



HAL
open science

Spoken language coding neurons in the Visual Word Form Area: Evidence from a TMS adaptation paradigm

Chotiga Pattamadilok, Samuel Planton, Mireille Bonnard

► To cite this version:

Chotiga Pattamadilok, Samuel Planton, Mireille Bonnard. Spoken language coding neurons in the Visual Word Form Area: Evidence from a TMS adaptation paradigm. *NeuroImage*, 2019, 186, pp.278 - 285. 10.1016/j.neuroimage.2018.11.014 . hal-01925708

HAL Id: hal-01925708

<https://hal.science/hal-01925708>

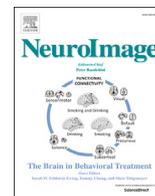
Submitted on 8 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Spoken language coding neurons in the Visual Word Form Area: Evidence from a TMS adaptation paradigm

Chotiga Pattamadilok^{a,*}, Samuel Planton^a, Mireille Bonnard^b

^a Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

^b Aix Marseille Univ, INSERM, INS, Inst Neurosci Syst, Marseille, France



ARTICLE INFO

Keywords:

Left-vOT
Cross-modal activation
Functional segregated neuronal populations
Reading
Speech processing

ABSTRACT

While part of the left ventral occipito-temporal cortex (left-vOT), known as the Visual Word Form Area, plays a central role in reading, the area also responds to speech. This cross-modal activation has been explained by three competing hypotheses. Firstly, speech is converted to orthographic representations that activate, in a top-down manner, written language coding neurons in the left-vOT. Secondly, the area contains multimodal neurons that respond to both language modalities. Thirdly, the area comprises functionally segregated neuronal populations that selectively encode different language modalities. A transcranial magnetic stimulation (TMS)-adaptation protocol was used to disentangle these hypotheses. During adaptation, participants were exposed to spoken or written words in order to tune the initial state of left-vOT neurons to one of the language modalities. After adaptation, they performed lexical decisions on spoken and written targets with TMS applied to the left-vOT. TMS showed selective facilitatory effects. It accelerated lexical decisions only when the adaptors and the targets shared the same modality, i.e., when left-vOT neurons had initially been adapted to the modality of the target stimuli. Since this within-modal adaptation was observed for both input modalities and no evidence for cross-modal adaptation was found, our findings suggest that the left-vOT contains neurons that selectively encode written and spoken language rather than purely written language coding neurons or multimodal neurons encoding language regardless of modality.

1. Introduction

The ventral part of the left occipito-temporal cortex (left-vOT) has been argued to play a key role in reading (see Dehaene and Cohen, 2011 for a review). Given that its responses to written words are consistent across individuals, tasks and scripts, the area has also been called the Visual Word Form Area, a label that explicitly attributes to it a functional specialization in written word processing (Cohen et al., 2000). Interestingly, several brain-imaging studies have shown that, under certain circumstances, this area also responds to spoken language input, for instance, when participants had to recognize spoken words, extract their meaning or analyze their phonological structures, and that the strength of the responses also depends on participants' reading skills (Booth et al., 2004; Burton et al., 2005; Burton et al., 2000; Dehaene et al., 2010; Desroches et al., 2010; Ludersdorfer et al., 2016; Yoncheva et al., 2010). The aim of the present study was to examine the neural mechanism of this cross-modal activation by looking at response patterns of neurons within this key area of the reading network to different modalities of

language input.

1.1. Hypotheses on the neural mechanism of left-vOT responses to spoken input

During the last 15 years, the main debate has centered around two opposing hypotheses (Dehaene and Cohen, 2011; Price and Devlin, 2011). According to the 'orthographic tuning hypothesis' proposed by Dehaene and colleagues (Cohen et al., 2004; Dehaene et al., 2005; Dehaene and Cohen, 2011), the left-vOT contains neuronal circuitry that becomes progressively specialized for orthographic coding through reading experience. Based on this hypothesis, left-vOT neurons are selectively tuned to written language input. Nevertheless, this theoretical position is not incompatible with the fact that the area also responds to speech. As argued by the authors, orthographic coding neurons could also be activated in a top-down fashion by spoken input once the input has been converted into an orthographic code (Dehaene and Cohen, 2011; Dehaene et al., 2015, see Fig. 1a).

* Corresponding author. Laboratoire Parole et Langage, Centre National de la Recherche Scientifique (UMR 7309), 5, Av. Pasteur, 13100, Aix-en-Provence, France.
E-mail address: chotiga.pattamadilok@lpl-aix.fr (C. Pattamadilok).

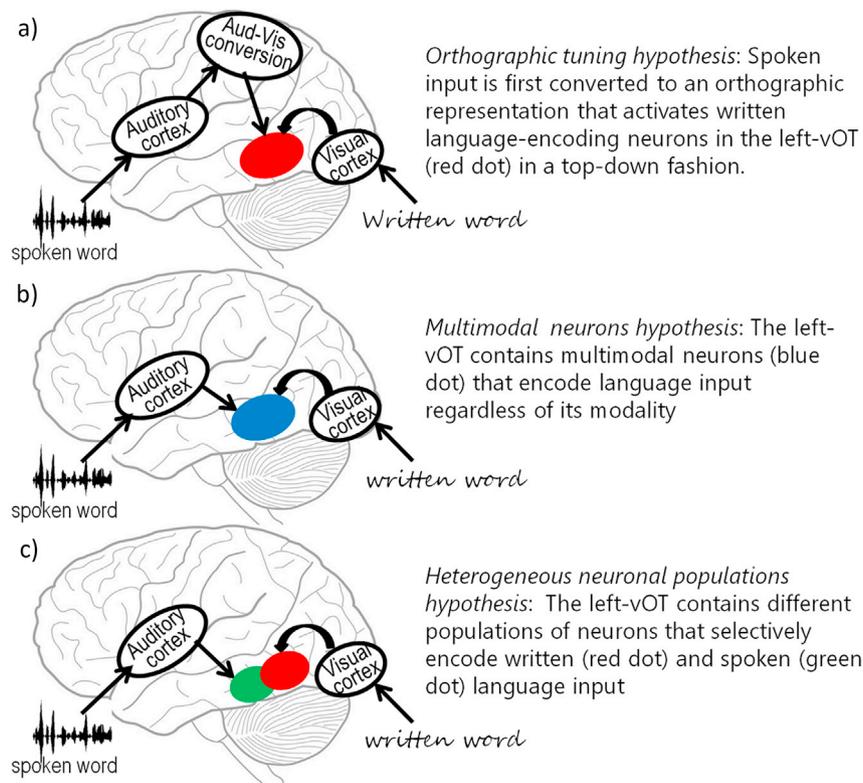


Fig. 1. Three hypotheses explaining the neural mechanism of left-vOT involvement in spoken language processing.

The idea that the left-vOT contains neurons that specifically encode written language input has been challenged mainly by the assumption that the area acts as an integration or a convergence zone that supports multiple functions depending on its interaction with other regions (Buchel et al., 1998; Price and Devlin, 2003, 2011). This assumption relies on number of observations showing that the area is responsive to different input modalities, including orthographic and non-orthographic ones (Amedi et al., 2001; Buchel et al., 1998; Ludersdorfer et al., 2016; Yoncheva et al., 2010). According to this view, the orthographic representation is a mere result of an integration of visual input with higher-level language representations such as phonology and/or semantics that occurs in this cortical area when one processes language. At the neural level, the integration mechanism would be supported by a population of ‘multimodal neurons’ that receive information from both spoken and written language inputs thanks to their connections with the auditory and visual cortices (see Fig. 1b).

In addition to these two main hypotheses, Price and Devlin (2003) also evoked the possibility that the left-vOT may contain different populations of neurons that are involved in different functions. In contrast to the previous account, here, the left-vOT is assumed to contain ‘heterogeneous neuronal populations’. Thus, its responses to written and spoken inputs would be supported by functionally distinct populations of unimodal neurons that are located within the same area. However, due, most likely, to the difficulty in differentiating responses from adjacent neuronal populations within a single cortical area, this hypothesis has been disregarded in most studies reporting left-vOT activation in non-visual tasks (see Fig. 1c).

1.2. The current study

Although these three hypotheses need not be mutually exclusive, they are difficult to disentangle on the basis of behavioral or brain activation patterns induced by task or stimulus manipulation. In the present study, we addressed this issue by combining noninvasive transcranial magnetic stimulation (TMS) with an adaptation protocol (Romei et al., 2016;

Silvanto et al., 2008; Silvanto et al., 2007). The rationale of the protocol is based on the fact that the effect of TMS applied over a cortical area depends on the initial state of the neurons being stimulated (Silvanto et al., 2007; Silvanto and Pascual-Leone, 2008). This state-dependent effect was illustrated in the seminal work of Silvanto and colleagues (Silvanto et al., 2007), taking, as an example, an area in the primary visual cortex that is assumed to contain neuronal populations that selectively respond to red and green. At the baseline level, all neuronal populations are expected to have the same level of excitability. However, if an individual is repeatedly exposed to red during the adaptation phase, the excitability of the neurons that encode this color will be reduced. Previous studies have shown that TMS applied over a cortical area has a selective facilitatory effect on the attribute encoded by the adapted (depressed) neurons (Cattaneo et al., 2009; Kadosh et al., 2010; Mazzoni et al., 2017). In the current example, TMS will facilitate the perception of red but not green. The opposite result is expected if the individual is exposed to green during the adaptation phase. Thanks to this neuronal adaptation mechanism, the TMS protocol can be used to selectively target specific or even spatially overlapping neuronal populations, thus revealing their existence within the stimulated area.

In the present study, we applied the same protocol to examine the sensitivity of neurons in the individually-defined left-vOT, at the location of the VWFA, to written and spoken language input (see also Silvanto and Pascual-Leone, 2008; Cattaneo et al., 2009; Mazzoni et al., 2017 for applications of the technique to examine the neural basis of high cognitive functions). Concretely, each experimental trial comprised an adaptation and a test phase. During the adaptation phase, participants were passively exposed to a series of written or spoken words. The aim of this phase was to adapt the initial state of left-vOT neurons to visual or auditory language modality prior to TMS application. During the test phase, TMS was applied to the left-vOT while participants performed a lexical decision on a stimulus (word or pseudoword) presented in either the visual or auditory modality. The critical manipulation was the congruency between the modalities of the stimuli presented during the adaptation and the test phase, which was orthogonally manipulated,

yielding two ‘within-modal’ (auditory adaptation/auditory task and visual adaptation/visual task) and two ‘cross-modal’ (auditory adaptation/visual task and visual adaptation/auditory task) conditions. At least one month after this session, the same participants underwent the same experimental protocol, except that TMS was applied to the right-vOT, which was considered as a control site. This part of the protocol allowed us to control for the anatomical specificity of neuronal adaptation as revealed by the TMS facilitatory effect.

The variable of interest was the effect of TMS on performance (reaction times and error rates) in the lexical decision task. Given that TMS has a selective facilitatory effect on the attribute encoded by the neurons depressed by adaptation, the three hypotheses mentioned above led to three distinct predictions:

Orthographic tuning hypothesis (Fig. 1a): If the left-vOT only contains a homogeneous population of neurons that encode written language (Dehaene and Cohen, 2011), then only visual adaptation can reduce the excitability of neurons in this area. As a result, a TMS facilitatory effect should be restricted to the ‘visual adaptation/visual task’ condition, i.e., when the stimuli presented during the adaptation and test phases were both in the visual modality.

Multimodal neurons hypothesis (Fig. 1b): If the left-vOT contains multimodal neurons that encode language input regardless of its modality (Buchel et al., 1998; Price and Devlin, 2003, 2011), then both visual and auditory adaptations can reduce the excitability of neurons in this area. TMS facilitatory effects should therefore be equally observed in both within-modal (auditory adaptation/auditory task and visual adaptation/visual task) and cross-modal (auditory adaptation/visual task and visual adaptation/auditory task) conditions.

Heterogeneous neuronal populations hypothesis (Fig. 1c): If the left-vOT contains distinct populations of neurons that selectively encode written and spoken language input (Price and Devlin, 2003), then the excitability of a neuronal population in this area can be reduced only by visual or by auditory adaptation but not both. TMS facilitatory effects should therefore be restricted to the within-modal conditions (auditory adaptation/auditory task and visual adaptation/visual task), i.e., whenever the stimuli presented during the adaptation and test phases were in the same modality. No facilitatory effect should be found in the cross-modal conditions.

2. Materials and methods

2.1. Participants

Seventeen native French speakers participated in the study (eight women, 20–27 yrs, mean = 23). All participants were right-handed and none reported any history of language or neurological disorders. The experiment was approved by the local ethics committee (Sud Méditerranée I). Informed consent was obtained from each participant after the experimental procedures were explained. Each participant participated to one fMRI (or MRI) session and two TMS sessions. They were paid for their participation.

2.2. Individual localization of the left-vOT

Prior to the TMS session, thirteen participants underwent a fMRI session which was used to functionally localize the left-vOT for neuro-navigated stimulation. Individual functional localization is currently the most powerful method since it provides precise information on the location of the target area, taking into account inter-individual anatomical and functional variability (Sack et al., 2009). fMRI data were collected on a 3-Tesla Siemens Prisma Scanner (Siemens, Erlangen, Germany) at the Marseille fMRI centre (Centre IRMF-INT-CERIMED, Aix-Marseille University) using a 64-channel head coil. It included the acquisition of a high-resolution structural T1-weighted image (MPRAGE sequence, resolution $1 \times 1 \times 1 \text{ mm}^3$) and of functional images with a T2*-weighted gradient-echo planar sequence (TR = 1224 msec., echo

time = 30 msec., flip angle = 66° , 54 slices with a thickness of 2.5 mm, FOV = $210 \times 210 \text{ mm}^2$, matrix = 84×84 , voxel = $2.5 \times 2.5 \times 2.5 \text{ mm}^3$, multiband factor = 3). The left-vOT was individually identified by combining the fMRI data obtained in a reading task (Pattamadilok et al., 2017) and the theoretical location of the left-vOT as reported in the literature (Jobard et al., 2003). Specifically, we first gathered individual activation maps from a “written word vs. fixation” contrast (with a voxel-wise threshold of $p < .001$ uncorrected). fMRI analyses were conducted without spatial normalization so that the functional maps remained in the participant's native space. Additionally, for each participant, a spherical volume of interest of 16 mm diameter, centered on the MNI coordinates, $x = -44$, $y = -58$, $z = -15$, was constructed; this corresponds to the occipito-temporal cluster reported in the meta-analysis of reading by Jobard et al. (2003). In order to match the individual anatomy, the VOI was transformed towards the participant's native structural space using the “deformations” module of the software SPM 12 running on Matlab R2015a (Mathworks Inc., Natick, MA, USA). The stimulation site was defined for each participant as the peak of activation at the intersection between the spherical VOI and the activation map. This target site was marked on each individual structural MRI before being imported into the neuronavigation software. For the four participants without activation maps, high-resolution anatomical scans were acquired. The stimulation site was localized using the spherical VOI and anatomical landmarks corresponding to the middle left fusiform gyrus, medial to the occipito-temporal sulcus on the crest of the gyrus, just superior to the principal sulcus of the cerebellum (coronal plane). Fig. 2 (top panel) illustrates the group functional map obtained during the fMRI session, the location of stimulation areas for each participant and the average coordinates for the group after MNI normalization ($x = -45$, $y = -60$, $z = -14$). The homologous areas in the right hemisphere are illustrated in the bottom panel of the Figure.

2.3. Transcranial magnetic stimulation

During the TMS session, a frameless stereotaxy system was used to position the TMS coil on the scalp to stimulate the precise anatomical region-of-interest. A Polaris infrared camera tracked the participant's head, and NBS software (Navigation Brain System 2.3 Nexstim, Helsinki, Finland) registered the participant's head to his/her MRI scan.

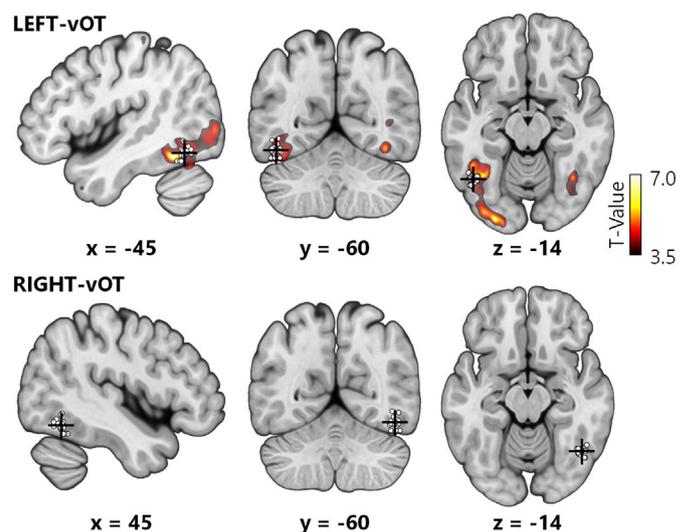


Fig. 2. Illustration of the stimulation sites. Top panel: Left-vOT locations (white circles) for 17 participants overlaid on the group functional map (written word > fixation, voxelwise $p < .001$ unc.), projected onto a MNI-space standard template. Bottom panel: Right-vOT locations (white circles) for 17 participants projected onto a MNI-space standard template. The crosses indicate the group average coordinates for the two stimulation sites.

Neuronavigation (Navigated Brain Stimulation system, Nexstim, Helsinki, Finland) was used to target and record stimulation sites. Stimulation was delivered using a MagStim figure-of-eight coil and a Magstim bi-stimulation 200 monophasic stimulator (Magstim Co., Whitland, UK). The stimulation intensity was individually adjusted based on the threshold necessary to observe a visible right hand twitch when stimulating the hand area of the primary motor cortex in the left hemisphere (M1). This area was identified anatomically according to the method of Yousry et al. (1997). The initial stimulation intensity applied to M1 was adjusted to the depth of the left-vOT in order to equalize the induced electric field value within the two areas (Spieser et al., 2013). The adjusted intensity was computed by the neuronavigation software (Navigation Brain System 2.3 Nexstim, Helsinki, Finland). The final intensity ranged from 40% to 60% of the maximum stimulator output (mean = 47%). This method of stimulation intensity setting has been successfully applied in our previous TMS studies that used “virtual lesion” protocols to reveal the functional role of the left-vOT in reading tasks (Duncan et al., 2010; Pattamadilok et al., 2015). During the lexical decision task, double-pulse stimulations (the two pulses were separated by 100 msec.) were delivered on half of the trials at 100 msec. post-stimulus onset. Due to the summation properties of TMS pulses, the double-pulse stimulation gives larger effects than single-pulse (Walsh et al., 2003) and, therefore, increases the possibility of an observable TMS effect at the behavioral level. The stimulation timing was chosen based on previous chronometric TMS studies showing that the left-vOT contributes to reading as early as 80–120 msec. after word onset (Duncan et al., 2010; Pattamadilok et al., 2015). Stimulation parameters were well within international safety guidelines (Rossi et al., 2009; Wassermann, 1998).

2.4. Stimuli

2.4.1. Adaptation phase: exposure to spoken or written words

This adaptation protocol aimed to tune neurons within the left-vOT to spoken or written language modality. To this end, prior to each lexical decision trial, participants were presented with a list of five words that were randomly selected from a list of 450 mono and bisyllabic French words (all were different from those presented during the lexical decision task described below). Within each list, the five words were presented in the same modality, either auditory or visual. A new randomized list was generated for each trial and for each participant to avoid any systematic phonological, orthographic or semantic overlap between the stimuli presented during the adaptation and the test phase.

2.4.2. Test phase: lexical decision task

The stimuli presented during the lexical decision task consisted of 200 words and 200 pseudowords. Each word was used to generate a pseudoword with the same number of letters, phonemes and syllables and that was closely matched for bigram frequency (www.lexique.org). Each type of stimuli was equally separated into eight subsets of 25 stimuli. Across the eight subsets, words were matched for objective written and spoken frequencies, number of phonological and orthographic neighbors, number of phonemes, number of letters, number of syllables and phonological and orthographic uniqueness points ($p > .10$) (New et al., 2004). Pseudowords were matched for number of letters, number of phonemes, number of syllables and bigram frequency ($p > .25$). For each participant, the stimuli from a subset were presented in one of eight experimental conditions corresponding to combinations of the following factors: adaptation modality (auditory vs. visual), task modality (auditory vs. visual) and TMS (with vs. without TMS). In this way, no stimulus was presented more than once within participants. Across participants, each stimulus was presented in all experimental conditions. The auditory version of the stimuli was recorded in a soundproof room and digitized at a sampling rate of 48 kHz with 32-bit analog-to-digital conversion. Their mean duration was 510 msec. During the experiment, auditory stimuli were presented via headphones at individually adjusted sound level. The

visual version of the stimuli was presented on a 16-Inch monitor at a viewing distance of ~50 cm. Each visual stimulus was presented for 510 msec. in order to match the average duration of the auditory input.

2.5. Procedure

At the beginning of the TMS session, the stimulation site was localized based on either the fMRI or the anatomical localizer. As illustrated in Fig. 3, each trial consisted of a sequence of five words presented at a rate of 1 s per word. During this adaptation phase, participants were required to either read or listen to the stimuli. For both visual and auditory modalities, a black background was used on the computer screen. In the visual modality, stimuli were written in white characters. To indicate the beginning of the test phase, the black background turned to grey and this blank screen with a grey background remained for 250 msec. This was followed by a word or a pseudoword presented either auditorily or visually and, regardless of modality, participants were instructed to indicate as early as possible whether this stimulus was a real French word or a pseudoword by pressing a button using either their right or left index. The association between response and finger was counterbalanced across participants.

The combination of stimulus modalities in the adaptation and the test phase led to four presentation conditions: auditory adaptation/auditory task; visual adaptation/visual task; auditory adaptation/visual task; visual adaptation/auditory task. Within each condition, double-pulse TMS was applied on half of the trials. At the most detailed level of description, the four factors of interest (adaptation modality, task modality, lexicality and TMS) yielded 16 within-subject conditions, with 25 trials each, which gave a total of 400 trials. These 400 trials were divided into 8 blocks of 50 trials (3 or 4 trials per condition) and trial order was randomized for each participant. The total duration of the task was about 40 min. Accuracy and reaction times (RTs) were recorded. Presentation, timing, and data collection were controlled by E-prime 2.0.10 software. The session started with a practice block that allowed the participants to familiarize themselves with the procedure and the task. Feedback was provided during the practice trials only.

The experiment was conducted again in a second TMS session, with the same group of participants and the same experimental protocol, but this time TMS was applied to the right-vOT (see Fig. 2, bottom), which was considered as a control site. At least one month separated the two experimental sessions for a given participant. Although there is no specific reason why neuronal adaptation effect should depend on session order, we acknowledge that not using a counter-balanced session order could be considered as a caveat.

3. Results

Preliminary inspection of the reaction times (RTs) on correct trials in the lexical decision task led us to discard deviant RTs, i.e., those longer or shorter than the mean RT observed on correct trials plus or minus 2SD (5% of the data), from further analyses. The remaining RTs were analyzed with the R software (R Core Team, 2014) using a linear mixed-effects model (LME). The model was fit with the lme4 package (Bates et al., 2014) and the p values (Type III: marginal sum of squares) were computed with the lmerTest package (Kuznetsova et al., 2013). To specifically test the three hypotheses described above, eight pairwise comparisons (see details below) were performed using multcomp package (Bretz et al., 2008) and Bonferroni corrections were applied for multiple comparisons. GLMER with a binomial link function was applied on error rates (modeling the same fixed factors and random intercepts as in the analysis performed on RT data, see details below). However, since our dataset was very large for the default fitting methods in glmer, it was necessary to implement the function options “nAGQ = 0” and “control = glmerControl (optimizer = “nloptwrap”)” for the model-fitting process to converge (Bates et al., 2016).

A LME model was performed on the RTs. Unless stated otherwise,

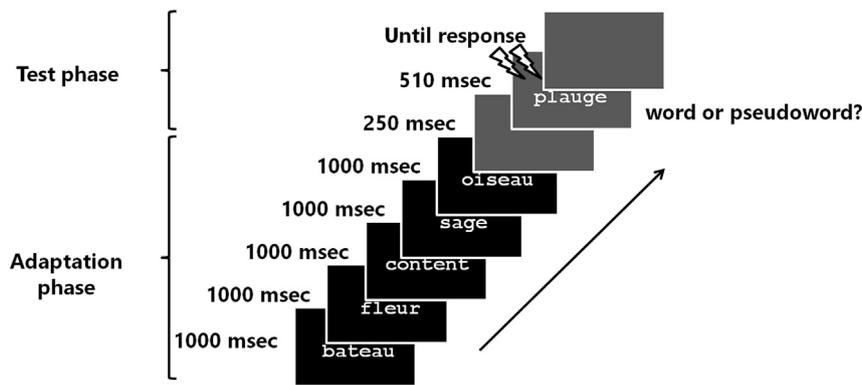


Fig. 3. Timeline of the adaptation and the test phase of an experimental trial from the visual adaptation/visual task condition. Double-pulse TMS was applied 100 msec. after the onset of the stimulus presented in the test phase.

lexicality (words vs. pseudowords), adaptation modality (auditory vs. visual), task modality (auditory vs. visual), TMS (trials with vs. without TMS), target area (left-vOT vs. right-vOT) and their interactions were considered as fixed factors. As recommended by Baayen et al. (2008), both participants and items were considered as random intercepts in all of the analyses presented below. However, we also ensured that including items as a random intercept (in addition to participants) significantly improved model fit [$\chi^2(1) = 922$, $p < .0001$]. The analysis showed significant main effects of lexicality [$F(1,764) = 160.7$, $p < .0001$], adaptation modality [$F(1,11575) = 58.2$, $p < .0001$], task modality [$F(1,764) = 2843.9$, $p < .0001$], TMS [$F(1,11518) = 11.1$, $p < .001$] and target area [$F(1,11547) = 240.3$, $p < .0001$]. The following two-way and three-way interactions were significant: task modality * adaptation modality [$F(1,11574) = 6.9$, $p < .01$], task modality * lexicality [$F(1,764) = 6.3$, $p = .01$], target area * lexicality [$F(1, 11472) = 14.8$, $p = .0001$], lexicality * TMS [$F(1,11463) = 3.9$, $p < .05$], target area * TMS [$F(1,11546) = 9.2$, $p < .005$], task modality * adaptation modality * target area [$F(1,11651) = 10.3$, $p = .001$], task modality * adaptation modality * TMS [$F(1,11355) = 8.6$, $p < .005$], and task modality * lexicality * TMS [$F(1,11463) = 4.1$, $p < .05$]. The latter interaction was due to the presence of a significant TMS facilitatory effect only on decisions made on visually presented pseudowords ($p < .001$; 650msec. vs. 665msec. on trials with and without TMS, respectively) but not on the other stimuli. This effect was nevertheless unrelated to the modality of stimuli presented during the adaptation phase and, thus, does not reflect the neuronal adaptation effect of TMS. Most critically, we observed a significant four-way interaction of adaptation modality * task

modality * TMS * target area [$F(1,11657) = 4.3$, $p < .05$]. As expected, further analyses performed separately on the data from the left- and right-vOT session confirmed that TMS applied to these two areas led to two different patterns of results. The critical three-way interaction of adaptation modality * task modality * TMS was significant for the left-vOT [$F(1,5624) = 10.2$, $p = .001$] but not for the right-vOT [$F(1, 5546) < 1$].

In order to test specifically the three hypotheses presented in the Introduction, pairwise comparisons examining the facilitatory effects of TMS on the four combinations of adaptation and task modality (i.e., auditory adaptation/auditory task; visual adaptation/visual task; auditory adaptation/visual task and visual adaptation/auditory task) were conducted for both target areas (the threshold was adjusted to $p < .0063$ for eight pairwise comparisons). The aim of these comparisons was to examine the existence of within-modal (auditory adaptation/auditory task; visual adaptation/visual task) and cross-modal (auditory adaptation/visual task; visual adaptation/auditory task) neuronal adaptation effects. As illustrated in Fig. 4 (left panel), left-vOT TMS facilitated lexical decisions on written stimuli when neurons within this area had been adapted to written ($z = -3.6$, $p < .0005$) but not spoken input ($z = -0.5$, $p > .5$). Likewise, TMS also facilitated lexical decisions on spoken stimuli when neurons within the left-vOT had been adapted to spoken ($z = -4.4$, $p < .0001$) but not written input ($z = -0.4$, $p > .5$). These results clearly indicate the presence of within-modal neuronal adaptation in both auditory and visual modalities and the absence of cross-modal neuronal adaptation.

Previous studies on TMS adaptation have argued that a behavioral

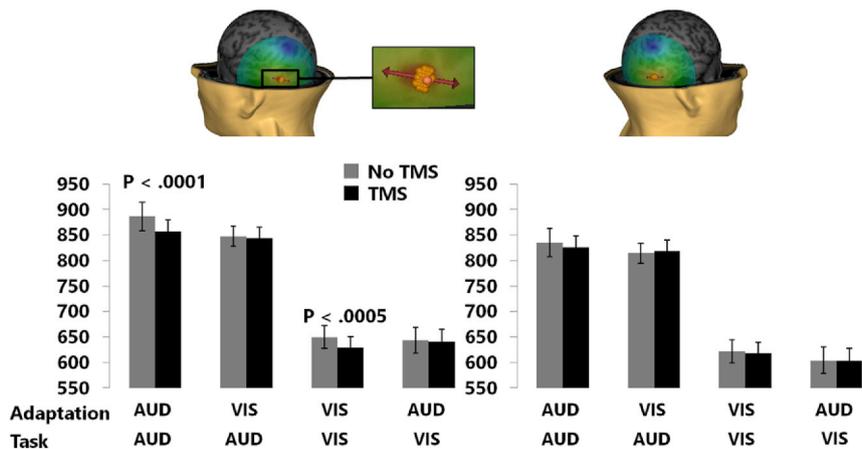


Fig. 4. Mean reaction times (with error bars representing standard error of the mean) obtained in the four experimental conditions on trials with (black bars) and without TMS (grey bars) applied to the left-vOT (left panel) and the right-vOT (right panel). Stimulation delivered to the left-vOT facilitated lexical decisions only in the within-modal adaptation conditions. No significant TMS effect was found in the other experimental conditions.

manifestation of adaptation on trials without TMS is not necessary for obtaining state-dependent effects of TMS (Cattaneo et al., 2009; Silvanto et al., 2007; Silvanto and Pascual-Leone, 2008). However, to ensure that the neuronal adaptation reported above was indeed independent of behavioral effect of adaptation, we conducted an additional analysis on 11 of the 17 participants as these participants showed an inhibitory effect of adaptation at the behavioral level. More precisely, this subgroup of participants corresponded to those who presented, on no-TMS trials, overall longer RTs for the conditions in which stimuli presented during the adaptation and the test phase were in the same modality (764 msec. on average, $SD = 90$) compared to the conditions in which the stimuli presented in the two phases were not in the same modality (740 msec. on average, $SD = 84$). The analysis of the neuronal adaptation effect performed on this restricted group of participants replicated the findings obtained on the entire group: When TMS was applied to the left-vOT, neuronal adaptation (corresponding to a facilitatory effect of TMS) was observed in the within-modal adaptation conditions (auditory adaptation/auditory task, $p < .0001$; visual adaptation/visual task, $p < .0005$) but not in cross-modal adaptation conditions (auditory adaptation/visual task, $p > .5$; visual adaptation/auditory task, $p > .5$).

In contrast to the above findings, applying TMS over the right-vOT did not modify the RTs in any experimental condition (all p values $> .10$, Fig. 4, right panel). This null result confirms that the neuronal adaptation to spoken and written language input was specific to the left-vOT.

Error rate analysis only revealed a significant effect of lexicality, showing higher error rates on pseudowords (7.6%) compared to words (5.3%) [$Z = -2.6$, $p < .01$]. No other main effects or interactions were significant (all p values $> .10$).

4. Discussion

The neural mechanism of left-vOT involvement in speech processing is subject to debate. Here, we investigated this issue by using a TMS-adaptation protocol to examine the modality of language input encoded by neurons within this specific brain area (Silvanto et al., 2007, 2008). The findings were clear-cut. The facilitatory effects of TMS were observed only in the within-modal adaptation conditions: Compared to trials without left-vOT stimulation, TMS reduced RTs in the visual lexical decision task when participants had been previously exposed to written but not spoken language during the adaptation phase. Likewise, it reduced RTs in the auditory lexical decision task when participants had been previously exposed to spoken but not written language during the adaptation phase. The fact that this pattern of results did not interact with stimulus lexicality further supports the idea that the neuronal adaptation reported here occurred at the level of input modality rather than at a higher linguistic level that is sensitive to the lexico-semantic properties of the input. Additionally, the absence of facilitatory effect when TMS was applied to the homologous area in the right hemisphere confirmed that the neuronal adaptation was specific to the left-vOT.

TMS facilitates detection of the adapted attribute only if the targeted area contains neurons that are tuned to the adapting stimulus (Cattaneo et al., 2009; Silvanto et al., 2007, 2008). As a result, the presence of the TMS facilitatory effects in both visual adaptation/visual task and auditory adaptation/auditory task conditions indicates that the left-vOT contains, not only neurons that respond to written language as typically assumed, but also those that respond to spoken language. At first glance, the presence of the latter neuronal populations in this brain area may seem at odds with the general properties of neurons that are part of the visual system (Dehaene et al., 2005; Gauthier and Palmeri, 2002; Malach et al., 2002; Price and Devlin, 2011). Thus, before adopting the above conclusion, it is important to rule out an alternative explanation, that is, the possibility that the effect of TMS applied to the left-vOT at the location of the VWFA might extend beyond the stimulated area and affect neurons in the anterior part of the left fusiform gyrus that responds to abstract representations (Cohen et al., 2004; Deng et al., 2013). Such a

case would imply that the TMS facilitatory effect observed in the auditory adaptation/auditory task condition was caused by the activity of these higher-order neurons. However, this explanation is not supported by our findings since it would also lead to facilitations in the cross-modal conditions.

The present experimental protocol provides a powerful tool for testing the three competing hypotheses regarding the properties of left-vOT neurons and how they contribute to speech processing. Among these hypotheses, the existence of different populations of neurons that selectively respond to written and spoken language is most coherent with the prediction made by the *Heterogeneous neuronal populations hypothesis* (Price and Devlin, 2003). Nevertheless, the implications of these findings for the *Orthographic tuning* (Dehaene et al., 2005; Dehaene and Cohen, 2011) and *Multimodal neurons hypotheses* (Price and Devlin, 2003) need to be discussed. Firstly, the existence of functionally segregated neuronal populations that selectively respond to written and spoken language modality does not preclude the possibility that neurons that respond to written language input may also be activated by speech in a top-down fashion, as claimed by the *Orthographic tuning hypothesis* (Dehaene and Cohen, 2011). Thus, left-vOT responses to spoken input could be supported by both mechanisms, although more research is needed to investigate whether they systematically operate in parallel and/or in a task-dependent manner. The second issue is related to the idea, put forward by Price and Devlin, that the left-vOT may act as an integration or convergence zone (Price and Devlin, 2003, 2011). Most multimodal integration or convergence phenomena reported in the literature could be supported by at least two neural mechanisms, that is, the same multimodal neurons respond to different input modalities and/or the area contains functionally segregated subpopulations of neurons, with each subpopulation responding only to one kind of unimodal input (King and Calvert, 2001; Meredith, 2002). The spatial resolution of most fMRI acquisitions is not sufficient to disentangle these two possibilities, especially when the data are averaged across participants (Gentile et al., 2017). This lack of spatial resolution may lead to an erroneous conclusion particularly when different subpopulations are spatially overlapped. As reviewed by Meredith (2002), the most reliable and efficient technique for addressing this issue is single-cell electrophysiological recording. Yet, this technique remains extremely invasive. The TMS-adaptation paradigm used here should be considered as a first step to address this issue at a coarser-grained level, using a non-invasive technique that can be safely applied to healthy human participants. Our findings suggest that, although the left-vOT responds to both written and spoken inputs which is suggestive of multimodal integration, at the neural level, the different inputs do not seem to terminate at the same neuronal population.

Recently, Zhao et al. (2016) examined the nature of representations processed in the left-vOT. Using Representation Similarity Analysis (RSA), the authors reported a significant association between the pattern of phonological similarity of written Chinese words and neural responses in the anterior and middle fusiform gyrus, the latter being located in the same area as the left-vOT targeted here. In line with the current observation, their RSA results indicate that both phonological and orthographic information are represented in this area. The authors further claimed that the emergence of this joint representation resulted from orthographic–phonological training. Although our study did not examine the role of reading acquisition, the emergence of neuronal populations that are sensitive to spoken language in the ventral visual pathway is likely to result from this activity. An important argument in favor of this idea is that the presence of left-vOT responses to speech was not observed in illiterates (Dehaene et al., 2010) and, in literates, the strength of the responses depended on participants' reading proficiency (Cone et al., 2008; Desroches et al., 2010). Thus, by showing that neuronal populations in the left-vOT are sensitive to spoken language, we speculate that the strong and repeated connections between the spoken and written language systems established during reading acquisition may lead to a modification of the functional properties of neurons within the visual

system. However, whether the sensitivity to spoken input reported here reflects a modification of intrinsic properties of left-vOT neurons or of their contextual circuit properties deserves further research.

Existing literature provides several examples of experience-dependent modifications of cortical functions. For instance, studies on the influence of orthographic knowledge on speech processing showed that the cortical origins of some orthographic effects observed in speech processing tasks were located in the auditory system. These effects could not be explained by a top-down activation of orthography in response to speech, but suggested that learning to read reorganized the phonological system such that it became sensitive to word spelling (Pattamadilok et al., 2010; Perre et al., 2009). In line with these observations, Dehaene et al. (2010) also reported an enhancement of brain activity in the auditory cortex in response to speech in literate compared to illiterate participants. Beyond the domain of reading acquisition, several studies also provided evidence of cortical reorganization in different sensory systems following lesions, deprivation or intensive training. One example of this is the engagement of ventral visual areas in auditory and tactile object-recognition tasks in the blind (Siuda-Krzywicka et al., 2016). Similarly, auditory areas were found to be recruited for visual and tactile inputs in the deaf (Bola et al., 2017). As discussed by Bola et al. (2017), such cortical reorganization questions the idea of the brain's sensory boundaries, which implies a strict association between a specific sensory cortex and the input being processed. In the present case, learning to read, which involves repeated association between spoken and written language, may not only cause certain neurons in the left-vOT to become highly sensitive to known scripts, but may also render other neurons within the same area sensitive to speech. At the functional level, the existence of neuronal populations that respond to written and spoken language input within the same cortical area could contribute to fluent reading by enabling rapid exchanges between different linguistic representations. Left-vOT responses to speech are typically observed in high-level speech processing situations and in individuals who have a high level of reading expertise (Brem et al., 2010; Dehaene et al., 2010; Desroches et al., 2010; Yoncheva et al., 2010). This task- and expertise-dependent activation profile suggests that left-vOT neurons that respond to spoken input may work in concert with neurons in the auditory cortex and contribute to more efficient speech processing in literates compared to illiterates (Castro-Caldas et al., 1998; Dehaene et al., 2010; Reis and Castro-Caldas, 1997) or in skilled compared to unskilled readers, especially in complex speech processing tasks (Monzalvo et al., 2012).

In conclusion, a combination of TMS with an adaptation protocol enhances the functional resolution of this technique. It allowed us to reveal different patterns of neural responses within the left-vOT. The finding is rather counter-intuitive. It suggests that this key area of the reading network contains, not only neuronal populations that are sensitive to written input, but also those that are sensitive to spoken input. This observation provides insightful information on the neural mechanism of left-vOT responses to speech and raises further questions about the development and the nature of cortical reorganization following reading acquisition that worth exploring in the future.

Acknowledgments

This work was supported by the French Ministry of Research: ANR-13-JSH2-0002. We warmly thank Dr. Agnès Trébuchon for taking medical responsibility during the study; and Mathilde Cans and Faustine Charignon for stimulus preparation.

References

Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Activation in the Ventral Visual Pathway, 4(3).
 Baayen, R.H., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59 (4), 390–412. <http://doi.org/10.1016/j.jml.2007.12.005>.

Bates, D., Bolker, B., Walker, S., 2016. Package 'lme4', version 1.1-12, linear mixed effects models using 'Eigen' and S4. Available: <https://cran.r-project.org/web/packages/lme4/lme4.pdf>.
 Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. Introduction to Statistical Analysis Using R. R package Version, pp. 1–23 (Mcmc). <http://doi.org/citeulike-article-id:7112638>.
 Bola, L., Zimmermann, M., Mostowski, P., Jednoróg, K., Marchewka, A., Rutkowski, P., Szwed, M., 2017. Task-specific reorganization of the auditory cortex in deaf humans. *Proc. Natl. Acad. Sci. Unit. States Am.* <http://doi.org/10.1073/pnas.1609000114>.
 Booth, J.R., Burman, D.D., Meyer, J.R., Gitelman, D.R., Parrish, T.B., Mesulam, M.M., 2004. Development of brain mechanisms for processing orthographic and phonologic representations. *J. Cognit. Neurosci.* 16 (7), 1234–1249. <http://doi.org/10.1162/0898929041920496>.
 Brem, S., Bach, S., Kucian, K., Guttorm, T.K., Martin, E., Lyytinen, H., et al., 2010. Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proc. Natl. Acad. Sci. Unit. States Am.* 107 (17), 7939–7944.
 Bretz, F., Westfall, P., Heiberger, R.M., Schuetzenmeister, A., Scheibe, S., 2008. Simultaneous inference in general parametric models. *Biom. J.* 50 (3), 346–363.
 Buchel, C., Price, C.J., Friston, K., 1998. A multimodal language region in the ventral visual pathway. *Nature* 394, 274–277.
 Burton, M.W., LoCasto, P.C., Krebs-Noble, D., Gullapalli, R.P., 2005. A systematic investigation of the functional neuroanatomy of auditory and visual phonological processing. *Neuroimage* 26 (3), 647–661.
 Burton, M.W., Small, S.L., Blumstein, S.E., 2000. The role of segmentation in phonological processing: an fMRI investigation. *J. Cognit. Neurosci.* 12 (4), 679–690.
 Castro-Caldas, A., Petersson, K.M., Reis, A., Stone-Elander, S., Ingvar, M., 1998. The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain* 121 (6), 1053–1063.
 Cattaneo, Z., Rota, F., Walsh, V., Vecchi, T., Silvanto, J., 2009. TMS-adaptation reveals abstract letter selectivity in the left posterior parietal cortex. *Cerebr. Cortex* 19 (10), 2321–2325. <http://doi.org/10.1093/cercor/bhn249>.
 Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M.A., Michel, F., 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 (2), 291–307.
 Cohen, L., Jobert, A., Le Bihan, D., Dehaene, S., 2004. Distinct unimodal and multimodal regions for word processing in the left temporal cortex. *Neuroimage* 23 (4), 1256–1270.
 Cone, N.E., Burman, D.D., Bitan, T., Bolger, D.J., Booth, J.R., 2008. Developmental changes in brain regions involved in phonological and orthographic processing during spoken language processing. *Neuroimage* 41 (2), 623–635.
 Dehaene, S., Cohen, L., 2011. The unique role of the visual word form area in reading. *Trends Cognit. Sci.* 15 (6), 254–262.
 Dehaene, S., Cohen, L., Morais, J., Kolinsky, R., 2015. Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* 16 (4), 234–244.
 Dehaene, S., Cohen, L., Sigman, M., Vinckier, F., 2005. The neural code for written words: a proposal. *Trends Cognit. Sci.* 9 (7), 335–341.
 Dehaene, S., Pegado, F., Braga, L.W., Ventura, P., Nunes Filho, G., Jobert, A., et al., 2010. How learning to read changes the cortical networks for vision and language. *Science* 330 (6009), 1359–1364.
 Deng, Y., Wu, Q., Weng, X., 2013. Unimodal and multimodal regions for logographic language processing in left ventral occipitotemporal cortex. *Front. Hum. Neurosci.* 7 (September), 1–9. <http://doi.org/10.3389/fnhum.2013.00619>.
 Desroches, A.S., Cone, N.E., Bolger, D.J., Bitan, T., Burman, D.D., Booth, J.R., 2010. Children with reading difficulties show differences in brain regions associated with orthographic processing during spoken language processing. *Brain Res.* 1356, 73–84.
 Duncan, K.J., Pattamadilok, C., Devlin, J.T., 2010. Investigating occipito-temporal contributions to reading with TMS. *J. Cognit. Neurosci.* 22 (4), 739–750. <http://doi.org/10.1162/jocn.2009.21207>.
 Gauthier, I., Palmeri, T.J., 2002. Visual neurons: categorization-based selectivity. *Curr. Biol.* 12 (8), 282–284. [http://doi.org/10.1016/S0960-9822\(02\)00801-1](http://doi.org/10.1016/S0960-9822(02)00801-1).
 Gentile, F., van Atteveldt, N., De Martino, F., Goebel, R., 2017. Approaching the ground truth: revealing the functional organization of human multisensory STC using ultra-high field fMRI. *J. Neurosci.* 37 (42), 10104–10113. <http://doi.org/10.1523/JNEUROSCI.0146-17.2017>.
 Jobard, G., Crivello, F., Tzourio-Mazoyer, N., 2003. Evaluation of the dual route theory of reading: a meta-analysis of 35 neuroimaging studies. *Neuroimage* 20 (2), 693–712.
 Kadosh, R.C., Muggleton, N., Silvanto, J., Walsh, V., 2010. Double dissociation of format-dependent and number-specific neurons in human parietal cortex. *Cerebr. Cortex* 20 (9), 2166–2171. <http://doi.org/10.1093/cercor/bhp273>.
 King, A.J., Calvert, G.A., 2001. Multisensory integration: perceptual grouping by eye and ear. *Curr. Biol.* 11 (8), R322–R325. [http://doi.org/10.1016/S0960-9822\(01\)00175-0](http://doi.org/10.1016/S0960-9822(01)00175-0).
 Kuznetsova, A., Brockhoff, B., Christensen, H.B., 2013. lmerTest: Tests for Random and Fixed Effects for Linear Mixed Effect Models (Lmer Objects of the Lme4 Package).
 Ludersdorfer, P., Wimmer, H., Richlan, F., Schurz, M., Hutzler, F., Kronbichler, M., 2016. Left ventral occipitotemporal activation during orthographic and semantic processing of auditory words. *Neuroimage* 124, 834–842.
 Malach, R., Levy, I., Hasson, U., 2002. The topography of high-order human object areas. *Trends Cognit. Sci.* 6 (4), 176–184. [http://doi.org/10.1016/S1364-6613\(02\)01870-3](http://doi.org/10.1016/S1364-6613(02)01870-3).
 Mazzoni, N., Jacobs, C., Venuti, P., Silvanto, J., Cattaneo, L., 2017. State-dependent TMS reveals representation of affective body movements in the anterior intraparietal cortex. *J. Neurosci.* 37 (30), 7231–7239. <http://doi.org/10.1523/JNEUROSCI.0913-17.2017>.

- Meredith, M.A., 2002. On the neuronal basis for multisensory convergence: a brief overview. *Cognit. Brain Res.* 14 (1), 31–40. [http://doi.org/10.1016/S0926-6410\(02\)00059-9](http://doi.org/10.1016/S0926-6410(02)00059-9).
- Monzalvo, K., Fluss, J., Billard, C., Dehaene, S., Dehaene-Lambertz, G., 2012. Cortical networks for vision and language in dyslexic and normal children of variable socioeconomic status. *Neuroimage* 61 (1), 258–274. <http://doi.org/10.1016/j.neuroimage.2012.02.035>.
- New, B., Pallier, C., Brysbaert, M., Ferrand, L., 2004. Lexique 2: a new French lexical database. *Behav. Res. Methods* 36 (3), 516–524.
- Pattamadilok, C., Bulnes, L.C., Devlin, J.T., Bourguignon, M., Morais, J., Goldman, S., Kolinsky, R., 2015. How early does the brain distinguish between regular words, irregular words, and pseudowords during the reading process? Evidence from neurochronometric TMS. *J. Cognit. Neurosci.* 27 (6), 1259–1274.
- Pattamadilok, C., Chanoine, V., Pallier, C., Anton, J.-L., Nazarian, B., Belin, P., Ziegler, J.C., 2017. Automaticity of phonological and semantic processing during visual word recognition. *Neuroimage* 149, 244–255. <http://doi.org/10.1016/j.neuroimage.2017.02.003>.
- Pattamadilok, C., Knierim, I.N., Kawabata Duncan, K.J., Devlin, J.T., 2010. How does learning to read affect speech perception? *J. Neurosci.* 30 (25), 8435–8444.
- Perre, L., Pattamadilok, C., Montant, M., Ziegler, J.C., 2009. Orthographic effects in spoken language: on-line activation or phonological restructuring? *Brain Res.* 1275, 73–80.
- Price, C., Devlin, J., 2003. The myth of the visual word form area. *Neuroimage* 19 (3), 473–481.
- Price, C.J., Devlin, J.T., 2011. The Interactive Account of ventral occipitotemporal contributions to reading. *Trends Cognit. Sci.* 15 (6), 246–253.
- R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Reis, A., Castro-Caldas, A., 1997. Illiteracy: a cause for biased cognitive development. *J. Int. Neuropsychol. Soc.* 3 (05), 444–450.
- Romei, V., Thut, G., Silvanto, J., 2016. Information-based approaches of noninvasive transcranial brain stimulation. *Trends Neurosci.* 39 (11), 1–14. <http://doi.org/10.1016/j.tins.2016.09.001>.
- Rossi, S., Hallett, M., Rossini, P.M., Pascual-Leone, A., 2009. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin. Neurophysiol.* 120, 2008–2039.
- Sack, A.T., Cohen Kadosh, R., Schuhmann, T., Moerel, M., Walsh, V., Goebel, R., 2009. Optimizing functional accuracy of TMS in cognitive studies: a comparison of methods. *J. Cognit. Neurosci.* 21 (2), 207–221. <http://doi.org/10.1162/jocn.2009.21126>.
- Silvanto, J., Muggleton, N.G., Cowey, A., Walsh, V., 2007. Neural adaptation reveals state-dependent effects of transcranial magnetic stimulation. *Eur. J. Neurosci.* 25 (6), 1874–1881. <http://doi.org/10.1111/j.1460-9568.2007.05440.x>.
- Silvanto, J., Muggleton, N., Walsh, V., 2008. State-dependency in brain stimulation studies of perception and cognition. *Trends Cognit. Sci.* 12 (12), 447–454.
- Silvanto, J., Pascual-Leone, A., 2008. State-dependency of transcranial magnetic stimulation. *Brain Topogr.* 21 (1), 1–10. <http://doi.org/10.1007/s10548-008-0067-0>. State-Dependency.
- Siuda-Krzywicka, K., Bola, K., Paplinska, M., Sumera, E., Jednorog, K., Marchewka, A., et al., 2016. Massive cortical reorganization in sighted braille readers. *ELife* 5 (MARCH 2016), 1–26. <http://doi.org/10.7554/eLife.10762>.
- Spieser, L., Aubert, S., Bonnard, M., 2013. Involvement of SMAp in the intention-related long latency stretch reflex modulation: a TMS study. *Neuroscience* 246, 329–341.
- Walsh, V., Pascual-Leone, A., Kosslyn, S.M., 2003. Transcranial Magnetic Stimulation: a Neurochronometrics of Mind. MIT press Cambridge, MA.
- Wassermann, E.M., 1998. Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the international workshop on the safety of repetitive transcranial magnetic stimulation, June 5–7, 1996. *Electroencephalogr. Clin. Neurophysiology Evoked Potentials Sect.* 108 (1), 1–16.
- Yoncheva, Y.N., Zevin, J.D., Maurer, U., McCandliss, B.D., 2010. Auditory selective attention to speech modulates activity in the visual word form area. *Cerebr. Cortex* 20 (3), 622–632.
- Yousry, T.A., Schmid, U.D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., Winkler, P., 1997. Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. *Brain* 120 (1), 141–157.
- Zhao, L., Chen, C., Shao, L., Wang, Y., Xiao, X., Chen, C., et al., 2016. Orthographic and phonological representations in the fusiform cortex. *Cerebr. Cortex* 1–14. <http://doi.org/10.1093/cercor/bhw300>.