How the brain processes complex words: an event-related potential study of German verb inflections

Martina Penke a, Helga Weyerts b, Matthias Gross c, Elke Zander c, Thomas F. Münte c,d, Harald Clahsen e,∗

a Department of Linguistics, University of Düsseldorf, 40225 Düsseldorf, Germany
b Department of Clinical Neurophysiology, University of Magdeburg, 39120 Magdeburg, Germany
c Department of Neurology, Medical School of Hannover, 30625 Hannover, Germany
d Department of Cognitive Science, University of California, San Diego, La Jolla, CA 92093-0515, USA
e Department of Language and Linguistics, University of Essex, Colchester CO4 3SQ, UK

Accepted 4 March 1997

Abstract

Event-related brain potentials (ERPs) were recorded as German-speaking subjects read verbs in correct and incorrect participle forms. The critical words were presented in three different versions to three different groups of subjects, as part of a simple sentence, in a word list, and embedded in a story; for each version separate ERPs were recorded. Three types of verbs were investigated, regulars, irregulars and nonce verbs. We compared correct regular and irregular participles with incorrect ones; the latter had -en endings on verbs that actually take -t participles (getanz-en), or -et on verbs that require -en (gelad-et). For the nonce verbs, we compared participles with the unexpected -en ending with the expected -et participle forms. The ERP responses were very consistent across the three versions of the experiment: i incorrect irregular participles gelad-et elicited a left frontotemporal negativity; ii incorrect regulars getanz-en produced no differences to the correct ones; iii nonce verbs were associated with an N400 component but did not show a difference between expected and unexpected endings. We will interpret these findings with respect to psycholinguistic models of morphological processing and argue that the brain processes regularly inflected words differently from irregularly inflected ones, the latter by accessing full-form entries stored in memory and the former by a computational process that decomposes complex words into stems and affixes.

Keywords: Event-related potential; Language processing; Left anterior negativity; Regular and irregular inflection; Psycholinguistics; Morphology of language

1. Introduction

One central issue in psycholinguistic research concerns the question of whether regularities of language involve internally represented symbolic rules or whether they can better be explained in terms of assemblies of stored units. Consider, for example, past tense formation in English. Under the latter view, both regular (fax-ed, walk-ed, etc.) and irregular past tense forms (went, came, saw) are said to be represented and processed in the same way, by the storage of items and the creation of connections among them. On the symbol-manipulation view, however, regular past tense forms are explained differently from irregulars, by a symbolic rule that instantiates a variable (the stem of a verb) with an instance, e.g. fax, and concatenates that variable with the -ed affix, yielding faxed. More generally, the controversy revolves around the question to what extent the recognition and production of inflected words involve parsing processes, i.e., the segmentation into stems and affixes, and to what extent it involves accessing whole-word based representations stored in memory.

To date, the processing of morphologically complex words has mainly been investigated by taking reaction-time measures in various kinds of behavioral tasks (e.g. lexical decision, priming, naming). The results of these experiments are discussed controversially with respect to three processing models, cf. Section 2. Another set of evidence that might bear on the theoretical debate on morphological...
processing comes from modern brain imaging techniques. These can be roughly divided in those techniques that provide high spatial resolution but very low temporal resolution, and those that provide a comparatively low spatial resolution but excellent temporal resolution. The former methods include functional MRI with a temporal resolution in the order of 10 s and stimulation PET using radioactively labelled water or butanol with a temporal resolution of about 1 min, while the latter consist mostly of electrophysiological techniques, in particular the event-related brain potentials (ERPs). The first published functional imaging study of processes underlying regular and irregular morphology was done by Jaeger et al. [18]. In this study, the PET technique was applied to a production task in which subjects had to generate past tense forms of regular, irregular and nonce verbs in English (see [8,45,46,48] for earlier PET studies on non-inflected words). Jaeger et al. [18] found differences in the activation patterns between regular and irregular verbs which, they argue, support the view that regular past tense forms are represented differently from irregular ones and involve the processing of symbolic rules. These conclusions are, however, not completely borne out, as we will show in Section 5.3.

For the present study, we adapted the technique of ERPs to investigate German verb inflection processes. Following a number of unsystematic attempts to apply the ERP method to language processing, Kutas and Hillyard [23] in their seminal paper described an ERP component, the N400, that was elicited by semantic (a semantically inappropriate terminal word) but not by physical violations (a sentence-final word written in another font). Subsequent research (summarized, e.g. in [7,17,20,21,40,51,61,62]) has demonstrated that a semantic violation is not a necessary prerequisite for the elicitation of the N400 response and that the N400 amplitude varies as a function of degree of semantic association, contextual constraint, position of the word within a sentence, priming to name, etc. However, the violation paradigm has proven to be extremely reliable in evoking N400 responses. It is not surprising, therefore, that investigators have relied on the presentation of violations in other language domains. Most research in the morpho-syntactic domain has been done on violations of case, number and tense in different languages, English [1,6,24,33,39,42], Dutch [14–16], Spanish [22], German [11,31,34,35], and Turkish [32]. Two ERP effects have been observed to these kinds of morpho-syntactic violations: a positivity with a latency of about 600 ms variably called *P600* [39,41] or *SPS* (syntactic positive shift [16]) and an earlier negativity with a left anterior frontal distribution called *LAN* (left anterior negativity [11,19,22]). While the status of the two effects is still hotly debated, it is safe to say that morpho-syntactic violations reliably elicit similar brain responses across different languages, see Section 5.2 for a discussion of LAN effects.

The violation paradigm has turned out to be one of the most promising avenues for ERP research on language. In our research, we have therefore adopted this paradigm to investigate ERP responses to morphological violations.

## 2. Models of morphological processing

Psycholinguists have suggested three broad approaches to capture the processing of morphologically complex words. First, proponents of *full-listing models* [2,27] do not attribute any significant role to the morphological structure of words in the way they are produced or perceived. Recent connectionist networks [13,26,50,53] are similar in spirit. In these models, all words are claimed to be stored in the mental lexicon irrespective of whether or not the linguist’s analysis yields a morphologically structured representation, e.g. a stem + affix combination.

Second, the *full parsing model* [57–60] claims that only stems have entries in the mental lexicon and that morphological variants need to be decomposed in processing before their stems can be accessed. This model assumes global affix-stripping mechanisms for processing purposes which do not necessarily correspond to the morphological structure of a complex word.

Third, *dual-route models* [49,25,9,56] claim that our mind/brain provides us with two qualitatively different clusters for treating morphologically complex words, each with its own representational basis and processing mechanism. In the Pinker and Prince model [49], these two clusters are linked to the linguistic distinction between regular and irregular inflection. In the so-called irregular cluster, inflected words are represented in terms of lexical entries, and they are processed by accessing full-form representations stored in memory, whereas in the regular cluster inflected words are represented through affixation rules, and in processing they are decomposed into stem + affix combinations.

To test these models of morphological processing, we have investigated two inflectional systems of German, noun plurals and past participles [4,5,44,65,66]. These two systems, we think, provide more appropriate grounds for examining the cognitive status of inflectional rules and the regular/irregular distinction than the English past tense. Notice, for example, that 95% of the verbs in English are regular. Hence, regular past tense forms are much more (type)-frequent than irregular ones (see [28]). Moreover, only regular past tense forms contain a segmentable affix. This means that potential differences between forms such as *walk-ed* versus *came* could be effects of frequency differences and/or effects of the presence or absence of an overt affix, rather than effects of morphological regularity versus irregularity. In German plurals and participles, however, regularity is not confounded with the presence of an overt affix and with type frequency. Regulars and irregulars have segmentable endings, and unlike in English, the regular forms do not represent the majority pattern in any
of the two systems. Regular -t participles (getanz-t “dance-d”) and irregular -n participles (gelad-en “load-ed”) have similar type frequencies¹ and the regular -s plural (Auto-s “car-s”) is less frequent than any of the irregular plurals.

The three morphological processing models mentioned above make different predictions for our ERP experiments. If morphologically complex words are stored in the mental lexicon, as argued by the full-listing model, then we would not expect to find any regular/irregular differences, i.e., violations of regular and of irregular inflection should elicit similar ERP responses. According to the full parsing model, all morphologically complex words are decomposed into stem + affix if possible. In German, regular as well as irregular participle and plural forms have segmentable endings. Thus, incorrect participle and plural forms of both regulars and irregulars might be expected to produce violation effects similar to those observed for morpho-syntactic anomalies in previous ERP studies, i.e., a P600 or a left anterior negativity (LAN). In contrast to the other two models, the dual-mechanism approach [49] would predict different ERP responses for regular and irregular inflection. Parsing violation effects should only occur for incorrect regular affixation, i.e., when the regular participle ending -t and the plural -s are applied incorrectly, and not for misapplications of irregular inflectional patterns.

In a previous study [66], we tested these predictions for German noun plurals. ERPs were recorded as 18 German-speaking subjects read sentences that contained German nouns in correct and incorrect plural forms as critical words. We found different ERP responses to regular and irregular plurals: incorrect applications of -s pluralization to nouns that normally take irregular plural forms, e.g. ’Bauer-s instead of the correct form Bauer-n “farmer-s”, were associated with a negativity that had an onset latency of about 200 ms and a maximum at left anterior temporal sites. In contrast to that, incorrect regulars, e.g. ’Auto-n instead of the correct form Auto-s “car-s”, did not produce a LAN, but rather a low amplitude negativity with a maximum at central sites. This dissociation, we argued, confirms the regular/irregular distinction of the dual-mechanism model in that violations of the -s plural rule elicited clearly different brain responses from misapplications of irregular plurals.

To determine whether these findings generalize to other inflectional systems, we will in the present study report results from three experiments on German participle formation. We will show that, parallel to the plurals, the distinction between regular and irregular participles is associated with different brain responses. Most importantly, violations of the regular -t participle rule elicited a LAN, just like incorrect -s pluralization did. The general implications of this finding for the LAN as a marker of rule/parsing violations will be discussed in Section 5.

3. Methods and materials

3.1. Subjects

Healthy young right-handed native speakers of German were recruited at the Medical School of Hannover, Germany. There were 20 subjects (11 women and 9 men, age range 21–30 years, mean 24.8) in the sentence experiment, 14 subjects (4 women and 10 men, age range 22–33 years, mean 25.6) in the story experiment and 14 subjects in the word-list experiment (4 women and 10 men, age range 22–37 years, mean 27.2). None of the subjects participated in more than one experiment. The data of several additional subjects had to be discarded because of a high artifact rate leading to the rejection of more than 30% of the trials.

3.2. Stimuli

The critical stimuli differed with respect to verb type (regular, irregular, nonce verb) and in the participle endings, -t vs. -n. This resulted in six conditions (see (1)); 50 items per condition were tested.

(1)  a. durchgetanzt “dance-d through” (regular verb, correct -t)
     b. * durchgetanzen (regular verb, incorrect -n)
     c. aufgeladen “load-ed on” (irregular verb, correct -n)
     d. * aufgeladet (irregular verb, incorrect -t)
     e. angerilten “rilf-ed at” (nonce verb, expected -t)
     f. angerilten (nonce verb, unexpected -n)

¹ Bybee [3] argued that -t participle suffixation applies to the largest number of verb stems and therefore has a higher type frequency than -n participle forms. Bybee’s frequency counts, however, are only based on a rather small set of verb types, the 1258 so-called basic verbs (“Grundverben”). Analyses of larger sets of verbs showed that -t and -n participles have similar type frequencies [4,28,63].
In order to minimize differences between participle forms of regular and irregular verbs, we did not include any irregular verbs with vowel changes in their participle forms. Instead, we only included irregular verbs, such as *aufladen* – *aufgeladen* ‘to load on – loaded on’ in which the participle stem (= lad- ‘load-’) is identical to the present tense stem. Hence, the only difference between the irregular and the regular participle forms used in the experiment were the two endings -t vs. -n. To generate 50 irregular verbs without vowel changes in the participle, we made use of prefix-verb formation, a highly productive process in German. This process combines a prefix such as *durch* ‘through’, *auf* ‘on’, *an* ‘at’, with a base verb, e.g. *tanzen* ‘to dance’, and *laden* ‘to load’ to form verbs such as those in (1). Note that all the 300 critical stimuli consist of such prefix/base verb combinations. The items used in the experiments were also controlled with respect to frequency and similarity. All existing participles and the corresponding base verbs have low token frequencies (participle forms < 30 [30]; participle forms of base verbs < 500 [30]; base form of the verbs used < 300 [54]). To minimize possible association effects, we made sure that the participles of nonce verbs used in the experiment did not rhyme with any existing regular or irregular verb. Two sets of critical stimuli with 300 participles each were recorded. To minimize possible association effects, we made sure that the participles of nonce verbs used in the experiment did not rhyme with any existing regular or irregular verb. Two sets of critical stimuli with 300 participles each were prepared; verbs that had the -t ending in the first set had the -n ending in the second set, and vice versa.

The 300 critical words were presented in three different versions, as elements of simple sentences of the following form:

(2) a. Sie haben die ganze Nacht *durchgetanzt*  
(they have danced the whole night through).

b. Sie haben ihre Möbel schon *aufgeladen*  
(cls have already loaded the furniture).

c. Er hat seine Schröge nicht *angerillet*  
(he has not *angerillet* (= nonce verb) his Schröge  
(= nonce noun)).

Each sentence consisted of six words. To exclude potential semantic associations, the direct object in sentences with nonce verbs such as (2c) was a nonce word too. The 300 sentences were randomized and presented one word at a time in yellow letters on a blue background. The words subtended 1.1° of visual angle in height and between 1.2° and 6.2° in width at the viewing distance of 140 cm. Throughout the whole experiment a yellow fixation dot was present in the center of the screen.

The words were shown for 300 ms and the stimulus onset asynchrony between successive words of a sentence was 700 ms. The interval between two sentences was 3500 ms. After 10 critical sentences a ‘‘test sentence’’ was shown in red letters, which either was an exact repetition of one of the 10 sentences shown before or was slightly modified, by exchanging one word. The task assigned to the subjects was to judge whether the test sentence was or was not an exact repetition of one of the previous 10 sentences and to press one of two buttons accordingly. The performance on this task was virtually perfect and will not be discussed further. Short breaks of 1–3 min were given after 40 sentences.

Subjects were tested in a single session while seated in an easy chair in a dimly lit experimental chamber. The subject was observed through an infra-red closed circuit video-system during the experiment. Instructions were given to minimize movements and eye-blinks during the experiment. The total experiment took an average of 150 min to complete including the placement of the electrodes.

#### 3.2.1. Sentence experiment

The critical words were presented as the terminal words within simple declarative sentences of the following form:

(2) a. Sie haben die ganze Nacht *durchgetanzt*  
(they have danced the whole night through).

b. Sie haben ihre Möbel schon *aufgeladen*  
(cls have already loaded the furniture).

c. Er hat seine Schröge nicht *angerillet*  
(he has not *angerillet* (= nonce verb) his Schröge  
(= nonce noun)).

Each sentence consisted of six words. To exclude potential semantic associations, the direct object in sentences with nonce verbs such as (2c) was a nonce word too. The 300 sentences were randomized and presented one word at a time in yellow letters on a blue background. The words subtended 1.1° of visual angle in height and between 1.2° and 6.2° in width at the viewing distance of 140 cm. Throughout the whole experiment a yellow fixation dot was present in the center of the screen.

The words were shown for 300 ms and the stimulus onset asynchrony between successive words of a sentence was 700 ms. The interval between two sentences was 3500 ms. After 10 critical sentences a ‘‘test sentence’’ was shown in red letters, which either was an exact repetition of one of the 10 sentences shown before or was slightly modified, by exchanging one word. The task assigned to the subjects was to judge whether the test sentence was or was not an exact repetition of one of the previous 10 sentences and to press one of two buttons accordingly. The performance on this task was virtually perfect and will not be discussed further. Short breaks of 1–3 min were given after 40 sentences.

Subjects were tested in a single session while seated in an easy chair in a dimly lit experimental chamber. The subject was observed through an infra-red closed circuit video-system during the experiment. Instructions were given to minimize movements and eye-blinks during the experiment. The total experiment took an average of 150 min to complete including the placement of the electrodes.

#### 3.2.2. Story experiment

The same 300 critical words as in the sentence experiment were used. This time the words were embedded in...
Table 2
Statistical results for the different word types

<table>
<thead>
<tr>
<th>Measure</th>
<th>Experiment</th>
<th>Temporal</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Irregular verbs</td>
<td>Early 250–500 ms Sentence</td>
<td>24.6, 0.0001</td>
<td>8.1, 0.008</td>
<td>15.1, 0.001</td>
<td>6.9, 0.003</td>
<td>20.0, 0.0003</td>
<td>1.46, –</td>
<td>4.1, –</td>
<td>1.44, –</td>
<td>15.7, 0.0008</td>
<td>0.13, –</td>
<td>Story</td>
<td>17.9, 0.001</td>
<td>1.0, –</td>
</tr>
<tr>
<td></td>
<td>Story</td>
<td>17.9, 0.001</td>
<td>1.0, –</td>
<td>17.1, 0.002</td>
<td>0.02, –</td>
<td>12.7, 0.004</td>
<td>0.8, –</td>
<td>3.9, –</td>
<td>0.4, –</td>
<td>8.3, 0.015</td>
<td>0.25, –</td>
<td>List</td>
<td>22.9, 0.0004</td>
<td>0.69, –</td>
</tr>
<tr>
<td></td>
<td>Late 500–750 ms Sentence</td>
<td>8.9, 0.01</td>
<td>3.78, 0.04</td>
<td>3.78, –</td>
<td>0.94, –</td>
<td>3.91, –</td>
<td>4.0, 0.015</td>
<td>2.1, –</td>
<td>0.9, –</td>
<td>1.93, –</td>
<td>0.49, –</td>
<td>Story</td>
<td>2.7, –</td>
<td>2.2, –</td>
</tr>
<tr>
<td></td>
<td>Story</td>
<td>2.7, –</td>
<td>2.2, –</td>
<td>10.9, 0.006</td>
<td>0.25, –</td>
<td>0.18, –</td>
<td>0.67, –</td>
<td>9.0, 0.01</td>
<td>0.29, –</td>
<td>0.13, –</td>
<td>0.16, –</td>
<td>0.92, –</td>
<td>0.98, –</td>
<td></td>
</tr>
<tr>
<td></td>
<td>List</td>
<td>3.9, –</td>
<td>2.4, –</td>
<td>0.7, –</td>
<td>1.9, –</td>
<td>2.11, –</td>
<td>0.8, –</td>
<td>1.3, –</td>
<td>1.8, –</td>
<td>0.92, –</td>
<td>0.98, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Regular verbs</td>
<td>Early 250–500 ms Sentence</td>
<td>0.69, –</td>
<td>2.0, –</td>
<td>0.47, –</td>
<td>0.07, –</td>
<td>0.32, –</td>
<td>3.84, 0.02</td>
<td>0.06, –</td>
<td>0.42, –</td>
<td>0.18, –</td>
<td>1.87, –</td>
<td>Story</td>
<td>0.57, –</td>
<td>1.03, –</td>
</tr>
<tr>
<td></td>
<td>Story</td>
<td>0.57, –</td>
<td>1.03, –</td>
<td>13.8, 0.003</td>
<td>1.22, –</td>
<td>0.05, –</td>
<td>0.7, –</td>
<td>1.2, –</td>
<td>0.3, –</td>
<td>0.05, –</td>
<td>0.8, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>List</td>
<td>0.00, –</td>
<td>2.25, –</td>
<td>0.41, –</td>
<td>0.33, –</td>
<td>0.16, –</td>
<td>0.26, –</td>
<td>0.01, –</td>
<td>0.18, –</td>
<td>0.02, –</td>
<td>0.7, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Late 500–750 ms Sentence</td>
<td>0.85, –</td>
<td>0.20, –</td>
<td>2.3, –</td>
<td>2.4, –</td>
<td>0.02, –</td>
<td>0.17, –</td>
<td>0.43, –</td>
<td>0.55, –</td>
<td>0.35, –</td>
<td>3.8, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Story</td>
<td>0.02, –</td>
<td>0.6, –</td>
<td>2.2, –</td>
<td>1.6, –</td>
<td>0.03, –</td>
<td>0.85, –</td>
<td>0.56, –</td>
<td>0.13, –</td>
<td>0.06, –</td>
<td>1.3, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>List</td>
<td>0.35, –</td>
<td>1.88, –</td>
<td>0.7, –</td>
<td>2.0, –</td>
<td>0.12, –</td>
<td>0.23, –</td>
<td>0.13, –</td>
<td>1.2, –</td>
<td>0.63, –</td>
<td>0.43, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Pseudo-verbs</td>
<td>Early 250–500 ms Sentence</td>
<td>1.27, –</td>
<td>0.20, –</td>
<td>2.91, –</td>
<td>3.1, –</td>
<td>1.85, –</td>
<td>3.71, 0.02</td>
<td>3.44, –</td>
<td>0.84, –</td>
<td>0.35, –</td>
<td>0.48, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Story</td>
<td>1.4, –</td>
<td>0.34, –</td>
<td>0.42, –</td>
<td>1.9, –</td>
<td>0.04, –</td>
<td>1.7, –</td>
<td>2.4, –</td>
<td>0.76, –</td>
<td>1.0, –</td>
<td>4.0, 0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>List</td>
<td>2.2, –</td>
<td>5.5, 0.01</td>
<td>4.87, 0.05</td>
<td>2.3, –</td>
<td>1.0, –</td>
<td>2.1, –</td>
<td>2.3, –</td>
<td>1.7, –</td>
<td>2.0, –</td>
<td>2.5, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Late 500–750 ms Sentence</td>
<td>3.84, –</td>
<td>1.85, –</td>
<td>0.66, –</td>
<td>0.58, –</td>
<td>0.68, –</td>
<td>3.5, 0.04</td>
<td>0.26, –</td>
<td>1.64, –</td>
<td>1.19, –</td>
<td>1.91, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Story</td>
<td>3.84, –</td>
<td>1.85, –</td>
<td>0.66, –</td>
<td>0.58, –</td>
<td>0.68, –</td>
<td>3.5, 0.04</td>
<td>0.26, –</td>
<td>1.64, –</td>
<td>1.19, –</td>
<td>1.91, –</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>List</td>
<td>2.2, –</td>
<td>5.5, 0.01</td>
<td>4.87, 0.05</td>
<td>2.3, –</td>
<td>1.0, –</td>
<td>2.1, –</td>
<td>2.3, –</td>
<td>1.7, –</td>
<td>2.0, –</td>
<td>2.5, –</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Correctness = C; Antpost = A: electrode site factor (anterior to posterior direction); Hem = H: hemisphere factor; A×B: interaction of factors and B. Other main effects and interactions omitted.

Degrees of freedom, sentence experiment: Correctness, Correctness×Hemisphere; 1, 19; Correctness×Antpost, C×A×H (temporal, midline) 2, 38; (parasagittal) 4, 76.

Degrees of freedom, story and list experiment: Correctness, Correctness×Hemisphere; 1, 14; Correctness×Antpost, C×A×H (temporal, midline) 2, 26; (parasagittal) 4, 52.
Fig. 1. ERPs for the irregular participles. Left side: grand average ERPs for correct and incorrect participles. The layout of the electrodes (see scheme in right lower corner) corresponds to the approximate positions on the scalp. On the right side of the figure the left anterior temporal site (F7) is depicted on a larger scale. In all three experiments incorrect irregular participles are associated with a more negative waveform. This negativity shows slight differences in distribution across experiments but the left anterior temporal site displays the effect in all cases.
three stories: an animal fable (1312 words total, 104/300 critical words), an episode from the Odyssey (1486 words total, 102/300 critical words) and the tale of the Cologne dwarfs (1349 words total, 94/300 critical words). Unlike the sentence study the critical words in this experiment did not always appear clause-finally; they were rather some-

![Figure 2. ERPs for regular participles. Same layout as Fig. 1. For the incorrect participles no consistent difference across experiment is found.](image)
times followed by a dislocated prepositional phrase, a grammatically well formed word-order variation in German. Two different versions of the stories were prepared with correct and incorrect participle forms exchanged. Half of the subjects were exposed to the first version, the other half to the second one. The mode of presentation (colours

Fig. 3. ERPs for participles of nonce verbs. Same layout as Fig. 1. No consistent differences between participles with the expected and the unexpected ending emerged across experiments.
of words and background, viewing distance was identical to the sentence experiment; each word was presented for 300 ms, the stimulus onset asynchrony between words was 800 ms. Subjects were told that they participated in a memory experiment and after each of the stories they received a questionnaire with 10 simple questions pertaining to the contents of the story. Performance on the questionnaires was flawless and will not be discussed further. The subjects were tested in a single session taking about 180 min including electrode placement.

3.2.3. Word-list experiment

The critical stimuli were the same as in the sentence and the story experiments. The 300 critical participle forms, 300 additional verb forms (100 infinitives, 100 present tense forms, 100 past tense forms) used as filler items, and 100 nouns and 20 nonce nouns were assembled in random order to yield a word list. The word list was presented one word at a time in yellow letters against a blue background (word duration 300 ms, stimulus onset asynchrony 2000 ms). The subjects' task was to press a button held in the right hand whenever they encountered a noun. While in written German the first letter of nouns has to be upper case, all stimuli of the list were presented in lower case letters to prevent subjects from making the noun/verb distinction simply on the basis of case. Performance under these conditions was very good (> 90% correct responses to nouns in each case) indicating that the subjects actually read the stimuli in order to make the required distinction. Subjects were tested in a single short (approx. 70 min) session.

3.3. Recording and analysis

EEG was recorded from all 19 standard scalp sites of the 10/20 system using tin electrodes mounted in an electrode cap with reference electrodes placed at the mastoid processes. Additional electrodes were affixed at the right external canthus and at the right lower orbital ridge to monitor eye movements for later off-line rejection. The biosignals were amplified with a bandpass from 0.01 to 70 Hz, digitized at 250 points/s and stored on magnetic disk. After artifact rejection by an automated procedure using an individualized amplitude criterion on the eye channel and the frontopolar channels, ERPs were averaged for 1024 epochs, including a 100 ms prestimulus interval. The waveforms were quantified by mean-amplitude measures in time windows 250–500 ms and 500–750 ms. These data were subjected to repeated measures analyses of variance. Since effects were differentially distributed over the scalp, separate analyses were done for the midline (ML; Fz, Cz, Pz), parasagittal (PS; Fp1/2, F3/4, C3/4, P3/4, O1/2), and temporal (TE; F7/8, T3/4, T5/6) electrodes with the latter two sets split into an electrode site (anterior to posterior, PS: 5 levels, TE: 3 levels) and a hemisphere factor (left vs. right hemisphere). The analysis was conducted in two steps: first, an overall ANOVA was computed with Word Type (regular vs. irregular vs. pseudo), Correctness (correct vs. incorrect), Site and Hemisphere as factors. As significant interaction effects involving the Word Type and Correctness factors were obtained in these analyses, additional ANOVAs were conducted separately for each word type to further delineate the pattern of results. The Greenhouse–Geisser correction for inhomogeneities of covariance was applied whenever applicable. Reported P values are corrected.

4. Results

The pattern of results was very similar across the three experiments; the data from these experiments will therefore be presented together organized according to the different participle forms, regulars, irregulars and nonce

Fig. 4. Comparison of participles of nonce verbs with the expected (-t) ending and correct regular and irregular participles for the Cz site. The nonce verb participles exhibit a negativity that corresponds to the N400 component.

Fig. 5. Comparison of regular and irregular participles according to ending for the F7 site: only the incorrect irregulars (those with the -t ending) are associated with the left anterior temporal negativity in all three experiments.
forms. Table 1 shows a summary of the results of the overall ANOVAs. For all experiments and for both time windows, interaction effects of the factors Type and Correctness were found, as well as interactions of these two factors with the topography factors Site and Hemisphere. These interactions warranted the computation of separate ANOVAs for the different word types (see Table 2A–C) to elucidate their nature; these will be assessed in the following.

4.1. ERPs to irregular words

The grand-average ERPs to the irregular words are shown in Fig. 1. The ERPs begin with an ensemble of early peaks and troughs (negativity at about 100 ms, positivity at about 200 ms) that differed considerably across the experiments. The differences in these early exogenous components are due to the physical conditions in which the stimuli were presented, most notably to the different inter-stimulus intervals. These were responsible, for example, for the higher amplitude occipital N1 in the word-list experiment.

More importantly, the ERPs in Fig. 1 show that in all three experiments, starting at approximately 150–200 ms, irregular verbs with incorrect -t endings were associated with a more negative waveform. This negativity differed somewhat in its distribution across the experiments but showed a preponderance at the left frontotemporal (F7) site, which is marked by boxes on the left of Fig. 1. At the right of Fig. 1, the F7 site is plotted on a larger scale. Statistically (see Table 2A), this negativity gave rise to a main effect of the factor Correctness for all three electrode sets across the three experiments for the mean amplitude measure in the 250–500 ms time window. For the temporal electrode set, interactions of Correctness with the factors Site (anterior to posterior) and Hemisphere were observed indicating a circumscribed topographical distribution of the negativity. While the effect appeared to extend into later portions of the waveform, statistical effects for the 500–750 ms time window were less pronounced.

4.2. ERPs to regular words

The ERPs to the regular verbs (Fig. 2) showed a similar general waveshape as the irregular verbs. However, no consistent differences between correct and incorrect verbs were seen in the three experiments. In the sentence experiment, the ERPs to incorrect verbs appeared to be slightly more positive starting at approximately 200 ms, an effect that was not present in the story and word-list experiments.

Fig. 6. Illustration of the scalp distribution of the negative effect: depicted are the mean differences between correct and incorrect participles in the 250–500 ms time window aligned for the prestimulus baseline (−100 to 0 ms). TEMP, temporal electrodes; PARA, parasagittal electrodes. A remarkable similarity across experiments emerges with the irregulars being associated with a negativity, which has a left sided preponderance, particularly for the temporal electrodes. The difference between correct and incorrect regular participles words is close to zero in all cases.
Statistically (see Table 2B), the positive-going waveform in the sentence study gave rise to a Correctness by Site (anterior to posterior) interaction for the parasagittal electrode set in the early time window. In the story experiment, however, a more negative-going ERP was found at T3 which led to a Correctness by Hemisphere interaction for the temporal electrode set in the early time window. Hence, there were no consistent ERP effects for incorrect regulars across the three experiments.

4.3. ERPs to nonce words

The ERPs to the nonce verbs are shown in Fig. 3. With regard to the two participle forms (expected -t versus unexpected -n), no consistent differences were obtained across the three experiments. Slightly more positive ERPs to the -n participles in the word-list experiment at the T5 and T6 sites gave rise to a Correctness by Site (anterior to posterior) interaction for the temporal electrode set in the

Fig. 7. A: difference waves (incorrect minus correct) for the three verb types (regular, irregular, nonce) averaged across all three experiments (n = 48). Only the difference wave for the irregulars shows a monophasic negativity with a left anterior maximum. B: isovoltage maps using spherical spline interpolation. Depicted is the mean amplitude in the 250–500 ms time window measured on the incorrect–correct difference waves for the irregulars. The isovoltage maps show a circumscribed left anterior negativity. Scaling: min, −2.7 μV; max, 1.25 μV. C: current source density maps on the same data as in B. These maps provide a reference-free estimate of current flow through the skull. They are characterized by a prominent (negative) sink at F7 and T3. Scaling: min, −17.75 μV/m²; max, 9.00 μV/m².
early time window. In the sentence experiment, on the other hand, a slightly more positive waveform for -n participles was observed at Fp1 and Fp2 causing a Correctness by Site interaction for the parasagittal electrode set.

Recall that the nonce word items used in our experiment are pronounceable words for German native speakers and that previous ERP studies found an N400 component especially for pronounceable nonce words [52,37]. To determine whether this also holds for our nonce word items, ERPs to correct irregular and regular participles were compared to nonce participle forms with the expected -t suffix for the Cz electrode (see Fig. 4). For all three experiments a negativity was observed for the nonce words that was more prominent in the sentence and story experiments than in the word-list study.

4.4. The distribution of the ERPs to regular and irregular words

In order to assess the distribution and the extent of the effects associated with incorrect participles, three analyses were performed. First, we compared the ERP responses to regular and irregular participle forms (1) at the F7 site (see Fig. 5). The waveforms show that only the incorrect irregular participles are associated with an enhanced negativity, whereas incorrect occurrences of the participle -n on regular verbs produced ERPs similar to those of the correct regular participle forms.

Second, we compared the difference in mean amplitude between correct and incorrect participles within the 250–500 ms time window for the temporal (upper row) and the parasagittal (lower row) electrode sets (see Fig. 6). A remarkable similarity between the three experiments emerges: the negativity associated with incorrect irregular verbs shows a left preponderance in all three experiments especially for the temporal set. Moreover, the negativity is most prominent for left frontotemporal and frontal sites. Note also, that the difference between correct and incorrect regular participles is close to zero in all cases.

Finally, to obtain a more precise localization on the scalp, the difference waves (incorrect–correct) were pooled from all three experiments. This averaging across experiments was done simply for illustrative purposes with the idea that the signal-to-noise ratio would be greatly enhanced, thus affording us a better idea of the generalizability of the effects. We felt justified to do this meta-analysis as the critical items were the same in all three experiments, the recording parameters were identical and the subjects were from the same pool. Fig. 7A presents the difference waves for regular, irregular and nonce verbs, Fig. 7B the isovoltage maps and Fig. 7C the current source density maps for the difference between correct and incorrect irregular participles. The difference waves show that it is only the incorrect occurrence of -t on irregulars that elicits a clear ERP response, namely a monophasic negativity with a left anterior maximum, whereas the incorrect regulars and the nonce verbs produce a flat line. The isovoltage maps indicate that the ERP response we found for incorrect irregular participles exhibits a clearly circumscribed distribution. This is further corroborated by the computation of the current source densities, which provides a reference-free estimate of the currents passing through the skull, and suggests a left frontal source for the negativity.

5. Discussion

The most salient and consistent result of the present study is that incorrect irregular participles were associated with a negative ERP with a left anterior temporal distribution starting at about 200 ms, an effect that has been called LAN. The LAN for incorrect irregulars was found in each of the three participle experiments presented in this study. Even more strikingly, this brain response to incorrect irregulars generalizes to an entirely different inflectional system, noun plurals. We found that incorrect irregular noun plurals of German elicited a LAN [66] just like incorrect irregular participles. Note that noun plurals differ from participle forms, both in terms of structural properties and in terms of vocabulary distribution. Whereas the regular participle ending -t is added to verb stems, e.g., "tanz-``danc-'', "kauf-``buy-'', leg-``put'’, the regular plural ending -s is added to word forms such as "Auto``car'', "Pizza, Bar", etc., i.e., to elements which exist as independent words in the language. Moreover, the frequency distribution and the similarity patterns in these two inflectional systems are quite different. Irregular verbs do not outnumber regular ones, independent of the method of frequency calculation chosen [64], but 92% of the German nouns form their plurals according to one of the various irregular patterns. Despite these linguistic differences, the ERP response to incorrect irregulars is very similar for noun plurals and participles, indicating the stability, reproducibility and generalizability of this negativity.

5.1. ERPs and morphological processing models

The current set of ERP results demonstrate clear electrophysiological differences between regular and irregular inflection for noun plurals and participles, a finding which is directly relevant to the competing models of morphological processing introduced in Section 2. Recall that according to the full-listing model regular and irregular inflection are processed in similar ways, i.e., by the storage of items and the creation of connections in the mental lexicon. From such a model, one would predict violations of regular and of irregular inflections to elicit similar ERP responses. Our results show that this prediction is not borne
out. The second variant of a single-mechanism approach is the full parsing model according to which all morphologically complex words are decomposed into stem + affix if possible. As German participle and plural forms have segmental endings, we would again expect to find similar brain responses to incorrect forms of both regulars and irregulars. This is not confirmed, however. Hence, our ERP results do not appear to support any of the single-mechanism models.

The dual-mechanism model, on the other hand, would be consistent with different ERP responses for regular and irregular inflections. This model distinguishes between two different clusters for morphologically complex words, the irregular cluster involving the retrieval of lexical entries stored in memory and the regular cluster involving morphological parsing. Consider, from the perspective of the dual-mechanism model, the specific linguistic conditions tested in our experiments. Two kinds of morphological anomalies were investigated: (i) regularizations, -t affixes appearing on irregular verbs (*geladet); (ii) irregularizations, the participle -n on regular verbs (*getanzen) or nonce verbs (*gerllben). The structure of these morphological anomalies is quite different. The first case is a parsing violation, i.e., a misapplication of the -t affixation rule to a verb that would normally block the rule, to produce illegal stem + affix combinations such as *gelad-et. The other two cases are incorrect or unexpected -n participle forms of verbs that would normally undergo the -t affixation rule. According to the model, there is no -s suffixation rule in German participle formation; irregular participles are rather said to be stored in memory together with the -n ending. Hence, it is only regularizations that involve rule violations. In such cases, we found a characteristic ERP response in the three participle experiments and the plural study [66]: the LAN occurred when the regular participle ending -t and the plural -s were applied incorrectly, but not in cases of irregularizations. Thus, the LAN found under these conditions can be interpreted as reflecting processes involved in (morphological) structure building.

5.2. A comparison with left anterior temporal negativities from other ERP studies

In previous studies, several kinds of morpho-syntactic anomalies have been claimed to elicit a LAN, e.g. violations of syntactic agreement, phrase-structure violations, illegal filler-gap constructions and certain kinds of grammatically well formed embedded clauses. Some researchers have interpreted the LAN as reflecting working memory operations ([6,19,22]). From a linguistic perspective, however, the LAN appears to vary as a function of processes involved in morpho-syntactic structure building. The LAN occurs when rules of affixation (e.g. subject-verb agreement) and phrase-structure building are incorrectly applied and in cases of complex syntactic dependencies that require extra parsing effort, such as filler-gap constructions and re-analysis of theta role assignment. Our findings on regular and irregular inflections are compatible with this interpretation; a LAN was found for violations of regular inflectional rules, i.e., for cases that involve morphological parsing processes, and it was not associated with violations of irregular inflectional patterns, which do not involve affixation rules.

Münte et al. [33] and Adamson et al. [1] found that violations of subject-verb agreement in English, e.g. *you goes and *your spend, were associated with a more anterior and left preponderant negativity, while in semantic judgements pairs of nouns such as sky–thief compared to robber–thief produced an N400 component. The negativity in the agreement condition was shown to be distinct from the N400. Most probably, it is related to the fact that agreement violations involve parsing processes and are structurally more complex than simple noun pairs. Osterhout and Mobley [38] also found a LAN for violations of subject-verb agreement but not for incorrectly inflected reflexive pronouns, such as himself instead of herself or themselves instead of himself. They take subject-verb agreement to be more syntactic than number or gender errors on pronouns and argue that the LAN reflects this difference.

Left anterior negativities were also found in cases of phrase-structure violations, such as *Ted’s about films America [36], and German sentences such as *Der Freund wurde im besucht “The friend was in visited” [10,12]. The latter authors argue that the LAN is associated with a relatively low-level early stage of structure building.

Some ERP studies have discovered a LAN in illegal filler-gap constructions, but results are less consistent across studies than for agreement and phrase-structure violations. Kutas and Klunder [22] found a LAN for ungrammatical wh questions such as *What can’t you remember who he advised .... This negativity was interpreted as reflecting working memory operations associated with the filler-gap construction. Such constructions were claimed to put a strain on the limited capacity of the working memory system, because the system must store the filler along with intermediate material. However, results from other ERP studies on the same phenomenon have produced different results. Neville et al. [36] found a LAN for so-called subjacency violations (*What, was sketch of —, admired by the man), but not for other violations of wh movement such as *What, did the man admire Don’s sketch of — , and McKinnon and Osterhout [29] elicited a P600, but no LAN in such cases.

Summarizing, we suggest a linguistic interpretation of the results mentioned above in terms of morpho-syntactic structure building, with the LAN being elicited by violations of morpho-syntactic rules or in cases of complex syntactic dependencies. The question, however, to what extent the LAN reflects processes at the level of the working memory requires further study of linguistic and non-linguistic stimuli; we will have to leave this question open.
5.3. Which brain structures are involved in the processing of inflected words?

If we accept the negativity obtained for incorrect -t participles and incorrect -s plurals as a correlate of the processing differences between regular and irregular words, the question arises as to which brain structures might be involved in its generation. Recall that the difference waves and the maps shown in Fig. 7 suggest a left frontal source of the negativity found for regularizations, i.e., for violations of the -t affixation rule. This might in turn fuel speculations that it is generated by the same brain areas which are traditionally said to cause agrammatism. Since we cannot be sure, however, that the activity measured at some electrode site is generated by the brain tissue directly below the electrode, caution must be exerted in advancing such an interpretation. Clearly, experiments using more electrode sites, a larger number of stimuli and the employment of the emerging new localization techniques [55,43] and/or magnetoencephalographic recordings are needed to warrant further reasoning along these lines.

The question as to whether similar or different brain structures are responsible for irregular and regular inflection has also been the subject of a recent PET study [18] on the English past tense. While lying in the PET camera, nine male subjects were performing two different tasks: (i) reading aloud of visually presented stems of regular, irregular and nonce verbs; and (ii) producing past tense forms of these verbs. Using radioactively labelled water it was possible to assess rCBF differences as a function of the tasks. For this purpose, subtractions were performed of the rCBF differences between the past-tense production and the stem-form reading tasks for the three kinds of verbs. Jaeger et al. [18] found different areas of cortical activation in the regular and the irregular tasks: the left dorsolateral prefrontal cortex (Brodmann’s area 46) and the left anterior congregate gyrus (area 24) showed activation in regular and nonce word past-tense production, while the mid-temporal gyrus (area 21) and the left superior frontal gyrus (area 10) were active in the irregular past tense condition, the latter also being active in the production of nonce past tense forms. Moreover, Broca’s area (44, 45) was active in all three production tasks. Jaeger et al. compared their activation patterns with other neuropsychological data and, on the basis of that, they made claims about the localization of past-tense formation processes in specific brain structures. Broca’s area, for example, is said to process the tense feature (+ past), area 46 computes regularly inflected forms, and area 21 is responsible for the retrieval of irregular forms. They conclude that the different activation patterns found for regulars and irregulars provide evidence against single-mechanism models such as the connectionist approach and support the dual-mechanism view that regulars are processed differently from irregulars.

Evaluating these far-reaching conclusions should be the subject of a further study involving a detailed assessment of the methods and materials used for their PET scans which we will not address here. One comment, however, seems to us to be pertinent with respect to the Jaeger et al. study [18]. Whereas most of their claims are about processing aspects of inflection, particularly the computation of regulars versus the retrieval of stored irregulars and the brain structures involved in these processes, the empirical evidence Jaeger et al. provide from their PET study does not bear on rapid processes, such as past tense formation, which take place within 200–400 ms. This is due to the poor temporal resolution of PETs. What Jaeger et al.‘s rCBF patterns show instead are cortical activations while subjects produce a list of past tense forms. Leaving this criticism aside, Jaeger et al.’s data do nonetheless suggest that regular and irregular inflections are represented in cortically distinct areas, thus providing support for the view that regulars and irregulars are based on different linguistic representations. To address issues of morphological processing, however, ERP experiments such as those presented here appear to be more revealing than PET studies.

6. Conclusion

We conclude that two complementary mechanisms co-exist in how our brain processes morphologically complex words: (i) accessing full-form entries stored in memory; and (ii) a computational system that decomposes complex words into stems and affixes. This distinction is reflected in different ERP responses to (i) and (ii).

Acknowledgements

Thanks are due to Berdieke Wieringa, Marta Kutas, Gary Marcus, Sönke Johannes, Mike Matzke, and Jonathan King for comments, Jobst Kilian for excellent technical assistance and Sam Featherston for correcting our English. T.F.M. was supported by a grant from the Hermann and Lilly Schilling Foundation (TS 013/177/96), and H.C. by a German Science Foundation grant (SFB 282/Project B7).

References

[47] S.E. Petersen, J.A. Fiez, The processing of single words studied with


