Neural dynamics of reading morphologically complex words

Johanna Vartiainen a,b,⁎, Silvia Aggjaroc,d, Minna Lehtonen a,e,f, Annika Hultén a,g, Matti Laine g, Riitta Salmelin a

a Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology, Finland
b Advanced Magnetic Imaging Centre, Helsinki University of Technology, Finland
c Department of Psychology, University of Milano-Bicocca, Italy
d Neuropsychology Unit, Villa Beretta Rehabilitation Center, Valduce Hospital, Costamasnaga, Italy
e Cognitive Neuroscience of Language Laboratory, Department of Linguistics, University of Maryland, USA
f Helsinki Collegium for Advanced Studies, University of Helsinki, Finland
g Department of Psychology, Åbo Akademi University, Finland

A R T I C L E   I N F O

Article history:
Received 11 February 2009
Revised 27 May 2009
Accepted 1 June 2009
Available online 8 June 2009

Keywords:
Decomposition
Inflected nouns
Language
MEG

A B S T R A C T

Despite considerable research interest, it is still an open issue as to how morphologically complex words such as “car + s” are represented and processed in the brain. We studied the neural correlates of the processing of inflected nouns in the morphologically rich Finnish language. Previous behavioral studies in Finnish have yielded a robust inflectional processing cost, i.e., inflected words are harder to recognize than otherwise matched morphologically simple words. Theoretically this effect could stem either from decomposition of inflected words into a stem and a suffix at input level and/or from subsequent recombination at the semantic–syntactic level to arrive at an interpretation of the word. To shed light on this issue, we used magnetoencephalography to reveal the time course and localization of neural effects of morphological structure and frequency of written words. Ten subjects silently read high- and low-frequency Finnish words in inflected and monomorphemic form. Morphological complexity was accompanied by stronger and longer-lasting activation of the left superior temporal cortex from 200 ms onwards. Earlier effects of morphology were not found, supporting the view that the well-established behavioral processing cost for inflected words stems from the semantic–syntactic level rather than from early decomposition. Since the effect of morphology was detected throughout the range of word frequencies employed, the majority of inflected Finnish words appears to be represented in decomposed form and only very high-frequency inflected words may acquire full-form representations.

© 2009 Elsevier Inc. All rights reserved.

Introduction

A fundamental feature of many languages is the use of complex word forms to convey semantic and syntactic information. An example from English would be, e.g., the word form “book + s” while in a morphologically very rich language like Finnish, word forms such as “koulu + i + ssa + mme + kin” (even in our schools’) are possible. Although this topic has received considerable research interest, it is still an open issue as to how such morphologically complex word forms are represented and processed in the brain. At first look, one might assume that they are either stored as a whole in the mental lexicon (e.g., having a separate lexical entry for books) or represented and accessed via their constituent morphemes (e.g., decomposing the input into book and -s, the representations of which are then recomposed to arrive at the meaning of the word form). These were in fact the two opposing views that were put forth in the early literature (e.g., Butterworth, 1983; Taft and Foster, 1975), while many recent models on morphological processing combine features from both models (Caramazza et al., 1988; Schreuder and Baayen, 1995). There is also evidence that an interplay between factors such as the frequency of a multimorphemic word (Alegre and Gordon, 1999), the type of morphological operation (inflection vs. derivation: Miceli and Caramazza, 1988), inflectional complexity of a language (Lehtonen et al., 2006a), and the language background of a speaker (Portin et al., 2008) determine whether a multimorphemic word is stored and recognized as a full form or is decomposed.

One would expect to find most consistent evidence for morphological decomposition in languages that are morphologically rich (Hankamer, 1989). A good example is Finnish, a non-Indo-European language with particularly rich inflectional morphology that, for example, yields about 2000 possible forms for each noun (Karlsén and Koskenniemi, 1985). Indeed, behavioral evidence consistently indicates a processing cost (longer reaction times and/or higher error rates) for inflected Finnish nouns when compared with otherwise matched monomorphemic nouns. This inflectional processing cost has been observed in visual lexical decision (e.g., Laine et al., 1999;
Niemi et al., 1994), progressive demasking where the exposure time to a word is gradually increased (Laine et al., 1999), eye movement patterns during reading (Hyönä et al., 1995), and reading errors in aphasia (Laine et al., 1995). This effect is modulated by the frequency of the word form so that it may vanish in the (very) high-frequency range (Laine et al., 1995; Lehtonen and Laine, 2003; Soveri et al., 2007), suggesting that massive exposure to an inflected form may result in a full-form representation for that form.

Theoretically, where is the inflectional processing cost thought to originate? Following general mental architectures proposed for lexical processing, one can divide the morpheme-based recognition route into the word form level and the lemma level (e.g., Laine et al., 1994). It has been argued that at the first level the input string is decomposed into a stem and affix(es) while at the second level the semantic–syntactic representations corresponding to the stem and the affix(es) are accessed and recombined to create a temporary representation for the meaning of the multimorphemic word. In principle, the inflectional processing cost could stem from either one or both of these levels. A behavioral study compared recognition of inflected vs. monomorphemic Finnish nouns presented in isolation or in a neutral sentence context (Hyönä et al., 2002). The sentence context was expected to influence the recombination stage, while decomposition at the word form level should be unaffected by the sentence context. The inflectional processing cost was found to vanish in sentence context, suggesting that it stemmed from the later recombination stage.

These theoretical levels of processing do not (necessarily) map directly onto distinct spatial or spatiotemporal activation patterns in the brain. Neuroimaging methods suggest their own functional landmarks in processing written words (e.g., Jobard et al., 2003; Salmelin and Helenius, 2004).

Hemodynamic studies provide high spatial accuracy but lack timing information. Functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) studies of reading have identified, most conspicuously, activation of the left occipito-temporal cortex and the fusiform gyrus (Brodmann area 37), thought to reflect analysis at word form level (cf. meta-analysis by Jobard et al., 2003). Lexical–semantic analysis in reading has been associated most consistently with activation of the left-hemisphere areas posterior middle temporal gyrus, basal temporal area, and inferior frontal gyrus (Jobard et al., 2003) and analysis of syntax with activation, again in the left hemisphere, of the middle temporal gyrus, temporal pole, posterior superior temporal sulcus, and inferior middle frontal gyrus (cf. meta-analysis by Vigneau et al., 2006). Most hemodynamic studies of morphology have focused on past tense inflection of regular and irregular verbs, typically using a verb generation task and the English language, although some experiments have also considered derived or inflected nouns and other languages. Effects were typically reported in activation of the left inferior frontal gyrus, interpreted to reflect analysis of grammatical features, and the left temporal regions, thought to denote access to the semantic representations of the stem and affix (Beretta et al., 2003; Bozic et al., 2007; Devlin et al., 2004; Jaeger et al., 1996; Sach et al., 2004; Sahin et al., 2006; Tyler et al., 2005; Vannest et al., 2005; Yokoyama et al., 2006).

Electrophysiological imaging provides another window into the neural processes of reading. Magnetoecephalography (MEG) studies, which usually report both the active areas and their time courses of activation, have identified a subset of functionally meaningful responses amongst the activation evoked by written words. The sequence proceeds from basic visual feature analysis in the midline occipital cortex at about 100 ms after stimulus onset, through letter-string analysis at the base of the left occipito-temporal cortex (fusiform gyrus) at 140–200 ms, to lexical–semantic activation in the left superior temporal cortex, reflected in a sustained activation that lasts from 200 ms until 600–800 ms (see Pykkänen and Marantz, 2003; Salmelin, 2007; for reviews). Electroencephalography (EEG) studies on reading have, for the most part, focused on the timing of brain signals measured on the scalp. They have typically found effects similar to those detected with MEG, i.e., a transient response on the occipital and posterior temporal electrodes to letter strings within 200 ms after stimulus onset, with left-hemisphere preponderance, and an effect of lexical–semantic manipulation more anteriorly from about 200 ms onwards on the central and parietal electrodes (e.g., Kutas and Hillyard, 1980; Schendan et al., 1998; for a review, see Barber and Kutas, 2007). Syntactic manipulations also influence activation of the left superior temporal cortex from 200 ms onwards (Barber and Carreiras, 2003; Service et al., 2007).

Given the nature of the research question (Do inflected words require additional processing early or late in the assumed sequence of cognitive operations?), time-sensitive brain imaging methods are particularly appropriate tools. EEG recordings have typically indicated effects of morphology at ∼250–700 ms after the word onset in Finnish, English, German, Spanish and Catalan (Dominguez et al., 2004; Lehtonen et al., 2007; Leinonen et al., 2009; Münte et al., 1999; Penke et al., 1997; Rodríguez-Fornells et al., 2001; Weyerts et al., 1996; Weyerts et al., 1997). MEG studies have revealed activation sensitive to English morphology at ∼350 ms over the left temporal cortex, interpreted to reflect decomposition of compound words (Fiorentino and Poeppel, 2007) and competition between morphologically related words (Pykkänen et al., 2004). A priming effect for inflected verbs was also detected on the timing of this activation (Stockall and Marantz, 2006). Diverging from this more common pattern, a recent MEG study (Zweig and Pykkänen, 2009), investigating the processing of derived vs. monomorphemic English words, found an early effect of morphological complexity at 170 ms after the stimulus presentation, primarily in the right hemisphere.

The present MEG study aims to shed light on the timing and location of neural involvement in processing inflected words. It builds on the two earlier studies that have sought to clarify the neural basis of Finnish morphology (EEG: Lehtonen et al., 2007; fMRI: Lehtonen et al., 2006b). Those studies used a lexical decision task on inflected vs. uninflected words. With fMRI, effects of morphology, i.e., stronger activation for inflected vs. monomorphemic words, were found in the left posterior superior temporal sulcus and left inferior frontal gyrus. The EEG study showed effects of morphology from 450 ms onwards: the N400 response, as well as a later positive component, was larger for inflected than monomorphemic words. Independent information on location (fMRI) and timing (EEG) suggested that effects of morphology in lexical decision stemmed from the semantic–syntactic processing stage. However, it remained unclear (i) whether the EEG effect in the specific time window and the fMRI activations in the specific brain areas reflected the same underlying neural process and (ii) if the decision process (Is the letter string a word or not?) influenced the observed activation pattern. In the present study, we used MEG and the same low- and high-frequency inflected vs. monomorphemic words that had been employed in the earlier EEG study to identify both the neural sources and their time courses of activation when the subjects read the words silently, knowing that they could be requested to read the word out loud. Source-level analysis both in individual subjects and at the group level provided a sensitive means to identify the point in the sequence of neural processing at which the effects of morphological complexity emerged, and to study the possible effect of word frequency.

Materials and methods

Subjects

Ten right-handed, Finnish-speaking subjects (5 females and 5 males; age 25–46 years, mean 30 years) participated in the experiment. They had normal or corrected-to-normal (3 subjects) vision. All participants gave their informed consent, in agreement with the prior approval of the Helsinki and Uusimaa Ethics Committee.
Partitive (e.g., "mately 75%) had a primarily WordMill Lexical Search program (Laine and Virtanen, 1999). The newspaper corpus with 22.7 million word tokens by using the high-frequency words 26.4 systems, Inc.) to picture display by the projector (VistaPro less than 6°). The MEG response timing was corrected for the 34-ms screen placed at a distance of 1 m in front of the subject (visual angle details on the stimulus materials.

The in locative cases (inessive, elative, illative, adessive, ablative and allative; rately for the low- and high-frequency words (see Table 1). The same lists of nouns comparable length, the frequency range was allowed to be broader large enough number of high-frequency monomorphemic items of tax+partitive case marker). In addition, there were six... 

Word stimuli were obtained from an unpublished Turun Sanomat newspaper corpus with 22.7 million word tokens by using the WordMill Lexical Search program (Laine and Virtanen, 1999). The low-frequency words had 0.04–4.23 occurrences per million and the high-frequency words 26.4–504 occurrences per million (see Table 1).

The inflected words were bimorphemic nouns that included nine different case suffixes. The majority of the inflected words (approximately 75%) had a primarily “syntactic” case suffix, genitive or partitive (e.g., “huivi + n” — scarf + genitive case marker; “vero + a” — tax + partitive case marker). In addition, there were six “semantic,” locative cases (inessive, elative, illative, adessive, ablative and allative; e.g., inessive “kassi + ssa” — in the bag) and a related case marker, esse, that typically denotes a role (e.g., “tyttö + n” — as a girl).

The inflected and monomorphemic words were matched according to the average word length in letters, lemma frequency, surface frequency, bigram frequency, and morphological family size, separately for the low- and high-frequency words (see Table 1). The experiment used inflected words, thus including at least six letters. However, high-frequency words are typically short. In order to find a large enough number of high-frequency monomorphemic items of comparable length, the frequency range was allowed to be broader than for the low-frequency range. Nevertheless, the high and low-frequency ranges remained distinctly separate. The same lists of nouns were used in a recent EEG study; see Lehtonen et al. (2007) for further details on the stimulus materials.

The words were presented visually one at a time, in capital letters, black Helvetica font on a white background, on a back-projection screen placed at a distance of 1 m in front of the subject (visual angle less than 6°). The MEG response timing was corrected for the 34-ms delay from stimulus trigger (Presentation software, Neurobehavioral Systems, Inc.) to picture display by the projector (VistaPro™, Electrohome Limited). Each noun was shown for 400 ms, and the stimulus onset asynchrony was 3000 ms. Stimuli from the different categories were presented in a random order and divided into 4 blocks (5 min each), with a short rest period in-between. The MEG recording lasted for about 25 min in total.

Subjects were instructed to read the words silently. To maintain the subjects’ attention on the task 20 additional target words were presented randomly among the actual 320 word stimuli. In these trials (6% of the presented items) the word was followed by a question mark (1500 ms after the word onset, duration 1500 ms) that prompted the subject to read out loud the preceding word. The next trial started after a delay (blank screen) of 1500 ms. The target words consisted of low- and high-frequency inflected and monomorphic words, five words per category, and they were matched similarly to the actual stimuli. The target trials as well as the trials immediately following the question marks were not included in the analysis.

MEG recording

Magnetic signals were recorded in a magnetically shielded room with a Vectorview™ whole-head MEG system (Elekta Neurolog Ltd., Helsinki, Finland). The device employs 102 triple sensor elements composed of two orthogonal planar gradiometers and one magnetometer. Planar gradiometers detect the maximum signal directly above a local current source. MEG signals were band-pass filtered between 0.03 and 200 Hz and digitized at 600 Hz. Vertical and horizontal electro-oculograms (EOG) were also recorded.

The MEG signals were averaged off-line across trials from 0.2 s before to 1 s after the stimulus onset. Averaged MEG responses were baseline corrected to the 0.2 s interval immediately preceding the stimulus onset and low-pass filtered at 40 Hz. Trials with MEG amplitude exceeding 3000 fT/cm were discarded automatically. Epochs contaminated by eye movements and blinks were rejected from the data by discarding trials during which EOG signal amplitude exceeded 150 μV. On average 67/80 (minimum 43/80) artifact-free responses were obtained for all subjects and for each word category.

Sensor-level analysis

For an initial overview of the data at the level of MEG sensors, areal mean signals (AMS) were calculated over ten regions: left and right frontal, left and right temporal, left and right rolandic, parietal, left and right occipito-temporal, and occipital area. For each region, 8–12 sensor pairs were included. Vector sums of each gradiometer pair were first obtained by squaring the MEG signals, summing them together, and calculating the square root of this sum. The areal mean signals were computed by averaging these vector sums for each region, individually for each subject. Finally, the areal mean signals were averaged across subjects. Because of the way the sensor-level areal mean signals are calculated (square root of sum of squared signals) they always have a positive value.

The AMS were characterized by measuring the mean signal strength in three time windows: 50–170 ms, 170–330 ms, and 330–500 ms after the word onset, based on the appearance of the grand mean AMS. The first time window contained the responses over the posterior regions at ~100–150 ms. The second time window contained the response at ~220 ms mainly visible over the right occipito-temporal region, and the third time window captured the sustained response peaking at ~400 ms over the left temporal region. The signal strengths were collected from individual subjects, separately for each stimulus category, and tested using a repeated-measures analysis of variance (ANOVA) in each area, with Word frequency (high- vs. low-frequency words) and Morphological

<table>
<thead>
<tr>
<th>Word category</th>
<th>Example</th>
<th>Translation</th>
<th>Word length in letters (SD)</th>
<th>Surface frequency (SD)</th>
<th>Lemma frequency (SD)</th>
<th>Bigram frequency (SD)</th>
<th>Morphological family size (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HighMono</td>
<td>ongelma</td>
<td>“problem”</td>
<td>6.0 (1.3)</td>
<td>84.4 (79)</td>
<td>399 (732)</td>
<td>1137 (374)</td>
<td>783 (668)</td>
</tr>
<tr>
<td>HighInf</td>
<td>aamu + îa</td>
<td>“morning” + adessive case = in the morning</td>
<td>6.1 (1.3)</td>
<td>83.8 (66)</td>
<td>423 (450)</td>
<td>1193 (458)</td>
<td>972 (663)</td>
</tr>
<tr>
<td>LowMono</td>
<td>kainalo</td>
<td>“armpit”</td>
<td>6.3 (1.2)</td>
<td>0.9 (0.9)</td>
<td>5.4 (5.1)</td>
<td>1161 (403)</td>
<td>53 (42)</td>
</tr>
<tr>
<td>LowInf</td>
<td>hiba + ssa</td>
<td>“sleeve” + inessive case = in the sleeve</td>
<td>6.4 (1.2)</td>
<td>0.8 (0.8)</td>
<td>5.5 (1.7)</td>
<td>1150 (487)</td>
<td>68 (67)</td>
</tr>
</tbody>
</table>

Surface and lemma frequencies are reported as frequencies per million.

Stimuli and experimental design

The stimuli consisted of low- and high-frequency Finnish nouns in the monomorphic and inflected form. The 320 stimuli were divided into four groups of 80 words (see Table 1): high-frequency monomorphic words (HighMono), high-frequency inflected words (HighInf), low-frequency monomorphic words (LowMono), and low-frequency inflected words (LowInf).

Table 1

<table>
<thead>
<tr>
<th>Word category</th>
<th>Example</th>
<th>Translation</th>
<th>Word length in letters (SD)</th>
<th>Surface frequency (SD)</th>
<th>Lemma frequency (SD)</th>
<th>Bigram frequency (SD)</th>
<th>Morphological family size (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HighMono</td>
<td>ongelma</td>
<td>“problem”</td>
<td>6.0 (1.3)</td>
<td>84.4 (79)</td>
<td>399 (732)</td>
<td>1137 (374)</td>
<td>783 (668)</td>
</tr>
<tr>
<td>HighInf</td>
<td>aamu + îa</td>
<td>“morning” + adessive case = in the morning</td>
<td>6.1 (1.3)</td>
<td>83.8 (66)</td>
<td>423 (450)</td>
<td>1193 (458)</td>
<td>972 (663)</td>
</tr>
<tr>
<td>LowMono</td>
<td>kainalo</td>
<td>“armpit”</td>
<td>6.3 (1.2)</td>
<td>0.9 (0.9)</td>
<td>5.4 (5.1)</td>
<td>1161 (403)</td>
<td>53 (42)</td>
</tr>
<tr>
<td>LowInf</td>
<td>hiba + ssa</td>
<td>“sleeve” + inessive case = in the sleeve</td>
<td>6.4 (1.2)</td>
<td>0.8 (0.8)</td>
<td>5.5 (1.7)</td>
<td>1150 (487)</td>
<td>68 (67)</td>
</tr>
</tbody>
</table>

Surface and lemma frequencies are reported as frequencies per million.
complexity (inflected vs. monomorphic words) as the within-subject factors.

Source analysis

An Equivalent Current Dipole (ECD) is an estimation of the center of an active cortical area, and the amplitude and orientation of electric current within that area (Hämäläinen et al., 1993). ECDs were determined separately for each subject, following analysis procedures used in previous MEG studies of reading (e.g., Helenius et al., 1998; Salmelin et al., 1996; Tarkkainen et al., 1999). The ECDs were identified one by one, at time points where each distinct magnetic field pattern was most salient. The time course of activation in each source area (source waveform) was estimated by introducing all ECDs of one individual simultaneously into a multidipole model. The ECD locations and orientations were kept fixed while their amplitudes were allowed to vary to best explain the measured MEG signals. The resulting multidipole models accounted for at least 77% of the total magnetic field variance at the activation peak in each condition. The final models were composed of 6–11 ECDs (mean 10).

The ECD locations were defined in the head coordinate system that was set by the nasion and two reference points anterior to the ear canals. Prior to the MEG recording, four Head Position Indicator (HPI) coils were attached to the subject’s head and their locations were measured with a 3D digitizer (Polhemus, Colchester, VT, USA). At the beginning of the MEG recording, the HPI coils were briefly energized to determine their location with respect to the MEG helmet. For visualization and comparison of the sources between subjects the ECDs were transformed to a standard brain (Roland and Zilles, 1996) using an elastic transformation (Schormann et al., 1996; Woods et al., 1998).

When sources from individual subjects clustered in the same general area, with similar time courses of activation, group-level statistical analysis was performed to detect possible differences between stimulus categories. The ECDs of individual subjects were classified according to location (nine areas: occipital cortex, left-hemisphere occipito-temporal, temporal, frontal, and parietal cortex, and right-hemisphere occipito-temporal, temporal, frontal, and parietal cortex) and the time at which the maximum activation was reached (peak latency; three time windows: 70–170 ms, 170–300 ms, and from 300 ms onwards). The clusters containing an ECD from at least 6 subjects were accepted for further statistical analysis. Transient activation, occurring typically within 300 ms after stimulus onset, was characterized by the maximum activation strength and the peak latency. Sustained activation was characterized by its maximum activation strength, mean strength from 200 to 800 ms, and its shape as a function of time: the peak latency and the points when the activation had reached/reduced to 50% of the maximum in the ascending and descending slopes (onset and offset latencies). All of these measures were collected from individual subjects, separately for each stimulus category, and tested using a repeated-measures analysis of variance (ANOVA) in each area, with Word frequency (high- vs. low-frequency words) and Morphological complexity (inflected vs. monomorphic words) as the within-subject factors.

We additionally estimated significant effects in individual subjects within the first 250 ms in order to facilitate detection of stimulus effects that might be too variable in location to be detected in group-level analysis. For each identified source area, the responses to inflected words were compared with responses to monomorphic words, separately for low- and high-frequency words. As a conservative test for significance, the difference between the two source waveforms of the same polarity had to exceed 2.58 times the standard deviation (SD) of the source waveforms during the 200-ms pre-stimulus baseline interval for at least 10 ms (corresponding to \( p < 0.01 \)). The SD was estimated from the standard deviations of the two source waveforms (SD1, SD2) as \( SD = \sqrt{(SD1^2 + SD2^2)} \) (cf. Tarkkainen et al., 1999).

Results

Areal mean signals

Fig. 1 shows the areal mean signals calculated over ten regions and averaged across all subjects for each of the four stimulus types. Within 200 ms after stimulus onset, signals were strongest over the occipital and occipito-temporal areas. A strong sustained response over the left temporal lobe reached the maximum at ~400 ms and showed unequal signal strength to the four stimulus categories.

An effect of morphology emerged over the left temporal and parietal regions. The response was stronger to inflected than to monomorphic words at 330–500 ms [left temporal region \( F(1,9) = 5.6, p < 0.05 \); parietal region \( F(1,9) = 5.1, p < 0.05 \)]. Effects of Frequency were detected over the same areas. The response was stronger to the low- than high-frequency words at 330–500 ms over the left temporal region \( [F(1,9) = 6.6, p < 0.05] \) and at 170–330 ms over the parietal region \( [F(1,9) = 12.8, p < 0.01] \). There were no significant effects in the 50–170 ms time window over any region.

Field patterns and source analysis

Fig. 2 displays the group-level sequence of activation elicited by the different word types. ECDs from individual subjects, each representing the center of an active cortical region, were grouped according to similarity in location and time course of activation. Occipital activation at ~100 ms was detected in all subjects (peak latency 119 ± 8 ms, mean ± standard error of mean). Thereafter, left occipito-temporal activation at ~150 ms (peak latency 151 ± 5 ms), with current flow oriented mostly dorsally, was detected in six subjects. Sources in the right occipito-temporal cortex, typically with horizontal orientation of current flow, were active at ~220 ms (peak latency 222 ± 9 ms, seven subjects). Sustained activation of the left superior temporal cortex at ~200–800 ms (peak latency 400 ± 11 ms), with the current flow perpendicular to the course of the Sylvian fissure, was detected in all ten subjects; in each subject, a single ECD sufficed to explain the signal over the entire time window. This was the only brain area where differences appeared between stimulus categories in the group-mean time course of activation (see Fig. 2). All subjects showed sustained...
activation in the right hemisphere as well, but the locus was markedly less consistent than in the left hemisphere, with the sources either in the temporal or frontal cortex.

Statistical tests on the source waveforms

Significant group-level effects emerged in the sustained left temporal activation only. As illustrated in Fig. 3, the activation was stronger for inflected than monomorphemic words [maximum activation: Morphological complexity $F(1,9)=16.6, p<0.005$, mean activation strength 200–800 ms: Morphological complexity $F(1,9)=25.9, p<0.005$] and stronger for the low- than for the high-frequency words [maximum activation: Word frequency $F(1,9)=9.9, p<0.05$, mean activation strength 200–800 ms: Word frequency $F(1,9)=8.2, p<0.05$].

A main effect of morphology was observed also in the duration of the response: the activation lasted longer for the inflected than for the monomorphemic words [offset latency: Morphological complexity $F(1,9)=5.5, p<0.05$]. The activation seemed to be strongest and of longest duration for the inflected low-frequency words and to be weakest and of shortest duration to the high-frequency monomorphemic words (high-frequency words: inflected-monomorphemic $\sim 10$ ms, low-frequency words: inflected-monomorphemic $\sim 35$ ms). Nevertheless, no significant interaction was found between Morphological complexity and Word frequency in any of the measures.

Similarly to the areal mean signals, there were no salient effects of morphology at the source level within the first 250 ms after the word onset either in the group-level analysis or when the data were examined at the level of individual subjects.
Discussion

The present study investigated neural processing of inflected written nouns in a morphologically rich language, Finnish, where chances of finding morphology-related effects should be particularly high. The aim was to identify the point in the sequence of neural processing at which the effects of morphological complexity emerged, and to study the possible effect of word frequency. The subjects read the words silently which was thought to probe the reading system in a more natural fashion than, e.g., the frequently employed lexical decision task. The cortical sequence of activation followed the usual pattern elicited by written words (see Pykkänen and Marantz, 2003; Salmelin, 2007 for reviews). The sensor-level data and the source analysis showed a consistent effect of morphological complexity in the sustained activity of the left superior temporal cortex from 200 ms onwards. The neural response was stronger and longer-lasting to the inflected than to the monomorphic words, suggesting decomposition of all the inflected words throughout the frequency range used in the study. Furthermore, the activation was stronger to low- than to high-frequency words, irrespective of morphological structure. No effects of morphology were found within the first 200 ms after the word onset.

The location (left superior temporal cortex) and timing (at ~200–800 ms) of the morphological effect found in the present MEG study are in general agreement with the previous EEG and fMRI neuroimaging studies on visual recognition of inflected vs. uninflected words in the Finnish language (Lehtonen et al., 2007; Lehtonen et al., 2006b; Lehtonen et al., in press; Leinonen et al., 2009). The present data together with these previous results, gathered altogether from over fifty individuals using two different reading tasks, provide compelling evidence for a late morphological effect in the temporal cortex and for the lack of early effects in the Finnish language.

Previous MEG and EEG studies of reading have associated activation in the ~200–800 ms time window in the left superior temporal cortex (“N400/N400m”) with lexical–semantic analysis (e.g., Helenius et al., 1998; Kutas and Hillyard, 1980; Pykkänen et al., 2006), morphosyntactic processing (Barber and Carreiras, 2003; Service et al., 2007) and analysis of phonological structure (e.g., Helenius et al., 1998; Rugg, 1984; Wydell et al., 2003). Thus, the present effects of morphology most likely reflect more demanding analysis of morphologically complex than monomorphic words at the semantic–syntactic or phonological level rather than pre-lexical visual or orthographic processing. This interpretation is concordant with fMRI data suggesting that the effect of morphology in the left temporal cortex reflects more demanding lexical–semantic access for complex words (e.g., Lehtonen et al., 2006b; Tyler et al., 2005).

In addition to the left temporal effect, the fMRI experiments on the Finnish language (Lehtonen et al., 2006b; Lehtonen et al., in press) revealed an effect of morphology in the left inferior frontal gyrus. Haemodynamic studies on morphological processing in other languages have also implicated the left inferior frontal gyrus (e.g. Jaeger et al., 1996; Yokoyama et al., 2006). Our MEG study revealed no activation in the frontal cortex that would have been sensitive to morphological complexity; such effects were located in the left superior temporal cortex, instead. This divergent pattern between fMRI/PET vs. MEG studies seems to parallel data on semantic processing: haemodynamic experiments (see Jobard et al., 2003 for review) typically suggest inferior frontal cortex to be involved in semantic processing whereas MEG experiments rather show activation in the temporal cortex (see Salmelin, 2007 for a review). One possible explanation for this difference is that MEG and fMRI/PET are sensitive to different aspects of neural function. The MEG signal tracks rapid changes of highly synchronized neural activation, which are not necessarily accompanied by changes in blood flow, oxygenation, and glucose uptake that are measured by fMRI and PET and indirectly related to overall changes in energy consumption between two conditions, integrated over a long time of ongoing processing. Another possibility is that the left inferior frontal gyrus is engaged in more controlled, post-lexical components of morphological processing as opposed to automatic, pre-lexical processing (Gold and Rastle, 2007). Controlled processing, such as assessment of the correctness of the word form, was not required in the present reading task but could have played a role in the lexical decision task employed by Lehtonen et al. (2006b; in press). Further experiments are clearly needed to clarify the discrepancy in the localization of activation between MEG and haemodynamic methods.

In the present experiment, we found no effects of morphology within the first 200 ms after the onset of the word, in line with previous EEG and MEG data on Finnish, English, German, Spanish and Catalan languages (Domínguez et al., 2004; Fiorentino and Poeppel, 2007; Lehtonen et al., 2007; Leinonen et al., 2009; Munte et al., 1999; Penke et al., 1997; Rodríguez-Fornells et al., 2001; Weyerts et al., 1996; Weyerts et al., 1997; but see Zweig and Pykkänen, 2009). The early visual processing of written words in the occipital cortex is typically (also in the present data) followed by activation of the left occipito-temporal region at ~150 ms that has been suggested to reflect processing of letter strings, as opposed to symbol strings and other visual stimuli (Tarkkaín et al., 1999). No effect of morphology should appear on this activation as it does not differentiate even between real words and consonant strings (MEG, Cornelissen et al., 2003; intracranial recordings, Nobre et al., 1994). Theoretically, the subsequent processing stages could differentiate between inflected vs. monomorphic words.

When Finnish inflected words were investigated using fMRI (Lehtonen et al., 2006b), effect of morphological complexity approached significance in the left occipito-temporal cortex. Functional MRI experiments have associated activation of the left BA 37, a region slightly anterior to the letter-string response detected with MEG (Salmelin and Helenius, 2004), with processing of real words (Visual Word Form area, VWFA, Cohen et al., 2000). This activation differs functionally from the letter-string response since it is stronger to real words than to consonant strings (Rüdell et al., 1998; Cohen et al., 2002; Rees et al., 1999). The VWFA activation seems to reflect a later processing stage than the letter-string response detected with MEG and would, thus, be a possible candidate for activation reflecting early morphological decomposition. Indeed, in the English language, an fMRI masked priming study has indicated sensitivity to morphological priming in this region (Gold and Rastle, 2007). However, several fMRI studies have failed to replicate the increased VWFA activation to words as opposed to consonant strings (Cohen et al., 2003; Joubert et al., 2004; Mayall et al., 2001; Tagamets et al., 2000), and it has been proposed that the VWFA is a functionally non-homogenous region with posterior parts sensitive to all letter strings and more anterior parts increasingly sensitive to word-like stimuli (Vinkkier et al., 2007). As an MEG-fMRI comparison of written word perception in the same subjects is currently lacking, it remains unclear why MEG and fMRI often seem to capture different subprocesses of reading in the left occipito-temporal cortex. At the moment it can be concluded, based on the present data and the literature, that in the Finnish language reliable effects of morphology have not been detected with any imaging modality in early time windows or in the regions assumed to be involved in pre-lexical analysis of written words.

Why does the assumed early decomposition not show in the early neural responses in Finnish? One possible explanation is that, irrespective of morphological complexity, the system automatically attempts to map orthographic input representations with both the whole input string and its parts (e.g., the monomorphic word canus might not only activate the orthographic representations for canvas but also for can- and -s). The present experimental design would, thus, not yield any clear differences in the corresponding early neural responses between inflected vs. monomorphic words.
Another possibility is that the neural processing cost of decomposition is compensated for by the high frequency of the letter combinations constituting the inflectional endings that may have facilitated processing of the inflected words. One could also argue that a presumably highly automatic activation of orthographic stem + suffix representations for inflected words does not call for significantly larger information processing resources than activation of a single monomorphemic representation, or that the increased processing is too weak to be detected with the current imaging techniques. At the level of the brain, it is currently not well known how visual information of written words is mapped onto more abstract representations, and which representational levels (orthographic, morphological, and/or possibly other) are involved.

In contrast to the present results and previous EEG (Dominguez et al., 2004; Lehtonen et al., 2007; Leinonen et al., 2009; Münte et al., 1999; Penke et al., 1997; Rodriguez-Fornells et al., 2001; Weyerts et al., 1996; Weyerts et al., 1997) and MEG (Fiorentino and Poeppel, 2007) findings, a recent MEG experiment on English derived words (Zweig and Pykkänen, 2009) found an effect of morphology in the right occipito-temporal cortex at ∼ 170 ms. In the present experiment, right occipito-temporal activation was detected in 7 subjects at ∼ 220 ms but these sources did not differentiate between the inflected and monomorphemic words. Activation in this time window and region has previously been reported to written words, without sensitivity to linguistic manipulations (Cornelissen et al., 2003). It is worth noting that, despite the early neural effect of morphology, Zweig and Pykkänen found no behavioral processing cost in their contrast between morphologically complex vs. simple words. One possible source of discrepancy is that, instead of inflected words, they employed derivations which have been found to behave differently from inflections and to not necessarily show a processing cost in the Finnish language, either (see e.g., Bertram et al., 1999). It thus remains to be resolved whether the differences in the early responses are due to language, word form, type, or some other factor.

A silent reading task was used here in order to probe the reading system in as natural a fashion as possible, and to avoid processes that are not involved in normal reading. As opposed to some other tasks that have been employed in imaging studies of morphological processing, for example lexical decision or generation of inflected forms, the present task did not require assessment of the correctness of the word form, or active, conscious manipulation of the morphemes or their meanings. Since the left inferior frontal cortex has been associated with such functions (e.g. Heim et al., 2005; Gold and Rastle, 2007), the use of the silent reading task in the present study may well explain the lack of frontal effects of morphology. Although the lack of early effects of morphology in the occipito-temporal region could, in principle, result from the choice of the task as well, this is unlikely since the previous experiments that employed the lexical decision task in the Finnish language using EEG or fMRI did not find effects of morphology in early time windows (Lehtonen et al., 2007; Leinonen et al., 2009) or in the occipito-temporal cortex (Lehtonen et al., 2008b; Lehtonen et al., in press) either.

In the present experiment, behavioral measures were not collected during the MEG recording to avoid motor-related activations that would temporally overlap with the responses to written words. However, the behavioral processing cost for inflected Finnish words is a robust finding that has been detected with several tasks including visual lexical decision (e.g., Laine et al., 1999; Niemi et al., 1994; Leinonen et al., 2009), progressive demasking (Laine et al., 1999), eye movement patterns during reading (Hyönä et al., 1995), and reading errors in aphasia (Laine et al., 1995). For the present stimulus set, the behavioral effect was verified in an earlier EEG experiment using a lexical decision task (Lehtonen et al., 2007). Therefore, it seems likely that a behavioral processing cost would be elicited in the present silent reading task as well if it could be measured in this type of task.

Theoretically, the present results are in line with those models that have suggested recognition of inflected words via decomposition (e.g., Taft, 2004), although it should be noted that this experiment was not designed to test any particular model. Recent behavioral and fMRI results on masked priming have provided compelling evidence for early decomposition preceding access to whole word representations and semantic information (Gold and Rastle, 2007; Longtin and Meunier, 2005; Longtin et al., 2003; Rastle et al., 2004). The stronger and longer-lasting responses to inflected words detected in the present experiment suggest that these words were recognized via decomposition route as opposed to monomorphemic words that presumably were recognized in full form. However, as discussed above, an early decomposition stage was not reflected in the MEG responses as increased activation to inflected vs. monomorphemic words. After the early decomposition stage, processing of an inflected word is assumed to continue by retrieval of the stem and affix meanings that then need to be combined to form a temporary representation of the meaning of the whole word form (e.g., Laine et al., 1994). Access to the meaning of the stem and affix separately, as opposed to meaning of the whole word alone, may require additional neural processing. Furthermore, semantic access to the stem may occur prior to access to the suffix (Laine, 1999; Lehtonen et al., 2007), causing some delay in the processing of inflected words. The combination of the meanings of the stem and affix into a temporary representation, and the possible check-up for the morphosyntactic legality of the stem and affix combination (coined licensing by Schreuder and Baayen, 1995), may additionally burden and delay the neural processing within the decomposition route. These assumed processing costs may explain the increased and lengthened neural responses to morphologically complex words detected in the present study.

Finally, earlier behavioral studies have suggested that morphological decomposition occurs for low- but not for high-frequency words (Laine et al., 1995; Lehtonen and Laine, 2003). In the present MEG data, the responses were strongest and of longest duration to the low-frequency inflected words. However, there was no statistically significant interaction between morphology and word frequency, which implies that morphological decomposition occurred for inflected words throughout the frequency range employed. This result is in line with the behavioral data in the earlier EEG study that used the same stimulus words in a lexical decision task (Lehtonen et al., 2007). At first glance, the present results seem most compatible with full-parsing accounts that assume decomposition of all inflected words (Stockall and Marantz, 2006; Taft and Foster, 1975). However, a model based on the morphologically rich Finnish language (Laine et al., 1994; Niemi et al., 1994) and recent behavioral findings (Soveri et al., 2007) suggest that only those inflected words that are of very high frequency in the Finnish language may acquire full-form representations. It remains to be seen whether such very high-frequency inflected words would elicit neural responses similar to monomorphemic items.

In conclusion, the present MEG data, together with previous EEG and fMRI results, provide strong evidence for morphological effects in the left temporal cortex beginning at ∼ 200 ms after the word onset and lack of early morphological effects in the morphologically rich Finnish language. This suggests that the inflectional processing cost stems from the semantic–syntactic level. Furthermore, the present results indicate across-the-board decomposition of inflected words in a morphologically rich language. However, earlier behavioral studies show that very high-frequency inflected words that were not tested here may be an exception to this rule.

Acknowledgments

This study was financially supported by the Finnish Cultural Foundation, the Academy of Finland (Centre of Excellence Program.


