and a subject but no object) suffices at that point (cf. [Gibson, 1998]). When a following item is clearly marked as the subject argument, then this structure has to be made transitive, in that an additional projection ("XP") has to be created preceding the subject in order to provide a position for the pronoun (that must then be object) as is illustrated in Fig. 5. By contrast, when the parser

[4]As the global main effect of PRONOUN in the second time window suggests, there seems to be an additional P600 effect for condition (24) compared to (23). Further analysis confirmed that this effect was present at anterior electrode sites and that it was probably due to a subgroup of participants preferring a subject interpretation for the pronoun in all instances. This, however, disproves neither our claim that *parallel function* is in principle able to override *subject preference* nor that the two positivities reflect functional differences in the revision of a preference. As has been widely discussed in the literature, strategies such as *subject preference* are also subject to interindividual differences (cf. [Mitchell, 1994]), and rarely found for *every* subject in a sample.

[5]The nomenclature follows standard X-Bar Theory [Haegeman, 1994]. "t" stands for "trace" and indicates a base position from which the respective element has been moved out of a canonical position. The labels denote lexical categories (C: complementizer, N: noun, V: verb) or phrasal categories (CP: complementizer phrase, IP: inflectional phrase, NP: noun phrase, VP: verb phrase). The "XP" projection in Fig. 6 is specific for pronouns in German which always have to precede non-pronominal nominal phrases, irrespective of their grammatical function (see [Schlesewsky *et al.*, 2003] for a detailed discussion of this issue). The important difference between the two trees is that in Fig. 6, an additional level (XP) has to be inserted in order to provide a position which the object can be moved into, thus increasing the complexity of the whole structure.

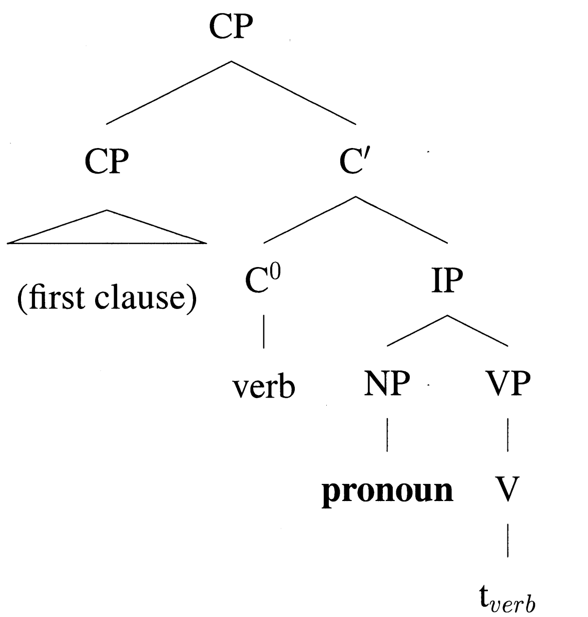


Fig. 5. Tree structure projected at the ambiguous pronoun on the basis of a subject preference, i.e. when either a general subject preference is effective or when parallel function is applied and the antecedent is subject. For further explanation see text.

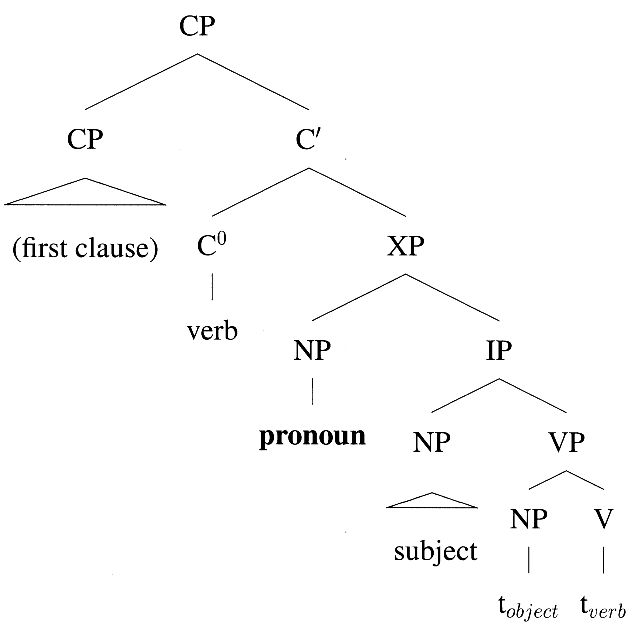


Fig. 6. Tree structure projected at the ambiguous pronoun on the basis of an object preference, i.e. when parallel function is applied and the antecedent is object. For further explanation see text.

(following *parallel function*) assumes that the pronoun is object, then the expected tree structure for the second clause cannot be intransitive. Seeing that each clause has to have a subject in order to be grammatical, a second argument (namely, a subject) has to be predicted and, therefore, a transitive structure with an additional node for the object preceding the subject, see Fig. 6. When this object preference of the pronoun proves false on the following argument, no additional nodes have to be created, but pronoun and following NP only have to be reassigned to already existing tree positions, that is, the pronoun to the subject and the nonpronominal NP to the object position. In sum, the data provide some interesting support for the proposal of Friederici and Mecklinger [1996] to refer to the P345 and the P600 component as reflecting different cases of reanalysis. We have to emphasize that this does not mean that the two revision processes are mutually exclusive. By contrast, seeing that a reassignment of pronoun and nonpronominal NP to other tree positions must also be performed in the case in which the antecedent is the subject, we propose that the longer latency of the P600 reflects the fact that *both* a new projection has to be inserted *and* the elements have to be reassigned, whereas the P345 reflects a revision including only the latter process.

4.3. *Symbolic dynamics of coherent half waves*

In the present study we have applied for the first time a newly developed algorithm of symbolic dynamics to ERP time series that detects up- and down-going (i.e. positive and negative) half waves, in the signal even in the case of additive and latency noise. The algorithm maps these half waves onto sequences of ones and zeros in a symbolic description whose lengths correspond to the durations of the half waves. After applying this approach an ERP epoch is represented by such symbolic sequence of "0"s and "1"s that can be further characterized by measures of complexity. In order to modify these concepts of information theory to analyze symbolically encoded ensembles of ERP time series, we introduced the word statistics and entropies of cylinder sets. The word statistics of the half wave encoding allows to determine the polarity of an ERP effect in an absolute way since it counts the number of trials from the ensemble having overlapping sequences of a certain kind (e.g. "1111" denotes a positive half wave of 4 samples duration). Thus, we were able to determine that the ERP difference between the two conditions with an object antecedent was indeed a positivity (P345) rather than a negativity. From the statistics of 4-words we have also computed the cylinder entropy as a measure of complexity. In the half wave encoding ERP components are revealed by decreasing entropy. Using longer words enhances the probability of overlapping words across trials. Therefore, high-order cylinder entropies are an appropriate tool to get rid of jittering latency times on the effects in the ERP.

We have developed the half wave encoding technique to treat some severe problems in analyzing ERP data. The voltage averaging technique rests on the baseline alignment of single trials and the subsequent ensemble averaging as applied in Sec. 2.6.1. The baseline correction is achieved by subtracting the time average of the prestimulus interval from the whole ERP epoch for each trial separately. This should provide a time series fluctuating around the zero Volt level. Unfortunately, for theoretical reasons this approach requires at least stationarity and ergodicity of the background EEG (see Sec. 1.2.1). However, it is commonly accepted that the EEG is neither stationary nor ergodic. As a consequence, the baseline alignment is only approximately justified for short ERP time series or for early ERP components, whereas late components which are relevant in cognitive sciences (e.g. P345, N400, P600) may be corrupted by the intrinsic nonstationarity of the EEG. The half wave encoding provides an opportunity to cope with this nonstationarity. It determines locally "jumping baselines" as secants between succeeding inflection points and it maps time intervals around maxima and minima of the EEG onto "words", i.e. sequences, of "0"s and "1"s. An ERP component is captured by a subset of trials having overlapping words of one kind, i.e. by a cylinder set. The cardinality of this cylinder is a straightforward measure of the coherence of the effect. As a global measure of intertrial coherence we have introduced the cylinder entropies.

Since the half wave encoding makes use of its locally "jumping baselines", it is immune to trends and drifts in the raw EEG such as slow waves, sweating artifacts, amplifier recalibrations, or noncontinuities resulting from the epoching procedure. Furthermore, the half wave encoding acts as a highly effective noise filter because whole half waves of the EEG are mapped onto sequences of consecutive "1"s or "0"s, thereby completely neglecting the noisy behavior between the inflection points of the signal. Using the half wave encoding in combination with high-order word statistics diminishes the impact of the latency jitter. Thus, symbolic dynamics resting on the half wave encoding turns out to be robust against observational and latency noise yielding a higher signal-to-noise ratio and hence more sensitivity than ERP voltage averaging [beim Graben, 2001]. As a consequence, symbolic dynamics of ERP data not only provides a way to determine the polarity of effects on methodological grounds alone when cognitive theories make opposing predictions. It also offers a promising perspective for experiments in which only a very limited number of trials can be analyzed, i.e. in studies testing children or neurological patients.

Acknowledgments

The work reported in this paper has been financially supported by the German Max Planck Society and by the Deutsche Forschungsgemeinschaft (research group "conflicting rules in cognitive systems"). We wish to thank Angela Friederici, Ina Bornkessel, Jürgen Kurths, Ralf Engbert and Douglas Saddy for helpful contributions, Sabine Koitzsch for her assistance in data acquisition, and Sylvia Jarick and Alexander Siebert for their assistance in data analysis.

References

Başar, E. [1980] *EEG–Brain Dynamics. Relations between EEG and Brain Evoked Potentials* (Elsevier/North Holland Biomedical Press, Amsterdam).

Başar, E. [1983] "EEG and synergetics of neural populations," in *Synergetics of the Brain*, eds. Başar, E., Flohr, H., Haken, H. & Mandell, A. J., Springer Series in Synergetics, Vol. 23 (Springer, Berlin), pp. 183–200.

Başar, E., Özgören, M., Karaka, S. & Başar-Eroğlu, C. [2004] "Super-synergy in brain oscillations and the grandmother percept," *Int. J. Bifurcation and Chaos* **14**, 453–491.

Bader, M. & Meng, M. [1999] "Subject-object ambiguities in German: An across-the-board comparaison," *J. Psycholing. Res.* **28**, 121–143.

Bader, M., Meng, M., Bayer, J. & Hopf, J.-M. [2000] "Syntaktische Funktions-Ambiguitäten im Deutschen — ein Überblick," *Zeitschrift für Sprachwissenschaft* **19**, 34–102.

beim Graben, P., Saddy, J. D., Schlesewsky, M. & Kurths, J. [2000] "Symbolic dynamics of event–related brain potentials, *Phys. Rev.* **E62**, 5518–5541.

beim Graben, P. [2001] "Estimating and improving the signal-to-noise ratio of time series by symbolic dynamics," *Phys. Rev.* **E64**, 051104.

beim Graben, P., Jurish, B., Saddy, D. & Frisch, S. [2004] "Language processing by dynamical systems," *Int. J. Bifurcation and Chaos* **14**, 599–621.

beim Graben, P. & Frisch, S. "Is it positive or negative? On determining ERP components," *IEEE Trans. Biomed. Eng.*, to appear.

Buchner, T. & Zebrowski, J. J. [1999] "Symbolic dynamics analysis of nonstationary data from a model of a magnetic system with solitons," *Phys. Rev.* **E60**, 3973–3981.

Callaway, E. & Halliday, R. A. [1973] "Evoked potential variability: Effects of age, amplitude and methods of measurement," *Electroenceph. Clin. Neurophysiol.* **34**, 125–133.

Coles, M. G. H. & Rugg, M. D. [1995] "Event-related brain potentials: An introduction," in *Electrophysiology of Mind: Event-Related Potentials and Cognition*, eds. Rugg, M. D. & Coles, M. G. H. (Oxford University Press, NY), pp. 1–26.

Coppola, R., Tabor, R. & Buchsbaum, M. S. [1978] "Signal to noise ratio and response variability measurements in single trial evoked potentials," *Electroenceph. Clin. Neurophysiol.* **44**, 214–222.

Engbert, R., Scheffczyk, C., Krampe, R. T., Rosenblum, M., Kurths, J. & Kliegl, R. [1997] "Tempo-induced transitions in polyrhythmic hand movements," *Phys. Rev.* **E56**, 5823–5833.

Fanselow, G., Schlesewsky, M., Cavar, D. & Kliegl, R. [1999] "Optimal parsing: Syntactic parsing preferences and optimality theory," *Rutgers Optimality Archive*, pp. 367–1299.

Flepp, L., Holzner, R., Brun, E., Finardi, M. & Badii, R. [1991] "Model identification by periodic-orbit analysis for NMR-laser chaos," *Phys. Rev. Lett.* **67**, 2244–2247.

Friederici, A. D. & Mecklinger, A. [1996] "Syntactic parsing as revealed by brain responses: First pass and second pass parsing processes," *J. Psycholin. Res.* **25**, 157–176.

Friederici, A. D., Steinhauer, K., Mecklinger, A. & Meyer, M. [1998] "Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses," *Biol. Psychol.* **47**, 193–221.

Frisch, S. [2000] *Verb-Argument-Struktur, Kasus und thematische Interpretation beim Sprachverstehen*, MPI Series in Cognitive Neuroscience, Vol. 12 (MPI of Cognitive Neuroscience, Leipzig).

Frisch, S., Schlesewsky, M., Saddy, D. & Alpermann, A. [2002] "The P600 as an indicator of syntactic ambiguity," *Cognition* **85**, B83–B92.

Gasser, T., Möcks, J. & Verleger, R. [1983] "SELAVCO: A method to deal with trial–to–trial variability of evoked potentials," *Electroenceph. Clin. Neurophysiol.* **55**, 717–723.

Gibson, E. [1998] "Linguistic complexity: Locality of syntactic dependencies," *Cognition* **68**, 1–76.

Grober, E. H., Beardsley, W. & Caramazza, A. [1978] "Parallel function strategy in pronoun assignment," *Cognition* **6**, 117–133.

Haberlandt, K. [1994] "Methods in reading research," in *Handbook of Psycholinguistics*, ed. Gernsbacher, M. A. (Academic Press, San Diego), pp. 1–31.

Haegeman, L. [1994] *Introduction to Goverment & Binding Theory*, Blackwell Textbooks in Linguistics, Vol. 1 (Blackwell Publishers Oxford).

Hao, B.-L. [1991] "Symbolic dynamics and characterization of complexity" *Physica* **D51**, 161–176.

Keller, K. & Wittfeld, K. [2004] "Distances of time series components by means of symbolic dynamics," *Int. J. Bifurcation and Chaos* **14**, 693–703.

Kelso, J. A. S., Bressler, S. L., Buchanan, S., DeGuzman, G. C., Ding, M., Fuchs, A. & Holroyd, T. [1992] "A phase transition in human brain and behavior," *Phys. Lett.* **A196**, 134–144.

Krieger, S., Timmer, J., Lis, S. & Olbrich, H. M. [1995] "Some considerations on estimating event–related brain signals," *J. Neural Trans.* **99**, 103–129.

Kurths, J., Voss, A., Witt, A., Saparin, P. I., Kleiner, H. J. & Wessel, N. [1995] "Quantitative analysis of heart rate variability," *Chaos* **5**, 88–94.

Kutas, M. & Petten, C. K. V. [1994] "Psycholinguistics electrified: Event-related brain potential investigations," in *Handbook of Psycholinguistics* ed. Gernsbacher, M. A. (Academic Press San, Diego), pp. 83–143.

Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E. & Sejnowski, T. J. [2002] "Dynamic brain sources of visual evoked responses," *Science* **295**, 690–694.

McMillan, B. [1953] "The basic theorems of information theory," *Ann. Math. Statist.* **24**, 196–219.

Mecklinger, A., Schriefers, H., Steinhauer, K. & Friederici, A. D. [1995] "Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials," *J. Mem. Lang.* **23**, 477–494.

Mischaikow, K., Mrozek, M., Reiss, I. & Szymczak, A. [1999] "Construction of symbolic dynamics from experimental time series," *Phys. Rev. Lett.* **82**, 1144–1147.

Mitchell, D. C. [1994] "Sentence parsing," in *Handbook of Psycholinguistics*, ed. M. A. Gernsbacher (Academic Press, San Diego), pp. 375–409.

Möcks, J., Gasser, T. & Tuan, P. D. [1984] "Variability of single visual evoked potentials evaluated by two new statistical tests," *Electroenceph. Clin. Neurophysiol.* **57**, 571–580.

Niedermeyer, E. & da Silva, F. L. (eds.) [1999] *Electroencephalography. Basic Principles, Clinical Applications, and Related Fields.* (Lippincott Williams and Wilkins Baltimore).

Puce, A., Berkovic, S. F., Cadusch, P. J. & Bladin,

P. F. [1994] "P3 latency jitter assessed using 2 techniques. I. Simulated data and surface recordings in normal subjects," *Electroenceph. Clin. Neurophysiol.* **92**, 352–364.

Quiroga, R. Q., Rosso, O. A., Başar, E. & Schürmann, M. [2001] "Wavelet entropy in event-related potentials: A new method shows ordering of EEG oscillations," *Biol. Cybern.* **84**, 291–299.

Rapp, P. E., Zimmerman, I. D., Vining, E. P., Cohen, N., Albano, A. M. & Jiménez-Montano, M. A. [1994] "The algorithmic complexity of neural spike trains increases during focal seizures," *J. Neurosci.* **14**, 4731–4739.

Regan, D. [1972] *Evoked Potentials in Psychology, Sensory, Physiology and Clinical Medicine* (Chapman and Hall London).

Regan, D. [1989] *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine* (Elsevier, NY).

Rényi, A. [1970] *Probability Theory* (North–Holland, Amsterdam).

Saddy, J. D. & beim Graben, P. [2002] "Measuring the neural dynamics of language comprehension processes," in *Basic Functions of Language, Reading and Reading Disorder*, eds. Witruk, Friederici, A. D. & Lachmann, T. (Kluwer Academic Publishers, Bostom), pp. 41–60.

Schack, B. [2004] "How to construct a microstate based alphabet for evaluating information processing in time," *Int. J. Bifurcation and Chaos* **14**, 793–814.

Schiek, M., Drepper, F. R., Engbert, R., Abel, H. H., & Suder, K. [1998] "Cardiorespiratory synchronization," in *Nonlinear Analysis of Physiological Data*, eds. Kantz, H., Kurths, J. & Mayer-Kress, G. (Springer, Berlin), pp. 191–213.

Schlesewsky, M. [1997] "Kasusphänomene in der Sprachverarbeitung," PhD thesis University of Potsdam, Potsdam.

Schlesewsky, M., Fanselow, G., Kliegl, R. & Krems, J. [2000] "Preferences for grammatical functions in the processing of locally ambiguous wh-questions in German," *Cognitive Parsing in German*, eds. Hemforth, B. & Konieczny, L. (Kluwer Dordrecht), pp. 65–93.

Schlesewsky, M., Bornkessel, I. & Frisch, S. [2003] "Local parallelism or the neurophysiological basis of word order variation in German," *Brain Lang.* **86**, 116–128.

Schriefers, H., Friederici, A. D. & Kühn, K. [1995] "The processing of locally ambiguous relative clauses in German," *J. Mem. Lang.* **34**, 499–520.

Schwarz, U., Benz, A. O., Kurths, J. & Witt, A. [1993] "Analysis of solar spike events by means of symbolic dynamics methods," *Astron. Astrophys.* **277**, 215–224.

Shannon, C. E. & Weaver, W. [1949] *The Mathematical Theory of Communication* (University of Illinois Press, Urbana).

Sharbrough, F., Chartrian, G.-E., Lesser, R. P., Lüders, H., Nuwer, M. & Picton, T. W. [1995] "American Electroencephalographic Society guidelines for standard electrode position nomenclature," *J. Clin. Neurophysiol.* **8**, 200–202.

Steuer, R., Ebeling, W., Russel, D., Bahar, S., Neiman, A. & Moss, F. [2001] "Entropy and local uncertainty of data from sensory neurons," *Phys. Rev.* **E64**, 061911.

Steuer, R., Ebeling, W., Bengner, T., Dehnicke, C., Hättig, H. & Meencke, H.-J. [2004] "Entropy and complexity analysis of intracranially recorded EEG," *Int. J. Bifurcation and Chaos* **14**, 815–823.

Tang, X. Z. & Tracy, E. R. [1998] "Data compression and information retrieval via symbolization," *Chaos* **8**, 688–696.

Wackerbauer, R., Witt, A., Atmanspacher, H., Kurths, J. & Scheingraber, H. [1994] "A comparative classification of complexity measures," *Chaos Solit. Fract.* **4**, 133–173.