Beyond the initial 140 ms, lexical decision and reading aloud are different tasks: An ERP study with topographic analysis

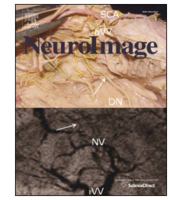
Mahe, Gwendoline; Zesiger, Pascal Eric; Laganaro, Marina

NeuroImage xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



Beyond the initial 140 ms, lexical decision and reading aloud are different tasks: An ERP study with topographic analysis

Gwendoline Mahé *, Pascal Zesiger, Marina Laganaro

FPSE, University of Geneva, 40 Bd Pont d'Arve, CH-1211 Genève 4, Switzerland

ARTICLE INFO

Article history:
Received 28 December 2014
Accepted 28 July 2015
Available online xxxx

Editor: Sonja Kotz

Keywords:
Lexical decision task
Reading aloud task
Event Related Potential

ABSTRACT

Most of our knowledge on the time-course of the mechanisms involved in reading derived from electrophysiological studies is based on lexical decision tasks. By contrast, very few ERP studies investigated the processes involved in reading aloud. It has been suggested that the lexical decision task provides a good index of the processes occurring during reading aloud, with only late processing differences related to task response modalities. However, some behavioral studies reported different sensitivity to psycholinguistic factors between the two tasks, suggesting that print processing could differ at earlier processing stages. The aim of the present study was thus to carry out an ERP comparison between lexical decision and reading aloud in order to determine when print processing differs between these two tasks. Twenty native French speakers performed a lexical decision task and a reading aloud task with the same written stimuli. Results revealed different electrophysiological patterns on both waveform amplitudes and global topography between lexical decision and reading aloud from about 140 ms after stimulus presentation for both words and pseudowords, i.e., as early as the N170 component. These results suggest that only very early, low-level visual processes are common to the two tasks which differ in core processes. Taken together, our main finding questions the use of the lexical decision task as an appropriate paradigm to investigate reading processes and warns against generalizing its results to word reading.

1. Introduction

Investigation of the cognitive processes involved in reading has become of particular interest over the last decades. The lexical decision task is one of the most common paradigms used to investigate print processing in behavioral (González-Nosti et al., 2014), functional magnetic resonance imaging (fMRI; Fiebach et al., 2002), and electrophysiological studies (Bentin et al., 1999; Simon et al., 2004). During the lexical decision task, participants are presented with letter strings and required to determine as quickly and accurately as possible whether they correspond to real words or not by pressing two key buttons. The time taken to make a lexical decision to a letter string is considered as an index of the operations needed to access the lexical representation. Overall, it is admitted that the lexical decision task and the reading aloud task rely on similar processes, with only late processing differences linked to planning and execution of different response modalities (oral versus button press; Carreiras et al., 2007; Grainger and Jacobs, 1996). The lexical decision task could thus be taken as an index of the processes occurring during reading aloud.

However, some behavioral studies reported differences between the two tasks in core linguistic processes. Indeed, different predictors have been associated with the performance of each task: lexical frequency in lexical decision and the first phoneme type in reading aloud (Ferrand et al., 2011). Other results suggest a marked sensitivity of the lexical decision task to lexical and semantic factors compared to the reading aloud task, with stronger lexical frequency effects (Balota and Chumbley, 1984; Balota et al., 2004; Schilling et al., 1998) and semantic effects (Balota et al., 2004; Yap et al., 2012). In contrast, the reading aloud task appears to be more strongly related to assembled phonological processing, with a strong relationship between reading aloud and phonological decoding ability and orthographic regularity effects limited to reading aloud and not found in lexical decision (Katz et al., 2012).

These task differences reported in behavioral studies have been supported by fMRI data, with a stronger activation of the Visual Word Form Area (VWFA) in lexical decision compared to reading aloud (Katz et al., 2005) and a stronger involvement of the inferior frontal gyrus in reading aloud compared to lexical decision (Katz et al., 2005; Valdois et al., 2006). Taken together, both behavioral and fMRI results suggest that the lexical decision task promotes larger orthographic unit processing whereas reading aloud promotes the generation of phonology. This distinction has been highlighted by Balota and Yap (2006), who postulated that print processing would be modulated through attentional control systems according to processes which are relevant to the response modality of specific tasks.

Together, the set of differences among tasks summarized above suggest differences in core processes between lexical decision and reading aloud which could not be limited to processes linked to the response mode. As the lexical decision task is the most used paradigm used in reading studies, it is essential to understand exactly at which processing

http://dx.doi.org/10.1016/j.neuroimage.2015.07.080
1053-8119/© 2015 Published by Elsevier Inc.

stages print processing differs between lexical decision and reading aloud. Without a clear understanding of task specificities and given that deciding whether a letter string corresponds or not to a known word does not reflect reading in real life, any conclusion taken from lexical decision data relative to reading abilities could be mistaken. The abovementioned studies suggest differences in orthographic, phonological and semantic processes between the two tasks. However, it should be noted that most of the studies comparing lexical decision and reading aloud are based on behavioral or fMRI studies. A more precise insight into the exact processing stages which differ between lexical decision and reading aloud could only be obtained with high temporal resolution techniques such as electrophysiological recordings. Event Related Potential (ERP) measurements have indeed allowed to determine online the time course of print processing. Overall, it has been reported in a large variety of tasks including visual orthographic or rhyme judgments, lexical or semantic decisions and silent reading, that orthographic and phonological processes respectively peak at about 200 ms and 300 ms (Bentin et al., 1999; Simon et al., 2004), followed by semantic processes at about 400 ms (Kutas and Hillyard, 1980; Simon et al., 2004). Some insights concerning the time course of the lexicality effect have also been provided with ERP data. Results from such studies have been somewhat inconsistent, with some experiments reporting a lexicality effect as early as around 150–200 ms following stimulus presentation (Dujardin et al., 2011; Hauk et al., 2012; Kim and Straková, 2012; Mahé et al., 2012, 2013; Shaul et al., 2012; but see Simon et al., 2007), while other experiments have reported later effects around 300–400 ms (Bentin et al., 1999; Pylkkännen and Marantz, 2003; Wydell et al., 2003).

It should be noted that electrophysiological studies using reading aloud tasks are lacking, with only a few EEG (Bentin et al., 1999; Simon et al., 2004) and MEG studies involving (silent) reading tasks (Chen et al., 2013; Cornelissen et al., 2003, 2009; Wydell et al., 2003). To our knowledge, only two electrophysiological studies investigated processing differences between lexical decision and reading but using either delayed (Yum et al., 2014) or silent reading (Chen et al., 2013). One of those studies was run in Chinese and compared regularity and consistency effects between lexical decision and delayed word reading (Yum et al., 2014). Results revealed that both effects were limited to the delayed word reading task and lacking in the lexical decision task, suggesting that phonological information would not be automatically accessed during character recognition in the lexical decision task. A second recent study has investigated word processing differences in English between lexical decision, semantic decision and silent reading using combined EEG–MEG and fMRI measurements (Chen et al., 2013). MEG data revealed significant word processing differences between lexical decision and silent reading as early as 150 ms following stimulus presentation, with stronger activation in areas involved in orthographic and semantic processes in lexical decision (i.e., left inferior temporal cortex and bilateral anterior temporal lobe) and stronger activation in areas involved in early phonological retrieval in silent reading (i.e., left precentral areas). This study allowed the identification of task differences early in the time course of print processing. However, it should be noted that this study used a silent reading task and that if this task is successful in avoiding potential speech movement related artifacts, one cannot be sure whether participants follow task instructions, since answers are only requested in a limited number of trials. In addition, the use of a silent reading task resulted in between task differences in the experimental design, with participants being requested to give a response to each trial in lexical decision and only to a limited number of trials in reading. Finally, the two tasks differed in the material, with an inclusion of pseudowords limited to the lexical decision task. This may affect task comparison, as performance in the lexical decision task has been shown to be especially affected by the difficulty of the word–pseudoword discrimination (Lupker and Pexman, 2010; Stone and van Orden, 1993). In order to determine which processing stages differ between lexical decision and reading aloud, it is absolutely necessary to compare the two tasks using strictly the same participants, material and experimental design. The aim of the present study was thus to determine exactly which processing stages differ between lexical decision and reading aloud when the same participants, material and design are involved in the two tasks. In addition to the classical amplitude waveform analysis, topographic analyses were performed in order to identify which periods of stable global electric fields at scalp, functional microstates reflecting particular periods in information processing, differ across tasks. If task differences are limited to late processing stages related to decisional and speech planning processes, ERP waveforms and topographic maps should only differ at late time intervals. In contrast, if the lexical decision task and the reading aloud task also differ in orthographic, phonological and/or lexical–semantic processes, different ERPs should be observed in time windows associated with these processes (i.e., respectively around 200, 300 and 400 ms; Bentin et al., 1999; Simon et al., 2004). Finally, as words and pseudowords were included in both lexical decision and reading aloud, task processing differences could be established for the two kinds of written stimuli.

2. Method

2.1. Participants

20 native French speakers (seven men), aged 20–35 years (mean 25.7 years) took part in the experiment. They were all right-handed as determined by the Edinburgh Handedness Scales (Oldfield, 1971; mean lateralization quotient index range: 96%; range: 80–100%). They were undergraduate students and reported having normal or corrected-to-normal vision and did not suffer from any neurological or motor problem. All participants gave their written informed consent to participate in the study and were paid for their participation. The study protocol was approved by the local research ethical committee at the University of Geneva.

2.2. Material

The same material was used for both the lexical decision task and the reading aloud task. A total of 120 mono- and bisyllabic words were selected from the French lexical database Lexique 3 (New et al., 2001). All words were four-to-six letter long with an average print lexical frequency of 89.4 per million. 100 orthographically legal and pronounceable pseudowords were created by changing at least two letters (which were never the first letter) in the set of words. Words and pseudowords were matched on a set of pertinent variables (see Table 1). Pseudowords were created for the requirement of the lexical decision task but were also displayed in the reading aloud task in order to compare the two tasks with exactly the same stimulus presentation. All stimulus material is provided in Appendix 1.

Table 1
Characteristics of the stimuli.

	Words	Pseudowords	t-Value
Number of letters	5 (4–6)	5.1 (4–6)	-1.2^{ns}
Number of phonemes	3.4 (3–5)	3.6 (3–5)	-1.4^{ns}
Number of syllables	1.3 (1–2)	1.3 (1–2)	-1^{ns}
Bigram frequency (per million)	826.3 (±452.7)	909.6 (±596.9)	-1.2^{ns}
Number of orthographic neighbors	2.5 (±2.5)	2.9 (±2.9)	-1.4^{ns}
Number of phonological neighbors	10.7 (±8.2)	10 (±8.7)	0.6^{ns}
First syllable frequency (per million)	481.1 (±311.4)	453.4 (±325.2)	-0.9^{ns}
Second syllable frequency (per million)	302 (±64.1)	289 (±62.6)	0.1^{ns}

Please cite this article as: Mahé, G., et al., Beyond the initial 140 ms, lexical decision and reading aloud are different tasks: An ERP study with topographic analysis, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.07.080

2.3. Procedure

The participants were tested individually in a soundproof dimly light room and sat at 60 cm in front of a computer screen. All participants performed both a lexical decision task and a reading aloud task in counterbalanced order, with an interval filled with an unrelated task in between. Each task was divided into 4 blocks of words and pseudowords presented in pseudorandom order (no more than 3 consecutive words or pseudowords).

The software E-prime (E-Studio) presented the trials and recorded the response latencies (RTs) and the errors for the lexical decision task. The procedure was the same in both lexical decision and reading aloud. Each trial began with a black fixation cross presented for 400 ms in the center of a gray screen. A gray screen was used to avoid extreme light exposition. The fixation cross was then replaced by a gray screen for 100 ms, followed by the stimulus for 800 ms. Stimuli were presented in Courier New font, with 18-point lower case letters and subtended approximately 3.6° of the visual angle. The next trial began after a random inter-trial interval of 900–1100 ms for the lexical decision task and 1400–1600 ms for the reading aloud task. Before each experimental task, the participants performed six practice trials.

For both tasks, participants were instructed that they would see words and pseudowords on the computer screen. In the lexical decision task, participants were asked to decide as quickly and accurately as possible whether the stimulus corresponded to a real word or not by pressing a YES response key or a NO response key with the right hand. In the reading aloud task, participants were required to read aloud as quickly and accurately as possible the stimulus displayed, whether it was a word or a pseudoword. The spoken responses were digitized and recorded for later response latency and accuracy check. After elimination of errors, latencies of vocal responses (i.e., the number of ms separating the stimulus onset from the articulation onset) were systematically checked with speech analysis software (Check Vocal; Protopapas, 2007).

2.4. EEG acquisition and pre-analyses

EEG was recorded continuously using the Active-Two Biosemi EEG system (Biosemi V.O.F. Amsterdam, Netherlands) with 128 channels covering the entire scalp. Signals were sampled at 512 Hz with bandpass filters set between 0.16 and 100 Hz.

Offline, ERPs were then bandpass-filtered to 0.2–30 Hz and notch-filtered to 50 Hz and reaveraged to average references. Averaging was computed with epochs of 500 ms starting at the stimulus onset using the Cartool software (Brunet et al., 2011). Epochs contaminated by eye blinking, movements or other noise were rejected and excluded from averaging after visual inspection performed in addition to an automated selection criterion rejecting epochs with amplitudes reaching ± 100 μV. In addition, only trials with correct responses and valid RTs were retained. As a result, a minimum of 61 averaged trials per participant in each condition and task entered the ERP analyses (words: 65–112 epochs [mean = 94] in lexical decision; 61–119 epochs [mean = 91] in reading aloud; pseudowords: 64–90 epochs [mean = 79] in lexical decision; 64–92 epochs [mean = 70] in reading aloud). Artifact electrodes were interpolated using 3-D spline interpolation (Perrin et al., 1987), with an average of 10 sites interpolated for each participant.

2.5. Waveform analyses

The ERPs were first subjected to standard waveform analysis to determine the time periods where amplitude differences were found between conditions. This analysis was performed on all electrodes and data-points. Repeated measure ANOVAs were computed on amplitudes of the evoked potentials between task (i.e., lexical decision task versus reading aloud task) and stimulus type (words versus pseudowords) using STEN toolbox (developed by Jean-François Knebel; http://www.unil.ch/fenl/home/menuguid/infrastructure/software–analysis-tools.html). Only differences over at least five clustered electrodes and extending over at least 20 consecutive time-frames (i.e., 40 ms) were retained with an alpha criterion of 0.01.

2.6. Global topographic ERP pattern analyses

While the waveform analysis will inform on whether and when differences in amplitudes appear between tasks, the global topographic pattern will further inform on whether different amplitudes are due to simple modulation in strength of the electric field or to different underlying neurophysiological mechanisms. Different periods of stable global electric fields (of topographies), likely correspond to particular periods in mental information processing (Changeux and Michel, 2004; Koukou and Lehmann, 1987; Lehman et al., 1998), thus indicating different processes across tasks.

We first run a topographic analysis on each sampling point to identify periods of significant topographic modulation between tasks. This procedure (called TANOVA, although it is not an analysis of variance) involves a non-parametric randomization test to the global dissimilarity between two electric fields (Lehmann and Skrandies, 1984; Murray et al., 2008). The permutation of the data is accomplished by reassigning randomly the topographic maps of single subjects to the different conditions. The global dissimilarity of these random group-averaged ERPs is compared time-point by time-point with the values of topographic dissimilarity of the actual conditions. A time-period criterion of 20 time-frame consecutive significant differences and an alpha criterion of 0.01 were applied.

Then a spatio-temporal segmentation was run. This spatio-temporal segmentation analysis (Brunet et al., 2011) allows summarizing ERP data into a limited number of topographic map configurations. This spatio-temporal segmentation allowed the identification of time periods during which different tasks (i.e., lexical decision and reading aloud) evoked different electric fields at scalp. This method presents the advantage of independence of the reference electrode (Michel et al., 2001; Michel et al., 2004) and insensitivity to pure amplitude modulations across conditions (topographies of normalized maps are compared). A spatio-temporal segmentation was applied to the grand-averages (i.e., from 60 ms to 500 ms after stimulus onset) to compare the two tasks for each kind of stimuli. To determine the most dominant map configurations, we used a modified hierarchical clustering analysis Q5 (Pascual-Marqui et al., 1995; Michel et al., 2001) the agglomerative hierarchical clustering (Murray et al., 2008). A modified cross-validation criterion was used to determine the optimal number of maps that best explained the group-averaged data sets across conditions. Statistical Q6 smoothing was used to eliminate temporally isolated maps with low strength. This procedure is described in detail in Pascual-Marqui et al. (1995). Additionally, a given topography had to be present for at least 20 time-frames. Then, the pattern of map templates observed in the averaged data was statistically tested by comparing each of these map templates with the moment-by-moment scalp topography of individual subjects' ERPs from each condition. Each time point was labeled according to the map with which it best correlated spatially, yielding a measure of map presence in milliseconds. This procedure referred to as 'fitting' allowed to establish how well a cluster map explained individual patterns of activity (GEV: Global Explained Variance) and its duration.

In order to analyze whether maps were more representative of one task compared to the other, the map presence, the GEV and the map duration observed in each subject's data were used for statistical analysis. Repeated measure ANOVAs were computed on both GEV and map duration with the factors Task and Lexicality. Concerning map presence, Pearson chi square tests were applied when relevant.

3. Results

3.1. Behavioral results

Incorrect responses, outliers (mean RT ± 3 SD) and contaminated epochs in ERPs data were excluded from the RT analysis. Latencies and error rates for each task and stimulus type are displayed in Table 2. Analysis of the RTs revealed a significant main effect of Task ($F(1,19) = 8.8$, $p < .01$), with faster RTs in the reading aloud task (562 ms ± 95.5) compared to the lexical decision task (596 ms ± 115.5). The effect of Lexicality, with pseudowords eliciting longer RTs than words ($F(1,19) = 313$, $p < .001$), and the Task ∗ Lexicality interaction ($F(1,19) = 9.7$, $p < .01$) were both significant. Planned comparisons revealed a larger Lexicality effect in lexical decision (61 ms; $F(1,19) = 215$, $p < .001$) compared to reading aloud (40 ms; $F(1,19) = 69.6$, $p < .001$). Concerning the error rate, the effect of Task was significant ($F(1,19) = 16.6$, $p < .001$), with a higher error rate for reading aloud (8.1% ± 4) than that for lexical decision (5.7% ± 4.2). The effect of Lexicality, with pseudowords generating more errors than words ($F(1,19) = 18.3$, $p < .001$), and the Task ∗ Lexicality interaction ($F(1,19) = 56.4$, $p < .001$) were both significant. Planned comparisons revealed that the Lexicality effect was significant in the reading aloud task (9.6%; $F(1,19) = 69.1$, $p < .001$) but not in the lexical decision task ($F < 1$).

3.2. ERP results: waveform analysis

Fig. 1 shows time points of significant amplitude differences between the lexical decision task and the reading aloud task (Fig. 1A) and between words and pseudowords (Fig. 1B) with an alpha criterion of .01 and a minimum duration of 20 time-frames, while the Task ∗ Lexicality interaction did not reveal any difference on amplitudes. Significant differences appeared between the two tasks between 180 and 500 ms following stimulus onset mostly on a cluster of 27 anterior and 16 central-posterior sites. Amplitude differences between words and pseudowords were observed between 400 and 500 ms mostly at the central-posterior sites.

3.3. ERP results: topographic analysis

Fig. 2 (left part) shows the time-window of task differences in the TANOVA analysis with an alpha criterion of .01, indicating different topographies between the two tasks from 150 ms to 500 ms following stimulus onset for words and mostly between 160–370 ms for pseudowords. The TANOVA analysis also revealed late topographic modulations (i.e., after 430 ms) related to the Lexicality effect, in both lexical decision and reading aloud (Fig. 2, right part).

As displayed in Fig. 2B, the spatio-temporal segmentation applied on the average data of lexical decision and reading aloud tasks from 60 ms to 500 ms after stimulus onset revealed eight different electrophysiological template maps accounting for 91.04% of the variance. From about 140 to 500 ms, different electrophysiological spatial configurations were observed between the two tasks for both words (on the left) and pseudowords (on the right). Only the first period of stable electrophysiological activity corresponding to the P100 component (Map A, from 60 to 140 ms in Fig. 2B) was common to the two tasks.

Table 2
Mean latencies in ms and error rate in percentage for each task and stimulus type (standard deviations into brackets).

Conditions	RTs	Error rate
Lexical decision word	565 (116)	5.8 (3.5)
Lexical decision pseudoword	626 (115)	5.6 (4.9)
Reading aloud word	542 (87)	3.3 (2.4)
Reading aloud pseudoword	582 (104)	12.9 (5.5)

The map templates observed in the grand averaged ERPs were validated by the results of the fitting procedure applied to individual ERP data in each task (i.e., lexical decision and reading aloud) and stimulus type (word and pseudoword) in two time-windows: from 140 to 240 ms for maps B and C and from 260 to 500 ms for maps D, E, F, G and H (see Table 3). Repeated measure ANOVAs were computed on both GEV and map duration with the factors Task and Lexicality. Pearson chi square tests were applied on map presence data when relevant. Within the first time window, the fitting procedure confirmed that the period of stable electric field labeled "map B" appeared more frequently in lexical decision compared to reading aloud (85% of map presence in lexical decision and 40% in reading aloud; Chi 2 on Task effect = 8.7, $p < .01$). Repeated measure ANOVAs revealed a significant Task effect for both GEV (GEV = 45.5% in lexical decision, and GEV = 17.5% in reading aloud; $F(1,19) = 15.7$, $p < .001$) and map duration ($F(1,19) = 14.4$, $p < .01$). The main effects of Lexicality and Task ∗ Lexicality interaction were not significant. The fitting procedure also confirmed that the period of stable electric field labeled "map C" appeared more frequently in the reading aloud task compared to the lexical decision task (77.5% of map presence in reading aloud, 32.5% of map presence in lexical decision; Chi 2 on Task effect = 8.9, $p < .01$). Repeated measure ANOVAs revealed a significant Task effect for both GEV (reading aloud: GEV = 14%, lexical decision: GEV = 2.5% GEV; $F(1,19) = 8.4$, $p < .01$) and map duration ($F(1,19) = 14.4$, $p < .01$). The main effects of Lexicality and Task ∗ Lexicality interaction were not significant.

The fitting procedure in the second time window revealed that map D tends to be specific to reading aloud compared to lexical decision (72.5% of map presence in reading aloud compared to only 40% of map presence in lexical decision; Chi 2 = 5.1, $p = .02$). Repeated measure ANOVAs revealed a significant Task effect for both GEV (44% and 15% in reading aloud and lexical decision respectively; $F(1,19) = 12.7$, $p < .01$) and map duration ($F(1,19) = 14$, $p < .01$). The main effects of Lexicality and Task ∗ Lexicality interaction were not significant. Map E did not differ significantly between the two tasks. Map F tended to be more specific to lexical decision compared to reading aloud (Chi 2 on map presence in the individual data = 5.6, $p = .02$). ANOVAs performed on GEV and map duration yielded only to a tendential Task effect on map F (GEV: $F(1,19) = 4$, $p = .06$; map duration: $F(1,19) = 4.9$, $p = .04$). Concerning map H, statistical analysis revealed a stronger map presence in lexical decision compared to reading aloud (Chi 2 on map presence in the individual ERPs = 8.6, $p < .01$). ANOVAs revealed only a tendential Task effect on map duration ($F(1,19) = 6.3$, $p = .02$). Finally, map G displayed a significant Lexicality effect on GEV ($F(1,19) = 17.3$, $p < .001$). Concerning map duration, statistical analysis revealed a tendential Task effect ($F(1,19) = 3.8$, $p = .06$) and a significant Lexicality effect ($F(1,19) = 15.9$, $p < .001$).

4. Discussion

The aim of the present ERP study was to determine exactly when print processing differs between lexical decision and reading aloud tasks. The same participants performed the tasks on the same material and design.

4.1. Early task differences

The core finding is the observation of early print processing differences between lexical decision and reading aloud. Indeed, waveform amplitude analysis revealed task differences from ~180 ms to the end of the analyzed interval. Topographic analysis revealed only one common period of topographic stability between lexical decision and reading aloud. This common microstate (between 90–100 and 140 ms) corresponds to the P100 component, which has been associated with visual feature analysis (Dien, 2009). After the P100, from ~140 ms to the end of the analyzed interval (i.e., 500 ms), the two tasks displayed completely distinct microstates. It should be noted that in both the amplitude waveform analysis and the topographic analysis, task differences did not vary between

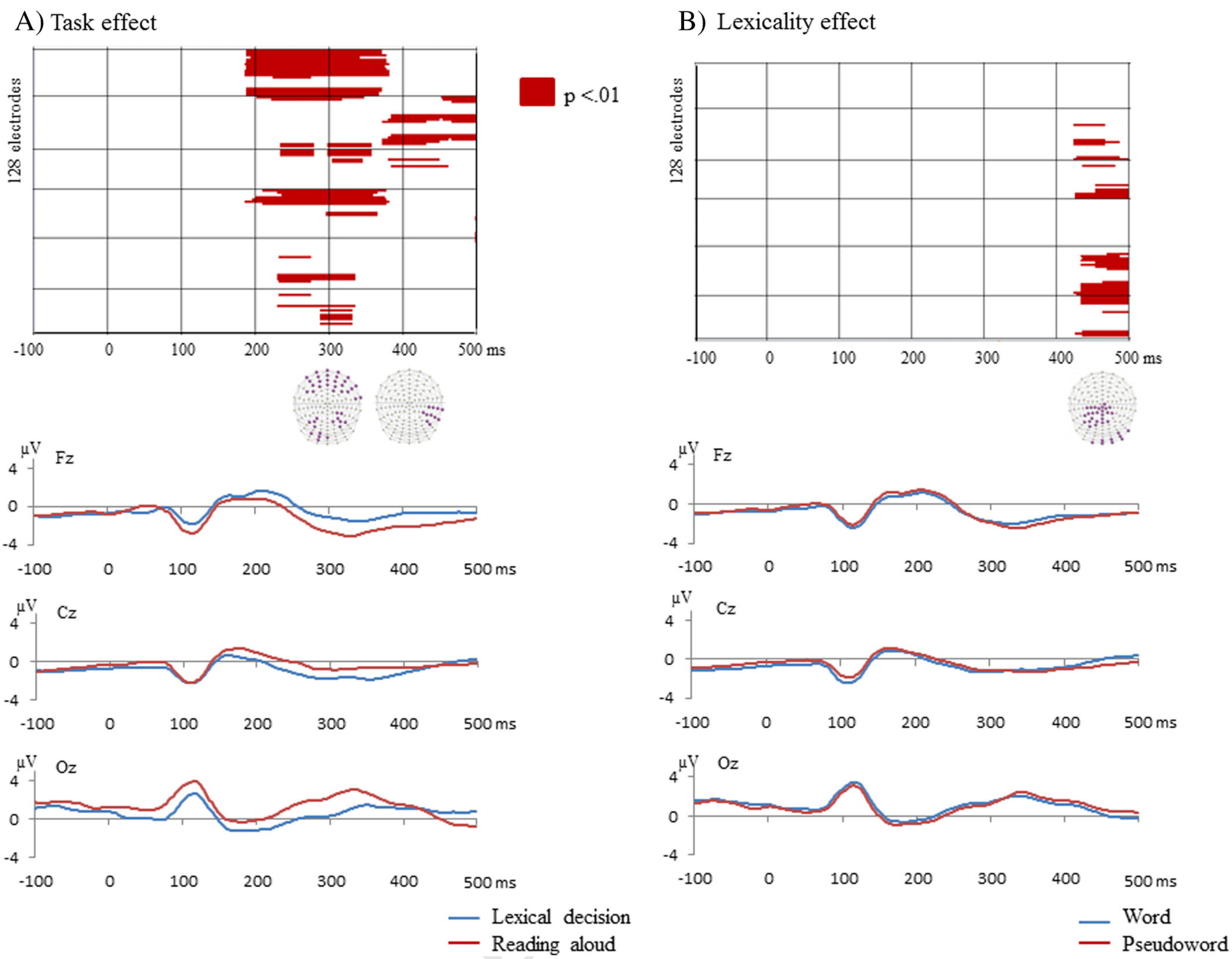


Fig. 1. Significant differences on ERP waveform amplitude on each electrode (Y axes) and time point (X axes) between lexical decision and reading aloud (left part, Fig. 1A) and between words and pseudowords (right part, Fig. 1B). Only differences over at least five clustered electrodes and 20 time frames, with an alpha criterion of .01 are displayed in red. The electrode sites yielding significant differences at 300 ms and 400 ms for the Task effect and at 450 ms for the Lexicality effect are displayed under each graphic. Examples of waveform are displayed below for the Task and Lexicality effect for Fz, Cz and Oz. (For interpretation of the reference to color in this figure legend, the reader is referred to the web version of this article.)

words and pseudowords. This suggests general differences in print processing between the two tasks which would not depend on lexicality. This finding, together with previous MEG results by Chen et al. (2013), showing activation differences between lexical decision and silent reading as early as 150 ms following stimulus onset, rules out the hypothesis of differences limited to late processing stages between lexical decision and reading. Our data rather suggest that after the low-level visual analysis, print stimuli are processed differently according to the task.

In the lexical decision task, the stable topographic configuration recorded between ~140–250 ms characterized by posterior negativity (i.e., map B in Fig. 2B) appeared to be predominant, with very strong map presence over participants and high GEV. In visual word recognition studies, the time window between 150 and 250 ms has been associated with the N170 component, reflecting the expert processing of print in the VWFA in a large variety of tasks including the lexical decision task (Brem et al., 2009; Dujardin et al., 2011; Maurer et al., 2005; Mahé et al., 2012). In addition, the topographic configuration recorded in our study corresponds to the classical topography of the N170 component, with the left posterior–temporal peak of negativity (Brem et al., 2009; Maurer et al., 2005; Maurer et al., 2005b). In contrast, a different stable electric field configuration, with anterior and posterior negativities surrounding central positivity, was recorded in the reading aloud task between ~140 and 240 ms (i.e., map C in Fig. 2B). It should be noted that in this time-window the period of topographic stability seems to be more stable across participants in the lexical decision task than that in the reading aloud task (larger GEV and presence for map B in lexical decision than for map C in reading aloud). Overall, our electrophysiological data revealed that the electrophysiological pattern previously associated with the expert processing of print appears to be a major processing stage of written strings only in lexical decision, a task promoting the orthographic pattern analysis in order to distinguish familiar words from pseudowords. Our data could thus be in accordance with previous results showing stronger left inferior temporal gyrus activation in lexical decision compared to silent reading in MEG (Chen et al., 2013) and to reading aloud in fMRI (Katz et al., 2005) at a location consistent with the VWFA. In fact, the N170 component has repeatedly been associated with the activation of the VWFA in studies using ERP source analysis (Brem et al., 2006, 2009) or combining EEG and fMRI measurements (Brem et al., 2006).

4.2. Late task differences and potential impact of the response mode

One might argue that the electrophysiological differences reported between the two tasks could be partly a consequence of the different response mode (oral vs. button press). Concerning responses requiring button press, data coming from studies measuring the Lateralized

Please cite this article as: Mahé, G., et al., Beyond the initial 140 ms, lexical decision and reading aloud are different tasks: An ERP study with topographic analysis, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.07.080

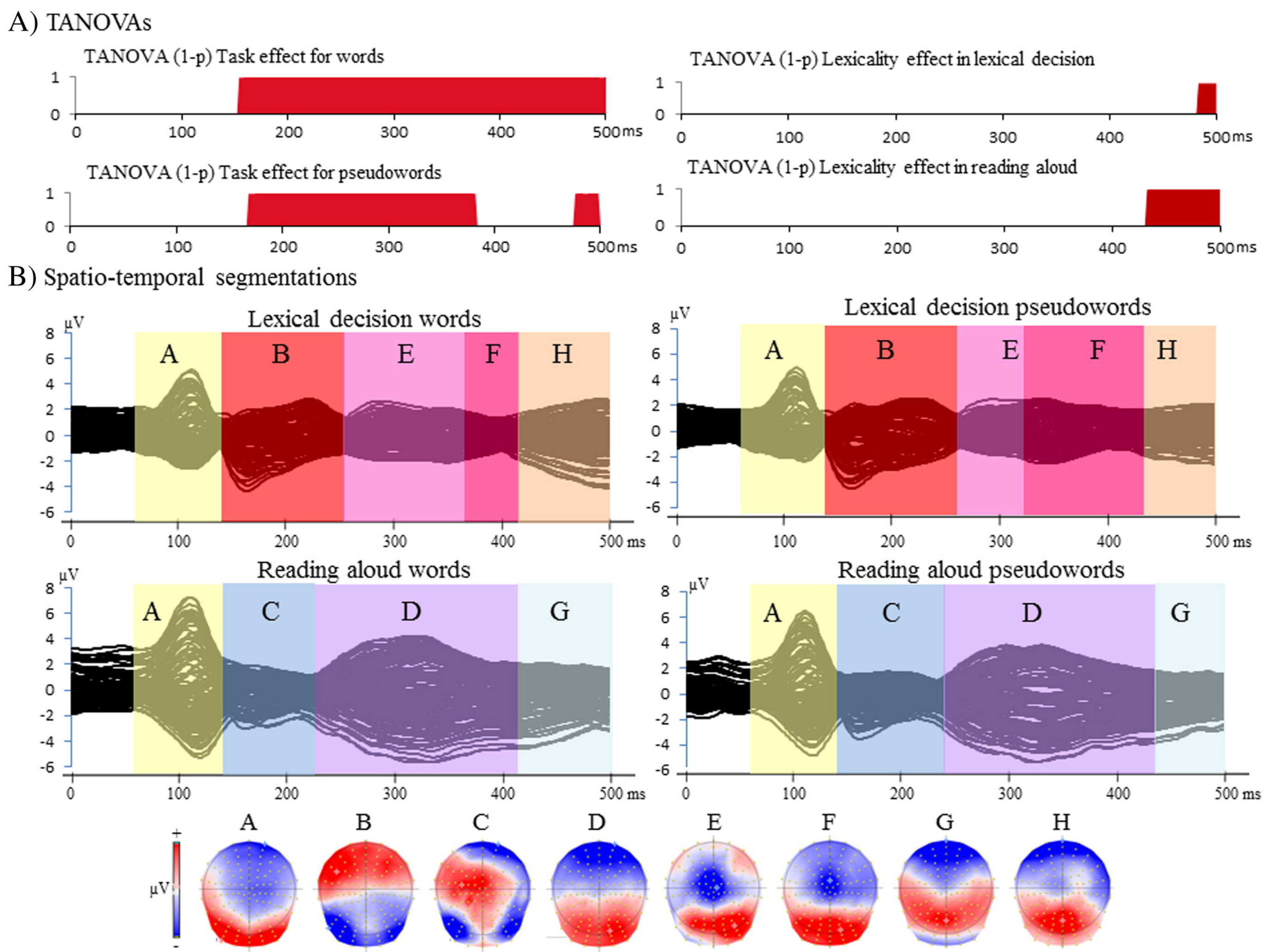


Fig. 2. A. Results of the TANOVA analysis (1 − p) for the Task effect (on the left) and for the Lexicality effect (on the right). Only differences over at least 20 time frames, with an alpha criterion of .01 are displayed in red. B. Grand average ERPs (128 electrodes) from each task (i.e., lexical decision on the top and reading aloud below) and each stimulus type (words on the left and pseudowords on the right) and temporal distribution of the topographic maps revealed by the spatio-temporal segmentation analysis in each data. Colors illustrate the time-window of each period of topographic stability. Corresponding map templates are displayed below. (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

Readiness Potential (LRP), taken as an index of specific response preparation in the motor cortex, indicate an occurrence of motor preparation occurring approximately 100 ms before response latency (i.e., with a peak of LRP; e.g., Killikely and Szücs, 2013; Müller and Hagoort, 2006). Here, given the distribution of RTs in the lexical decision task (596 ms ± 115), an impact of response preparation processes should be limited to the end of the second time interval analyzed (i.e., 260–500 ms). This is in line with the observation that topographic consistency across participants was lower in this time window, probably due to variability in response latencies and motor preparation (see GEV data and map presence for maps E, F and H).

Concerning the reading aloud task, the last period of topographic stability (i.e., map G, appearing between 420–500 ms) displayed very low topographic stability across participants. Given its proximity to

Table 3
For each topography, presence in the individual ERPs, GEV and duration in time frames (TF) for each task and stimulus type.

	Map presence				GEV				Map duration			
	Lexical decision		Reading aloud		Lexical decision		Reading aloud		Lexical decision		Reading aloud	
	W	PW	W	PW	W	PW	W	PW	W	PW	W	PW
140–240 ms												
Map B	**85%**	**85%**	**35%**	**45%**	**43%**	**48%**	**16%**	**19%**	**38**	**39**	**14**	**18**
Map C	**35%**	**30%**	**80%**	**75%**	**2%**	**3%**	**13%**	**15%**	**13**	**12**	**37**	**33**
260–500 ms												
Map D	**45%**	**35%**	**70%**	**75%**	**13%**	**17%**	**42%**	**46%**	**19**	**21**	**56**	**64**
Map E	50%	60%	20%	40%	6%	5%	1%	3%	36	34	19	26
Map F	45%	45%	10%	25%	12%	12%	4%	6%	23	26	8	13
Map G	45%	15%	65%	35%	4%	1%	6%	1%	15	8	30	12
Map H	50%	60%	10%	15%	14%	14%	6%	3%	29	32	9	5

Please cite this article as: Mahé, G., et al., Beyond the initial 140 ms, lexical decision and reading aloud are different tasks: An ERP study with topographic analysis, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.07.080

mean RTs (562 ms ± 96) it is likely due to variability in response latencies and articulatory programming. Finally, it should be noted that the map displaying the most consistence across participants in reading aloud (with especially high GEV and map presence) was located in the 230–420 ms time interval (map D in Fig. 2). This suggests that processes specific to reading aloud would occur during this time interval. The interpretation of the cognitive processes related to this time window can only be hypothetical, as electrophysiological data concerning the time course of reading aloud are limited. The interval of our map D partly encompassed the time window previously associated to early phonological activation in reading aloud (i.e., around 150–330 ms; Timmer and Schiller, 2014 for a review) or to grapheme-to-phoneme mapping in silent reading (i.e., 270–336 ms for the N320 component; Simon et al., 2004). Consequently one may hypothesize that the time interval associated to map D could integrate several cognitive processes specific to reading aloud, including phonological processes.

4.3. Lexicality effect across tasks

Behavioral data revealed a stronger lexicality effect in lexical decision than in reading aloud on RTs and a lexicality effect on accuracy limited to reading aloud. Concerning electrophysiological data, the waveform amplitude and the topographic analysis converged on a very late lexicality effect, occurring after 430 ms following stimulus presentation. This late lexicality effect was similar across tasks. The time window of the lexicality effect is inconsistent in the literature: some studies reported a lexicality effect as early as the N170 time window (Dujardin et al., 2011; Mahé et al., 2012; Shaul et al., 2012) while other studies described later effects, around 300–400 ms following stimulus presentation (Bentin et al., 1999; Pylkkänen and Marantz, 2003; Wydell et al., 2003). It should be noted that even if the present electrophysiological results do not support an early lexicality effect nor an interaction with task, the experimental design may not be the ideally suited for the investigation of lexicality as each participant saw the stimuli twice (once in each task). This design was chosen to optimize the between tasks comparison, which was the main focus of the experiment. However, stimulus repetition may have reduced the lexicality effect in the ERP data, as this effect has been reported to be affected by stimulus repetition in previous electrophysiological studies (Almeida and Poeppel, 2013; Bermúdez-Margaretto et al., 2014). The use of different lists of stimuli across tasks would be best suited in future to compare the time course of the lexicality effects across reading and lexical decision tasks.

5. Conclusion

Overall, our electrophysiological data revealed early processing differences between the lexical decision task and the reading aloud task, with a predominance of the expert processing of print in lexical decision only. As it has been suggested (Balota and Yap, 2006), the task used impacts on the processing stages involved in print processing. As a consequence, one could question the use of the lexical decision task in the assessment of reading abilities. This question is of particular relevance as the lexical decision task is extensively used in ERP studies (Bentin et al., 1999; Simon et al., 2004; Hauk et al., 2012) whereas the use of reading aloud tasks is limited.

In conclusion, even if the lexical decision task provides a good index of visual word recognition processes, the present findings warns against an overgeneralization of lexical decision data to the processes occurring during reading aloud. Future studies should thus interpret very carefully lexical decision data, especially in the case of studies designed to better understand reading disorders.

Acknowledgment

This research was supported by Swiss National Science Foundation grant no. 100014_149595.

The STEN toolbox (http://www.unil.ch/fenl/Sten) has been programmed by Jean-François Knebel, from the Laboratory for Investigative Neurophysiology (the LINE), Lausanne, Switzerland, and is supported by the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne and by National Center of Competence in Research project "SYNAPSY — The Synaptic Bases of Mental Disease"; project no. 51AU40_125759.

Appendix 1. Stimulus material

Words

absent	coeur	huile	poing
accent	corps	idole	point
achat	court	image	poker
adulte	crabe	index	porc
allure	cube	iris	pouce
arche	cyan	jade	quoi
archet	cygne	jardin	rhum
argent	dehors	jockey	rythme
aspect	demi	joli	sabot
asthme	dieu	joyeux	sauf
aube	dinde	jupon	second
auburn	doigt	meute	sept
avoir	faisan	mythe	seul
bœuf	famine	nerfs	sinon
cake	faon	nuque	soeur
cerf	femme	oeil	taupe
chacun	fils	oeuf	temps
chaise	foetus	oeuvre	timbre
chauve	foin	oignon	toast
chlore	foot	onde	toit
chœur	fougue	orme	toux
chose	fugue	ours	trou
chou	fusil	paon	truffe
chrome	geai	pardon	type
chute	gentil	parfum	vaccin
cinq	ghetto	peau	veau
clef	guidon	pied	viande
cloche	hall	pirate	vieil
clown	heure	plante	vivre
club	hiver	poids	volcan

Pseudowords

achon	cluge	guippo	ponve
adolge	coum	halic	pude
agoif	crag	hoise	puldot
albaux	crige	hule	quaive
albeux	cronse	idave	quaux
aleibe	crou	iluge	quin
ancrou	crour	inar	rancte
anste	cune	inse	rardou
anterm	cunge	intax	raup
anve	cynne	jave	remon
archou	dappe	jayaux	rori
astint	dehage	jophau	sebau
ausbet	dian	lugis	seum
bence	diso	mauce	sipe
caige	dunge	munne	sipun
caime	falone	neube	sonve
chage	fantou	nouse	soun
chande	feume	oude	surous
charou	fipuge	oulphe	taige
choune	fome	ouphau	tausbe
cibe	fonque	pastou	tibe
cife	fouc	piruge	toust
cinve	funce	plonve	vistou
cleu	gantin	plunne	voble
climpe	ginpas	pomal	vuclet

References

Almeida, D., Poeppel, D., 2013. Word-specific repetition effects revealed by MEG and the implications for lexical access. Brain Lang. 127, 497–509.

Balota, D.A., Chumbley, J.I., 1984. Are lexical decisions a good measure of lexical access: the role of word frequency in the neglected decision stage. J. Exp. Psychol. Hum. Percept. Perform. 10, 340–357.

Balota, D.A., Yap, M., 2006. Attentional control and flexible lexical processing: explorations of the magic moment of word recognition. In: Andrews, S. (Ed.), From Inkmarks to Ideas: Current Issues in Lexical Processing. Psychology Press, New York, N.Y., pp. 229–258.

Balota, D.A., Cortese, M.J., Sergent-Marshall, S.D., Spieler, D.H., Yap, M.J., 2004. Visual word recognition of single-syllable words. J. Exp. Psychol. Gen. 133, 283–316.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M.H., Echallier, J.F., Pernier, J., 1999. ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. J. Cogn. Neurosci. 11, 235–260.

Bermúdez-Margaretto, B., Beltrán, D., Domínguez, A., Cuetos, F., 2014. Repeated exposure to "meaningless" pseudowords modulates LPC, but not N(FN)400. Brain Topogr.

Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., Brandeis, D., 2006. Evidence for developmental changes in the visual word processing network beyond adolescence. NeuroImage 29, 822–837.

Brem, S., Halder, P., Bucher, K., Summers, P., Martin, E., Brandeis, D., 2009. Tuning of the visual word processing system: distinct developmental ERP and fMRI effects. Hum. Brain Mapp. 30, 1833–1844.

Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. Comput. Intell. Neurosci.

Carreiras, M., Mechelli, A., Estévez, A., Price, C.J., 2007. Brain activation for lexical decision and reading aloud: two sides of the same coin? J. Cogn. Neurosci. 19, 433–444.

Changeux, J.P., Michel, C.M., 2004. Mechanisms of neural integration at the brain scale level: the neuronal work space and microstate models. In: Grillner, S., Grabyel, A.M. (Eds.), Microcircuits: The Interface Between Neurons and Global Brain Function. The MIT Press, Cambridge, M.A., pp. 347–370.

Chen, Y., Davis, M.H., Pulvermüller, F., Hauk, O., 2013. Task modulation of brain responses in visual word recognition as studied using EEG/MEG and fMRI. Front. Hum. Neurosci. 7 (376).

Cornelissen, P., Tarkiainen, A., Helenius, P., Salmelin, R., 2003. Cortical effects of shifting letter position in letter strings of varying length. J. Cogn. Neurosci. 15, 731–746.

Cornelissen, P.L., Kringelbach, M.L., Ellis, A.W., Whitney, C., Holliday, I.E., Hansen, P.C., 2009. Activation of the left inferior frontal gyrus in the first 200 ms of reading: evidence from magnetoencephalography (MEG). PLoS One 4 (4).

Dien, J., 2009. The neurocognitive basis of reading single words as seen through early latency ERPs: a model of converging pathways. Biol. Psychol. 8, 10–22.

Dujardin, T., Etienne, Y., Contentin, C., Bernard, C., Largy, P., Mellier, D., Lalonde, R., Rebaï, M., 2011. Behavioral performances in participants with phonological dyslexia and different patterns on the N170 component. Brain Cogn. 75, 91–100.

Ferrand, L., Brysbaert, M., Keuleers, E., New, B., Bonin, P., Méot, A., Augustinova, M., Pallier, C., 2011. Comparing word processing times in naming, lexical decision, and progressive demasking: evidence from Chronolex. Front. Psychol. 2, 306.

Fiebach, C.J., Friederici, A.D., Müller, K., von Cramon, D.Y., 2002. fMRI evidence for dual routes to the mental lexicon in visual word recognition. J. Cogn. Neurosci. 14, 11–23.

González-Nosti, M., Barbón, A., Rodríquez-Ferreiro, J., Cuetos, F., 2014. Effects of the psycholinguistic variables on the lexical decision task in Spanish: a study with 2,765 words. Behav. Res. Methods 46, 517–525.

Grainger, J., Jacobs, A.M., 1996. Orthographic processing in visual word recognition: a multiple read-out model. Psychol. Rev. 22, 696–713.

Hauk, O., Coutout, C., Holden, A., Chen, Y., 2012. The time-course of single word reading: evidence from fast behavioral and brain responses. NeuroImage 60, 1462–1477.

Katz, L., Lee, C.H., Tabor, W., Frost, S.J., Mencl, W.E., Sandak, R., Rueckl, J., Pugh, K., 2005. Behavioral and neurobiological effects of printed word repetition in lexical decision and naming. Neuropsychologia 43, 2068–2083.

Katz, L., Brancazio, L., Irwin, J., Katz, S., Magnuson, J., Whalen, D.H., 2012. What lexical decision and naming tell us about reading. Read. Writ. 25, 1259–1282.

Killikelly, C., Szücs, D., 2013. Asymmetry in stimulus and response conflict processing across the adult lifespan: ERP and EMG evidence. Cortex 49, 2888–2903.

Kim, A.E., Straková, J., 2012. Concurrent effects of lexical status and letter-rotation during early stage visual word recognition: evidence from ERPs. Brain Res. 1468, 52–62.

Koukou, M., Lehmann, D., 1987. An information processing perspective of psychophysiological measurements. J. Psychophysiol. 1, 109–112.

Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.

Lehmann, D., Skrandies, W., 1984. Spatial analysis of evoked potentials in man—a review. Prog. Neurobiol. 23, 227–250.

Lehmann, D., Strik, W.K., Henggeler, B., Koenig, T., Koukkou, M., 1998. Brain electric microstates and momentary conscious mind states as building blocks of spontaneous thinking: I. Visual imagery and abstract thoughts. Int. J. Psychophysiol. 29, 1–11.

Lupker, S.J., Pexman, P.M., 2010. Making things difficult in lexical decision: the impact of pseudohomophones and transposed-letter nonwords on frequency and semantic priming effects. J. Exp. Psychol. Learn. Mem. Cogn. 5, 1267–1289.

Mahé, G., Bonnefond, A., Gavens, N., Dufour, A., Doignon-Camus, N., 2012. Impaired visual expertise for print in French adults with compensated dyslexia as shown by N170 tuning. Neuropsychologia 50, 3200–3206.

Mahé, G., Bonnefond, A., Doignon-Camus, N., 2013. Is the impaired N170 print tuning specific to developmental dyslexia? A matched reading level study with poor readers and dyslexics. Brain Lang. 127, 539–544.

Maurer, U., Brem, S., Bucher, K., Brandeis, D., 2005a. Emerging neurophysiological specialization for letter strings. J. Cogn. Neurosci. 17, 1532–1552.

Maurer, U., Brandeis, D., McCandliss, B., 2005b. Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. Behav. Brain Funct. 1, 13.

Michel, C.M., Thut, G., Morand, S., Khateb, A., Pegna, A.J., Grave de Peralta, R., 2001. Electric source imaging of human brain functions. Brain Res. Rev. 36, 108–118.

Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. Clin. Neurophysiol. 115, 2195–2222.

Müller, O., Hagoort, P., 2006. Access to lexical information in language comprehension: semantics before syntax. J. Cogn. Neurosci. 18, 84–96.

Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. Brain Topogr. 20, 249–264.

New, B., Pallier, C., Ferrand, L., Matos, R., 2001. Une base de données lexicales du français contemporain sur internet: Lexique. Annee Psychol. 101, 447–462.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans. Biomed. Eng. 42, 658–665.

Perrin, F., Pernier, J., Bertrand, O., Giard, M.H., Echallier, J.F., 1987. Mapping of scalp potentials by surface spline interpolation. Electroencephalogr. Clin. Neurophysiol. 66, 75–81.

Protopapas, A., 2007. Check Vocal: a program to facilitate checking the accuracy and response time of vocal responses from DMDX. Behav. Res. Methods 39, 859–862.

Pylkkännen, L., Marantz, A., 2003. Tracking the time course of word recognition with MEG. Trends Cogn. Sci. 7, 187–189.

Schilling, H.H., Rayner, K., Chumbley, J.I., 1998. Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. Mem. Cogn. 26, 1270–1281.

Shaul, S., Arzouan, Y., Goldstein, A., 2012. Brain activity while reading words and pseudo-words: a comparison between dyslexic and fluent readers. Int. J. Psychophysiol. 84, 270–276.

Simon, G., Bernard, C., Largy, P., Lalonde, R., Rebaï, M., 2004. Chronometry of visual word recognition during passive and lexical decision tasks: an ERP investigation. Int. J. Neurosci. 114, 1401–1432.

Simon, G., Petit, L., Bernard, C., Rebaï, M., 2007. N170 ERPs could represent a logographic processing strategy in visual word recognition. Behav. Brain Funct. 3, 21.

Stone, G.O., Van Orden, G.C., 1993. Strategic control of processing in word recognition. J. Exp. Psychol. Hum. Percept. Perform. 19, 744–774.

Timmer, K., Schiller, N.O., 2014. Neural correlates reveal sub-lexical orthography and phonology during reading aloud: a review. Front. Psychol. 5, 884.

Valdois, S., Carbonnel, S., Juphard, A., Baciu, M., Ans, B., Peyrin, C., Segebarth, C., 2006. Polysyllabic pseudo-word processing in reading and lexical decision: converging evidence from behavioral data, connectionist simulations and functional MRI. Brain Res. 1085, 149–162.

Wydell, T.N., Vuorinen, T., Helenius, P., Salmelin, R., 2003. Neural correlates of letter-string length and lexicality during reading in a regular orthography. J. Cogn. Neurosci. 15, 1052–1062.

Yap, M.J., Pexman, P.M., Wellsby, M., Hargreaves, I.S., Huff, M.J., 2012. An abundance of riches: cross-task comparisons of semantic richness effects in visual word recognition. Front. Hum. Neurosci. 6, 1–10.

Yum, Y.N., Law, S.P., Su, I.F., Lau, K.Y., Mo, K.N., 2014. An ERP study of effects of regularity and consistency in delayed naming and lexicality judgment in a logographic writing system. Front. Psychol. 14, 315.

Please cite this article as: Mahé, G., et al., Beyond the initial 140 ms, lexical decision and reading aloud are different tasks: An ERP study with topographic analysis, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.07.080