



## Registered report

# The consonant/vowel pattern determines the structure of orthographic representations in the left fusiform gyrus<sup>☆</sup>

Fabienne Chetail<sup>a,b,\*</sup>, Mariagrazia Ranzini<sup>a,b</sup>, Xavier De Tiège<sup>b,c,d</sup>,  
Vincent Wens<sup>b,c,d,1</sup> and Alain Content<sup>a,b,1</sup>

<sup>a</sup> LCLD, CRCN, Université Libre de Bruxelles (ULB), Belgium

<sup>b</sup> UNI – ULB Neuroscience Institute, Université Libre de Bruxelles (ULB), Belgium

<sup>c</sup> Laboratoire de Cartographie Fonctionnelle du Cerveau (LCFC), Université Libre de Bruxelles (ULB), Belgium

<sup>d</sup> Magnetoencephalography Unit, Department of Functional Neuroimaging, Service of Nuclear Medicine, CUB Hôpital Erasme, Brussels, Belgium

## ARTICLE INFO

## Article history:

Protocol received: Feb 05, 2015

Protocol accepted: May 06, 2016

Received 2 June 2017

Reviewed 29 August 2017

Revised 14 October 2017

Accepted 15 January 2018

Action editor Pia Rotshtein

Published online 2 February 2018

## Keywords:

Visual word recognition

Letter processing

Consonant/vowel structure

Same/different matching task

MEG

VWFA

## ABSTRACT

Recent findings demonstrated readers' sensitivity to the distinction between consonant and vowel letters. Especially, the way consonants and vowels are organised within written words determines their perceptual structure. The present work attempted to overcome two limitations of previous studies by examining the neurophysiological correlates of this perceptual structure through magnetoencephalography (MEG). One aim was to establish that the extraction of vowel-centred units takes place during early stages of processing. The second objective was to confirm that the vowel-centred structure pertains to the word recognition system and may constitute one level in a hierarchy of neural detectors coding orthographic strings. Participants performed a cross-case matching task in which they had to judge pairs of stimuli as identical or different. The critical manipulation concerned pairs obtained by transposing two letters, so that the vowel-centred structure was either preserved (FOUVERT-*fouvert*, two vowel letter clusters) or modified (BOUVRET-*bovuret*). Mismatches were detected faster when the structure was modified. This effect was associated with a significant difference in evoked neuromagnetic fields extending from 129 to 239 msec after the stimulation. Source localization indicated a significant effect in the visual word form area around 200 msec. The results confirm the hypothesis that the vowel-centred structure is extracted during the early phases of letter string processing and that it is encoded in left fusiform regions devoted to visual word recognition.

© 2018 Elsevier Ltd. All rights reserved.

<sup>☆</sup> Registered Reports Editor: Chris Chambers. Email address: [ChambersC1@cardiff.ac.uk](mailto:ChambersC1@cardiff.ac.uk).

\* Corresponding author. Laboratoire Cognition Langage Développement (LCLD), Research Centre in Cognition & Neuroscience (CRCN), ULB Neuroscience Institute (UNI), Université libre de Bruxelles (ULB), Av. F. Roosevelt, 50/CP 191, 1050, Brussels, Belgium.

E-mail address: [fchetail@ulb.ac.be](mailto:fchetail@ulb.ac.be) (F. Chetail).

<sup>1</sup> These authors contributed equally to the study.

<https://doi.org/10.1016/j.cortex.2018.01.006>

0010-9452/© 2018 Elsevier Ltd. All rights reserved.

## 1. Introduction

Alphabets are made of symbols, which strongly vary in shape from one script to another (e.g., compare β and ε in Greek to ๒ and ๓ in Thai). Yet despite the disparity in visual shapes, current alphabets share basic features. The most straightforward one is the classification of symbols into two classes: consonants and vowels, although the signs for vowels appeared later in history than consonants and can be optional in Semitic alphabets such as Arabic or Hebrew (Shimron, 1993). Alphabetic scripts by no means constitute the sole solution that cultures invented to encode spoken languages in visual form, as attested from syllabaries and logographic/morphographic scripts, but in alphabetic scripts only characters are categorized as consonants and vowels. This formal categorization of letters is primarily driven by the relation between writing systems and spoken language. Speech is characterized by the quasi-regular alternation between relatively stable complex periodic waves –the vowels– and brief friction or explosion noises produced by constrictions or occlusions and release of articulators, the consonants (MacNeilage, 1998). This categorization of speech sounds has been extended to the symbols used to represent spoken forms, although the two sets of elements are not strictly isomorphic. In French for example, the N in *pente* (/pɑ̃t/) belongs to the class of consonant letters although it is included in a letter cluster (EN) that codes for a vowel phoneme. Despite some mismatches between letters and phonemes, the consonant/vowel (CV) categorization in the written modality is unambiguous, it is part of the basic knowledge of any child learning to read, and many recent findings have demonstrated readers' sensitivity to the CV distinction (e.g., Carreiras, Duñabeitia, & Molinaro, 2009; Chetail & Content, 2012, 2014; Duñabeitia & Carreiras, 2011).

The present magnetoencephalography (MEG) study aims at examining the spatiotemporal dynamics of the neural correlates of the sensitivity to the CV distinction and of its influence on the structure of the mental representation of orthographic strings. Contrary to Semitic writing systems where most vowels are specified by diacritics marks, there is no obvious visual distinction between consonant and vowel letters in languages transcribed with the Latin alphabet (e.g., French, English, Spanish). Any effect of CV structure in these languages therefore provides strong evidence that experience with print makes the visual word processing system tuned to letter category.

### 1.1. CV effects in written word processing

The role of consonants and vowels in written word processing has been discussed for decades, with a special focus on written word segmentation (e.g., Adams, 1981; Hansen & Rodgers, 1965) and on phonological assembly (e.g., Berent & Perfetti, 1995), but it has recently triggered a renewed interest. Several recent priming studies manipulated the preservation of consonant or vowel letters with isolated word recognition paradigms. For example, in a go/no-go semantic categorization task, Carreiras et al. (2009) presented participants with target words briefly preceded either by consonant-sharing primes (e.g., *frl* – FAROL) or by vowel-sharing primes (e.g.,

*aeo* – ACERO) while recording event-related potentials (ERPs). They found a dissociation of priming effects between the two conditions in time windows including the N250 and the N400 components. Consonant-sharing primes elicited a smaller negativity amplitude than control primes, whereas vowel-sharing primes elicited amplitudes as large as those elicited by control primes. The behavioural results produced a pattern consistent with neurophysiological observations, as word recognition was facilitated when targets were preceded by consonant-sharing primes (e.g., *frl*) compared to an unrelated control prime (e.g., *tsb*) whereas no such difference was found for vowel-sharing primes (e.g., Duñabeitia & Carreiras, 2011). These results were interpreted in terms of lexical selection constraints. Vowel letters constrain the identity of words to a lesser extent than consonant letters, presumably because the former are more frequent and less informative (e.g., more English words can be generated from *\_A\_E* than from *F\_C\_*). Word identification processes would therefore more strongly rely on consonants than on vowel letters to select a target among possible lexical candidates (see Duñabeitia & Carreiras, 2011; see also Nespor, Peña, & Mehler, 2003, for a similar argument in the spoken modality).

In addition, observations from dysgraphic patients' production errors suggest that the CV status of letters would not only play a role in lexical selection, but would also be used to structure spelling production. Cubelli (1991) described the spelling of two acquired dysgraphic patients who showed a striking selective deficit for vowels, and Caramazza and Miceli (1990) reported a patient whose spelling errors systematically preserved the letter category (i.e., consonants were substituted for consonants and vowels for vowels). They further observed that the CV structure was equally kept intact, that is the alternation of consonant and vowel letters (see also Buchwald & Rapp, 2006), leading to the conclusion that lexical orthographic representations specify the letter category separately from letter identity.

The notion of CV structure has been extended from production to word perception by Chetail and Content. They assembled several strands of behavioural evidence to support the idea that the organization of consonant and vowel letters determines the perceptual structure of letter strings in written word recognition (e.g., Chetail & Content, 2012, 2014; Chetail, Drabs, & Content, 2014; Chetail, Scaltritti, & Content, 2014). One study showed that readers rely on the CV structure to determine the number of syllables in written stimuli. The congruence between orthographic and phonological structure was manipulated by using words such as *client*, which have two syllables (/kli.jɑ̃/) but include only one orthographic vowel cluster (CCVVCC). Participants were slower and more error prone for such words compared to words with the same number of vowel clusters and syllables (e.g., *média*./me.dja/, CVCVV). Further, when they failed to give the correct response, participants usually underestimated the number of syllables, meaning that they were more prone to count the number of vowel clusters than the number of syllables (Chetail & Content, 2012). Such results suggest that the way consonant and vowel letters are organized within strings determines the number of units perceived, with each vowel or vowel cluster being the core of an orthographic unit. Thus, a word like *réunion* would be structured as two orthographic

units because its CV structure entails two vowel clusters (e.g., *éu* and *io*).

One essential corollary of this claim is the requirement of a fundamental revision of current models of orthographic encoding, as none of them formally incorporates the distinction between consonants and vowels. Indeed, the need to introduce the CV distinction in word recognition models has been ruled out until now since most relevant effects were found in tasks involving lexical processing and could thus be interpreted in terms of shared neighbours (see Chetail, Drabs et al., 2014; Perea & Acha, 2009, for discussions). However, recent results indicate that readers' sensitivity to the CV status of letters is associated with early, pre-lexical, involuntary processing mechanisms. Indeed, Chetail, Drabs et al.'s (2014) study with the same-different matching task directly demonstrated that the CV structure of words plays a role during word processing specifically at a sublexical level. In the sequential variant of the same-different task, a first item (the referent) is presented, and is replaced by the target. The task is to decide whether the target is identical or different from the referent, independently of the case. Combined with letter transposition, the task has been used to examine the coding of letter position (see Norris & Kinoshita, 2008). Pseudowords built from words by the transposition of two adjacent letters (e.g., *gadren* from *garden*) are frequently misperceived as the corresponding base words (e.g., Bruner & O'Dowd, 1958; Chambers, 1979). The counterpart result in the same-different task is a greater difficulty to detect a mismatch produced by transposition (e.g., *RFCV-rcfv*) than one based on substitution (e.g., *RDTV-rcfv*, Duñabeitia, Dimitropoulou, Grainger, Hernández, & Carreiras, 2012; Massol, Duñabeitia, Carreiras, & Grainger, 2013), providing evidence against a strict letter position coding scheme. Chetail, Drabs et al. (2014) used this paradigm to show that the perceptual discrimination is influenced by the organisation of consonants and vowels. If readers are sensitive to the CV structure at a sublexical level of processing, then a transposition that disrupts the CV structure (e.g., *BOUVRET – bovuret*) should be easier to discriminate than a transposition that preserves the structure (*FOUVERT – fouvert*) because in the former case only, the number of vowel clusters is different between the referent (*BOUVRET*: two vowel clusters) and the target (*bovuret*: three vowel clusters). The results fully confirmed this prediction, both with words and with pseudowords.

### 1.2. The present study

The critical hypothesis, which we aim to test in the present study, is that the impact of the organization of consonants and vowels occurs at an early stage during word identification. Despite converging evidence in favour of the influence of the CV structure, none of the previous results directly supports this claim. This is primarily due to the fact that in previous behavioural experiments, reaction times and error rates only reflect the cumulative effect of the successive stages of processing and not the detailed time course of each of them. The main aim of the study is therefore to assess how early CV structure effects occur, an issue that is critical to establish whether CV structure needs to be considered in future models of visual word recognition. To do so, we will use the same-

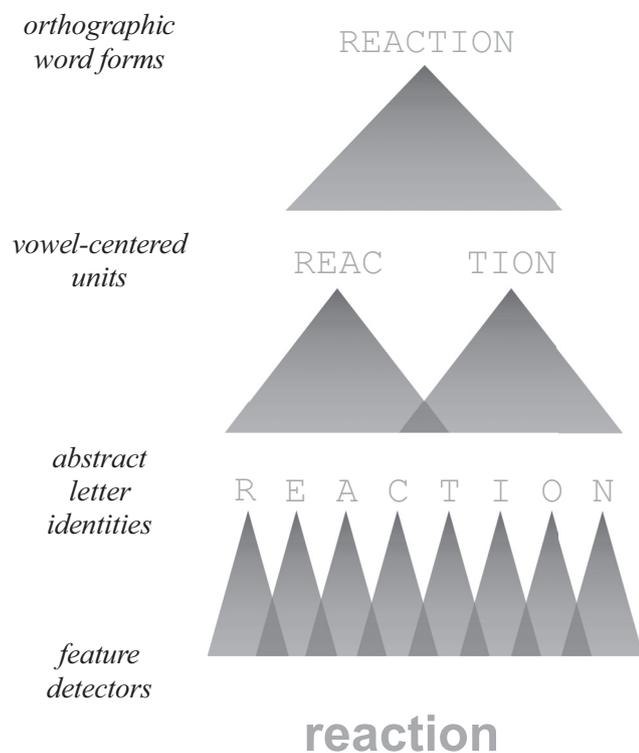
different matching task with the same design as in Chetail, Drabs et al. (2014), combined with MEG recording. We propose to use the simple variant of the task rather than the masked priming variant (see Norris & Kinoshita, 2008), as the former seems more sensitive to stimulus manipulations. For example, García-Orza, Perea, and Muñoz (2010) found no difference in letter transposition effects between letters, digits, and symbols when they combined masked priming with the same-different task, whereas Duñabeitia et al. (2012) did find an effect with the simple version of the task. Additionally, Duñabeitia et al. (2012) argued that the same-different task with letter transpositions allows for the in-depth study of basic perceptual processes, which is required to assess the time course of CV structure effects.

When comparing the processing of transposed letters pairs (e.g., *NDTF – NTDF*) to the usual substitution baseline (e.g., *NDTF – NSBF*), larger negativities were elicited in the 100–200 msec and 200–325 msec temporal windows (Duñabeitia et al., 2012) in the baseline condition compared to the transposition condition. The difference in the first time window was interpreted as an N/P150 component effect, reflecting early perceptual processes of mapping between visual features and higher-level representations. The difference in the second time window was taken as an N2 component effect, produced by a deviation between the –mentally stored– referent and the target. In the present study, if the arrangement of consonant and vowel letters constrains the perceptual and orthographic processing of letter strings at an early stage, we should find differences in these two windows. Thus the modified-structure condition (e.g., *BOUVRET – bovuret*) should elicit a larger waveform than the preserved-structure condition (e.g., *FOUVERT – fouvert*) since the CV structure is different in the former condition only (i.e., different number of vowel clusters), despite an identical letter manipulation (i.e., transposition of the two same letters, at the same position). As a control, based on Duñabeitia et al.'s (2012) ERP results, we assume that the baseline conditions (e.g., *BOUVRET – bociret* and *FOUVERT – fociert*) obtained by the replacement of the two focal letters should elicit larger waveforms than the corresponding transposition conditions in the same time windows. No difference between the preserved-structure and modified-structure replacement conditions was found in the behavioral study (Chetail, Drabs et al., 2014, experiment 5b), presumably because the replacement of two letters facilitates discrimination enough to bypass the influence of structure. Hence, we expect no difference between the two replacement conditions in cerebral activity either.

In addition to the time course of CV effects, the present experiment will examine the brain structures associated with the sensitivity to letter category. Visual word recognition is assumed to be subserved by two neural pathways, namely the left temporo-occipital ventral stream and the left temporo-parietal dorsal stream. The ventral occipitotemporal cortex is believed to be involved in fast and skilled processing of written words (see Dehaene & Cohen, 2011; Price & Devlin, 2011, for discussions). Especially, reading-related activations are located in the visual word form area (VWFA), a small region within the occipitotemporal sulcus lateral to the left fusiform gyrus. Dehaene, Cohen, Sigman, and Vinckier (2005) proposed that visual word recognition reflects the

simultaneous activity of a complex hierarchy of detectors in the ventral stream, each one being responsible for the coding of a certain element of information. Actually, comparing the activation caused by stimuli entailing increasing orthographic redundancy (e.g., from random strings of infrequent letters to words), Vinckier et al. (2007) reported fMRI evidence for such a hierarchy in the left occipitotemporal cortex, at the site of the VWFA. However, the exact nature of higher-order orthographic constituents is still under scrutiny (e.g., Dehaene et al., 2005; Lupker, Acha, Davis, & Perea, 2012; Taft & Krebs-Lazendic, 2013). As proposed by Chetail, Drabs et al. (2014), we assume that the structure determined by the arrangement of consonants and vowels shapes one level in this hierarchy of detectors, so that higher-order elements in the hierarchy correspond to groups of contiguous letters, centred on a vowel or vowel cluster (see Fig. 1). In our experiment, target stimuli in the two critical conditions (preserved- vs modified-structure letter transpositions) will differ in terms of the number of orthographic units based on vowel clusters (i.e., compare *bouret* and *fouert*). If the CV structure of words is processed by local detectors in the ventral stream, we expect a difference in brain activation between the two conditions specifically in the VWFA.

However, despite the doubtless involvement of the ventral occipitotemporal cortex in skilled reading, the mandatory status of this pathway and its exact role in written word processing is currently under debate (e.g., Carreiras, Quiñones, Hernández-Cabrera, & Duñabeitia, 2015; Price & Devlin, 2011), since the left parietal cortex has also been reported to be activated for letter identity and letter position



**Fig. 1** – Example of hierarchy of detectors involved in visual word processing (features, letters, vowel-centred units, words).

encoding in the early phases of visual word recognition. Thus, fMRI studies with the same-different task showed that transposed letter stimuli produced more activation than a baseline condition in the left inferior parietal gyrus and in the left superior parietal gyrus (e.g., Carreiras et al., 2015; see also; Reilhac, Peyrin, Démonet, & Valdois, 2013), whereas no difference was found in the VWFA. If the CV structure of letter strings is processed in the left dorsal stream, we expect the condition with preserved-structure transpositions to elicit more activation than the condition with modified-structure transpositions in the left inferior parietal gyrus and in the left superior parietal gyrus.

Contrary to previous fMRI and ERP studies (e.g., Carreiras et al., 2015), the MEG approach used in the present study will enable us to combine temporal and spatial information. Given that the hypothesis of the involvement of the left ventral pathway during the processing of CV structure does not exclude the potential involvement of the left dorsal pathway (and vice versa), tracking neural activity in these regions in different time windows will enable us to test the hypothesis that both pathways are recruited for the processing of visual stimuli, but at different stages of processing. Furthermore, it is worth mentioning that the targets of the two critical conditions in the experiment differ in terms of number of syllables in addition to the number of vowel clusters (preserved structure: *bouret*, /bɔ.vy.ʁe/ vs modified structure: *fouert*, /fɔ.vʁɛʁ/). Previous results showed, however, that effects of CV structure are independent from the phonological structure of letter strings. For example, the facilitatory effect for negative responses found in the same/different discrimination task is present whether the addition of one vowel cluster changes the number of syllables (e.g., *POIVRER* – *povirer*, /pwa.vʁe/-/pɔ.vi.ʁe/) or does not (e.g., *PEUPLIER* – *peupiler*, /pø.pli.je/-/pø.pi.le/) (Chetail, Drabs et al., 2014, Experiments 1–2). On the contrary, a mere change of number of syllables (i.e., without change of number of vowel clusters) does not modify the performance compared to a baseline condition (i.e., no change in the number of vowel clusters nor in the number of syllables, Chetail, Leberre, & Content, 2013). More generally, previous studies support the fact that CV letter effects are not confounded with phonological structure effects in behavioral tasks (e.g., Buchwald & Rapp, 2006; Caramazza & Miceli, 1990; Chetail & Content, 2014). The present study enables us to further test this claim, at a neuro-anatomical level. If orthographic structure effects are merely phonological structure effects, we should find an effect of number of syllables located in the left superior temporal cortex, between 200 and 600 msec (Wydell, Vuorinen, Helenius, & Salmelin, 2003) and/or in the left supramarginal gyrus (e.g., Church, Balota, Petersen, & Schlaggar, 2011) rather than effects located in the left fusiform gyrus and/or in the left inferior parietal/left superior parietal gyri in earlier time windows.

## 2. Method

### 2.1. Participants

All the participants were native French speakers and were right-handed (as confirmed by the French version of the

Edinburgh Handedness Inventory, Oldfield, 1971). They had normal or corrected-to-normal vision and they reported no prior history of neurological or language disorder. They gave informed consent to participate to the study and they received a financial compensation for their participation. The overall study was approved by the Biomedical Ethics Committee of the CUB—Hôpital Erasme. As declared prior to data collection (see Appendix), we recruited participants until a sample of 24 participants reaching the inclusion criteria was obtained. This led us to recruit and test 36 participants. One was excluded straight away (he/she did not feel well in the scanner and stopped the experiment before the end). Out of the 35 remaining participants, 11 could not be kept for analyses (and were therefore replaced gradually). Seven of them reached at least one of the exclusion criteria (i.e., more than 20% of MEG epochs rejected, more than 33% of errors in any condition, or rejection of more than 33% of trials in any condition due to the combination of errors and invalid epochs; these thresholds were set so that a sufficient number of valid trials was still available by design cell for each participant) and four of them had MEG data which were not usable due to technical problems during acquisition (i.e., head-tracking coils failure, problems in digitalization data).

## 2.2. Materials

Seven hundred and twenty pseudowords of seven letters were devised as referents (Table 1). Half were used for positive responses (fillers) and half for negative responses. Following the design of Chetail, Drabs et al. (2014, Experiment 5b), 120 triplets of referents were created for the negative responses. One referent included a –VVCC– or –CCVV– internal sequence so that the target constructed by the transposition of a consonant and a vowel had an additional vowel cluster (e.g., BOUVRET leading to *bovuret*, /buvrɛ/ - /bovyrɛ/: modified structure, after

letter transposition). A baseline target was devised by the replacement of the two transposed letters by two new letters (*fociert*, /fɔsjɛr/). The second referent of the triplet included a –VVCV– or –VCVV– sequence, so that transposing a consonant and a vowel did not alter the number of vowel clusters (e.g., FOUVERT-*fovuert*, /fuvrɛ/ - /fovyɛr/: preserved structure, after letter transposition). Again, a baseline target was devised by the replacement of the two transposed letters by two new letters (*bociret*, /bosirɛ/). A third set of referents used as fillers was devised (Chetail, Drabs et al., 2014). They had the same characteristics as the two other sets of referents (half with a VVCC/CCVV structure, half with a VVCV/VCVV structure) and were followed by an unrelated target with no identical letter at a same position (e.g., TOUVIRT-*lariaud*, /tuvir/ - /larjɔ/). In the experimental conditions (Table 1), the two transposed letters and their position were identical in the two letter transposition conditions (U and V here), as well as the two replaced letters in the letter replacement conditions (C and I here). Referents for the baseline, preserved structure, and modified structure conditions were matched on the number of phonemes, number of letters, number of vowels, orthographic similarity (OLD20), and summed bigram frequency (see Table 2). These variables were also matched across the two experimental conditions of targets (transposed letters vs replaced letters). Three hundred and sixty referents for positive responses were devised so that they were matched with the referents for negative responses on these factors, while sharing similar CV structures. Two lists of stimuli were used, with every referent appearing once in each list and an equal number of trials of the experimental target conditions.

## 2.3. Procedure

Participants were tested individually in a light-weight magnetically shielded room (MSR) (MaxShield, Elekta Oy, Finland). The stimuli were delivered via a DLP projector (Model PT-D7700E, Panasonic, New Jersey, USA) located outside the room. Items were back-projected on a screen inside the MSR via an optical periscope. They were presented in white against a black background on a computer screen at a distance of one meter. Once installed, participants performed a cross-case same-different task programmed in Matlab using the Psychtoolbox extension (Brainard, 1997). Each trial began with a centred string of three asterisks for 1,000 msec, followed by a blank period during 500 msec and then the referent in

**Table 1 – Type of items presented in the experiment.**

	Response		
	Yes		No
Type	Fillers	Experimental	Fillers
N	360	240	120
Example (REFERENT – target)	HOUPEUX – houpeux	FOUVERT – fovuert	TOUVIRT – lariaud

**Table 2 – Characteristics of the experimental items.**

	Preserved structure			Modified structure		
	Referent	Target with letter transposition	Target with letter replacement	Referent	Target with letter transposition	Target with letter replacement
Example	FOUVERT	fovuert	fociert	BOUVRET	bovuret	bociret
N	120	120	120	120	120	120
Number of vowels	3.12	3.12	3.12	3.12	3.12	3.12
Number of phonemes	5.92	6.01	5.95	5.86	6.09	6.05
OLD20	2.72	2.73	2.72	2.72	2.75	2.80
Summed bigram frequency <sup>a</sup>	20,482	13,100	12,796	19,238	12,846	13,326

Notes. <sup>a</sup>Based on token bigram frequencies (subtitles) computed on Lexique 3.80 (New, Pallier, Brysbaert, & Ferrand, 2004). To avoid repetition of referent, each participant will see half the targets with either transposition or replacement, making 60 observations per participant per cell.

uppercase for 500 msec. After a blank period ranging from 400 to 600 msec (random value), the target appeared for 500 msec. Participants were instructed to decide as rapidly and accurately as possible whether the referent and the target comprised the same sequence of letters (response ‘same’) or not (response ‘different’), by pressing the corresponding keys of a response pad within maximum 3,000 msec. Reaction times were measured from target onset until the keypress. The presentation of the three asterisks was also used to allow participants to blink. There was a 500 msec inter-trial interval (see Fig. 2). This timing of events was chosen to be as close as possible to the procedure used by Chetail, Drabs et al. (2014), but also to be suited for MEG recording. Especially, a limited presentation of targets was used here (500 msec vs until response in Chetail et al.) to avoid ocular saccades (see Reilhac et al., 2013). All participants performed twelve practice trials before receiving the 720 trials in a variable random order. Breaks of at least 1 min were introduced after each 80 trials. At this moment, a feedback on performance was provided to the participant to maintain motivation and accuracy. The experiment was approximately 60 min long.

#### 2.4. MEG acquisition

Visual evoked magnetic fields (VEFs) were recorded using a whole-scalp MEG (Triux, Elekta Oy, Helsinki, Finland) installed in a light-weight MSR (MaxShield; for more details about MaxShield, see De Tiège et al., 2008). The MEG system contains 102 triple sensor elements, each composed of a magnetometer (measuring the component of the magnetic field perpendicular to the helmet) and two orthogonal planar gradiometers (measuring the gradient of this component along two orthogonal directions tangential to the helmet), which exhibit different sensitivity profiles to the location, orientation, and strength of neural generators and different sensitivities to noise. In order to monitor the subjects' head position inside

the MEG helmet during the recording, four head-tracking coils were placed on subjects' head and localized with respect to anatomical fiducials (nasion and preauricular points) using an electromagnetic tracker (Fastrak, Polhemus, Colchester, VT, USA). Additionally, about 500 head-surface points (scalp, nose and face) were taken on average to ensure co-registration between subjects' head and cerebral magnetic resonance images (MRIs). Eye movements and blinks were monitored with vertical and horizontal electro-oculograms (EOGs). Electrocardiogram (ECG) was recorded using bipolar electrodes placed below the clavicles. EOGs and ECG signals were recorded synchronously with MEG signals. All signals were recorded using a bandpass-filter between .1 and 330 Hz and sampled at 1 kHz. Subjects' high-resolution 3D-T1 structural MRIs were acquired on a 1.5 T MRI scan (Intera, Philips, The Netherlands).

#### 2.5. MEG analyses

##### 2.5.1. Data preprocessing

Continuous MEG data were first preprocessed off-line with the signal space separation (SSS) method to subtract external interferences and correct from head movements (Taulu, Simola, & Kajola, 2005). Each subject's head was realigned into a common sensor space (mean of participants) using the SSS method. Second, the raw MEG, EOGs and ECG signals were band-pass filtered in the .1–30 Hz frequency range. Ocular, cardiac and electronic artefacts were removed from the raw MEG signals using an independent component analysis (ICA, see Vigario, Sarela, Jousmiki, Hämäläinen, & Oja, 2000), with artefactual components to be chosen based on their spatial topography and temporal correlation with EOG and ECG signals. To avoid large electronic jumps in the ICA, the continuous data were scanned using nonoverlapping 1,000 msec-long windows and each window presenting MEG signal amplitude exceeding predefined thresholds (3 pT for

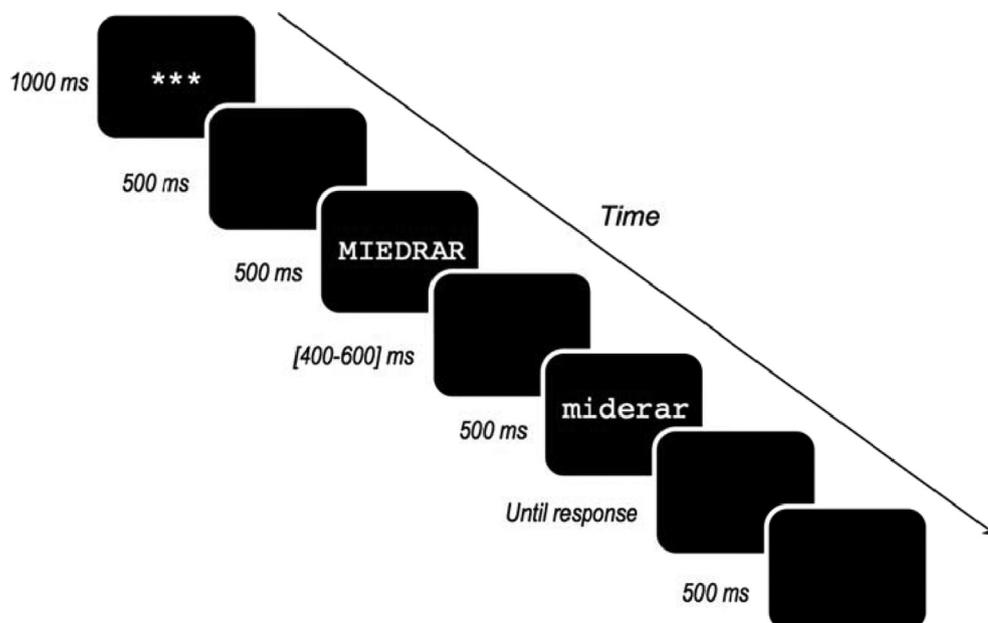


Fig. 2 – Sequence of events in each trial.

magnetometers, 7 pT/cm for gradiometers) was excluded from the ICA. Artefactual components were then projected out of the continuous MEG data (mean  $m = 2.75$  components per participant, standard deviation  $SD = .61$ ). Last, the cleaned MEG signals were epoched into 700 msec intervals, from 100 msec before to 600 msec after stimulus onset. Epochs exceeding the amplitude thresholds defined above were rejected from further analysis ( $m = 2.90\%$  of the data,  $SD = 4.40\%$ ). As declared at pre-registration, trials (epochs) with extreme reaction times or leading to error at the behavioural level were also removed from the MEG analyses ( $m = 6.64\%$ ,  $SD = 2.43\%$ ). Residual epochs were then averaged after baseline correction (from  $-100$  msec to  $0$  msec) within each participant and condition (i.e., preserved structure with transposition, modified structure with transposition, preserved structure with replacement, and modified structure with replacement), leaving four data sets of MEG evoked responses per subject.

To reconstruct the neural activity underlying the MEG data, individual MRIs were also preprocessed. First, each individual MRI was segmented using the Freesurfer software (Martinos Center for Biomedical Imaging, Massachusetts, USA) and manually co-registered to each subject's head using the three anatomical fiducial points for initial estimation and the head-surface points to manually refine the surface coregistration. Second, to investigate the neural correlates at the group level, a common source-space between subjects was produced using a 5 mm homogeneous source grid matching the Montreal Neurological Institute (MNI) brain and transformed onto individual MRIs using the non-linear spatial-normalization algorithm implemented in Statistical Parametric Mapping (SPM12, Wellcome Department of Cognitive Neurology, London, UK). Last, the MEG forward solution was computed for each subject's brain using the one-layer boundary element method implemented in the MNE software suite (Martinos Center for Biomedical Imaging, Massachusetts, USA).

### 2.5.2. Contrast analyses

As we had clear predictions about the time course of the expected effects, we restricted our analysis to the time window between 100 and 500 msec post-stimulus onset. Previous electrophysiological research has indeed shown that this period covers the N/P150, N2, and N400m effects (e.g., Almeida & Poeppel, 2013; Duñabeitia et al., 2012).

In order to deal with the issue of multiple spatiotemporal comparisons, statistical differences between two conditions were computed at the group level in the sensor space using permutation-based maximum cluster statistics (Maris & Oostenveld, 2007). Because magnetometers and planar gradiometers provide different physical measures with distinct sensitivity profiles, the two sets of sensors were treated separately. Furthermore, since planar gradiometers come in orthogonal pairs, we grouped each pair of sensors into a single effective gradiometer derived as their Euclidean norm (hence measuring the amplitude of the field's tangential gradient, independently of orientation). This allows applying a spatiotemporal clustering algorithm with the same neighbour structure for both the magnetometers and the effective gradiometers.

More precisely, spatiotemporal clusters were first derived by applying a t-test independently at each sensor and time point and thresholded at uncorrected significance level  $\alpha = .05$ . The weight of each cluster was computed as its summed t-values. The maximum cluster weight was taken as statistic and its null distribution was generated in a similar fashion from 10,000 simulated data built from random subject-wise permutations of the compared conditions (Maris & Oostenveld, 2007). The significance threshold at  $\alpha = .05$  was derived as the 95th percentile of this distribution, and the p-value of the statistic was estimated as the fraction of null samples exceeding the original maximum cluster. The effect size of the test was assessed by standardizing the maximum statistic into a Z-score based on the mean and standard deviation of the null distribution. This approach reveals spatiotemporal locations of cluster(s) showing statistically significant difference(s) within the time window of interest. Only temporal bins corresponding to significant clusters identified were used for further source-level analyses. Following this method, we performed two contrasts at the group level: preserved versus modified structure with transposition, and preserved versus modified structure with replacement.

To identify the underlying neural correlates, we applied an approach that has been used in previous MEG studies investigating the neural correlates of various cognitive processes; namely, we focused the source-space contrast analysis on the statistically significant periods produced by the sensor space cluster analysis (see, e.g., Furl, van Rijsbergen, Treves, Friston & Dolan, 2007; Furl, van Rijsbergen, Kiebel, Friston, Treves & Dolan, 2010; Furl et al., 2011; Galer et al., 2015; Litvak et al., 2011; Urbain et al., 2013, 2016). Specifically, individual MEG signals corresponding to statistically significant time periods in the sensor space were projected on each individual brain using Dynamic Statistical Parametric Mapping (dSPM, Dale & Sereno, 1993). The noise covariance matrix used to produce the inverse matrix was computed from an empty room measurement acquired on the acquisition day. The regularization parameter was set automatically using the prior consistency condition of Wens et al. (2015). To identify the source location of statistically significant differences, we contrasted the reconstructed dSPM maps of power activity temporally averaged within the windows of interest determined in the sensor space analyses. Significance was assessed using a non-parametric permutation test (with maximum statistics to avoid the multiple comparisons issue, Nichols & Holmes, 2002) in the windows of interest determined in the sensor space analyses. The significance threshold, p-value and effect size were estimated similarly to those of the sensor space maximum cluster.

## 3. Results

### 3.1. Behavioural results

Trials with extreme reaction times (below 250 msec or above the deadline of 3,000 msec, see Chetail, Drabs et al., 2014) were excluded (.10% of the data). The mean correct reaction times and mean error rates averaged over the 24 participants are presented in Table 3. A repeated measures analysis of

**Table 3 – Mean reaction times (in msec) for correct responses and percentage of errors (brackets) on target words.**

	Preserved structure	Modified structure	Difference
Transposition	892 (15.1%)	850 (13.6%)	42 msec (1.5%)
Replacement	644 (1.3%)	631 (1.5%)	13 msec (-.2%)

Notes. Mean reaction times for the ‘same’ responses and the ‘different’ responses in the baseline condition were 707 msec and 580 msec respectively (7.7% and 1.1% in error rates respectively).

variance (ANOVA) with modification type as factor showed a significant difference between items with a letter replacement and those with a letter transposition [reaction times:  $F(1, 23) = 110.20, p < .001, \eta_p^2 = .83$ , error rates:  $F(1, 23) = 83.77, p < .001, \eta_p^2 = .78$ ], guaranteeing that participants were sensitive to the experimental manipulation (letter transposition).

The data presented in Table 3 were submitted to separate two-way ANOVAs with structure (preserved or modified) and modification type (transposition vs replacement) as main factors. Two simple effects were tested. The conditions of preserved and modified structure were compared in the transposition condition on the one hand, and in the replacement condition on the other hand. As planned, the interaction between the two factors was to be tested only if both simple effects were significant.

The reaction time analysis in the transposition condition showed that the modified-structure items were processed more rapidly than preserved-structure items,  $F(1, 23) = 16.86, p < .001, \eta_p^2 = .42$ . The effect was also present in the replacement condition,  $F(1, 23) = 8.41, p = .008, \eta_p^2 = .27$ . Critically, however, the interaction between structure and modification type was significant, indicating that effect of structure was more pronounced in the transposition condition than in the replacement condition,  $F(1, 23) = 6.81, p = .016, \eta_p^2 = .23$ .

In the error analyses, there was no effect of structure, be it in the transposition condition,  $F(1, 23) = 1.20, p = .29, \eta_p^2 = .01$ , or in the replacement condition,  $F(1, 23) = .21, p = .65, \eta_p^2 = .05$ .

### 3.2. MEG results

#### 3.2.1. Outcome-neutral tests for neural activity

Regarding neural activity, a first positive control was provided by examination of the presence of clear signal-to-noise peaks in the three time windows of interest (100–200, 200–325, 350–500 msec, e.g., Chen, Davis, Pulvermüller, & Hauk, 2013; Chen, Davis, Pulvermüller, & Hauk, 2015; Hauk, Coutout, Holden, & Chen, 2012). Fig. 3 illustrates the signal-to-noise ratio of the MEG evoked responses averaged over all subjects and conditions. Clear evoked activity in the expected time windows emerged from the baseline (maximum evoked amplitude > 10 times baseline amplitude) mainly over the occipital, parietal, and temporal sensors. Three distinct phases of activity can be seen in the butterfly plots (Fig. 3b and c), roughly ranging respectively from 80 to 130 msec, 130–350 msec and 350–600 msec.

A second marker of the sensitivity of the study is the presence of differences between the replacement and

transposition conditions. As mentioned previously, based on Duñabeitia et al.’s (2012) ERP results, we assumed that the baseline conditions obtained by the replacement of the two focal letters should elicit larger waveforms than the corresponding transposition conditions. The sensor-space comparison of the replacement and the transposition conditions consistently demonstrated significant differences for magnetometers (i.e., replacement > transposition),  $Z = 12.54, p < .001$ , in the respective group-level evoked responses. The statistical analysis indeed disclosed one significant spatiotemporal cluster from 204 to 500 msec post-stimulus onset. No significant contrast was detected at the level of gradiometer amplitudes ( $Z = .17, p = .29$ ).

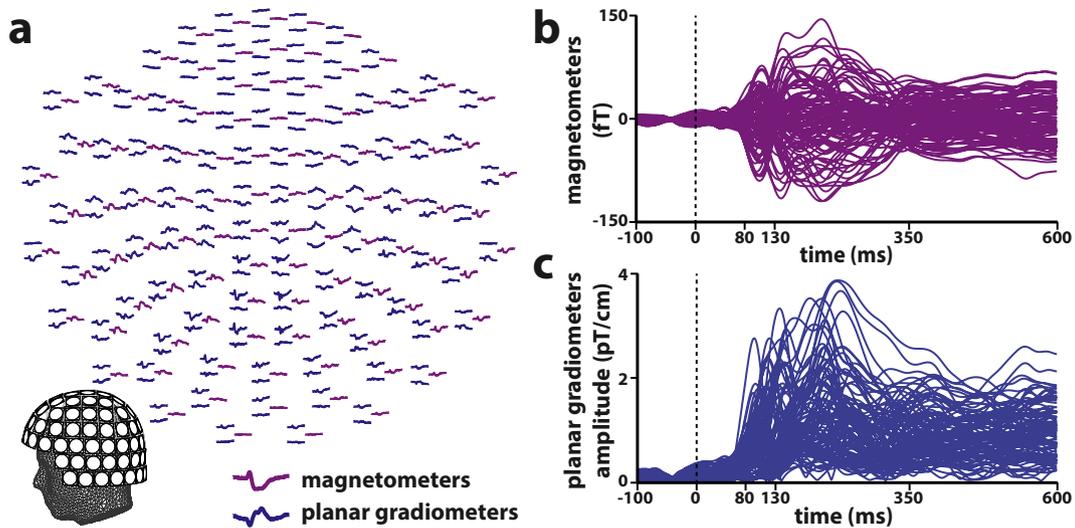
#### 3.2.2. Sensor- and source-space results

Fig. 4 depicts the sensor- and the source-space results obtained from the comparison of preserved and modified structures with transposition. Significant sensor-space differences were observed for gradiometer amplitudes (i.e., preserved > modified with transposition),  $Z = 2.30, p = .034$ , with a single spatiotemporal cluster emerging from 129 to 239 msec post-stimulus onset (Fig. 4a). No significant contrast was detected at the level of magnetometers ( $Z = -.47, p = .64$ ). At the source level, however, no significant difference was observed when considering mean source power over the entire time window disclosed by this cluster ( $Z = .63, p = .24$ ).

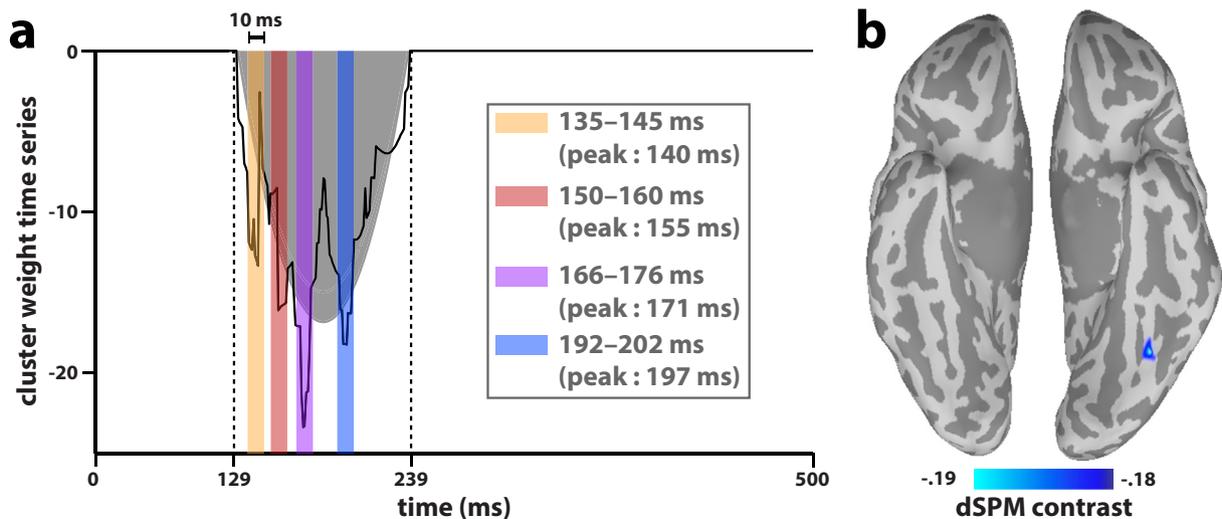
When comparing the preserved and modified structure with replacement, no significant sensor-space difference was observed ( $Z = .30, p = .25$  for magnetometers;  $Z = -.79, p = .87$  for gradiometer amplitudes). As planned in the registration procedure, the interaction between the structure (preserved vs modified) and modification (transposition vs replacement) factors was therefore not tested further.

#### 3.2.3. Post-hoc analysis

Coming back to the comparison of preserved and modified structures with transposition and in view of the discordance between the presence of sensor-space contrast and absence of source-level difference, we performed further analyses to increase the sensitivity of our source-space contrast test, which may be conservative (Nichols & Holmes, 2002). Furthermore it is plausible that the neural activity underlying the sensor-space difference is short-lived, so that averaging over the 110 msec-long cluster-based window (129–239 msec post-stimulus onset) would hamper the detection of significant differences. With this hypothesis in mind, we revisited the sensor-level cluster weight time series (Fig. 4a) and repeated our source analysis within short, 10 msec-long time windows centred over the main cluster weight peaks. We selected those peaks that exceeded the 95% confidence interval of a parabolic fit of the cluster weight time course so as to detect transient deviations from the global dynamics of the cluster whereby it first grows and then shrinks monotonically. This yielded four periods of interest (135–145 msec, 150–160 msec, 166–176 msec, and 192–202 msec post-stimulus onset, see Fig. 4a). A significant contrast in source dSPM power activity emerged within the late 192–202 msec time window at the left VWFA (Fig. 4b, MNI coordinates of maximum peak:  $[-42 -64 -20]$  mm,  $Z = 2.28, p = .023$ ). The earlier windows did not disclose significant differences (135–145 msec:  $Z = 1.78,$



**Fig. 3** – Illustration of the signal-to-noise ratio of MEG evoked fields averaged across all subjects and conditions. (a) Topographical view of the response in all sensors (magnetometers and gradiometers, units set as in parts b and c) from 100 msec pre-stimulus to 600 msec post-stimulus onset. The sensor array is viewed from the top with the nose pointing upwards. For reference regarding sensors localization, the average head position used for data realignment is shown in the bottom left. (b) Butterfly plot of the 102 magnetometers evoked responses. (c) Butterfly plot of the associated 102 gradiometers amplitudes (i.e., Euclidean norm of each pair of orthogonal planar gradiometers). The stimulus onset at time zero is marked by a vertical line in both butterfly plots.



**Fig. 4** – Sensor- and source-space statistical comparison of evoked responses in the preserved- and modified-structure items with letter transposition. (a) Time development of the significant sensor-space cluster, as measured by its weight time series (computed as the total  $t$ -value summed over all sensors in the cluster as a function of time). The 129–239 msec time window of significance appears as its deviation from zero and is delimited by two vertical lines. The one-sided 95% confidence region of a parabola fitted to the cluster weight time series and used to detect the main peaks, and the 10 msec-long periods of interest around the resulting peaks used for the subsequent source-space localizations are emphasized by shaded areas. (b) Localization of significant source-space differences in mean dSPM power activity within the 192–202 msec time window projected on the MNI brain and statistically masked via non-parametric maximum statistic at  $p < .05$ . The inflated MNI brain is viewed from its bottom with the anterior part pointing to the top.

$p = .052$ ; 150–160 msec:  $Z = .70$ ,  $p = .22$ ; 166–176 msec:  $Z = 1.24$ ,  $p = .11$ ). Of notice, the trend for significant contrast in the early 135–145 msec window localized at the left primary sensorimotor (SM1) cortex (MNI coordinates of maximum peak:  $[-40 -11 30]$  mm).

#### 4. Discussion

Although most psycholinguists would agree that long letter strings need to be parsed into smaller letter groups, and much

research effort has been devoted to delineate the nature of the units involved in visual word recognition, no consensus has yet been reached. In a series of recent studies, Chetail and Content argued that the organisation of consonant and vowel letters constitutes one determinant of orthographic parsing. They offered convergent behavioural evidence to the hypothesis that each group of adjacent vowel letters serves as the core of one perceived orthographic unit (e.g., Chetail & Content, 2012, 2014; Chetail, Drabs et al., 2014). The present work attempted to overcome two limitations of these previous studies. First, we wanted to establish that the extraction of vowel-centred units takes place during early stages of letter string processing. The second objective was to confirm that the vowel-centred structure pertains to the word recognition system and hence that it may constitute one level in a hierarchy of neural detectors coding orthographic strings. Recording MEG data made it possible to go beyond the limited interpretations that can be generated from behavioural data, and enabled us to decide on both the timing and the cortical source location of the activity underpinning the CV structure effects, thus addressing the two limits of previous work.

The present experiment was based on the same design and logic as Chetail, Drabs et al. (2014, Experiment 5b). A new and larger set of pseudowords was constructed with stringent controls (orthographic neighbourhood, number of phonemes, number of vowels, summed bigram frequency, morphological structure). The behavioural outcome replicates the major finding reported in the previous study. Participants were significantly faster to detect a mismatch when the number of orthographic units was modified by a letter transposition than when this number remained identical (e.g., BOUVRET – *bovuret* vs. FOUVERT – *fouvert*). However, a significant difference between the preserved-structure and modified-structure was also found in the replacement condition (e.g., BOUVRET - *bociret* vs. FOUVERT - *fociert*), but the difference was much smaller than in the transposition condition. No significant effect was found on error rates. The absence of any effect on error rates is not surprising, given that instructions explicitly emphasized accuracy over speed, and that participants producing more than 33% of errors in any of the conditions were excluded. Although it was not observed in Chetail and Content (2014), the finding of a small response time difference in the replacement condition might suggest that the structure change facilitates discrimination, despite the saliency of letter substitutions. Nevertheless, the presence of a significant interaction clearly indicates the influence of structure, over and above the letter changes.

Regarding the MEG results, we found a significant difference between the letter replacement and letter transposition conditions, starting 204 msec after target onset. This effect thus reflects the accumulation of neural information indicating a mismatch between the memorized referent and the target stimulus and provides a temporal reference point for the onset of orthographic processing in the brain. The observation brings further neurophysiological evidence to the view that letter identity extraction and precise position coding have distinct time courses and that letter identity information is more salient (see Perea & Lupker, 2003, for a discussion). It is concordant with Duñabeitia et al. (2012) who reported larger negativities in ERPs for the replacement condition than for the

transposition condition in the 200–325 msec post-stimulus temporal window. More generally, it is also in agreement with numerous MEG and EEG studies showing specific activity for letter strings relative to non-linguistic symbols from around 150 msec on (e.g., Cohen et al., 2000; Rey, Massol, Dufau, & Grainger, 2009; Nobre, Allison, & McCarthy, 1994; Tarkiainen, Cornelissen, & Salmelin, 2002; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999).

More importantly, the critical comparison in the present study concerns the effect of vowel-centred orthographic structure and involves the preserved- and modified-structure trials in the transposition condition. A significant effect of structure was observed from 129 msec to 239 msec post-stimulus onset. This finding confirms the hypothesis that the orthographic structure based on consonant/vowel organization is already available during the earliest phases of letter string processing. In fact, given the reference point offered by the transposition/replacement comparison, it suggests that the extraction of vowel-centred structure starts extremely early with regard to letter identification and orthographic processing. Interestingly, the data might even be taken to indicate that structure information accrues *earlier* than precise letter identification.

It is noteworthy that the mismatch between the letter replacement and letter transposition conditions appeared significant at the level of magnetometers, whereas the effect of structure in the transposition condition reached significance at the level of gradiometer amplitudes. This may be due to the fact that these two sensor types exhibit different sensitivity profiles, so that one may be better suited to detect the neural activity underlying the mismatch in one case or the other, and behave differently regarding noise and averaging. In particular, using gradiometer amplitudes in the clustering analysis may increase the group-level signal-to-noise ratio by eliminating the inter-subjects variability of source orientation, but at the cost of missing directional information.

Regarding the localization of the structure effect in the transposition condition, no significant source could be reconstructed when the full 129–239 msec time window was considered. Methodologically, the discrepancy with sensor-space analysis relates to the distinct statistical approaches used. The clustering approach represents a standard way of handling sensor-space spatiotemporal data and naturally deals with the problem of multiple comparisons by taking into account both the magnetic field spatial spread and the temporal smoothness of evoked responses (Maris & Oostenveld, 2007). As such, the identification of differences appears both spatially and temporally extended and leads to large spatiotemporal clusters, because their formation was based on a lenient (i.e., uncorrected) statistical threshold. At the source level, on the other hand, a typical statistical strategy taken from functional imaging consists in examining the maximum difference of mean power activity (Nichols & Holmes, 2002). This approach deals with the spatial multiple comparisons issue and takes into account the spatial interdependencies of the reconstructed source data. It typically identifies less extended regions of significance because the statistical threshold is intrinsically corrected, which may also lead to lesser sensitivity than the cluster-based analysis (Maris & Oostenveld, 2007). Furthermore, it can also be overly conservative, especially over long

time windows, because temporally averaging source-level maps prior to their comparison may hamper the identification of temporally localized mismatch generators. For this reason, and notwithstanding the differences in dealing with spatial smoothness, we hypothesized that working with smaller time windows would allow identifying one or multiple transient effects that would underlie the sensor-level cluster of interest. Accordingly, the post-hoc contrast analysis focusing on selected time windows disclosed a short-lived significant contrast around 197 msec post-stimulus in the left fusiform gyrus (MNI coordinates:  $[-42 -64 -20]$  mm). This result was found with a limited number of post-hoc tests carried out when considering the four 10 ms-wide time windows centred on the cluster weight peaks that deviated from a parabolic fit, which idealized the situation where the cluster merely grows, reaches a maximum size, and then shrinks. Indeed the presence of multiple peaks deviating from this global dynamics would suggest different phases of cluster activity, and we reasoned that they would be associated with distinct short-lived neural generators whose intensity profiles would mix at the sensor-level, drive the cluster weight dynamics, and induce the transient peaks observed.

The localization of this significant contrast peak in the left fusiform gyrus around 200 msec corresponds very precisely to the region identified as the VWFA in many studies (e.g., Cohen et al., 2000; Cohen et al., 2002; Dehaene et al., 2004; Polk & Farah, 1997; Vinckier et al., 2007). This region is assumed to subserve the representation of the abstract sequence of elements composing letter strings, encompassing a hierarchy of detectors from simple letter features to multiletter substrings (e.g., frequent bigrams, trigrams, quadrigrams; Dehaene et al., 2005). In support of this view, Vinckier et al. (2007) identified a region of the fusiform gyrus extending along the anteroposterior axis from  $y = -80$  to  $-40$  mm (MNI coordinates) with a functional gradient of selectivity from random letter strings to strings made of recurrent letter clusters. The fact that the source we identified was located midway along this axis ( $y = -64$  mm) supports the hypothesis that vowel-centred structure constitutes an integral part of the representational code calculated by the visual word processing system.

The source localization analysis did not provide any indication of activity in either the left superior temporal cortex (Wydeell et al., 2003) or the left supramarginal gyrus (e.g., Church et al., 2011), two regions associated with phonological coding in previous MEG studies, nor in the dorsal stream (left inferior and superior parietal gyri, Carreiras et al., 2015; Reilhac et al., 2013), which is often believed to reflect top-down attentional factors. In the present experiment the modified- and preserved-structure items differed systematically in the number of syllables. Notwithstanding the possible sensitivity limitations associated with the spatial multiple comparisons involved when contrasting functional images (as discussed above), we argue that the failure to detect these regions is consistent with previous findings. The absence of indication of phonological activity reinforces previous demonstrations that vowel cluster effects cannot be accounted for by a phonological mismatch (Chetail et al., 2013; Chetail, Drabs et al., 2014; see also; Buchwald & Rapp, 2006; Caramazza & Miceli, 1990). Moreover, the observation that the manipulation of structure is associated with the ventral

rather than with the dorsal pathway is consistent with the view that the extraction of structure pertains to the fast automatic perceptual processes involved in fluent reading (e.g., Dehaene et al., 2005; Tarkiainen et al., 1999).

In sum, the results support the hypothesis that orthographic structure, as defined in terms of the CV pattern, is extracted during the very earliest phases of letter string processing and is encoded in left fusiform regions devoted to visual word recognition. This conclusion is in line with the general scheme of Dehaene et al.'s (2005) proposal and offers a more precise definition of the intermediate units between letters and word forms.

---

## Authors' note

This study was funded by the National Belgian Fund for Scientific Research (F.R.S.-FNRS) (grant no 19536975 allocated to AC and FC). XDT is post-doctorate clinical master specialist at the National Belgian Fund for Scientific Research (F.R.S.-FNRS, Brussels, Belgium). FC was a post-doctoral researcher at the F.R.S.-FNRS at the time the study started. MR was a post-doctoral researcher funded by the Interuniversity Attraction Poles Program of the Belgian Science Policy Office (Project P7/33) during the project. The MEG project at CUB — Hôpital Erasme is financially supported by the Fonds Erasme (convention “Les voies du savoir”, Brussels, Belgium).

The behavioural raw data and scripts for analyses are available on Open Science Framework, as well as links to access the raw and pre-processed MEG data (<https://osf.io/c8s5q/>).

---

## Appendix

### *Power analyses to define the size of the sample a priori*

For the behavioral data, a prospective power analysis was run with G\*power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) based on the results of Chetail, Drabs et al. (2014), Chetail, Scaltritti et al. (2014). The critical RT difference between preserved-structure transpositions and modified-structure transpositions was significant in all six experiments, with effect sizes between .45 and 1.35 (mean = .76) across participants. Assuming that the real effect is at least .70, a sample of 24 participants would be sufficient to detect it with a power of .90. However, if we consider the lower bound estimate of the effect size (.45), 54 participants would be necessary to detect it with a power of .90. On the other hand, we also estimated the number of participants required to reach a value of the Bayes factor demonstrating adequate sensitivity (i.e.,  $B > 3$ , see Dienes, 2014), based on Experiment 5b of Chetail, Drabs et al. (2014), Chetail, Scaltritti et al. (2014) which entailed exactly the same design as the present experiment. In that experiment, the mean behavioral difference was 40 msec (903 vs 863 msec for the preserved vs modified-structure conditions, respectively), and the standard deviation of the difference was 87 msec. Note that the effect in that experiment was the smallest of the whole series. Using a uniform distribution from 0 to 90 msec (conservative values, 90 msec

corresponding to the largest magnitude of the effect in Chetail, Drabs et al., 2014; Chetail, Scaltritti et al., 2014), and assuming a sample mean of 40 msec, we used Dienes' Bayes factor calculator (Dienes, 2014) to estimate the sample standard error required to obtain a Bayes factor above 3 or below 1/3. A Bayes factor of 3.05 was reached with a sample standard error of 22 msec, which corresponds to 16 participants ( $sd_2/se_2 = 15.6$ ). Based on these different analyses, 24 participants were initially recruited. If necessary for the behavioral analyses, additional participants did the experiment (outside of the scanner) until the Bayes factor reaches 3 or 1/3, with an upper limit of 54 participants. For the error rates, the effect was significant in 5 out of 6 experiments and ranged between .36 and 1.16 (mean = .68). As instructions will be designed to minimize errors in order to maintain the largest possible number of trials per subject, no specific sample size constraint was computed from error rates.

To our knowledge there is no published MEG study using a similar task, design and analysis. To estimate the number of participants required for the MEG analyses, we therefore relied on the study of Duñabeitia et al. (2012), which is the closest experiment to ours (same task, similar conditions, and analyses in three similar time windows). To calculate the Bayes factor, we estimated a plausible standardized effect size from the reported F values for the pairwise letter replacement and letter-substitution comparisons (respectively 12.63, 11.87, and 7.00 for the time windows 100–200, 200–325, and 350–500 msec, with 23 participants). The resulting Fisher z transformed r values (see Dienes, 2014, Notes 3 and 4) are respectively .699, .681 and .538. Based on these values, we therefore consider it reasonable to expect an effect around .40. Using a uniform distribution from 0 to .80 (twice the size of the expected standardized effect), and a sample standard error of .213 (based on  $N = 24$ ), we obtain a Bayes Factor of 3.66. For the MEG analyses, 24 participants were therefore be recruited. Note however that this number of participants planned is larger than what is typically reported in recent visual word recognition MEG studies (i.e., 12–20, Almeida & Poeppel, 2013; Assadollahi & Pulvermüller, 2003; Chen et al., 2013, 2015; Cornelissen et al., 2009; Hauk et al., 2012; Pyllkänen & Okano, 2010; Simos et al., 2009; Tsigka, Papadelis, Braun, & Miceli, 2014).

## REFERENCES

- Adams, M. J. (1981). What good is orthographic redundancy? In O. J. L. Tzeng, & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 197–221). Hillsdale: Lawrence Erlbaum Associates.
- Almeida, D., & Poeppel, D. (2013). Word-specific repetition effects revealed by MEG and the implications for lexical access. *Brain and Language*, 127, 497–509. <https://doi.org/10.1016/j.bandl.2013.09.013>.
- Assadollahi, R., & Pulvermüller, F. (2003). Early influences of word length and frequency: A group study using MEG. *NeuroReport*, 14, 1183. <https://doi.org/10.1097/01.wnr.0000075305.76650.60>.
- Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review*, 102, 146–184. <https://doi.org/10.1037/0033-295X.102.1.146>.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436. <https://doi.org/10.1163/156856897X00357>.
- Bruner, J. S., & O'Dowd, D. (1958). A note on the informativeness of parts of words. *Language and Speech*, 1, 98–101. <https://doi.org/10.1177/002383095800100203>.
- Buchwald, A., & Rapp, B. (2006). Consonants and vowels in orthographic representations. *Cognitive Neuropsychology*, 23, 308–337. <https://doi.org/10.1080/02643290442000527>.
- Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition*, 37, 243–297. [https://doi.org/10.1016/0010-0277\(90\)90047-N](https://doi.org/10.1016/0010-0277(90)90047-N).
- Carreiras, M., Duñabeitia, J. A., & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cerebral Cortex*, 19, 2659.
- Carreiras, M., Quiñones, I., Hernández-Cabrera, J. A., & Duñabeitia, J. A. (2015). Orthographic coding: Brain activation for letters, symbols, and digits. *Cerebral Cortex*, 25(12), 4748–4760. <https://doi.org/10.1093/cercor/bhu163>.
- Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Verbal Behavior*, 18, 225–241. [https://doi.org/10.1016/S0022-5371\(79\)90136-1](https://doi.org/10.1016/S0022-5371(79)90136-1).
- 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. *Frontiers in Human Neuroscience*, 7. <https://doi.org/10.3389/fnhum.2013.00376>.
- Chen, Y., Davis, M. H., Pulvermüller, F., & Hauk, O. (2015). Early visual word processing is flexible: Evidence from spatiotemporal brain dynamics. *Journal of Cognitive Neuroscience*, 1–14. [https://doi.org/10.1162/jocn\\_a\\_00815](https://doi.org/10.1162/jocn_a_00815).
- Chetail, F., & Content, A. (2012). The internal structure of 'chaos': Letter category determines visual word perceptual units. *Journal of Memory and Language*, 67, 371–388. <https://doi.org/10.1016/j.jml.2012.07.004>.
- Chetail, F., & Content, A. (2014). What is the difference between OASIS and OPERA? Roughly five pixels orthographic structure biases the perceived length of letter strings. *Psychological Science*, 25, 243–249. <https://doi.org/10.1177/0956797613500508>.
- Chetail, F., Leberre, J., & Content, A. (2013). The role of the consonant/vowel pattern in the perceptual discrimination of letter strings. In *Oral communication at the 18th conference of the European Society for Cognitive Psychology (ESCoP)*. Budapest, August 29th-September 1st.
- Chetail, F., Drabs, V., & Content, A. (2014). The role of consonant/vowel organization in perceptual discrimination. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 938–961. <https://doi.org/10.1037/a0036166>.
- Chetail, F., Scaltritti, M., & Content, A. (2014). Effect of the consonant–vowel structure of written words in Italian. *The Quarterly Journal of Experimental Psychology*, 67, 833–842. <https://doi.org/10.1080/17470218.2014.898668>.
- Church, J. A., Balota, D. A., Petersen, S. E., & Schlaggar, B. L. (2011). Manipulation of length and lexicality localizes the functional neuroanatomy of phonological processing in adult readers. *Journal of Cognitive Neuroscience*, 23, 1475–1493. <https://doi.org/10.1162/jocn.2010.21515>.
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M., et al. (2000). The visual word form area – Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, 123, 291–307. <https://doi.org/10.1093/brain/123.2.291>.
- Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, 125(5), 1054–1069.
- 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, e5359. <https://doi.org/10.1371/journal.pone.0005359.t001>.
- Cubelli, R. (1991). A selective deficit for writing vowels in acquired dysgraphia. *Nature*, 353, 258–260. <https://doi.org/10.1038/353258a0>.
- Dale, A. M., & Sereno, M. I. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience*, 5, 162–176. <https://doi.org/10.1162/jocn.1993.5.2.162>.
- De Tiège, X., de Beeck, M. O., Funke, M., Legros, B., Parkkonen, L., Goldman, S., et al. (2008). Recording epileptic activity with MEG in a light-weight magnetic shield. *Epilepsy Research*, 82, 227–231. <https://doi.org/10.1016/j.eplepsyres.2008.08.011>.
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15, 254–262. <https://doi.org/10.1016/j.tics.2011.04.003>.
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, 9, 335–341. <https://doi.org/10.1016/j.tics.2005.05.004>.
- Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., et al. (2004). Letter binding and invariant recognition of masked words. *Psychological Science*, 15(5), 307. <https://doi.org/10.1111/j.0956-7976.2004.00674.x>.
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Quantitative Psychology and Measurement*, 5, 781. <https://doi.org/10.3389/fpsyg.2014.00781>.
- Duñabeitia, J. A., & Carreiras, M. (2011). The relative position priming effect depends on whether letters are vowels or consonants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1143–1163. <https://doi.org/10.1037/a0023577>.
- Duñabeitia, J. A., Dimitropoulou, M., Grainger, J., Hernández, J. A., & Carreiras, M. (2012). Differential sensitivity of letters, numbers, and symbols to character transpositions. *Journal of Cognitive Neuroscience*, 24, 1610–1624. [https://doi.org/10.1162/jocn\\_a\\_00180](https://doi.org/10.1162/jocn_a_00180).
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>.
- Furl, N., Kumar, S., Alter, K., Durrant, S., Shawe-Taylor, J., & Griffiths, T. D. (2011). Neural prediction of higher-order auditory sequence statistics. *NeuroImage*, 54(3), 2267–2277. <https://doi.org/10.1016/j.neuroimage.2010.10.038>.
- Furl, N., van Rijsbergen, N. J., Kiebel, S. J., Friston, K. J., Treves, A., & Dolan, R. J. (2010). Modulation of perception and brain activity by predictable trajectories of facial expressions. *Cerebral Cortex*, 20(3), 694–703. <https://doi.org/10.1093/cercor/bhp140>.
- Furl, N., van Rijsbergen, N. J., Treves, A., Friston, K. J., & Dolan, R. J. (2007). Experience-dependent coding of facial expression in superior temporal sulcus. *Proceedings of the National Academy of Sciences*, 104(33), 13485–13489. <https://doi.org/10.1073/pnas.0702548104>.
- Galer, S., Op De Beeck, M., Urbain, C., Bourguignon, M., Ligot, N., Wens, V., et al. (2015). Investigating the neural correlates of the stroop effect with magnetoencephalography. *Brain Topography*, 28, 95–103. <https://doi.org/10.1007/s10548-014-0367-5>.
- García-Orza, J., Perea, M., & Muñoz, S. (2010). Are transposition effects specific to letters? *The Quarterly Journal of Experimental Psychology*, 63, 1603–1618. <https://doi.org/10.1080/174702109.03474278>.
- Hansen, D., & Rodgers, T. (1965). An exploration of psycholinguistic units in initial reading. In *Proceedings of the symposium on the psycholinguistic nature of reading process*. Detroit: Wayne State University Press.
- 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. <https://doi.org/10.1016/j.neuroimage.2012.01.061>.
- Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., et al. (2011). EEG and MEG data analysis in SPM8. *Computational Intelligence and Neuroscience*, 2011, 852961. <https://doi.org/10.1155/2011/852961>.
- Lupker, S. J., Acha, J., Davis, C. J., & Perea, M. (2012). An investigation of the role of grapheme units in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 1491–1516. <https://doi.org/10.1037/a0026886>.
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *The Behavioral and Brain Sciences*, 21, 499–511.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- Massol, S., Duñabeitia, J. A., Carreiras, M., & Grainger, J. (2013). Evidence for letter-specific position coding mechanisms. *PLoS One*, 8, e68460. <https://doi.org/10.1371/journal.pone.0068460>.
- Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingua E Linguaggio*, 2, 203–230. <https://doi.org/10.1418/10879>.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36, 516–524. <https://doi.org/10.3758/BF03195598>.
- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, 15, 1–25. <https://doi.org/10.1002/hbm.1058>.
- Nobre, A. C., Allison, T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature*, 372(6503), 260–263. <https://doi.org/10.1038/372260a0>.
- Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General*, 137, 434–455. <https://doi.org/10.1037/a0012799>.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4).
- Perea, M., & Acha, J. (2009). Does letter position coding depend on consonant/vowel status? Evidence with the masked priming technique. *Acta Psychologica*, 130, 127–137. <https://doi.org/10.1016/j.actpsy.2008.11.001>.
- Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition*, 31(6), 829–841. <https://doi.org/10.3758/BF03196438>.
- Polk, T. A., & Farah, M. J. (1997). A simple common contexts explanation for the development of abstract letter identities. *Neural Computation*, 9(6), 1277–1289. <https://doi.org/10.1162/neco.1997.9.6.1277>.
- Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, 15, 246–253. <https://doi.org/10.1016/j.tics.2011.04.001>.
- Pylkkänen, L., & Okano, K. (2010). The nature of abstract orthographic codes: Evidence from masked priming and magnetoencephalography. *PLoS One*, 5, e10793. <https://doi.org/10.1371/journal.pone.0010793>.
- Reilhac, C., Peyrin, C., Démonet, J.-F., & Valdois, S. (2013). Role of the superior parietal lobules in letter-identity processing within strings: FMRI evidence from skilled and dyslexic readers. *Neuropsychologia*, 51, 601–612. <https://doi.org/10.1016/j.neuropsychologia.2012.12.010>.

- Rey, A., Dufau, S., Massol, S., & Grainger, J. (2009). Testing computational models of letter perception with item-level event-related potentials. *Cognitive Neuropsychology*, 26(1), 7–22. <https://doi.org/10.1080/09541440802176300>.
- Shimron, J. (1993). The role of vowels in reading: A review of studies of english and Hebrew. *Psychological Bulletin*, 114, 52.
- Simos, P. G., Pugh, K. R., Mencl, E., Frost, S., Fletcher, J. M., Sarkari, S., et al. (2009). Temporal course of word recognition in skilled readers: A magnetoencephalography study. *Behavioural Brain Research*, 197, 45–54. <https://doi.org/10.1016/j.bbr.2008.07.038>.
- Taft, M., & Krebs-Lazendic, L. (2013). The role of orthographic syllable structure in assigning letters to their position in visual word recognition. *Journal of Memory and Language*, 68, 85–97. <https://doi.org/10.1016/j.jml.2012.10.004>.
- Tarkiainen, A., Cornelissen, P. L., & Salmelin, R. (2002). Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. *Brain*, 125(5), 1125–1136. <https://doi.org/10.1093/brain/awf112>.
- Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain*, 122(11), 2119–2132. <https://doi.org/10.1093/brain/122.11.2119>.
- Taulu, S., Simola, J., & Kajola, M. (2005). Applications of the signal space separation method. *IEEE Transactions on Signal Processing*, 53, 3359–3372. <https://doi.org/10.1109/TSP.2005.853302>.
- Tsigka, S., Papadelis, C., Braun, C., & Miceli, G. (2014). Distinguishable neural correlates of verbs and nouns: A MEG study on homonyms. *Neuropsychologia*, 54, 87–97. <https://doi.org/10.1016/j.neuropsychologia.2013.12.018>.
- Urbain, C., Bourguignon, M., de Beeck, M. O., Schmitz, R., Galer, S., Wens, V., et al. (2013). MEG correlates of learning novel objects properties in children. *PLoS One*, 8(7), e69696. <https://doi.org/10.1371/journal.pone.0069696>.
- Urbain, C., De Tiège, X., Op De Beeck, M., Bourguignon, M., Wens, V., Verheulpen, D., et al. (2016). Sleep in children triggers rapid reorganization of memory-related brain processes. *NeuroImage*, 134, 213–222. <https://doi.org/10.1016/j.neuroimage.2016.03.055>.
- Vigario, R., Sarela, J., Jousmiki, V., Hämäläinen, M., & Oja, E. (2000). Independent component approach to the analysis of EEG and MEG recordings. *IEEE Transactions on Biomedical Engineering*, 47, 589–593. <https://doi.org/10.1109/10.841330>.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron*, 55, 143–156. <https://doi.org/10.1016/j.neuron.2007.05.031>.
- Wens, V., Marty, B., Mary, A., Bourguignon, M., Op de Beeck, M., Goldman, S., et al. (2015). A geometric correction scheme for spatial leakage effects in MEG/EEG seed-based functional connectivity mapping. *Human Brain Mapping*, 36(11), 4604–4621. <https://doi.org/10.1002/hbm.22943>.
- Wydell, T. N., Vuorinen, T., Helenius, P., & Salmelin, R. (2003). Neural correlates of letter-string length and lexicality during reading in a regular orthography. *Journal of Cognitive Neuroscience*, 15, 1052–1062. <https://doi.org/10.1162/089892903770007434>.