



HAL
open science

Learning to read transforms phonological into phonographic representations: Evidence from a Mismatch Negativity study

Chotiga Pattamadilok, Shuai Wang, Deidre Bolger, Anne-Sophie Dubarry

► To cite this version:

Chotiga Pattamadilok, Shuai Wang, Deidre Bolger, Anne-Sophie Dubarry. Learning to read transforms phonological into phonographic representations: Evidence from a Mismatch Negativity study. 2024. hal-04831380

HAL Id: hal-04831380

<https://hal.science/hal-04831380v1>

Preprint submitted on 11 Dec 2024

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 - NonCommercial - NoDerivatives 4.0 International License

1 Running title: From phonological to phonographic representations

2 **Learning to read transforms phonological into phonographic representations:**
3 **Evidence from a Mismatch Negativity study**

4 Chotiga Pattamadilok^{a*}, Shuai Wang^a, Deidre Bolger^{a,b}, & Anne-Sophie Dubarry^c

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

6 ^b Institute of Language, Communication and the Brain (ILCB)

7 ^c Aix Marseille Univ, CNRS, CRPN, Marseille, France

8

9

10 *Corresponding author*

11 Chotiga Pattamadilok, PhD.

12 Laboratoire Parole et Langage

13 Centre National de la Recherche Scientifique (UMR 7309)

14 5, Av. Pasteur

15 13100 Aix-en-Provence, France

16 Email: chotiga.pattamadilok@univ-amu.fr

17 Tel : +33 601323435

18

19

20

Abstract

21 Learning to read changes the nature of speech representation. One possible change consists
22 in transforming phonological representations into phonographic ones. However, evidence
23 for such transformation remains surprisingly scarce. Here, we used a novel word learning
24 paradigm to address this issue. During a learning phase, participants were exposed to
25 unknown words in both spoken and written forms. Following this phase, the impact of
26 spelling knowledge on spoken input perception was assessed at two time points through an
27 unattended oddball paradigm, while the Mismatch Negativity component was measured by
28 high density EEG. Immediately after the learning phase, no influence of spelling
29 knowledge on the perception of the spoken input was found. Interestingly, one week later,
30 this influence emerged, making similar sounding words with different spellings more
31 distinct than similar sounding words that also share the same spelling. Our finding provides
32 novel neurophysiological evidence of an integration of phonological and orthographic
33 representations that occurs once newly acquired knowledge has been consolidated. These
34 novel phonographic representations may characterize how known words are stored in
35 literates' mental lexicon.

36

37

38 **Introduction**

39 The ability to process spoken language is considered to be universal. This intuitive
40 assumption implies that literate and illiterate individuals share at least the most
41 fundamental speech processing skills that allow them to recognize spoken words. In line
42 with this view, *all* existing psycholinguistic models of spoken word recognition have been
43 built upon data that were exclusively collected in literate populations. In other words, any
44 potential differences in how literate and illiterate listeners recognize spoken words and in
45 the nature of representations stored in their mental lexicons have been disregarded (see ^{1,2}
46 for general discussions on the impact of literacy acquisition on human language and
47 cognition).

48 The present study investigated the nature of the representations stored in the
49 phonological lexicon of literate listeners. It addressed a specific question of whether these
50 representations are purely phonological. The question emerged from the existing literature
51 which suggests that the impact of learning to read is far beyond a simple connection
52 between the two language codes. More specifically, this connection, which is promptly
53 established after being exposed to the spoken and written input, could induce more
54 profound and long-term changes within the spoken language system itself. For instance, it
55 has repeatedly been reported that spoken words that had inconsistent spellings (e.g., /aɪt/
56 can be spelled ‘ite’, ‘ight’, or ‘yte’) took more time to be recognized than spoken words
57 that had consistent spellings e.g., /ʌg/ is always spelled ‘ug’)³⁻⁶. This increase in reaction
58 time was associated with an increase in amplitude of a negative-going ERP waveform
59 around 320-350 msec after stimulus presentation; a time-window that coincided with the
60 onset of lexical access^{7,8}. Interestingly, further examination of the cortical source of this
61 ERP signal showed that the so-called “orthographic effect” was located in the left
62 supramarginal gyrus (SMG), a classic phonological area. No evidence for activation in the

63 left-ventral occipital cortex (vOT, also known as the Visual Word Form Area) coding
64 orthographic information was found⁷. The same conclusion was obtained in a study that
65 used transcranial magnetic stimulation to interfere with the orthographic consistency effect
66 by interrupting the function of either the left-SMG or the left-VOT⁹. The result showed that
67 the reaction time advantage of orthographically consistent over inconsistent words was
68 removed when the stimulation was delivered to the left SMG whereas it was maintained
69 when delivered to the left-vOT.

70 Another type of evidence for a modulation of brain activity within spoken language
71 areas following reading acquisition comes from comparisons of brain activity in
72 participants with different levels of reading expertise. Using this approach, Dehaene and
73 colleagues¹⁰ compared brain activity (measured by fMRI) of literate, ex-illiterate and
74 illiterate Brazilian/Portuguese participants during spoken word recognition and sentence
75 listening. The authors found that the activity within the planum temporale increased
76 gradually with participants' reading skill. However, it is worth noting that Hervais-
77 Adelman et al.¹¹ who addressed a similar issue in Hindi speakers who were either illiterate
78 or literate in a non-alphabetic script did not observe the same result. These diverging results
79 further suggest that the impact of literacy on the modulation of brain activity should not be
80 interpreted as an all-or-none phenomenon. In addition to inter-individual variability such
81 as reading expertise, the characteristics of the writing system itself might also play an
82 important role.

83 Despite the various origins of variability, the reported modulations of brain activity
84 within the spoken language system following reading acquisition seems to fit with the
85 general idea that phonological knowledge is progressively restructured through language
86 experience^{12,13}. As discussed by Muneaux and Ziegler^{14,15}, although the lexical
87 restructuring framework mainly focuses on the role of vocabulary size, it is plausible that

88 reading acquisition, which is one of the most critical language experiences, also contributes
89 to restructuring the spoken language system to some extent^{16–18}. One form of lexical
90 restructuring is the emergence of abstract *phonographic* representations. These abstract
91 representations reflecting an integration of phonological and orthographic features of
92 language input have been mentioned or alluded to in the literature on several occasions
93 ^{14,19–22}. Yet, direct evidence of such transformation remains surprisingly scarce. One
94 noteworthy finding was reported by Bakker et al.²³ who conducted behavioral experiments
95 to investigate the formation of novel word memories across modalities. In their
96 experiments, participants learned either the spoken or written version of novel words. After
97 the learning phase, within- and cross-modal integration of the novel words to the mental
98 lexical was behaviorally assessed by measuring the degree to which visually or auditorily
99 acquired novel words engaged in a lexical competition with existing phonological or
100 orthographic neighbors. The authors found that the novel words learned in the spoken form
101 needed 24h before entering into phonological and orthographic neighbor competitions.
102 Novel words learned in the written form needed 24h before entering into orthographic
103 neighbor competition but one week before entering into phonological neighbor
104 competition. This behavioral evidence of a spontaneous generation of cross-modal
105 representations from language input that had been learned in a specific modality indicates
106 that, once an individual had become literate, the information contained in his/her mental
107 lexicon might become multimodal. Interestingly, this claim is compatible with
108 neurophysiological evidence reported by Pattamadilok et al.²⁴. In this
109 electroencephalography (EEG) study, the authors examined the nature of the
110 “phonological” lexical representations in literate participants by recording auditory
111 Mismatch Negativity (MMN) responses in an unattentive oddball paradigm where
112 participants were exposed to sequences of standard and deviant spoken words. Standard

113 and deviant stimuli always ended with the same phonological rimes. Critically, some
114 deviant stimuli also shared the same rime spelling with the standard stimuli while others
115 did not (e.g., standard: /tRi/, spelled “tri”; deviant with the same rime spelling: /kRi/,
116 spelled “cri”, deviant with different rime spelling: /pRi/, spelled “prix”). Although the
117 manipulation of orthographic congruency remained implicit in this auditory oddball
118 protocol, the amplitude of the MMN was found to be modulated by participants’ knowledge
119 of the spoken words’ spelling: A higher MMN amplitude was observed in the situation
120 where the rime spellings of the standard and the deviant spoken words were mismatched
121 than when they were matched. Since MMN is an index of automatic detection of changes
122 in both concrete and abstract features of the sensory input²⁵⁻²⁷, its sensitivity to the
123 orthographic feature of spoken input suggests that this feature might be integrated into the
124 memory traces of spoken words.

125 However, the stimuli used in Pattamadilok et al.²⁴ were known spoken words with
126 known spellings, and thus, did not allow the authors to reveal any transformation of the
127 nature of the underlying representations at different learning stages. In other words, even
128 if there was an integration of phonological and orthographic features, it would have already
129 taken place well before the experiment. The present study aimed to fill this gap by
130 conducting a two-phase study. In the *novel word learning phase*, participants learned the
131 spoken and written forms of unknown words. Contrary to past research, this protocol
132 allowed us to control two critical factors, i.e., the participants’ knowledge about
133 pronunciation and spelling of the stimuli and the precise moment at which this knowledge
134 was presented to the participants. In the second phase, the nature of the mental
135 representations associated with the newly acquired words was examined through an
136 *unattended oddball paradigm*. In this phase, participants were passively exposed to the
137 spoken version of the novel words that they had learned, and their brain activity was

138 recorded with high density EEG. As in the previous study²⁴, our critical manipulation was
139 the orthographic congruity of the learned words presented as standard and deviant in the
140 oddball paradigm (e.g., standard: /izo/, spelled “izôt”; deviant with the same rime spelling:
141 /ivo/, spelled “ivôt”, deviant with different rime spelling: /i3o/, spelled “ijaux”). The
142 difference between the auditory MMNs obtained in the orthographically incongruent and
143 orthographically congruent situation was examined at two time points, i.e., immediately
144 after the learning phase and one week later²³, without additional learning between the two
145 experimental sessions. The second EEG recording session was added to the protocol
146 following the predictions made by the Complementary Learning Systems model of
147 memory consolidation^{28–30}. According to this framework, lexical acquisition operates in
148 two stages. The first stage corresponds to a rapid formation of episodic memory. The
149 episodic representations are stored in the hippocampal system as independent, context-
150 specific, and non-integrated sparse codes. This initial learning stage was followed by a
151 slower neocortical learning during which novel information is gradually integrated into
152 existing memory networks by forming connections between the novel words and
153 orthographically, phonologically, or semantically related existing words. Hence, the
154 resulting consolidated representations are assumed to be encoded in widely distributed
155 networks and are no longer modality specific²³. A number of studies provided evidence for
156 the role of overnight sleep in the transfer of newly acquired knowledge stored as episodic
157 representations into lexicalized long-term memory, with the duration of the consolidation
158 period varying from a few hours to several days^{23,31–34}. So far, this framework has mainly
159 been used to explain how new knowledge is integrated into the existing body of knowledge
160 ^{32,35–37}. Here, we argued that the same integration process might also apply to the combining
161 of information from different input modalities associated with the new knowledge.

162 The present study aims at revealing the nature of the underlying representations
163 associated with the newly learned spoken words, and how these representations evolved
164 across the two learning phases. We hypothesized that if the memory traces of the novel
165 spoken words only contained the phonological features, the MMN would not be sensitive
166 to the orthographic congruency between the standard and the deviant stimuli. If, in contrast,
167 the orthographic features were integrated into these memory traces, as reported in
168 Pattamadilok et al.²⁴, MMN amplitude should increase in the situation where the standard
169 and deviant spoken stimuli were orthographically incongruent. Importantly, the evolution
170 of the MMN across the two time points should provide insightful information about the
171 dynamic of the underlying representations. Specifically, if the impact of spelling
172 knowledge resulted from an activation of the episodic memories of word spellings acquired
173 during the learning phase, it should be detected during the first EEG session which took
174 place immediately after the learning phase, when the episodic traces were still recent.
175 However, if the impact of spelling knowledge resulted from an integration of the
176 phonological and orthographic features that occurred during the lexicalization process, it
177 should be detected during the second EEG session, once the newly acquired knowledge
178 had been consolidated and lexicalized.

179

180

181

182

183

184 **Results**

185 **Behavioral data**

186 The letter detection task applied during the learning phase showed that, after the initial
187 exposure and the first dictation task, all participants had successfully learned the spellings
188 of all novel words. The averaged accuracy score obtained in the task was 96.3% (range:
189 87.5% - 98.7%). All participants obtained 100% correct responses in the subsequent
190 dictation tasks conducted on Day 1 and Day 8 after the EEG recording session, which
191 suggested that the spelling knowledge was correctly learned and remained accurate one
192 week after the exposure to the novel words.

193 **Electroencephalographic data**

194 *ERP analysis*

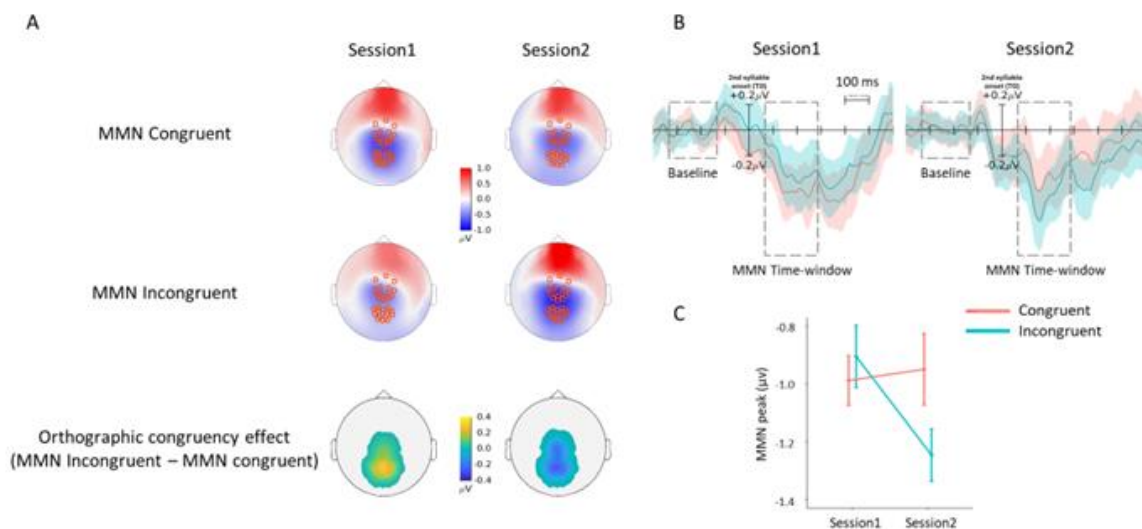
195 For each experimental session and each participant, peak amplitude, Area Under Curve
196 (AUC) and peak latency of the MMNs obtained for the orthographically congruent and
197 incongruent conditions were extracted from three electrode clusters (anterior central,
198 central and posterior central) selected based on extant literature. Prior to analyzing the
199 effects of orthographic congruency and experimental session, we assessed whether the
200 grand average MMN (which included all experimental conditions and sessions) was
201 observed as expected in each electrode cluster. The comparison of the MMN peak
202 amplitudes extracted from the time-window of interest with the peak amplitude extracted
203 from the baseline period showed significant differences in the central [$t(127) = -7.63, p <$
204 $.0001$] and posterior central clusters [$t(127) = -5.03, p < .0001$] but not in the anterior

205 central cluster [$t(127) = 1.76, p > .08$]. Based on this observation, only the central and
206 posterior central clusters were retained for further analyses.

207 Following the above preliminary analysis, we examined the effects of
208 orthographic congruency and session. The analyses conducted on MMN peak amplitude
209 and AUC showed the same pattern of results. The ANOVA summary of the LME showed
210 a significant interaction between orthographic congruency and session for both dependent
211 variables [$F(1, 224) = 4.47, p = .036$ and $F(1, 224) = 4.47, p = .036$ for peak amplitude and
212 AUC, respectively]. No other main effects or interactions showed any significance (p
213 values $> .094$). Based on the significant interaction, we specifically tested our main
214 hypothesis on the influence of orthographic knowledge on spoken language perception by
215 comparing the data obtained in the orthographic congruent and orthographic incongruent
216 conditions within each experimental session (i.e. Day 1 and Day 8). As illustrated in Figure
217 1, in the first experimental session, MMNs obtained in the orthographically congruent and
218 incongruent conditions were statistically equivalent [$t(231) = 0.64, p > .51, 95\%$ CI of the
219 difference between the two MMNs $-0.34, 0.17$ and $t(231) = 0.80, p > .42, 95\%$ CI $-6.35,$
220 2.69 for peak amplitude and AUC, respectively]. Interestingly, when the MMNs were
221 measured one week later, we found that MMN amplitude induced by the orthographically
222 incongruent deviant was larger than that induced by the orthographically congruent deviant
223 [$t(231) = -2.30, p = .022, 95\%$ CI $-0.55, -0.042$ and $t(231) = -2.14, p = .033, 95\%$ CI $-9.43,$
224 -0.40 for peak amplitude and AUC, respectively]. The difference between the pattern of
225 results obtained in the first and the second experimental session was caused by a significant
226 increase of MMN amplitude generated by the orthographically incongruent deviant
227 measured one week after the learning session [$t(231) = -2.64, p = .009, 95\%$ CI $-0.60 -$
228 0.087 and $t(231) = -2.54, p = .012, 95\%$ CI $-10.33, -1.30$ for peak amplitude and AUC,
229 respectively]. No difference between the MMNs generated by the orthographically

230 congruent deviant was found when the two sessions were compared [$t(231) = 0.30, p >$
231 $.76, 95\% \text{ CI } -0.29, 0.22$ and $t(231) = 0.40, p > .68, 95\% \text{ CI } -5.44, 3.59$ for peak amplitude
232 and AUC, respectively].

233 Finally, the analysis conducted on MMN latency only revealed a significant main
234 effect of cluster [$F(1, 224) = 5.30, p = .022$]. The latency of the MMN recorded in the
235 central cluster was slightly shorter than that observed in the posterior central cluster
236 (139 ± 20 ms vs. 145 ± 26 ms). No other main effect or interaction presented a significant
237 result (p values $> .13$).



238

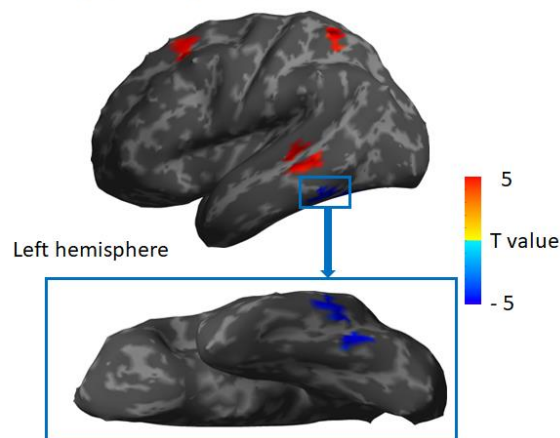
239 **Figure 1.** MMN topographic maps, waveforms, and peak amplitude. **a.** Full scalp
240 topographies of the MMNs obtained in the orthographically congruent and incongruent
241 conditions in session 1 and session 2 (the electrodes from the central and posterior-central
242 clusters highlighted in red circles). Topographical view of the difference between the two
243 conditions at the central and posterior central clusters showing the impact of orthographic
244 congruency on the MMN amplitudes. **b.** Averaged ERP waveforms (T_0 = second syllable)
245 extracted from the central and posterior central clusters, for each condition in each session.
246 **c.** MMN peak amplitudes extracted from the central and posterior-central electrode

247 clusters, for each condition in each session. Error ribbons in B and error bars in C represent
248 standard error of the means.

249 *Source localization analyses*

250 To explore the generators of the orthographic congruency effect observed in the ERP
251 analysis, we compared the EEG reconstructed sources in the incongruent condition
252 compared to the congruent condition recorded during the second experimental session. The
253 statistical analysis of the full cortex revealed stronger neural responses in the incongruent
254 condition, located in the left middle temporal gyrus (MTG), left middle frontal gyrus
255 (MFG) and left superior parietal lobule (SPL), whereas a decreased neural response in the
256 incongruent condition was observed in the left fusiform (FUS) (p values < 0.001 unc. with
257 cluster size > 10; see Figure 2 and Table 1S in supplementary material).

Orthographic congruency effect observed in Session 2



258

259 **Figure 2.** Group level statistical source map of the orthographic congruency effect
260 (incongruent - congruent) observed in the second experimental session. The computation
261 was conducted in the MMN time-window. Left middle temporal cortex, left middle frontal
262 gyrus and left superior parietal lobule showed stronger neural responses in the incongruent

263 compared to the congruent condition while the left fusiform showed the opposite pattern
264 of result. P values < 0.001 unc. with cluster size > 10.

265 **Discussion**

266 Phonological lexicon is the core element of spoken word recognition models. Yet, despite
267 a significant amount of work devoted to understanding the nature of the representations
268 stored in the mental lexicon, no consensus has been reached regarding the format and the
269 level of abstractness of the representations. The absence of such a consensus could be due,
270 in part, to the plasticity of the speech recognition system itself which, under some
271 circumstances, enables newly acquired knowledge or language experience to modulate the
272 nature of the existing representations and interactions between them (see ^{38,39} for extensive
273 reviews on this topic). Based on the current literature on the impact of reading acquisition
274 on speech processing, the present study further explored the consequences of this plasticity
275 by testing the hypothesis that, once individuals learn to read, the nature of spoken words'
276 representations might no longer reflect the phonological properties of spoken words alone
277 but would also be contaminated by how spoken words are spelled.

278 The novel word learning protocol adopted here allowed us to manipulate the
279 phonological and orthographic representations associated with new linguistic inputs, as
280 well as the moment at which these inputs were learned by the participants. After the
281 learning phase, the nature of the novel words' representations was assessed in a low-level
282 and strategy-free processing context by examining the modulation of MMN recorded in an
283 unattentive oddball paradigm where only the spoken form of the words was presented. The
284 analyses that we conducted aimed to address two questions: First, whether the MMN
285 elicited by the spoken form of the new words was also sensitive to implicit knowledge of
286 the words' spellings as previously reported on known words²⁴. Second, if such sensitivity

287 exists, would it emerge immediately after being exposed to the linguistic inputs or only
288 after a consolidation period.

289 The analysis of the MMN recorded immediately after the learning phase showed
290 no sensitivity to the manipulation of word spelling even though the participants had
291 correctly acquired and memorized this information. Interestingly, the sensitivity to spoken
292 word' spelling spontaneously emerged one week later without additional exposure to the
293 stimuli between the two sessions. This sensitivity was reflected by a significant increase of
294 the MMN amplitude (indexed by peak and AUC) in the condition where the standard and
295 the deviant spoken words were orthographically incongruent compared to when they were
296 orthographically congruent. Given that this difference was not observed in the first session
297 with the same stimuli, it could not be attributed to potential uncontrolled differences in the
298 acoustic features of the stimuli used in the different experimental conditions. Further
299 exploration of the cortical sources of the orthographic congruency effect obtained in the
300 second session showed a stronger brain activity in distributed brain areas including the left
301 MTG, left MFG, and left SPL. Interestingly, we also observed a reduction of brain activity
302 in the left FUS.

303 What does this change in the sensitivity to spoken word spelling across the two
304 experimental sessions tell us about the nature of the representations stored in the mental
305 lexicon? As mentioned in the Introduction, our finding is considered within the
306 Complementary Learning Systems (CLS) framework²⁸⁻³⁰. The behavioral data collected in
307 both letter detection and dictation tasks clearly indicated that all participants correctly
308 learned the spoken and written forms of the novel words after a short learning period.
309 According to the CLS, at this initial learning stage, these two input modalities would be
310 stored as separate entities. Our finding is coherent with this claim: The MMN observed
311 immediately after the learning phase was sensitive only to the phonological aspect of the

312 spoken inputs, thus suggesting that the episodic memory traces of the spoken words
313 presented during the unattentive oddball task remained modality specific at the initial stage.
314 Most importantly, the pattern of MMN response recorded one week later revealed a change
315 in the nature of the representations. Although our protocol did not allow us to identify the
316 precise moment when the change occurred, it clearly suggests that the modality-specific
317 episodic memory traces of the spoken word forms (and most probably of the written word
318 forms) had been consolidated and transferred to long-term memory within the interval of
319 one week. The consolidation process led to an integration of the phonological and
320 orthographic features of the inputs, thus, rendering the underlying representations of
321 spoken words multimodal (or amodal, see²³) as revealed by the MMN response. This
322 observation is coherent with both the MMN pattern reported in Pattamadilok et al.'s study²⁴
323 conducted on known words and a behavioral study reported by Bakker and colleagues²³
324 who argued that “...*in the process of consolidation, novel words acquire an amodal lemma*
325 *representation which no longer contains detailed episodic information about initial*
326 *exposure conditions. This amodal representation would be able to engage in competition*
327 *both in spoken and printed word recognition, regardless of modality consistency between*
328 *training and test phases.*” (p. 127). Although no cross-modal lexical competition was
329 assessed in the present protocol, the fact that the MMN elicited by spoken words became
330 sensitive to their spellings is coherent with the presumed transformation of modality-
331 specific word forms to multimodal or amodal representations. The claim regarding
332 converging representations of words that have been learned in the spoken and written
333 modality is also coherent with the conclusion proposed by Partanen and colleagues⁴⁰ who
334 argued in favor of similar cortical dynamics that are involved in lexicalization of novel
335 word forms learned in auditory and visual modalities. Regarding the underlying cortical
336 sources activity related to the consolidation of new linguistic knowledge, Bakker-Marshall

337 and colleagues⁴¹ reported in an EEG study a modulation of theta oscillations, which was
338 argued to reflect a consolidation process, occurring in the left posterior middle temporal
339 gyrus (pMTG), a region known to be involved in lexical storage^{42–45}. While EEG-based
340 source localization does not allow us to draw a strong conclusion on the cortical sources of
341 the effect of interest, the fact that our analysis revealed the left-MTG as one of the regions
342 that showed an increase of activity in the orthographically incongruent condition in the
343 second session is coherent with Bakker-Marshall et al.'s finding. The two other brain
344 regions that showed an increase of activity, i.e., the left-MFG and the left-SPL, are
345 generally considered as parts of the domain-general multiple demand network that supports
346 high-level cognition^{46–48}. Their recruitment has also been reported in various linguistic
347 tasks, along with other areas of the phonological network such as the left inferior frontal
348 gyrus and the left temporoparietal region^{49–51}. For the time being, one could speculate that
349 the effect observed in these regions could be linked to additional cognitive resources
350 allocated to processing incoming inputs that differed from the existing memory traces in
351 both phonological and orthographic dimensions. Interestingly, in addition to the increase
352 of activity in the regions mentioned above, we also observed a decrease of activity in the
353 left-FUS at the approximate location of the Visual Word Form Area⁵². This activation
354 pattern, which is opposite to the direction of the effect captured by the MMN component,
355 argued against the possibility that the orthographic knowledge might be activated during
356 the task. On the contrary, it seems to reflect a cross-modal inhibition of the visuo-
357 orthographic system during spoken word perception⁵³. The observed decrease of left-FUS
358 activity in the orthographically incongruent compared to congruent condition could result
359 from a stronger top-down inhibition from the spoken language system to the visuo-
360 orthographic system in the presence of conflicting information. While awaiting
361 confirmation from future research, the findings from this exploratory source localization

362 point to a consolidation process that recruits brain areas in distributed networks within and
363 outside the spoken language system.

364 **Conclusion**

365 So far, the potential impact of reading acquisition on speech processing has mainly been
366 explained by the online activation of orthographic representations that occurs while speech
367 is being processed^{52,54}. While the evidence for the online activation of orthographic
368 representations is particularly robust in high-level speech processing tasks (see ⁵⁵ for a
369 review), the findings obtained in speech processing tasks that involve lower perceptual
370 stage are less consistent⁵⁶⁻⁵⁸, thus leading to a debate about whether the impact of learning
371 to read on speech processing truly exists or simply reflects experimental or task-dependent
372 artifacts. Here, we avoid this pitfall by investigating the impact on orthographic knowledge
373 in a low-level speech processing context where spoken inputs were processed in the
374 absence of focussed attention. Our finding clearly indicates that the impact of reading
375 acquisition on speech processing does not necessarily reflect an online mapping or
376 interaction between the two language codes. In fact, orthographic knowledge has already
377 exerted its influence on speech representations well before participants performed the task,
378 i.e., by rendering speech representations multimodal. This transformation occurs
379 spontaneously outside individuals' awareness or strategic control. Hence, this new
380 evidence raises the possibility that, once we learn to read, the representations stored in our
381 "phonological" lexical may no longer be purely phonological.

382

383 **Methods**

384 **Participants**

385 Sixteen right-handed native French speakers aged 18–32 years (mean 23.3 years)
386 participated as paid volunteers. They were randomly separated into two groups of four men
387 and four women. All had normal hearing and no history of neurological or language
388 disorders. The experiment was performed in accordance with the Declaration of Helsinki
389 and was approved by the local ethics committee (Sud Méditerranée I: RCB 2015-A00845-
390 44). Informed consent was obtained from all participants.

391 **Stimuli**

392 The novel words to be learned by the participants corresponded to two triplets of disyllabic
393 pseudowords having $V_1-C_2V_2$ and $V_1-C_2V_2C_2$ phonological structures. Within each triplet,
394 one novel word was used as standard stimulus and two as deviant stimuli in the auditory
395 oddball paradigm. The novel words within each triplet differed at the position of the onset
396 of the second syllable (C_2). Otherwise, they always shared the first syllable (V_1) and ended
397 with the same phonological rime (V_2 or V_2C_2).

398 To manipulate the orthographic congruency between the standard and the deviant
399 stimuli, for each triplet, the standard novel word was associated with one spelling while
400 each deviant novel word was associated with two spellings. One of the two spellings was
401 congruent with the spelling of the standard stimulus while the other one was incongruent.
402 As shown in Table 1, for participants in Group 1, the rime spelling of Deviant 1 was
403 orthographically congruent with that of the standard novel word, while the rime spelling of
404 Deviant 2 was orthographically incongruent. The associations were reversed for the

405 participants in Group 2. Overall, the standard stimulus remained constant and each deviant
406 stimulus was associated with both congruent and incongruent spellings across the two
407 groups of participants. This manipulation allowed us to ensure that any differences that
408 might be observed between the MMNs obtained in the orthographically congruent and
409 incongruent conditions were not due to the physical properties of the deviant stimuli (e.g.,
410 amplitude, duration, intensity and spectral characteristics).

411 Table 1: Two novel word triplets and their associated spellings presented to the two
412 groups of participants.

	Triplet 1		Triplet 2	
	spoken form	Written form Group 1 Group 2	spoken form	Written form Group 1 Group 2
Standard	/izo/	izôt	/ødõt/	eudonte
Deviant 1	/ivo/	ivôt ivaux	/øbõt/	eubonte eubompte
Deviant 2	/izo/	ijaux ijôt	/øgõt/	eugompte eugonte

413

414 The spoken pseudowords were each recorded several times using a female native
415 French speaker voice in a soundproof room. All recorded stimuli were digitized at a
416 sampling rate of 44.1 kHz with 16-bit analog-to-digital conversion. The final token of each
417 stimulus was selected such that the three final stimuli within a triplet were as acoustically
418 similar as possible except for C₁. The final stimuli were normalized to the same loudness
419 and matched for peak intensity. The acoustic durations of the entire stimulus and the
420 approximate duration from the stimulus onset to C₁ onset judged by two phoneticians are
421 as follow: /izo/ = 398 ms (stimulus duration), 134 ms (first syllable duration) ms., /ivo/ =
422 394 ms, 134 ms., /izo/ = 394ms., 134ms., /ødõt/= 667 ms., 131 ms., /øbõt/ = 659 ms., 114
423 ms., /øgõt/ = 657 ms., 114 ms.

424 Procedure

425 Each participant was individually tested over two sessions separated by one week. The
426 first session began with the novel word learning phase followed by the unattentive EEG
427 oddball paradigm recording. Only the unattentive oddball task was repeated in the second
428 session without additional learning.

429 **First session (Day 1)**

430 Novel word learning phase: The learning phase was separated into four parts described
431 below and lasted a total of 30 min.

432 *Part 1*: At the beginning of the learning phase, participants were simultaneously exposed
433 to the spoken and written forms of the six novel words from the two triplets. The spoken
434 inputs were presented through earphones while the written inputs were visually presented
435 on a computer screen. The associations between the spoken and written forms
436 corresponded to the description in Table 1. Each novel word was presented eight times,
437 leading to a total of 48 stimuli presented in a random order.

438 *Part 2*: After the initial exposure, participants performed a dictation task to strengthen the
439 learning of novel words' spellings. During this task, the six novel words were orally
440 presented one by one in a random order and the participants were required to write down
441 each novel word using the spelling that they had learned. The list of six novel words was
442 repeated three times.

443 *Part 3*: Following the dictation task, the association between sounds and spellings were
444 further reinforced using a letter detection task. On each trial, a letter was presented on-
445 screen for 750 ms and immediately after, one of the six novel words was presented aurally.
446 Participants had to decide as fast and as accurately as possible whether the letter was
447 present in the novel word's spelling. Feedback (correct vs. incorrect response) and the
448 correct spelling of the novel word were provided on the computer screen after each trial.

449 The task comprised six blocks of 36 trials (six repetitions per novel word within each
450 block). On half of the trials, the letter was present in novel word's spelling.

451 *Part 4:* At the end of the learning phase, the dictation task was conducted again on the same
452 six novel words to ensure that their spellings had been correctly acquired.

453 *Unattentive oddball paradigm with EEG recording:* The two triplets of novel words were
454 presented in two separated blocks. Each block contains 960 repetitions of the standard
455 stimulus and 120 repetitions of each of the deviant stimuli, corresponding to 11.1% of all
456 trials. The stimuli were randomly presented with the constraint that there were at least three
457 standard stimuli between two deviants. The inter-stimulus interval varied between 500 and
458 700 ms. Participants were informed that they would hear the novel words presented
459 repeatedly via headphones. Throughout the auditory task the participants watched silent
460 documentaries (without subtitles to avoid activation of orthographic representations) and
461 were asked to ignore the auditory stimuli. EEG was recorded during the entire session and
462 participants were instructed to remain as still as possible. The presentation order of the two
463 blocks was counterbalanced across participants. The session lasted approximately 45
464 minutes.

465 At the end of the session, the participants performed the dictation task on the six
466 novel words to ensure that their spelling knowledge remained accurate after the oddball
467 paradigm phase.

468 **Second session (Day 8)**

469 For each participant, the second session took place one week after the first one. The
470 participants carried out the same unattentive oddball task with EEG recording. No
471 additional novel word learning was performed. After the oddball paradigm, the dictation
472 task was conducted.

473 For all tasks, stimulus presentation and behavioral data acquisition were
474 controlled by E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

475 **Data recording and analysis**

476 *EEG recording*

477 During the experiment, participants were seated in a soundproof and dimly lit experimental
478 booth at 100 cm from a 13-inch laptop. This configuration allowed participants to visualize
479 the full screen without inducing important eye movements. The EEG system (256-channel
480 EGI Geodesic HydroCel) was continuously recording during the unattended oddball task
481 at a sampling rate of 500Hz using the EGI software Net Station 5. This system uses the
482 vertex (Cz) as the recording reference. The impedance of the electrodes was checked prior
483 to recording and was kept below 40 kOhm. The two triplets of novel words were recorded
484 in two separate runs and the order of presentation was counterbalanced across participants.
485 After EEG acquisition, the electrode positions were localized by photogrammetry via
486 EGI's GPS sensor digitization system and their GPS 3.0 Solver .

487 *Preprocessing of EEG data*

488 Offline preprocessing was conducted using Brainstorm⁵⁹ ([neuroimage.usc.edu/](http://neuroimage.usc.edu/brainstorm)
489 brainstorm). Continuous data were bandpass filtered between 0.3 and 40Hz using a FIR
490 Kaiser filter (order 7252) with a stopband attenuation set to 60dB. The EEG electrodes
491 located on the face and on the neck (75 channels) were excluded from further analysis. For
492 each participant. Power spectrum density was computed to control the overall quality of
493 the recording and to identify bad channels. After bad channel rejection, the signals were
494 re-referenced to the average of the mastoids. For one participant, one mastoid channel was
495 identified as bad, therefore a neighboring channel was used for re-referencing.

496 The continuous data were visually inspected in order to identify segments of signal
497 containing artifacts (e.g. muscle artifacts, movements, except eye blinks) and remove them
498 from subsequent analysis. This visual inspection was performed on non-labelled signals.
499 So, it was not biased by the events of interest. Eye blinks artifacts were first automatically
500 detected and corrected using signal space projection (SSP). It should be noted that in order
501 to create the SSP projector, eye blinks which co-occurred with events of interest or with
502 other types of artifacts (+-200 ms) were removed before constructing the SSP projector.

503 Epochs from -200 to 800 ms relative to the onset of the first syllable were extracted
504 and baseline corrected by DC-offset subtraction (-200 to 0 ms). Epochs containing any
505 segment of artifacts were rejected. A total of 11.04% of the trials were excluded (deviant:
506 11.11%, standard: 11.03%).

507 *ERP analysis*

508 One critical issue for the present study was to ensure that any difference that might be
509 observed between the MMNs generated by orthographically congruent and by
510 orthographically incongruent deviants was due to novel words' spellings rather than any
511 acoustic or phonetic differences that might exist between the two deviant stimuli. To this
512 end, the data obtained in the two groups of participants were combined in all subsequent
513 analyses such that the spoken forms of both deviant stimuli were presented in both
514 orthographically congruent and incongruent conditions (Table 1) and only the spelling
515 form differed.

516 Since the first syllable of the standard stimulus from Triplet 2 was 17ms longer
517 than the first syllable of the deviants, the signals of the deviants were shifted by 17 ms prior
518 to computing the MMN.

519 Three subsets of electrodes were defined by selecting three clusters of 15
520 electrodes around FCz, Cz and Pz. The anterior-central cluster contains E036, E029,
521 E022, E014, E005, E224, E030, E023, E015, E006, E215, E024, E016, E007, E207. The
522 central cluster contains E008, E198, E185, E144, E131, E090, E080, E053, E044, E017,
523 E009, E186, E132, E081, E045. The posterior-central cluster contains E100, E101, E129,
524 E099, E110, E119, E128, E141, E109, E118, E127, E140, E117, E126, E139.

525 In the first stage, we functionally identified the peak of the MMN. For each triplet,
526 the grand average signal was computed across all subjects, conditions (i.e. incongruent,
527 congruent), sessions and electrodes clusters, separately for standard and deviant. The grand
528 average MMN was obtained by subtracting the averaged signal of the standard stimuli from
529 the averaged signals of the deviant stimuli. Then, for each triplet, the peak was detected on
530 the grand average MMN over a theoretical time-window (200 to 400 ms, taking into
531 account the acoustic duration of the first syllable).

532 In the second stage, at the individual level, the MMN peak for each triplet falling
533 within a 100ms time-window centered around the group-level MMN peak was identified.
534 This procedure was conducted separately for each condition, each session and each cluster.
535 For each detected peak, three measures were extracted: 1) amplitude, 2) latency, 3) Area
536 Under Curve (AUC) corresponding to 40 ms around the peak. The same procedure was
537 conducted at both group and individual levels within the baseline period, relative to the
538 first syllable onset, from -200 to 0 ms; this allowed us to compare the presence of the MMN
539 against the real baseline signal rather than against zero.

540 In order to combine the data from the two triplets in the same analysis, we applied
541 a temporal shift of the T0 to realign the signals to the onset of the second syllable (this
542 resulted in a 134 ms and 114 ms shift for the Triplet 1 and Triplet 2, respectively). Hence,
543 the MMN latency reported here is relative to the onset of the second syllable.

544 Statistical analyses were conducted in two stages. First, we verified the existence
545 of the MMN within each of the three electrode clusters. To this end, we conducted t-tests
546 comparing the amplitude of the averaged MMN (all triplets, conditions and sessions
547 combined) extracted from the time-window of interest to the signal extracted from the
548 baseline period. The procedure was applied separately on the signal obtained from the three
549 electrode clusters. Only those clusters presenting a significant MMN were retained for the
550 subsequent analyses. Second, we examined the effects of orthographic congruency and
551 experimental session, using R software (R Core Team, 2014), with a linear mixed-effects
552 model (LME). The model was fitted with the lme4 package (Bates et al., 2014) and the p
553 values (Type III: marginal sum of squares) were computed with the lmerTest package
554 (Kuznetsova et al., 2013). The model was applied to peak amplitude, peak latency and
555 AUC. Orthographic congruency, session, cluster and their interactions were considered as
556 fixed factors. Participants, triplets and their interactions were considered as random
557 intercepts.

558 *Source localization analysis*

559 To identify the cortical origin of the orthographic congruency effect observed at the scalp
560 electrodes, an exploratory source reconstruction analysis was conducted using Brainstorm
561 software⁵⁹. To this end, the coordinates of the 256 electrodes and three landmarks (LPA,
562 RPA and Nasion) were extracted for each participant using the GPS 3.0 Solver. For each
563 participant the ICBM152 MNI template anatomy was warped such that the scalp matched
564 the head shape defined by the digitized head points. To estimate the forward (head) model,
565 the boundary element method (BEM) from OpenMEEG was applied within Brainstorm
566 with no constraint on the orientation of the sources. Noise covariance matrices were
567 calculated on a pre-stimulus baseline (i.e., -300 to 0 ms) of all trials from all conditions for
568 each recording run (two per participant). The dynamic statistical parametric mapping

569 (dSPM) method was used to estimate the neural activity at the cortical level with 15,002
570 vertices both for the averages of standard and deviant stimuli and the results obtained on
571 the standard stimuli were subtracted from those of the deviant stimuli from the congruent
572 and incongruent condition, respectively. In order to identify the cortical source of the
573 orthographic congruency effect observed at the sensor level, the peak latency of the MMN
574 detected at the sensor level was used for the source level analysis. Hence, within the time-
575 window of interest, a source reconstruction was obtained for each triplet, subject, condition,
576 and session.

577 Since the source model was computed with unconstrained orientation, each vertex
578 presented three time series (x,y,z). To reduce the three dimensions to one single dimension,
579 a PCA was computed for each vertex. The absolute value was then output to SPM12
580 (<https://www.fil.ion.ucl.ac.uk/spm/>) for further statistical comparisons. In SPM12, a paired
581 t-test examining the effect of orthographic congruency observed at the sensor level was
582 applied to source activity maps with cluster and triplet as covariates.

583 **Acknowledgements**

584 This work was supported by the French Ministry of Research: ANR-13-JSH2-0002 and
585 ANR-19-CE28-0001-01 (to C.P.), ANR-16-CONV-0002 (ILCB), ANR-11-LABX- 0036
586 (BLRI) and the Excellence Initiative of Aix-Marseille University (A*MIDEX). It was
587 performed in the Centre d'Expérimentation de la Parole (CEP), Laboratoire Parole et
588 Langage. We warmly thank Dr. Agnès Trébuchon for taking medical responsibility during
589 the study. Center de Calcul Intensif d'Aix-Marseille is acknowledged for granting access
590 to its high-performance computing resources.

591 **Additional Information**

592 *Conflict-of-interest*: None declared.

593 **Data availability**

594 The datasets generated and/or analyzed during the current study are not publicly available
595 since the ethical approval for this study does not include permission to share data in a public
596 data repository but they are available from the corresponding author
597 (chotiga.pattamadilok@univ-amu.fr) on reasonable request.

598

References

1. Kolinsky, R. & Morais, J. The worries of wearing literate glasses. *Annee. Psychol.* **118**, 321–347 (2018).
2. Huettig, F. & Pickering, M. J. Literacy Advantages Beyond Reading: Prediction of Spoken Language. *Trends Cogn. Sci.* **23**, 464–475 (2019).
3. Ziegler, J. C., Petrova, A. & Ferrand, L. Feedback Consistency Effects in Visual and Auditory Word Recognition: Where Do We Stand After More Than a Decade? *J. Exp. Psychol. Learn. Mem. Cogn.* **34**, 643–661 (2008).
4. Pattamadilok, C., Morais, J., Paulo, V. & Kolinsky, R. The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Lang. ...* 37–41 (2007).
5. Ventura, P., Morais, J., Pattamadilok, C. & Kolinsky, R. The locus of the orthographic consistency effect in auditory word recognition. *Lang. Cogn. Process.* **19**, 57–95 (2004).

6. Peereman, R., Dufour, S. & Burt, J. S. Orthographic influences in spoken word recognition: The consistency effect in semantic and gender categorization tasks. *Psychon. Bull. Rev.* **16**, 363–368 (2009).
7. Perre, L., Pattamadilok, C., Montant, M. & Ziegler, J. C. Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Res.* **1275**, 73–80 (2009).
8. Pattamadilok, C., Perre, L., Dufau, S. & Ziegler, J. C. On-line orthographic influences on spoken language in a semantic task. *J. Cogn. Neurosci.* **21**, 169–179 (2009).
9. Pattamadilok, C., Knierim, I. N., Kawabata Duncan, K. J. & Devlin, J. T. How does learning to read affect speech perception? *J. Neurosci.* **30**, 8435–8444 (2010).
10. Dehaene, S. *et al.* How Learning to Read Changes the Cortical Networks for Vision and Language. *Science (80-.)*. **330**, 1359–1364 (2010).
11. Hervais-Adelman, A. *et al.* How does literacy affect speech processing? Not by enhancing cortical responses to speech, but by promoting connectivity of acoustic-phonetic and graphomotor cortices. *J. Neurosci.* **42**, 8826–8841 (2022).
12. Stanovich, K. E. & Cunningham, A. E. Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Mem. Cognit.* **20**, 51–68 (1992).
13. Metsala, J. L. & Walley, A. C. Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. in *Word Recognition in Beginning Literacy* (eds. Metsala, J. L. & Ehri, L. C.) 89–120 (Lawrence Erlbaum Associates Publishers, 1998).

14. Muneaux, M. & Ziegler, J. C. Locus of orthographic effects in spoken word recognition: Novel insights from the neighbour generation task. *Lang. Cogn. Process.* **19**, 641–660 (2004).
15. Ziegler, J. C. & Goswami, U. Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychol. Bull.* **131**, 3–29 (2005).
16. Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A. & Eden, G. F. Development of neural mechanisms for reading. *Nat. Neurosci.* **6**, 767–773 (2003).
17. Kolinsky, R. *et al.* The impact of alphabetic literacy on the perception of speech sounds. *Cognition* 104687 (2021) doi:10.1016/j.cognition.2021.104687.
18. Morais, J. The phoneme: A conceptual heritage from alphabetic literacy. *Cognition* 104740 (2021) doi:10.1016/j.cognition.2021.104740.
19. Han, J. I. & Oh, S. The role of phonetic similarity and orthographic information in asymmetrical lexical encoding in second language. *J. Psycholinguist. Res.* **47**, 1015–1033 (2018).
20. Veivo, O. & Järvikivi, J. Proficiency modulates early orthographic and phonological processing in L2 spoken word recognition. *Biling. Lang. Cogn.* **16**, 864–883 (2013).
21. Taft, M. Orthographically influenced abstract phonological representation: Evidence from non-rhotic speakers. *J. Psycholinguist. Res.* **35**, 67–78 (2006).
22. Taft, M. & Hambly, G. The influence of orthography on phonological representations in the lexicon. *J. Mem. Lang.* **24**, 320–335 (1985).

23. Bakker, I., Takashima, A., Hell, J. G. Van, Janzen, G. & McQueen, J. M. Competition from unseen or unheard novel words : Lexical consolidation across modalities. *J. Mem. Lang.* **73**, 116–130 (2014).
24. Pattamadilok, C., Morais, J., Colin, C. & Kolinsky, R. Unattentive speech processing is influenced by orthographic knowledge: Evidence from mismatch negativity. *Brain Lang.* **137**, 103–111 (2014).
25. Shtyrov, Y., Pulvermuller, F., Naatanen, R. & Ilmoniemi, R. J. Grammar Processing Outside the Focus of Attention: an MEG Study. *J. Cogn. Neurosci.* **15**, 1195–1206 (2003).
26. Pulvermüller, F., Shtyrov, Y., Kujala, T. & Näätänen, R. Word specific cortical activity as revealed by the mismatch negativity. *Psychophysiology* **41**, 106–112 (2004).
27. Shtyrov, Y., Hauk, O. & Pulvermuller, F. Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words. *Eur. J. Neurosci.* **19**, 1083–1092 (2004).
28. McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).
29. Davis, M. H. & Gaskell, M. G. A complementary systems account of word learning: Neural and behavioural evidence. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 3773–3800 (2009).

30. Fenn, K. M., Nusbaum, H. C. & Margoliash, D. Consolidation during sleep of perceptual learning of spoken language. *Nature* **425**, 614–616 (2003).
31. Dumay, N. Sleep not just protects memories against forgetting, it also makes them more accessible. *Cortex* **74**, 289–296 (2016).
32. Gaskell, M. G. & Dumay, N. Lexical competition and the acquisition of novel words. *Cognition* **89**, 105–132 (2003).
33. Palma, P. & Titone, D. A review of the literature on sleep-related lexicalization of novel words in adults. *Psychon. Bull. Rev.* **28**, 96–121 (2021).
34. Shtyrov, Y. Neural bases of rapid word learning. *Neuroscientist* **18**, 312–319 (2012).
35. Dumay, N. & Gaskell, M. G. Sleep-associated changes in the mental representation of spoken words. *Psychol. Sci.* **18**, 35–39 (2007).
36. Leach, L. & Samuel, A. Lexical configuration and lexical Engagement: When adults learn new words. *Cogn. Psychol.* **55**, 306–353 (2007).
37. Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J. & Gaskell, M. G. Sleep spindle activity is associated with the integration of new memories and existing knowledge. *J. Neurosci.* **30**, 14356–14360 (2010).
38. McQueen, J. M. Eight questions about spoken word recognition. in *The Oxford Handbook of Psycholinguistics* (ed. Gaskell, G. M.) 37–53 (Oxford university press, 2007).
39. Weber, A. & Scharenborg, O. Models of spoken-word recognition. *Cogn. Sci.* **3**, 387–401 (2012).

40. Partanen, E. J., Leminen, A., Cook, C. & Shtyrov, Y. Formation of neocortical memory circuits for unattended written word forms: neuromagnetic evidence. *Sci. Rep.* **8**, 1–10 (2018).
41. Bakker-Marshall, I. *et al.* Theta-band Oscillations in the Middle Temporal Gyrus Reflect Novel Word Consolidation. *J. Cogn. Neurosci.* **30**, 621–633 (2018).
42. Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).
43. Gow, D. W. The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. *Brain Lang.* **121**, 273–288 (2012).
44. Bakker-Marshall, I. *et al.* Overlapping and distinct neural networks supporting novelword learning in bilinguals and monolinguals. *Biling. Lang. and Cognition* **24**, 524–536 (2021).
45. Howard, D. *et al.* The cortical localization of the lexicons: Positron emission tomography evidence. *Brain* **115**, 1769–1782 (1992).
46. Duncan, J. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn. Sci.* **14**, 172–179 (2010).
47. Davis, M. H. & Johnsrude, I. S. Hierarchical processing in spoken language comprehension. *J. Neurosci.* **23**, 3423–3431 (2003).
48. Erb, J., Henry, M. J., Eisner, F. & Obleser, J. The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *J. Neurosci.* **33**, 10688–10697 (2013).

49. Kuo, W.-J. *et al.* Orthographic and phonological processing of Chinese characters: an fMRI study. *Neuroimage* **21**, 1721–31 (2004).
50. Malik-Moraleda, S. *et al.* An investigation across 45 languages and 12 language families reveals a universal language network. *Nat. Neurosci.* **25**, 1014–1019 (2022).
51. Banaszkiwicz, A. *et al.* The role of the superior parietal lobule in lexical processing of sign language: Insights from fMRI and TMS. *Cortex* **135**, 240–254 (2021).
52. Dehaene, S. & Cohen, L. The unique role of the visual word form area in reading. *Trends Cogn. Sci.* **15**, 254–262 (2011).
53. Laurienti, P. J. *et al.* Deactivation of sensory-specific cortex by cross-modal stimuli. *J. Cogn. Neurosci.* **14**, 420–429 (2002).
54. Grainger, J. & Ziegler, J. C. Cross-code consistency effects in visual word recognition. in *Single-word reading: Biological and behavioral perspectives* (eds. Grigorenko, E. L. & Naples, A. J.) 129–157 (2007).
55. Kolinsky, R., Pattamadilok, C. & Morais, J. The impact of orthographic knowledge on speech processing. *Rev. Ilha do Desterro A J. English Lang. Lit. English Cult. Stud.* 161–186 (2012).
56. Pattamadilok, C., Morais, J. & Kolinsky, R. Naming in Noise: The Contribution of Orthographic Knowledge to Speech Repetition. *Front. Psychol.* **2**, 1–12 (2011).
57. Mitterer, H. & Reinisch, E. Letters don't matter: No effect of orthography on the perception of conversational speech. *J. Mem. Lang.* **85**, 116–134 (2015).

58. Cutler, A. & Davis, C. An orthographic effect in phoneme processing, and its limitations. *Front. Psychol.* **3**, doi: 10.3389/fpsyg.2012.00018 (2012).
59. Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D. & Leahy, R. M. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput Intell Neurosci*, **2011**, 879716 (2011).

599

600 **Author contributions**

601 C.P. A-S.D., and D.B. designed research, collected data; A-S.D and S.W. analyzed and
602 visualized data; C.P., A-S.D. and S.W. wrote the original manuscript, All authors reviewed
603 the manuscript.

604

605 **Legends**

606 **Figure 1.** MMN topographic maps, waveforms, and peak amplitude. **a.** Full scalp
607 topographies of the MMNs obtained in the orthographically congruent and incongruent
608 conditions in session 1 and session 2 (the electrodes from the central and posterior-central
609 clusters highlighted in red circles). Topographical view of the difference between the two
610 conditions at the central and posterior central clusters showing the impact of orthographic
611 congruency on the MMN amplitudes. **b.** Averaged ERP waveforms (T0 = second syllable)
612 extracted from the central and posterior central clusters, for each condition in each session.
613 **c.** MMN peak amplitudes extracted from the central and posterior-central electrode
614 clusters, for each condition in each session. Error ribbons in B and error bars in C represent
615 standard error of the means.

616 **Figure 2.** Group level statistical source map of the orthographic congruency effect
617 (incongruent - congruent) observed in the second experimental session. The computation
618 was conducted in the MMN time-window. Left middle temporal cortex, left middle frontal
619 gyrus and left superior parietal lobule showed stronger neural responses in the incongruent
620 compared to the congruent condition while the left fusiform showed the opposite pattern
621 of result. $p < 0.001$ *unc. with cluster size > 10*.

622 **Table 1:** Two novel word triplets and their associated spellings presented to the two
623 groups of participants.

624

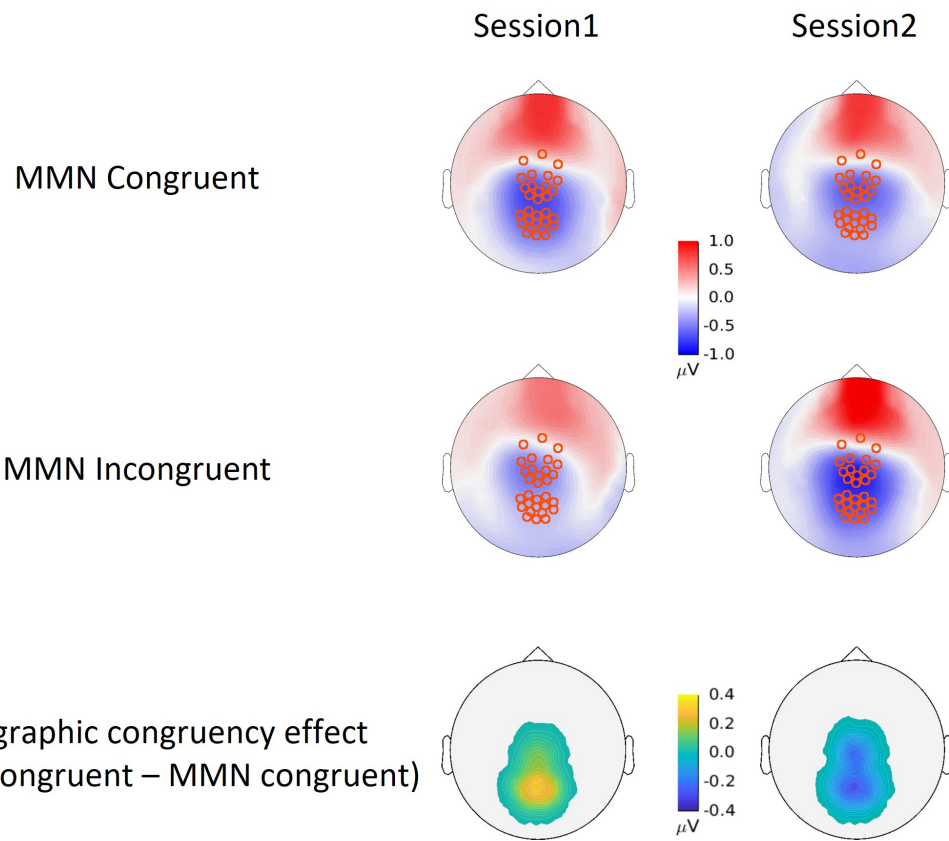
625

626

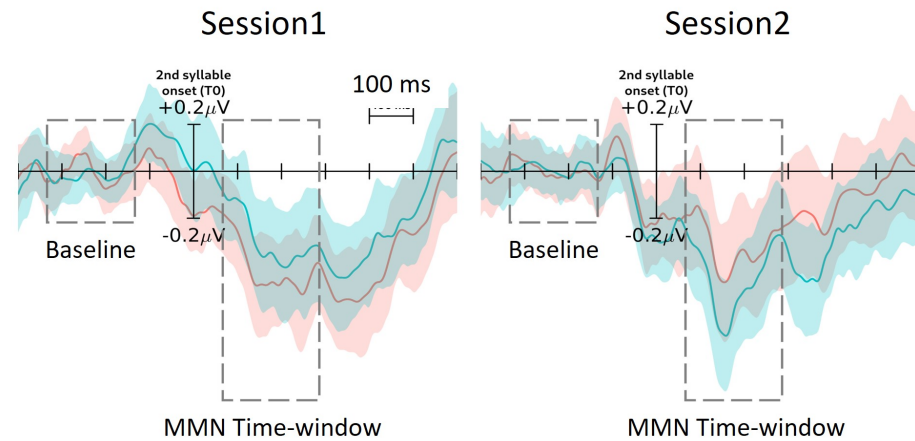
627

628

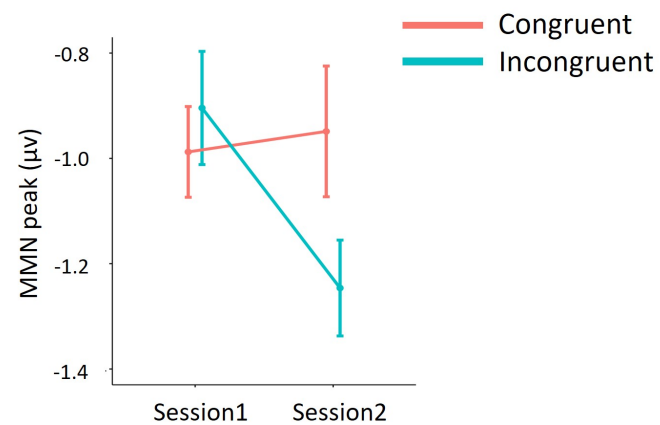
a



b



c



Orthographic congruency effect observed in Session 2

