

1 **Functional characterisation of letter-specific responses in time, space and**
2 **current polarity using magnetoencephalography.**

5 *Gwilliams, L.,^{1,2} Lewis, G. A.² & Marantz, A.^{1,2, 3}

7 ¹*Department of Psychology, New York University*

8 ²*NYUAD Institute, New York University Abu Dhabi*

9 ³*Department of Linguistics, New York University*

14 RUNNING TITLE:

15 FUNCTIONAL CHARACTERISATION OF LETTER-SPECIFIC RESPONSES

19 * Corresponding author:

21 Laura Gwilliams
22 10 Washington Place
23 6th Floor
24 New York, NY
25 10003

27 Email: laura.gwilliams@nyu.edu
28 Tel: 1-347-725-5635

30

31

Abstract

32 Recent neurophysiological evidence suggests that a hierarchical neural network of
33 low-to-high level processing subserves written language comprehension. While a
34 considerable amount of research has identified distinct regions and stages of
35 processing, the relations between them and to this hierarchical model remain
36 unclear. Magnetoencephalography (MEG) is a technique frequently employed in
37 such investigations; however, no studies have sought to test whether the
38 conventional method of reconstructing currents at the source of the magnetic field is
39 best suited for such across-subject designs. The present study details the results of
40 three MEG experiments addressing these issues. Neuronal populations supporting
41 responses to low-level orthographic properties were housed posteriorly near the
42 primary visual cortex. More anterior regions along the fusiform gyrus encoded
43 higher-level processes and became active ~80ms later. A functional localiser of
44 these early letter-specific responses was developed for the production of functional
45 regions of interest in future studies. Previously established response components
46 were successfully grouped based on proximity to the localiser, which characterised
47 location, latency and functional sensitivity. Unconventional anatomically constrained
48 signed minimum norm estimates of MEG data were most sensitive to the primary
49 experimental manipulation, suggesting that the conventional unsigned unconstrained
50 method is sub-optimal for studying written word processing.

51

52

Keywords: *fROI; LCD Model; source estimation; VWFA; visual word processing*

53
54

1. Introduction

55

56 1.1 *Primary responses to written words*

57

58 Tarkiainen *et al.* (1999) was one of the first studies to successfully disassociate the
59 neural dynamics of visual feature analysis and letter-specific recognition, in terms of
60 both neuronal location and response timing. Employing magnetoencephalography
61 (MEG), the authors identified two primary neural responses to written words,
62 interpreted as an early linguistically insensitive response (Type One), and a later
63 letter-specific response (Type Two – not to be confused with Type I & II type errors
64 in the statistical sense). They interpret these responses as reflecting the role of the
65 left inferior occipital-temporal cortex in supporting the processing of letter strings
66 during visual word processing, and propose that the Type Two signals may act as
67 the first port of call when filtering valid letter strings for further lexical processing.

68

69 A wealth of research utilising a wide range of methodologies has since been
70 conducted on the neural processes underlying written language comprehension.
71 Much of this research has focused on identifying distinct regions and stages of
72 processing that are specifically responsive to linguistic content. Table 1 below
73 outlines a number of such response components, which overlap to different extents
74 in latency, spatial location and lexical sensitivity.

75

76

77

Response / Region	Latency	Location(s)	Sensitivity	Method(s)	Representative Studies
Type One	~100 ms	V1; occipital cortex	Luminosity, visual complexity; non-linguistic properties	MEG	Tarkiainen et al., 1999; Helenius et al., 1999
Type Two	~140 ms	Temporal-occipital junction	Symbol strings vs. Letter strings, legibility of letter strings	MEG	Tarkiainen et al., 1999; Helenius et al., 1999
M130	~130 ms	Occipital lobe	Orthographic/surface properties (e.g., bigram freq., orthographic affix freq.)	MEG	Solomyak and Marantz, 2009; 2010; Lewis et al., 2011
M170	~170 ms	Occipital-temporal cortex and fusiform gyrus	Morphological properties (e.g., lemma freq., morphological affix freq., transition probability)	MEG	Solomyak and Marantz, 2009; 2010; Lewis et al., 2011; Zweig and Pylykkänen, 2009; Fruchter et al., 2013
VWFA	~180 ms	Fusiform gyrus; Talairach co-ordinates ($x = -43$, $y = -54$, $z = -12$)	High level features such letter shapes; Real words vs. Consonant strings	EEG & fMRI	Cohen et al., 2000; 2002; Dehaene et al., 2001; 2002
LFA	~160 ms	Occipital-temporal cortex; MNI co-ordinates (-40, -78, -18)	Consonants vs. False fonts	ECoG, MEG & fMRI	Thesen et al., 2012
WFA	~225 ms	Posterior fusiform gyrus; MNI co-ordinates (-46, -52, -20)	Real words vs. Consonants	ECoG, MEG & fMRI	Thesen et al., 2012

78

79

80 **Table 1.** Summary of response components and regions specifically associated with early written
 81 word processing. VWFA = “Visual Word Form Area”; LFA = “Letter Form Area”; WFA = “Word Form
 82 Area”; EEG = “Electroencephalography”; fMRI = “Functional Magnetic Resonance Imaging”; ECoG =
 83 “Electrocorticography”.

84

85 Recent investigation has suggested that rather than a single “lexically-sensitive
 86 region” of the brain, there exists a collection of hierarchical networks set along the
 87 occipital-temporal lobe and fusiform gyrus. The Local Combination Detector (LCD)
 88 model (Dehaene et al., 2005), for example, proposes graded sensitivity along the
 89 entire occipital-temporal cortex to abstract visual stimuli (i.e., written word forms)
 90 coded in a posterior-to-anterior progression. Under this account, posterior regions
 91 closest to the primary visual cortex are less selective and become active with less
 92 proximity to real words, while the most anterior portion of the fusiform overlapping
 93 with the Visual Word Form Area (VWFA) has greatest selectivity for high-level visual
 94 features.

95

96 Initial neuropsychological support for the LCD model was gathered from hierarchical
97 neural detectors in macaque monkeys (Booth and Rolls, 1998; Rolls, 2000), finding
98 neurons located anteriorly to support more abstract processing than populations
99 located posteriorly. The generalisation of these results to humans was directly tested
100 later by Vinckier *et al.* (2007). Using Functional Magnetic Resonance Imaging (fMRI),
101 they compared responses along the occipitotemporal cortex to visually presented
102 items differing systematically in their similarity to valid words: 1) false fonts; 2) strings
103 of infrequent letters; 3) infrequent bigrams; 4) infrequent quadrigrams; 5) frequent
104 quadrigrams; 6) real words. Consistent with the LCD model, they found gradient
105 sensitivity to real words in a posterior-to-anterior progression. The authors
106 characterised the specialisation of responses along the “visual word form system” as
107 supporting evidence for a graded sensitivity anterior from the occipital lobe towards
108 the defined location of the VWFA.

109

110 Recent work by Lewis *et al.* (2011) examined the influence of linguistic variables of
111 varying “abstractness” on an MEG component (putatively overlapping with the
112 VWFA) associated with morphological detection (the MEG M170; Solomyak and
113 Marantz, 2009, 2010). Restricted ROI analyses of posterior and anterior portions of
114 the M170 yielded distinct sensitivities to linguistic variables; the anterior ROI showed
115 an effect of transition probability from stem to suffix of apparently morphologically
116 complex words, whereas the posterior portion was only sensitive to surface (high-
117 level n-gram) frequency. Transition probability effects were associated with activation
118 of abstract morphemic representations, and surface frequency effects with activation
119 of concrete n-gram representations. The sources set more anterior upon the fusiform
120 gyrus therefore appeared to code more abstract representations. This finding is in

121 corroboration to the LCD model as well as Vinckier *et al.*'s (2007) results, and
122 suggests that even within a single evoked response component, defined by the
123 timing of a peak response over MEG or Electroencephalography (EEG) sensors, not
124 all sources may display the same sensitivity.

125

126 Later work by Thesen *et al.* (2012) employed fMRI and time-sensitive methodologies
127 including MEG and intracranial EEG recordings to identify distinct temporal and
128 spatial attributes of a “letter-form area” (LFA) and a “word-form area”. Findings
129 included letter-selective responses peaking 160 ms post-stimulus onset, around 60
130 ms earlier than activation of the VWFA (Cohen *et al.*, 2000; 2002). The authors
131 propose a feed-forward structure of responses, whereby the system assesses
132 information for word-likeness at different stages of processing, and subsequently
133 carries it forward through the system. This implies that access to the visual word
134 form (VWF: Warrington and Shallice, 1980) is first fed by identification of valid letters,
135 and then valid word-shapes, suggesting that the first stage of visual word recognition
136 is not VWF access but rather identification of valid letters in posterior regions.

137

138 Taken together, these findings strongly suggest that the visual system is composed
139 of neuronal populations that support graded low-to-high level processing in a
140 posterior-to-anterior arrangement. As such, established response components and
141 regions can be placed along this graded visual processing system, both to index the
142 complexity of processing they support and to assess their similarity to other
143 response components.

144

145 *1.2 Methods of source estimation*

146

147 As illustrated, investigation into the neurophysiological underpinnings of visual word
148 processing has focused on categorising stages of processing based on information
149 such as timing, location, and sensitivity of particular neural responses. An additional
150 dimension that has been used in methodologies that measure electrical current
151 directly, either from the cortical surface (Electrocorticography: ECoG) or through the
152 skull (EEG), is the polarity of the electrical current (in this case, relative to the
153 reference electrode). For example, the N400 component is defined by negative
154 current potential, and the P600 by a positive current potential. Although polarity
155 relative to a reference electrode is arbitrary, activity with opposing polarities reflect
156 distinct response components. Directionality is therefore an important dimension in
157 characterising responses in these methodologies.

158

159 MEG does not measure the electrical current in the brain directly, but rather the
160 magnetic field generated around the brain from which location, amplitude and
161 direction of neuronal currents can be estimated. In many methods of source
162 localisation, including the MNE software used in this paper, source estimation is
163 achieved by placing a current dipole at each source to be estimated, spread across
164 the brain's volume or across the cortical surface. Six co-ordinates are required to
165 characterise a dipole vector within this MRI co-ordinate space: three to position the
166 origin and three to position the vector tip. The method of deriving a single-value
167 current estimate depends on the orientation constraint applied. "Free orientation"
168 applies no constraint on the dipoles' orientation, and amplitude is estimated from the
169 vector's norm (length). Consequently, amplitude always has a positive value, and is
170 referred to as "unsigned". In "fixed orientation" the dipole is cortically constrained by

171 being projected onto the axis normal to the cortical surface. The current estimate is
172 then calculated from the norm of the projected vector - assigned a positive value
173 when the dipole is directed out of the cortical mass, and negative when oriented into
174 the cortical mass. As cortically constrained estimates can realise either positive or
175 negative values, they are referred to as "signed". The vast majority of MEG studies
176 have adopted the convention of using free orientation in source reconstruction, thus
177 removing directional information.

178

179 Cortically-constrained, signed estimates model the physiology of the cortical mass,
180 thus making it the most anatomically motivated method given the hypothesised
181 source of the MEG signal in pyramidal cells of the cortex, which are oriented
182 perpendicular to the cortical surface (Hämäläinen *et al.*, 1993). Furthermore, current
183 flowing from the base of a pyramidal cell to the pial surface is likely to support
184 distinct processes from neuronal populations whose current flows from the surface
185 into the cortical mass; however, the exact differences between polarities of the
186 current are poorly understood (da Silva, 2010). Assuming that switches in polarity
187 mirror distinct responses, unsigned free orientation loses one dimension of
188 discrimination. As a consequence, neuronal populations whose current shifts from
189 one polarity to another may be incorrectly characterised as displaying just one rather
190 than two separate responses. Furthermore, the timing of distinguished response
191 peaks may be misidentified due to erroneous averaging (see Fruchter and Marantz,
192 2015).

193

194 One complication of using fixed orientation, where sources are fit perpendicular to
195 the surface, is caused by the convolution of the cortex. Fixed orientation results in

196 the localisation of the reconstructed current as outgoing activity on one side of a
197 sulcus and ingoing on the other side. One of the localisations and directionalities
198 represents the “true” source of activity, while the other represents a reconstruction
199 “bleed”. Averaging positive and negative sources together would lead to a shallower
200 estimate of source amplitude, and thus reduce sensitivity to experimental
201 manipulations. Care must therefore be taken not to cancel out activity by selecting
202 both positive and negative sources when averaging over a ROI; for example,
203 anatomical regions often encapsulate sources from both sides of a cortical fold,
204 which would include both positive and negative data estimates if utilising a cortically
205 constrained method of signing data. Instead, a region needs to be selected in such
206 as way that the inclusion of opposite polarity is minimised.

207

208 Despite extensive investigation into the different methods of generating source
209 estimates of EEG and MEG data, there is little study into the consequences of using
210 signed versus unsigned estimates in terms of sensitivity to experimental
211 manipulations across subjects. Comparisons have instead focused on the spatial
212 accuracy of locating a source in data simulations when using free, fixed and loose
213 dipole constraining parameters, often in a single subject. Such work has advocated
214 the advantages of incorporating anatomical constraints on source orientations to
215 avoid over-estimating the spatial extent of activity (Chang *et al.*, 2013; Dale, *et al.*,
216 2000; Hauk, 2004; Lin *et al.*, 2006; Lui *et al.*, 1998; 2002). To our knowledge, only
217 one study has considered the implications of activity polarity relative to experimental
218 sensitivity (Fruchter & Marantz, 2015, Appendix B), although comparing the
219 functional differences between methods was not the main purpose of the
220 experiment. The current study therefore focused on the relative experimental

221 sensitivity of two methods of source reconstruction (free unsigned, fixed signed)
222 rather than spatial accuracy. Based on this comparison we determined whether the
223 sign of MEG data is a useful dimension to retain, and characterised the directionality
224 of Type Two responses relative to the cortical surface. The dataset employed a
225 FreeSurfer average brain to localise responses, and did not combine MEG data with
226 individual subject structural MRIs – information that is not always available to
227 researchers. It must therefore be stressed that present results aimed to compare
228 experimental sensitivity of signed and unsigned estimates and *not* accuracy in
229 localisation. If, in spite of this inherently imperfect estimation, using cortically
230 constrained methods proved to be a superior method, it would suggest that the sign
231 is an important aspect of the signal to retain.

232

233 *1.3 Letter-specific functional ROI*

234

235 The use of functionally defined regions of interest (fROIs) is frequent with
236 methodologies such as fMRI, but is not common practice with MEG. This technique
237 involves running an orthogonal experimental task known to robustly evoke a
238 functionally specific region of the brain, with subsequent analysis of the same region
239 for a critical manipulation of interest in a separate experiment (see Poldrack, 2007
240 for a discussion of fROIs in fMRI). This “localiser”, which is typically run on the same
241 sample of participants during the same experimental session, allows for a motivated
242 method of selecting a particular region of the brain to analyse.

243

244 Importantly for the present purposes, if the fROI is identified using threshold-based
245 cluster tests (explained in detail below), it avoids the complication of polarity striping

246 that comes along with using fixed orientation. This is because the fROI will consist of
247 a uni-directional set of sources, which will match the response of interest *if* the
248 processes underlying the localiser and the critical experiment are the same.
249 Furthermore, it has advantages over anatomically defined regions in avoiding
250 analysing larger areas than necessary (a problem when correcting for multiple
251 comparisons across space, and when averaging activity within a given ROI), and is
252 not constrained by borders between parcellations. Thus, if the localiser truly
253 identifies the same underlying response, statistical power should be increased, and
254 the analysis will be less susceptible to Type I errors.

255

256 Given the statistical and theoretical advantages of utilising a localiser of functionally
257 specific activity, we assessed whether fROIs of early letter-specific responses could
258 be created by results of a reduced version of the Tarkiainen paradigm. To achieve
259 this, we replicated a study known to evoke both lower- and higher-level lexical
260 processing - Solomyak and Marantz (2010: henceforth "S&M") involving a visual
261 lexical decision task of mono-morphemic and bi-morphemic words. The original
262 study's main finding was that activity around 170 ms in the fusiform gyrus (M170)
263 was modulated by transition probability from stem to suffix of the morphologically
264 complex words, as well as by morphological but not orthographic affix frequency.
265 The M170 response component appeared to index high-level processes sensitive to
266 sub-lexical (*i.e.*, morphological) structure. Earlier and more posterior processing at
267 the M130 was sensitive only to surface properties of the stimulus, whereby activity
268 was modulated by orthographic affix frequency.

269

270 Testing the utility of Tarkiainen *et al.*'s paradigm as a localiser importantly allows us
271 to assess whether cluster-based fROIs can overcome the issue of activity
272 cancellation of positive and negative sources when using cortically constrained
273 estimates. Further, if lexical variables from the S&M replication modulate activation
274 of the region(s) shown to display Type Two responses in the abridged Tarkiainen
275 design, it would suggest that the Type Two response can be used to localise the
276 lower and higher level lexical processing as identified in S&M. For the present
277 purposes, this would enable a functional comparison between neuronal populations
278 underlying the Type Two component and other established responses and regions in
279 the literature.

280

281 *1.4 Aims*

282

283 The principle aim was to establish the location, timing, and functionality of the Type
284 Two response relative to associated regions (e.g., VWFA, letter-form area, word-
285 form area) and responses (e.g., M130, M170) defined in the literature, as well as the
286 posterior-to-anterior progression of higher-level processes more broadly. The original
287 Tarkiainen *et al.* study employed multiple dipole modelling analysis, which does not
288 provide information regarding the spatial coverage of response-specific activation in
289 a region. To fully ascertain the spatio-temporal extent of Type Two activity, we ran
290 an English-adaptation of Tarkiainen *et al.*'s (1999) study while recording neural
291 responses with MEG. We then conducted a distributed source analysis to locate
292 brain regions most sensitive to our critical manipulations.

293

294 The second aim was to develop a localiser of early letter-specific activity for use as a
295 fROI in future experiments. To do so, we identified a manageable abridgement of the
296 Tarkiainen paradigm to be administered as an orthogonal task within the same
297 recording session as a replication of the S&M study – consisting of a visual lexical
298 decision task of mono-morphemic and bi-morphemic words. We examined the
299 influence of the lexical variables from S&M on activity in the regions identified by the
300 localiser to determine whether the Type Two response taps into the same neuronal
301 populations as the M130 or M170 components. Determining the functional sensitivity
302 of these regions would allow us to: 1) make stronger connections with other MEG
303 responses, following aim one above; 2) position the regions supporting the Type Two
304 response relative to the hierarchical LCD model; and 3) determine whether
305 underlying neural populations are shared between components, allowing the Type
306 Two responses to be used as a localiser of letter-specific activity more broadly
307 construed.

308

309 In aiming to characterise the Type Two response, the final goal of this study was to
310 ascertain the direction of the neural current at the source of the recorded magnetic
311 field with respect to the cortical surface. This purpose is built upon the assumption
312 that a shift in polarity indexes a distinct neural response, and that retaining the sign
313 of activation avoids loss of evoked responses due to averaging. Ultimately, we
314 assessed whether using signed estimates allows for greater ease in discriminating
315 neural responses in comparison to unsigned source estimates. To avoid the issue of
316 averaging over positive and negative estimates, all analyses employed a threshold-
317 based permutation cluster test to identify spatial regions that elicit activation of the
318 same polarity.

319

320 To address these questions, we conducted three MEG experiments: 1) A full
321 Tarkiainen *et al.* (1999) replication, 2) An abridged localiser, and 3) An S&M
322 replication.

323

324 **2. Method**

325

326 *2.1 Participants*

327

328 Participants in all three experiments were right-handed native English speakers with
329 normal or corrected-to-normal vision and were recruited from the NYU Abu Dhabi
330 community. Written informed consent was provided by all participants prior to data
331 collection. The Tarkiainen replication experiment included 16 participants (6 females,
332 mean age = 23.8, $SD = 4.5$, median = 22, range = 18-31). The abridged localiser and
333 S&M replication experiment included 24 participants (17 females, mean age = 21.9,
334 $SD = 6.18$, median = 20, range = 19-50)

335

336 *2.2 Materials*

337

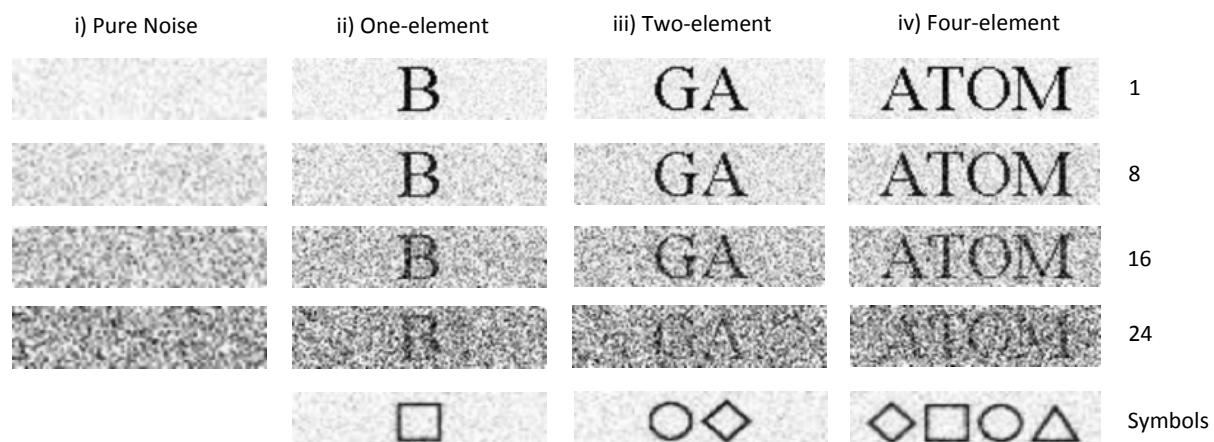
338 *2.2.1 Tarkiainen Replication*

339

340 Stimuli were an English adaptation of materials originally developed by Tarkiainen *et*
341 *al.* (1999). There were 950 items in total, consisting of four different categories of
342 stimulus: i) pure Gaussian noise; ii) single element: one letter (total of 25 letters,
343 without the letter “O”) or one geometrical symbol (a triangle, diamond or square); iii)
344 two-element: two-letter English syllables (25 different syllables) or two-element

345 symbol strings (four possible combinations of a triangle, diamond, square or circle):
346 and iv) four-element: four-letter English words (50 different words) or four-element
347 symbol strings (four combinations of triangle, diamond, square or circle). All four-
348 letter words two-syllable common English nouns (e.g., *SODA*, *PONY*, *ATOM*).
349 Letters, syllables and words were embedded in four different levels of Gaussian
350 noise, increasing from 1, 8, 16, to 24. The Gaussian noise was zero centered, and
351 the variance of the noise was set to 0.0234, 0.0938, 0.375 and 1.5, respectively.
352 Symbols of all lengths were always presented in the lowest noise, and served as
353 controls for the low-noise letters, syllables and words. Items in noise levels 1 and 8
354 were clearly visible, whereas noise level 16 made identification more difficult, and
355 items at noise level 24 were extremely difficult to identify. Figure 1 shows examples
356 of these stimuli.

357



358
359 **Figure. 1.** Full set of stimuli used in the experiment. All are English adaptations of those used by
360 Tarkiainen *et al.* (1999).

361

362 2.2.2 Abridged Localiser

363

364 A subset of the original stimuli was used for the localiser: the four-element (four letter
365 words and length-matched symbols) and one-element items (single letter and single
366 symbol), in the lowest and highest noise levels. This resulted in a subset of 300 trials
367 (50 trials x 6 conditions).

368

369 *2.2.3 S&M Replication*

370

371 The S&M replication experiment was a lexical decision task of 530 items, including
372 265 English words that served as critical stimuli. These words were split into a series
373 of five conditions, with 53 items per condition. 1) *Truly Complex Free* items
374 contained a stem morpheme whose orthography is unchanged when combined with
375 a suffix (e.g., “mile” is orthographically identical when appearing in “mileage”).
376 Furthermore, the meaning of the stem when combined with a suffix was the same as
377 its meaning in isolation – mileage means the number of miles. 2) *Truly Complex*
378 *Bound* items contained a stem morpheme that changes its orthography when
379 combined with a suffix (e.g., “social” loses the final ‘L’ to become “sociable”). But
380 again, the meaning of the stem is maintained in the complex form. 3) *Pseudo-*
381 *Complex* items such as “trolley” contain a stem morpheme whose use in isolation
382 has an unrelated meaning to that of the complex word (e.g., “troll” is semantically
383 unrelated to “trolley”). 4) *Unique Stem* items such as “excursion” do not contain a
384 stem morpheme that appears in isolation (i.e., there is no such word as “excuse”).
385 However, their form and meaning matched that of productively formed words with
386 the same suffix – for example, “excursion” means “the event of going on a trip” in the
387 same way that “explosion” means “the event of blowing up”. 5) *Pseudo-Unique Stem*
388 items such as “winter” also did not contain an attested stem morpheme, but their

389 form and meaning does not match productively formed words with the same suffix –
390 “winter” does not mean someone who does something as an occupation, in the way
391 that “baker” means someone who bakes.

392

393 The collection of items varied in their values on a range of lexical variables known to
394 modulate early letter-specific activity, presented in Table 2 below.

395
396

Condition	Bigram LM		Affix OLF		Affix MLF		Surface LF.		Lemma. LF.		Log TP	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Truly Complex Free <i>Mileage</i>	3.63	0.14	5.51	0.59	4.68	0.39	1.8	0.55	3.08	0.56	1.28	0.7
Truly Complex Bound <i>Sociable</i>	3.53	0.13	5.07	0.4	4.45	0.37	1.75	0.55	2.49	0.72	0.75	0.66
Pseudo-Complex <i>Trolley</i>	3.54	0.12	5.16	0.56	4.1	0.83	1.65	0.63	2.78	0.97	1.29	0.71
Unique Stem <i>Excursion</i>	3.59	0.15	5.16	0.46	4.13	0.65	1.69	0.72	NA	NA	NA	NA
Pseudo-Unique Stem <i>Winter</i>	2.54	0.15	5.31	0.57	4.22	0.84	1.61	0.58	NA	NA	NA	NA

397
398

399 **Table 2.** Lexical statistics of S&M replication experiment. LM = Log of the Mean; OLF = Orthographic
400 Log Frequency; MLF = Morphological Log Frequency; LF. = Log Frequency; TP = Transition
401 Probability.

402

403 2.3 Procedure

404

405 All participants’ head shapes were digitised using a hand-held FastSCAN laser
406 scanner (Polhemus, VT, USA) to allow for co-registration during data preprocessing.

407 Five points on each participant’s head were also digitised: the nasion, just anterior of
408 the left and right auditory canal, and three points on the forehead. Marker coils were
409 later placed at the same five positions to localise each participant’s skull relative to
410 the sensors. These marker measurements were recorded just before and after the
411 experiment in order to track the degree of movement during the recording.

412

413 At the beginning of each of the three experiments, participants completed a practice
414 session outside of the machine to ensure full understanding of the task prior to
415 entering the magnetically shielded room. The Tarkiainen replication was recorded in
416 one session, while the S&M replication and localiser were recorded together in a
417 separate session.

418

419 MEG data were recorded continuously using a 208 channel axial gradiometer
420 system (Kanazawa Institute of Technology, Kanazawa, Japan), with a sampling rate
421 of 1000 Hz and applying an online low-pass filter of 200 Hz.

422

423 Stimuli in all three experiments were displayed using Presentation software. Stimuli
424 were projected onto a screen 85 cm away from the individual's face.

425

426 *2.3.1 Tarkiainen Replication*

427

428 Stimuli were organised into four blocks, so that all of the items within a single block
429 contained the same stimuli type; i.e., all of the single-element items were within a
430 single block, and all of the four-element items were displayed in a separate block.

431 There were 8 possible block orders, and two participants were allocated to each
432 block-order. The order of stimulus presentation was randomised within
433 conditions/blocks but was the same for all subjects. It was ensured that the same
434 word did not appear within 10 trials of presentation to avoid repetition effects.

435

436 Each item was presented in a centrally placed rectangular patch (~5 X 2 inches),
437 and was displayed on the screen for 60 ms with a 2 second inter-stimulus interval.

438 Participants were instructed to focus on the images and to verbally report the item if
439 a question mark appeared. The prevalence of these question-mark trials was 5% (40
440 out of 950 trials) and served to aid concentration. There were no question-mark trials
441 during the pure noise block. As this block was 20% shorter, it was easier to maintain
442 concentration. The whole experiment lasted around 40 minutes.

443

444 *2.3.2 Abridged Localiser*

445

446 Stimuli were presented using the same parameters as the full paradigm. Unlike the
447 full Tarkiainen replication, stimulus order was fully randomised within and between
448 six blocks of presentation so that each block contained a mixture of the six types of
449 stimuli; this was done in order to make the task more engaging. Each block lasted
450 around 60 seconds, and participants were asked not to blink during stimulus
451 presentation to minimise artifacts in the MEG recording. No overt task was employed
452 due to the brevity of the paradigm; participants were simply asked to pay attention to
453 the items as they appeared on-screen. The localiser took around 6 minutes.

454

455 *2.3.3 S&M Replication*

456

457 Each trial began with a fixation cross (“+”) for 400 ms, followed by the critical item for
458 2 seconds. Participants were asked to indicate whether the item was a word or not
459 by pressing one of two buttons with their left hand. No feedback was provided. Order
460 of stimulus presentation was fully randomised, and each participant received a
461 unique randomisation. The experiment was split into 4 blocks and lasted around 20
462 minutes.

463

464 *2.4 Analysis*

465

466 Data from all three experiments underwent the same preprocessing steps. The
467 continuous MEG data were first noise reduced by utilising eight magnetometer
468 reference channels located away from the participant's head, using the Continuously
469 Adjusted Least Squares Method (CALM; Adachi, Shimogawara, Higuchi, Haruta, &
470 Ochiai, 2001), with MEG160 software (Yokohawa Electric Corporation and Eagle
471 Technology Corporation, Tokyo, Japan). The noise-reduced MEG data was imported
472 into MNE-Python (see Gramfort *et al.*, 2014), low pass filtered at 40 Hz, and
473 epoched from 200 ms pre-stimulus onset to 800 ms post-stimulus onset. In order to
474 clean the data, we automatically rejected all trials whose amplitude exceeded a +/-
475 2000 femtotesla threshold; additional artifact rejection was performed through
476 manual inspection of the data, removing noisy trials that were contaminated with
477 movement artifacts or extraneous noise. These clean epochs were averaged across
478 conditions to produce an evoked signal at each MEG sensor.

479

480 To reconstruct the location of MEG sensors relative to the individuals' heads, the
481 neuromagnetic data were co-registered with the FreeSurfer average brain (CorTechs
482 Labs Inc., Lajolla, CA). This involved scaling and orienting the average brain to the
483 participant's head-shape. The digitised scan was imported, the digital fiducial points
484 were aligned to the coil markers' position, and the average brain was expanded or
485 reduced to fit the size of the digital scan. Next, an ico-4 source space was created,
486 consisting of 2562 vertices per hemisphere, each corresponding to a potential
487 electrical sources. At each vertex, activity was computed for the forward solution with

488 the Boundary Element Model (BEM) method, which provides an estimate of each
489 MEG sensor's magnetic field. For each subject, the inverse solution was computed
490 from the forward solution and the grand average activity across all trials.

491

492 Two different orientation parameters were employed in the inverse solution and
493 applied to the data: 1) signed fixed orientation, which defines the direction of the
494 current normal to the cortex by projecting dipoles perpendicular to the cortical
495 surface and estimating activity from the magnitude of the current dipole normal to the
496 cortex; 2) unsigned free orientation, which allows the fitted dipole at each potential
497 electrical source to orient in any direction. Estimates are calculated from the
498 magnitude (absolute length) of the current dipole fitted at the source. The inverse
499 solution was applied to each condition average employing an SNR value of 3, which
500 produced a conversion into noise-normalised Dynamic Statistical Parameter Map
501 (dSPM) units (see Dale *et al.*, 2000).

502

503 **3. Results**

504

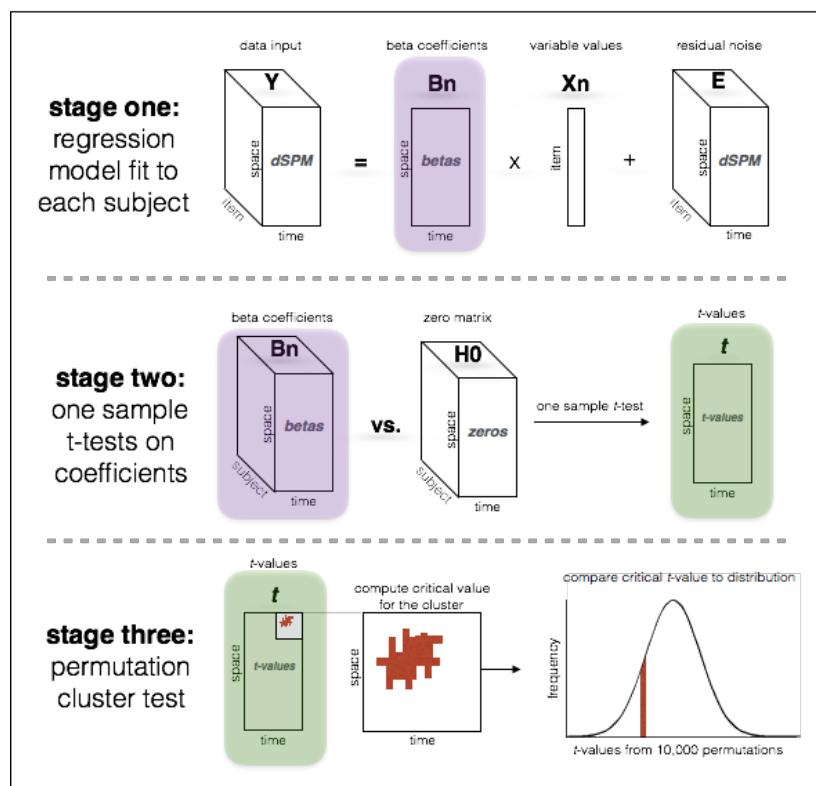
505 In order to localise the Type One and Type Two responses in our data, we ran
506 spatio-temporal permutation cluster tests over the time-windows (80-130 ms, 130-
507 180 ms) and regions (occipital and temporal lobes) reported in Tarkiainen *et al.*'s
508 (1999) results using the *Eelbrain* module in Python
509 (<https://pythonhosted.org/eelbrain/>). This established the location of sources
510 sensitive to the experimental manipulation, and when these sensitivities arose. All
511 cluster-based permutation tests reported here followed the procedures detailed in
512 Maris and Oostenveld (2007).

513

514 In order to maximise statistical power and account for extraneous variation, we
515 coded conditions as binary variables within a regression analysis. Combining
516 regression with spatio-temporal permutation cluster tests involved three stages
517 (summarised in Figure 2). First, the source estimates of each trial within a selected
518 region were used in turn as the dependent measure of an ordinary least squares
519 regression. The data were in the form of N number of 3-dimensional matrices with
520 the shape space (number of vertices in the tested ROI) x time (number of
521 milliseconds within the window interest) x item (number of trials). The design matrix
522 of this regression model included variables of interest (e.g., noise level, string type
523 and stimulus type), a random intercept, and nuisance variables such as number of
524 elapsed trials. This was applied to each subject's data separately, resulting in a beta
525 coefficient for each vertex in the source space, for each millisecond within the
526 selected time window, for each variable of interest, for each subject. Second, a one-
527 sample t-test was performed on the distribution of beta values across subjects for
528 each variable separately, again at each vertex and time-point, in order to test if their
529 value was significantly different from zero. This resulted in a matrix of *t*-values, with a
530 dimension for each vertex and time-point. Third, all *t*-values exceeding a $p < .05$
531 threshold were clustered based on spatio-temporal adjacency. As only *t*-values with
532 the same polarity are clustered together, clusters of different underlying polarity are
533 identified as separate regions. If a cluster consisted of a minimum of 10 vertices and
534 lasted for at least 20 ms, the *t*-values within this cluster were summed, resulting in a
535 cluster-level statistic, for comparison with test statistics of 10,000 random
536 permutations of the data. Each permutation involved shuffling predictor values at
537 random, and re-computing the cluster statistic of the permuted data to form a

538 distribution of cluster-level t -values. If the original critical test statistic fell at or below
 539 the 2.5th percentile, or beyond the 97.5th percentile of this distribution, the cluster was
 540 considered significant at a corrected level of $p < .05$. This p -value was then corrected
 541 for multiple comparisons over time and space following Maris and Oostenveld
 542 (2007). This procedure was followed identically for both fixed and free source
 543 estimates.

544



545
 546 **Figure 2.** Stages of cluster test based on mixed effects regression. n = place holder for each variable
 547 of interest. Depiction illustrates computations for a single variable – the same procedure is carried out
 548 to form clusters for each variable. H_0 = null hypothesis.

549

550 3.1 Type One response: *Distributed source analysis*

551

552 Tarkiainen *et al.* (1999) localised the Type One response bilaterally in the occipital
553 lobe, with peak latency of dipole activity between 95-115 ms post stimulus onset. All
554 stimulus items displayed a correlation of activity with noise, whereby more noise
555 elicited higher amplitude of activity, and there was no fall-off in activation between
556 level 8 noise and level 24 noise. Activity also positively correlated with stimulus
557 length, whereby longer items elicited more activation.

558

559 To capture this response, we included Noise Level and Stimulus Type as variables
560 of interest. A nuisance variable, the number of elapsed trials, was also included to
561 account for variance due to fatigue, but will not be reported here. The regression
562 analysis was run over the 50 ms time-window of 80-130 ms, and vertices were
563 restricted to the occipital and temporal lobes bilaterally (see Figure 3A), merged into
564 one region per hemisphere. This region included the FreeSurfer *aparc* parcellation
565 labels: lateral-occipital, cuneus, lingual, pericalcarine, fusiform, middle-temporal and
566 inferior-temporal (available for download at <http://surfer.nmr.mgh.harvard.edu>).

567

568 Analyses on both free and fixed estimates identified the Type One response (see
569 Table 3). For Noise Level, the signed fixed data yielded three significant clusters in
570 each hemisphere (p 's < .05); for Stimuli Type, two left lateralised clusters and three
571 right lateralised clusters were identified (p 's < .012). The free unsigned estimates
572 identified one significant cluster in each hemisphere for Noise Level, both p < .01;
573 one cluster was identified for Stimulus Type in the left hemisphere (p < .014), while
574 two non-significant clusters were identified in the right hemisphere (p > .2). Note that
575 more clusters are found for signed estimates because sources with opposing polarity
576 are not identified as adjacent neighbours and therefore form separate clusters.

Variable	Data	Cluster size	Hemisphere	Cluster start (ms)	Cluster Stop (ms)	Cluster t-statistic	Corrected p-value
Noise Level (24 > 8)	fixed	77	left	80	130	-2035.61	<.0001
		52	left	80	130	-853.324	0.0185
		20	left	80	130	-455.428	0.0504
		20	left	80	130	386.917	0.0634
		76	right	80	130	-1978.87	0.0001
		57	right	80	130	1583.32	0.0019
		48	right	80	130	-984.855	0.0133
		15	right	80	130	363.749	0.0703
	free	23	right	80	130	315.11	0.0902
		224	left	80	130	3303.5	0.0003
		209	right	80	115	2569.49	0.0048

577

578 **Table 3.** Summary statistics: Results of the spatio-temporal regression analysis of the Type One
 579 Noise Level variable, for both fixed and free source estimates. Cluster size is measured in number of
 580 vertices. The green highlighted row corresponds to the most significant “fixed” cluster.

581

582 3.2 Type Two response: *Distributed source analysis*

583

584 The Type Two response reported by Tarkiainen *et al.* (1999) was characterised by a
 585 fall-off in activity at the highest noise level (24) eliciting greater activity for the lower
 586 noise letter strings than the strings with the highest level of noise. It was also
 587 reported that greater activity was observed for letter strings over length-matched
 588 symbol strings. The peak latency of the fitted dipoles was between 140-170 ms post
 589 stimulus onset, and was localised to the left-hemisphere occipital-temporal junction.

590

591 Type Two sources were identified with two spatio-temporal cluster tests. The first
 592 was run only on responses to letter strings (not symbol strings or blank stimuli), and
 593 only on noise levels 8 and 24. This is because, unlike the Type One response, the
 594 Type Two response of Tarkiainen was not characterised by a linear relationship
 595 between noise-level and amplitude. Rather, there was a drop-off in activation at the
 596 highest noise level (24) when the letters became illegible. To capture this

597 relationship we selected just the level of noise that was found to elicit the largest
598 response (8) and the smallest response (24). The design matrix included Noise
599 Level and Stimulus Type as predictors. We restricted time-points to a 50 ms window
600 between 130-180 ms, and vertices to the same temporal-occipital region utilised in
601 the Type One analysis. A summary of these results is presented in Table 4.

602

Variable	Data	Cluster size	Hemisphere	Cluster start (ms)	Cluster Stop (ms)	Cluster t-statistic	Corrected p-value
Noise-Level (8 > 24)	fixed	58	left	130	180	-967.281	0.002
		21	left	130	180	283.296	0.0384
		16	left	130	175	267.156	0.0449
		19	left	130	170	244.938	0.0591
		12	left	130	175	167.875	0.1485
String-Type	free	no clusters					
		fixed	23	left	130	180	315.442
			15	left	140	180	-157.317
		free	124	left	130	180	1378.6
							0.014

603

604 **Table 4.** Summary statistics: Results of the spatio-temporal regression analysis of the two Type Two
605 variables, for both fixed and free source estimates. Cluster size is measured in number of vertices.
606 The green highlighted rows correspond to the most significant “fixed” clusters for each variable.

607

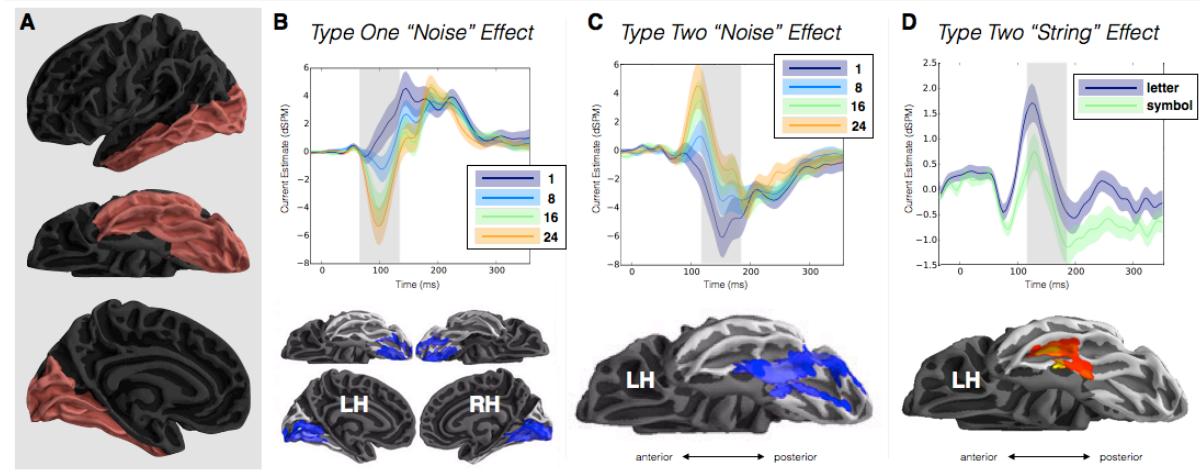
608 For the signed fixed data, three significant clusters were sensitive to the Noise Level
609 variable (8 > 24): p 's < .05. Clusters formed for the Stimulus Type variable were not
610 significant (p 's > .1). When testing unsigned free data, no clusters were formed
611 either for the Noise Level variable or Stimulus Type.

612

613 The second cluster test was run just on responses to letter and symbol strings at the
614 lowest noise level. The design matrix included String Type and Stimulus Type as
615 predictors. Spatial and temporal restrictions were the same as the first Type Two
616 analysis above. For the signed fixed data, one cluster survived corrections for
617 multiple comparisons both for the String Type variable $p = .046$ and the Stimulus

618 Type variable ($p = .037$). When using unsigned free source estimates, one significant
619 cluster was formed for String Type ($p = .014$), and no clusters were formed for
620 Stimulus Type.

621



622
623 **Figure 3.** Summary of distributed source analysis used signed fixed estimates. (A) Location of spatial
624 vertices used in the analysis. (B) Activation averaged over the most significant cluster of the Type
625 One “Noise” response. Brains presented in ventral (above) and medial (below) view. (C) Activation
626 averaged over the most significant cluster of the Type Two “Noise” response. Brain presented in
627 ventral view. (D) Activation averaged over the most significant cluster of the Type Two “Symbol”
628 response. Grey shadowing indicates the time during which the cluster was significant – cluster
629 location illustrates position when averaged over this significant time-window. LH = left hemisphere,
630 RH = right hemisphere.

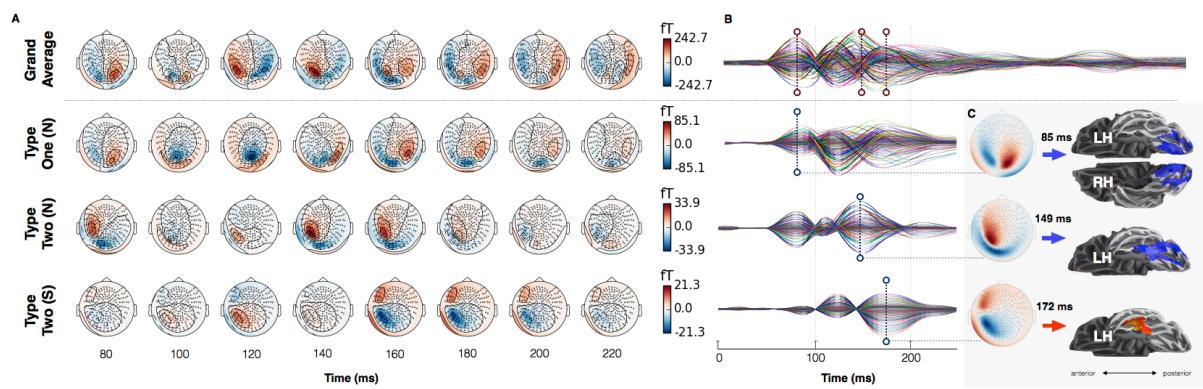
631

632 3.3 Linking responses to M100, M130 and M170

633
634 The next aim was to directly link the responses found in source space to the MEG
635 response components M100, M130 and M170, which are defined based on peaks in
636 sensor data. Using the fixed orientation signed data and results from the full dataset,
637 we projected activation in each cluster back into sensor space to determine which
638 peak (and therefore, component) it corresponded to. Note that this is only possible

for data created with fixed orientation, as polarity of the magnetic field cannot be reconstructed using unsigned current estimates. The location of the cluster was used to mask the grand average source estimate, setting all vertices outside the cluster to zero. The forward solution, which is a matrix that characterises the mapping from source space to sensor space, was then applied to the masked grand average, translating current estimates at each source to femtotesla amplitude at each sensor. Examining the sensor-level responses of each cluster allows us to directly infer to which component it corresponds. Figure 4 illustrates the results of the cluster projection for a single representative subject.

648



649 **Figure 4.** Projecting clusters identified in source-space back into sensor-space: data for one representative subject. (A) Sensor topographies averaged over 20 ms bins. Top panel: Grand average evoked response across all trials. Lower three panels: Evoked responses across sensors that contributed to the significant cluster of each response. N = “Noise Level Response”, S = “String-Type Response”. (B) Butterfly plots for the grand average (top panel) and each projected cluster (lower three panels). Dashed vertical lines indicate peaks of the M100, M130 and M170 respectively. (C) Topography at the sensor peak for each of the three responses, and their corresponding clusters in source space. All brains presented in ventral position. LH = left hemisphere, RH = right hemisphere.

659

660 From inspecting the grand average response, three peaks were identified, with a
661 latency of 85 ms, 149 ms and 172 ms (marked with vertical dashed lines in the top
662 panel of Figure 4B). These timings fall within the range of what has previously been
663 found for the peak latency of the M100, M130 and M170, respectively (Solomyak
664 and Marantz, 2009; 2010; Simon et al., 2011; Lewis et al., 2011; Fruchter and
665 Marantz, 2015). Each of the three projected clusters was found to correspond to one
666 of these evoked peaks, consistent both in terms of the sensor topography and the
667 timing of increased sensor amplitude. These results therefore link the Type One
668 response to the M100 component, the Type Two Noise Level response to the M130
669 component, and the M170 response to the Type Two String Type component.

670

671 *3.4 Functional localiser*

672

673 3.4.1 Designing the abridged paradigm

674

675 We next ran an exploratory analysis to identify stimuli from the Tarkiainen replication
676 that would yield significant effects of the Type Two response. These stimuli would be
677 used in the abridged Tarkiainen adaptation for the purpose of identifying an fROI for
678 use in future studies. We found that using just the four-element (four letter words and
679 length-matched symbols) and one-element items (single letter and single symbol), in
680 the lowest and highest noise levels achieved this goal. This resulted in a subset of
681 300 trials (50 trials x 6 conditions), which would take a participant around 6 minutes
682 to complete.

683

684 The posterior Type Two-Noise response was identified based on responses to the
685 one and four element letter string items (recall that the noise level for all symbol
686 strings was 1). A regression on Noise Level created a beta value for each vertex and
687 each millisecond from stimulus onset to 500 ms after onset. The same threshold-
688 based spatio-temporal cluster tests reported above were performed from 130-180
689 ms on the same temporal-occipital region analysed in the full paradigm. This resulted
690 in one significant cluster of 64 vertices in the lateral-occipital lobe for the entirety of
691 our time-window of interest ($p < .001$). An identical regression was run over the
692 same time-window and spatial extent on the String Type variable, comparing low-
693 noise letter strings to low-noise symbol strings. This identified the anterior Type Two
694 String-Type response: one significant cluster of 19 vertices in the inferior temporal
695 lobe from 145-180 ms ($p < .001$).

696

697 3.4.2 Results of localiser

698

699 To assess the suitability of the reduced stimulus set as a localiser of early lexical
700 processing, we ran the abridged Tarkiainen paradigm during the same recording
701 session as the replication of the S&M visual lexical decision task. The S&M stimuli
702 included nonwords and an equal number of mono-morphemic and bi-morphemic
703 words falling along continua of linguistic variables tested in the original S&M study,
704 including transition probability from stem to suffix, and orthographic and
705 morphological affix frequency. Also included were bigram frequency and lemma
706 frequency, which were found to be significant determiners of M130 and M170 activity
707 in later studies (Lewis *et al.*, 2011; Simon *et al.* 2012). We later tested the

708 modulation of activity in localised regions as a function of these continuous lexical
709 variables.

710

711 Data analysis procedures were identical to those used with the full dataset: Type
712 Two sources were identified by running regression-based spatio-temporal cluster
713 tests on a 130-180 ms time-window in the temporal-occipital region. Tests on fixed
714 signed data revealed clusters located very close to, and with the same current
715 polarity as, those found in the full dataset: an effect of Noise ($p = .01$), and of String
716 Type 200 - 250ms ($p = .009$). Unsigned free data did not identify any clusters where
717 responses to noise level 8 were greater than noise level 24, and no clusters were
718 formed for String Type. Only Type One clusters were identified with free unsigned
719 estimates.

720

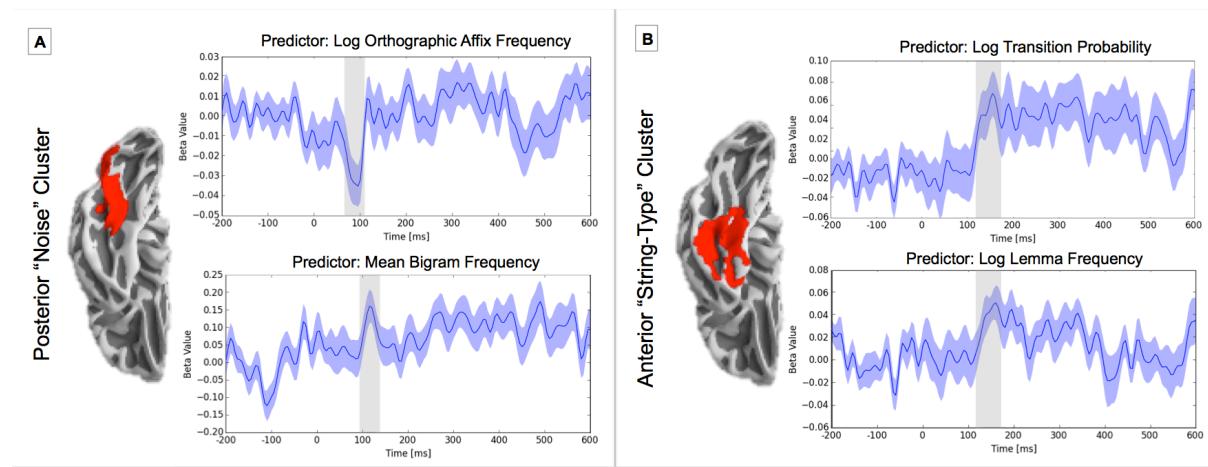
721 Using signed fixed estimates, the next step was to extract regions found to be
722 sensitive to the Tarkiainen localiser experiment, and to test the same regions'
723 sensitivity to lexical variables in the lexical decision task. Note that the localiser
724 experiment contained an orthogonal set of stimuli, which insured that the technique
725 was statistically sound and not "double dipping". To do this, we ran temporal cluster
726 tests on the S&M data from 80-180 ms (chosen to encompass potential M130 and
727 M170 effects) in each Tarkiainen-identified region, assessing the significance of the
728 five continuous variables. The regression included all items with meaningful values
729 on a given variable (*i.e.*, not listed as "NA" in Table 2).

730

731 Activity in the posterior Noise-Level region significantly correlated with orthographic
732 affix frequency ($p = 0.008$) from 80-100 ms; morphological affix frequency ($p = 0.03$)

733 from 80-100 ms and log transform of mean bigram frequency ($p = 0.01$) from 100-
734 130 ms. No clusters surpassed the threshold for either of the other two variables.
735 The anterior String-Type region correlated with lemma frequency from 135 – 170 ms
736 ($p = .001$) and log transition probability from 130 – 170 ms ($p = .008$). No other
737 variables surpassed the cluster-forming threshold.

738



739
740 **Figure 5.** Beta values of the regression analysis between (A) the posterior localiser and log
741 orthographic affix frequency and mean bigram frequency; (B) the anterior localiser and log transition
742 probability and log lemma frequency. Beta coefficients are averaged over the spatial extent of the
743 cluster. Grey shading indicates when the beta values were significantly different from zero.

744
745 Finally we wanted to test whether running the abridged paradigm was more
746 beneficial than simply extracting the cluster location from the full dataset. When
747 running a temporal regression test on the S&M data (as detailed above), but
748 averaging activity over the two Type Two clusters found in the full dataset rather than
749 the localiser, no clusters were formed for the lexical variables of interest. This
750 suggests that applying a localiser at the group level aids the identification of
751 functionally specialised regions.

752

753 **4. Discussion**

754

755 The primary goal of the present study was to compare the relative localisation and
756 timing of the letter-specific Type Two response (Tarkiainen *et al.*, 1999) with other
757 established MEG responses (e.g., M130, M170). We achieved this by conducting a
758 replication study of Tarkiainen *et al.*'s original MEG experiment, combining
759 distributed source analysis, spatio-temporal cluster tests and the projection of
760 responses back into sensor space. The second goal was to assess the functionality
761 of these responses. For this purpose, we designed an abridged version of the
762 Tarkiainen experiment that may be used to localise letter-specific responses in future
763 studies. We assessed sensitivity of these fROIs by testing the influence of
764 continuous variables from the S&M replication on neural activation of the fROIs; in
765 doing so, we further linked the fROIs' sensitivity to other reported components in the
766 literature. Our final goal was to spell out differences between cortically constrained
767 source estimates and unconstrained methods of MEG source reconstruction and to
768 characterise the Type Two responses in terms of current polarity.

769

770 **4.1 Type One**

771

772 The Type One response was localised bi-laterally in the occipital lobe between 80-
773 130 ms post stimulus onset, beginning in V1 / BA 17 and extending up to V3 / BA 19.
774 This response encapsulates the primary visual evoked response at around 100 ms,
775 eliciting the field pattern over sensors typically associated with the M100. Consistent
776 with Tarkiainen *et al.*, signal amplitude of the Type One response increased
777 systematically as a function of noise (greater Gaussian noise correlated with greater

778 amplitude of activity), and increasing stimulus length, suggesting that the neural
779 populations underlying this response are sensitive to the visual complexity of the
780 given stimulus.

781

782 *4.2 Type Two*

783

784 The sensitivity to both String-Type (Symbol Strings vs. Letter Strings) and Noise
785 level ($8 > 24$) were originally reported by Tarkiainen *et al.* to originate at the same
786 place and time. However, we found two main effects for the Type Two response,
787 each localising to different patches of cortex at different time-points, realising
788 opposing polarity of the current with fixed orientation (see Figure 3). The main effect
789 of Noise ($8 > 24$) (henceforth the “Type Two-Noise” response) was found at the
790 junction between the occipital cortex and temporal lobe, as was the original location
791 reported by Tarkiainen *et al.* The main effect of String-Type (henceforth “Type Two-
792 Letter” response), however, was localised at the anterior-most portion of the fusiform
793 gyrus, more anterior than the Type Two-Noise response. The posterior Noise
794 response had negative polarity, whereas the anterior Letter response had positive
795 polarity.

796

797 When projecting the two Type Two clusters back into sensor space (Figure 4), the
798 observed switch in cluster polarity was also present at the sensor level (i.e., the
799 magnetic field over sensors shifted direction). Furthermore, each response was
800 found to have a distinct spatial and temporal profile that corresponded to peak
801 responses of the grand average, and was distinct from the other projected clusters.
802 The Type One response corresponded to the M100 peak; the Type Two-Noise

803 response corresponded to the M130 peak, and the Type Two-Letter response
804 corresponded to the M170 peak. Interestingly then, our results do not support the
805 previous links that have been made between the M130 and the earlier Type One
806 response (Lewis et al., 2011). Instead, they suggest that the Type One response is
807 driven by much lower-level properties of the stimulus. Further, they suggest that the
808 Type Two response as identified by Tarkiainen et al. should be analysed as two
809 functionally distinct components.

810

811 4.2.1 Functional ROI

812

813 Functionally localised ROIs have a number of advantages over anatomically defined
814 parcellations or ROIs based on peaks in grand-average sensor or source data. In
815 particular, fROIs are not constrained by borders between regions, and do not require
816 analysing larger regions than necessary – a complication when correcting for
817 multiple comparisons in cluster-based analyses. Provided that the localiser and
818 critical experiment tap into the same neural sensitivities, a fROI should yield the least
819 variance between the location of the effect and the region being tested, thus
820 providing the greatest level of statistical power. Functional ROIs also remove the
821 experimenter's role in selecting a region based on visual inspection of the data, and
822 also do not employ arbitrary parameters to localise regions. Further, fROIs are
823 particularly pertinent for cortically constrained source estimates; as the localiser is
824 based on a threshold-based cluster analysis, it identifies a region of homogenous
825 polarity, and thus fully overcomes the issue of activity cancellation - the principal
826 motivation against using signed source estimates with MEG data.

827

828 In testing just a subset of the experimental materials from Tarkiainen *et al.* (1999),
829 we found a very robust effect for both Type Two-Noise and –Letter responses when
830 using just the four-element block (four letter words, noise levels 1 and 24 and length-
831 matched symbols), and one-element block (one letter, noise levels 1 and 24 and one
832 symbol) which takes participants only ~6 minutes to complete. The location and
833 timing of each Type Two response corresponded with the full dataset, supporting
834 that the same neural sensitivities can be localised even when using one quarter of
835 the stimulus materials. This set is therefore an ideal candidate for a localiser
836 paradigm.

837

838 Tests of the posterior Type Two-Noise response revealed significant modulation of
839 activity by orthographic and morphological affix frequency between 80-100ms, and
840 log mean bigram frequency between 100-130 ms. Numerous studies have
841 associated both of these variables with the M130 component (Lewis *et al.*, 2011;
842 Simon *et al.*, 2012; Solomyak and Marantz, 2009; 2010), suggesting that the
843 posterior localiser identified spatio-temporal regions associated with lower-level
844 orthographic processing.

845

846 Analysis of the more anterior Type Two-Letter response showed that log lemma
847 frequency and log transition probability significantly modulated activity from 130-170
848 ms (morphological affix frequency was not a significant determiner). Both of these
849 variables have previously been associated with the M170 response (Lewis *et al.*,
850 2011; Simon *et al.*, 2012), suggesting that regions supporting the anterior Type Two-
851 Letter response are shared with the M170, which is responsible for more abstract

852 lexical processing. Regression coefficients of the continuous variables are presented
853 in Figure 5.

854

855 This functional disassociation between posterior and anterior responses corresponds
856 to recent results employing cortically constrained MEG data from a lexical decision
857 task (Chen *et al.*, 2015). The authors found that activity in posterior portions of the
858 fusiform correlated with bigram frequency ~100 ms post onset, while anterior regions
859 correlated with word surface frequency ~160 ms, thus strongly corroborating the
860 present results. Lewis *et al.* (2011) also reported similar results using unconstrained
861 signed estimates of MEG data, finding that only posterior portions of the anterior
862 M170 ROI displayed surface frequency effects. This was interpreted as indexing a
863 “high-ngram” effect, and the activation of more concrete representations.

864

865 Our findings therefore support that the abridged Tarkiainen paradigm can be used to
866 successfully localise early posterior orthographic processing, as well as later anterior
867 sub-lexical processing. More specifically, finding activity in these regions to
868 significantly correlate with variables shown to modulate the M130 and M170 is strong
869 functional evidence that the Type Two-Noise and -Letter responses are analogous to
870 those MEG components, in line with the clusters’ peaks in the sensor data reported
871 above.

872

873 The stimuli used in the localiser of the present study are available to download,
874 either on github (<https://github.com/LauraGwilliams/TarkiainenLocaliser.git>) or by
875 contacting the first author.

876

877 4.2.2 Graded lexical sensitivity of Type Two

878

879 Each linguistic variable included in the localiser analysis is associated with a certain
880 level of processing complexity and has been linked to a specific response
881 component in previous literature. Orthographic affix frequency and bigram frequency
882 are linked to surface orthographic properties entailing relatively low-level processing
883 and access to concrete representations. Both variables modulated activity in the
884 Type Two-Noise region, which is in line with the LCD model prediction that posterior
885 cortical regions subserve lower level processes. The posterior sensitivities of this
886 region are consistent with the “letter-form” area (Thesen *et al.*, 2012) and the M130
887 component (Solomyak & Marantz, 2010), suggesting that the responses and the
888 localiser share underlying neural mechanisms.

889

890 Variables such as lemma frequency and transition probability are linked to more
891 abstract processing involving the connection between input and stored word forms.
892 The LCD model accordingly positions neurons tuned to higher-level processes along
893 anterior portions of the fusiform gyrus, in agreement with our results. The sensitivity
894 of the Type Two-Letter response to these variables, in addition to its anterior
895 location, are consistent with equating the Type Two-Letter response with the M170
896 as identified by S&M, with the VWFA (Cohen *et al.*, 2000) and with the “word-form”
897 areas (Thesen *et al.*, 2012) that appear to involve higher-level processing.

898

899 The posterior-to-anterior progression of abstract processing is thus supported by our
900 results, whereby lower-level sensitivities such as that to letter frequency arise
901 posteriorly, and higher order variables such as lemma frequency appear to be

902 encoded more anteriorly. Our results are in full corroboration with the LCD model as
903 tested by Vinckier *et al.* (2007), and offer striking similarities in localisation and
904 functionality between the two Type Two responses and the “letter-form” and “word-
905 form” regions identified by Thesen *et al.* (2012). Together our findings support the
906 hypothesis that words are first processed through the orthographic properties of
907 letter strings, followed by the processing of word forms and sub-lexical structure.

908

909 *4.3 Comparing source estimate constraints*

910

911 The final issue to discuss is the implication of applying different methods of source
912 reconstruction. We directly assessed whether, in comparison to signed source
913 estimates, conventional unsigned source estimates of MEG data would reduce the
914 ability to discriminate between spatio-temporally neighbouring responses or “lose”
915 evoked peaks by taking absolute strength of activation.

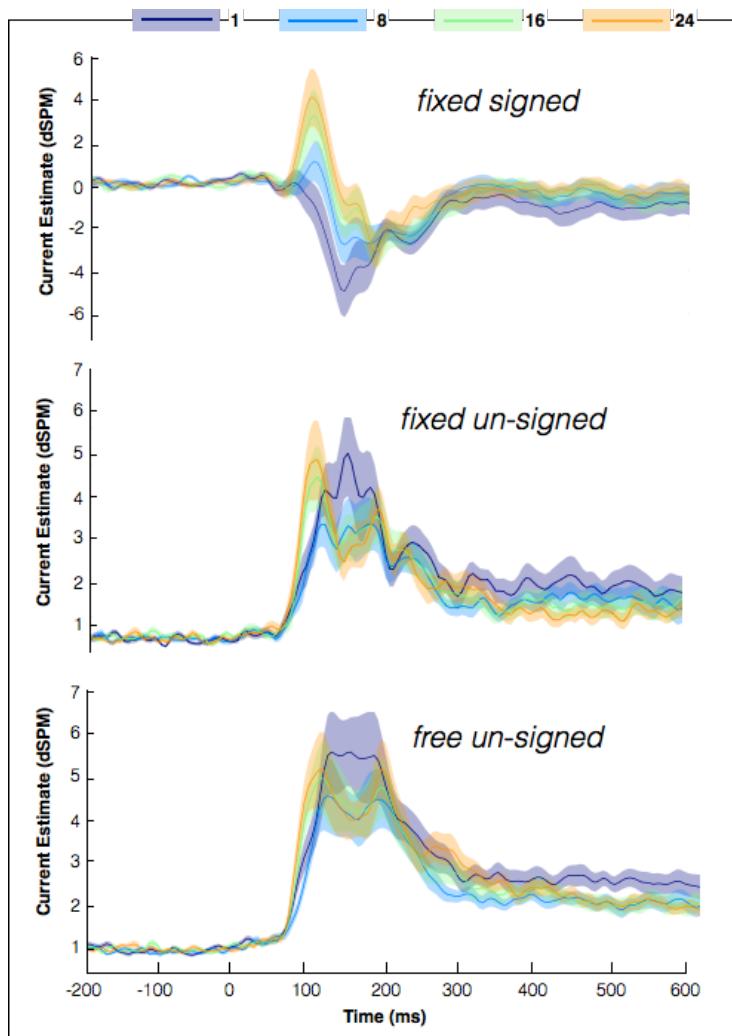
916

917 The present study consistently found unsigned free orientation to be less sensitive to
918 experimental manipulations than signed fixed estimates. For the full Tarkiainen
919 dataset, analyses on free unsigned data did not form any clusters above the $p < .05$
920 threshold for the Type Two-Noise response, and for the abridged paradigm, free
921 unsigned data failed to uncover both the Type Two-Noise response and the Type
922 Two-Letter response. Analyses using fixed signed data straightforwardly identified all
923 of these responses, even in the absence of structural MRIs for participants – the
924 presence of which should only serve to improve accuracy further.

925

926 Figure 6 presents the source estimates of the Type Two-Noise response when
927 reconstructing the source of activity using three methods: 1) signed data fixed
928 normal to the cortical surface, whereby negative activity corresponds to current
929 flowing into the cortical mass, and positive to current flowing out of the cortical mass;
930 2) unsigned fixed data, taking the norm of the dipole fitted perpendicular to each
931 vertex; 3) unsigned free data, which does not retain the direction of the source and
932 allows the dipole to freely orient in any direction. Recall that free unsigned data did
933 not elicit a reliable Type Two-Noise response in the full dataset. All sources are from
934 the same spatial cluster identified with the signed fixed estimates, shown in Figure
935 3C.

936



937

938 **Figure. 6.** Timecourse of activation averaged over the Type Two-Noise cluster found in the full
939 dataset (shown in Figure 3C), for three methods of source estimation. Values 1, 8, 16 and 24
940 correspond to noise level of the stimulus.

941

942 As can be seen in Figure 6, the separation between peaks of activity is much less
943 clear for the unsigned estimates, as compared to the signed methodology. Because
944 there is temporal overlap between one response and the next in this region, and
945 because the two responses have opposing polarity, averaging these values distorts
946 the actual relationship between activation and the given experimental conditions.
947 This erroneous averaging across polarity is therefore likely to explain why Type Two-
948 Noise responses were not identified using unsigned estimates.

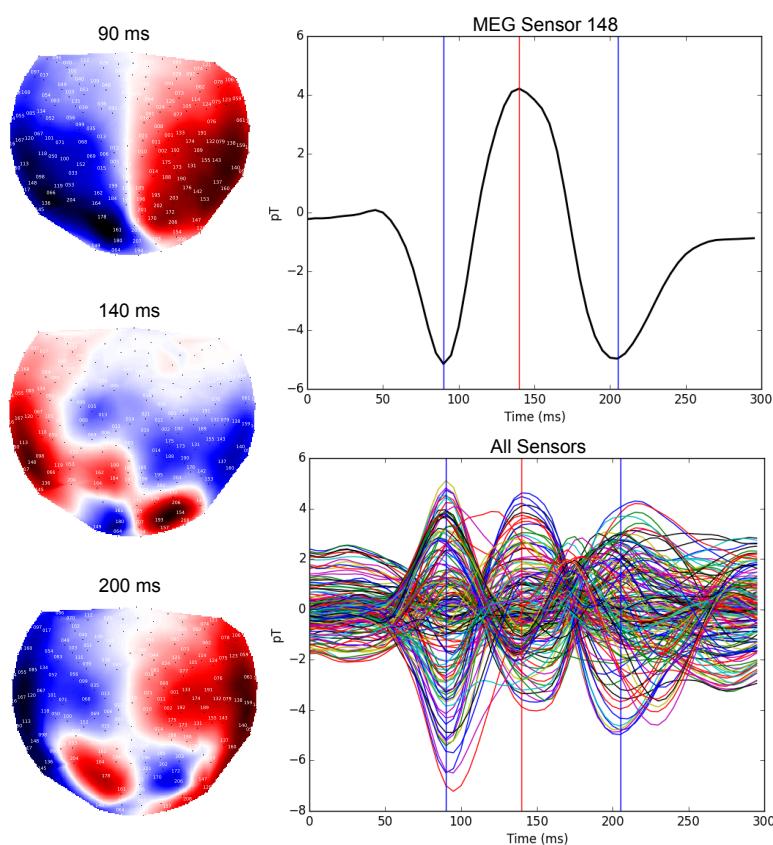
949

950 In the study conducted by Fruchter and Marantz (2015, Appendix B), the authors
951 analysed their data employing the two methods of source reconstruction discussed
952 above. They found that when comparing source estimation methods, unsigned data
953 greatly reduced activation peaks. This dampening of evoked responses can also be
954 observed when comparing the unsigned MEG and signed ECoG measurements in
955 Thesen *et al.*'s (2012) study. Figures 4h and 4i from their paper represent the local
956 field potential recorded from the cortical surface directly, and the MEG source
957 reconstruction from the same region. There we can see very clear polarity shifts in
958 the ECoG data, allowing for discrimination between response components in this
959 dataset; however, in the unsigned free orientation MEG data, it is unclear whether
960 the “bumps” of activity reflect separate responses. One level of discrimination is
961 therefore lost when removing this polarity information, as compared to the method
962 that records the electric potentials generating the MEG signal.

963

964 Such shifts in polarity are not negligible and are clearly observed at the sensor level:
965 A single channel will shift between a positive and negative magnetic field depending
966 upon the orientation of the dipole at the source. Figure 7 below illustrates the
967 strength of the magnetic field at a left lateralised sensor averaged across all subjects
968 and items. In line with the timing we observed for the source analysis of the Type
969 Two responses, a field reversal is also apparent, indicating that the current dipole(s)
970 in that region flip direction at different time-points.

971



972
973 **Figure 7.** Polarity of magnetic field at a single sensor (above) and across all sensors (below)
974 averaged over subjects and items. Topographic plots show the polarity of the magnetic field at each
975 of the three peaks in the sensor data.

976

977 This field reversal was also observable when projecting the activity of clusters into
978 sensor space for a single subject. The direction of the dipole for the Type Two

979 responses was clearly oriented differently (Figure 4C), corresponding to the polarity
980 differences also observed at the source level. Even without structural MRIs for
981 individual subjects, the polarity of the clusters found with fixed orientation lined up
982 with the sensor data, such that the field pattern over sensors corresponded to
983 polarity of sources. This is consistent with the polarity difference between the M130
984 and M170 when using a free signed methodology (assigning polarity based on
985 whether the dipole is oriented up or down with respect to the head) observed in the
986 original Solomyak and Marantz (2010) analysis and in follow-up studies (Lewis *et al.*,
987 2011; Simon *et al.*, 2011).

988

989 This is only a limited investigation into the question of current reconstruction. Further
990 study should aim to ascertain whether all negative response components reflect
991 functionally different computations from all positive components, or if the
992 directionality of the current is an arbitrary dimension of discrimination (similar to
993 polarity in EEG). The present study illustrates that polarity of the reconstructed
994 sources is an important element of MEG data, and can be used to disassociate
995 functionally discrete (in this case, Type Two) neighbouring responses. Furthermore,
996 using signed estimates appears to be more experimentally robust when analysing a
997 reduced dataset, as well as a more sensitive method, especially when current
998 dipoles rapidly switch direction. Finding evidence, both at the sensor and source
999 level that the neural generators underlying the identified responses are alternating in
1000 polarity highlights the importance of current dipole directionality for MEG data.

1001

1002 **5. Summary**

1003

1004 Using distributed source analysis of MEG data, we localised the Type Two response
1005 of Tarkiainen *et al.*'s (1999) study for comparison with other established response
1006 components and regions of lexical-specific activity. The Type Two response
1007 localised to two different regions with different preferences: 1) preference for visible
1008 over obscured letter strings in the lateral-occipital lobe with negative activity with
1009 respect to the cortical surface; 2) preference for letter over symbol strings in the
1010 anterior fusiform gyrus with positive activity. When testing the lexical sensitivities of
1011 these regions as part of an abridged paradigm, functional responses were shared
1012 between the posterior Type Two-Noise response, the M130 and “letter-form” area,
1013 and between the Type Two-Letter response, the M170 and “word-form” area. These
1014 results suggest that each case evokes the same underlying processes, and crucially
1015 that a subset of the stimuli materials is sufficient to localise these response
1016 components with notable accuracy.

1017

1018 In order to address the issue of source reconstruction with MEG data, we applied
1019 two methods to the current dataset and compared the results to the findings of
1020 Tarkiainen *et al.* (1999). In the presence of rapidly alternating polarity, utilising
1021 cortically constrained estimates was the most sensitive approach, ensuring the
1022 preservation of evoked response components. By contrast, cortically unconstrained
1023 unsigned estimates were susceptible to the loss of discrimination between activation
1024 peaks. In this regard, our findings directly indicate that, arbitrary or not, retaining the

1025 sign of MEG data can allow for greater sensitivity to experimental manipulations and
1026 an additional level of discrimination.

1027

1028 Bringing our results together, we are able to characterise two localisers of letter-
1029 sensitive responses for future studies in time, space and current directionality with
1030 respect to the cortical surface. We propose that the posterior Type Two sensitivity to
1031 visible letter strings can be used to localise orthographic processing, and the anterior
1032 Type Two sensitivity to letter strings over symbol strings can localise higher-level
1033 processing of sub-lexical structure, such as morphological composition. Our results
1034 promote the use of cortically constrained signed estimates of MEG data, in unison
1035 with functional ROIs when investigating letter-specific neurophysiological responses
1036 in future studies.

1037

1038 **Acknowledgements:**

1039 This work was supported by the NYU Abu Dhabi Institute under Grant G1001 (AM)
1040 and NSF Grant DGE-1342536 (GL).

1041

References

- 1042
1043
1044 Booth, M. C., & Rolls, E. T. (1998). View-invariant representations of familiar objects
1045 by neurons in the inferior temporal visual cortex. *Cerebral Cortex*, 8(6), 510-523.
1046
1047 Chang, W. T., Ahlfors, S. P., & Lin, F. H. (2013). Sparse current source
1048 estimation for MEG using loose orientation constraints. *Human brain mapping*,
1049 34(9), 2190-2201.
1050
1051 Chen, Y., Davis, M. H., Pulvermüller, F., & Hauk, O. (2015). Early Visual Word
1052 Processing Is Flexible: Evidence from Spatiotemporal Brain Dynamics. *Journal of*
1053 *cognitive neuroscience*.
1054
1055 Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff,
1056 M. A., & Michel, F. (2000). The visual word form area. *Brain*, 123(2), 291-307.
1057
1058 Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002).
1059 Language-specific tuning of visual cortex? Functional properties of the Visual Word
1060 Form Area. *Brain*, 125(5), 1054-1069.
1061
1062 Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., &
1063 Halgren, E. (2000). Dynamic statistical parametric mapping: combining fMRI and
1064 MEG for high-resolution imaging of cortical activity. *Neuron*, 26(1), 55-67.
1065
1066 da Silva, F. L. (2010). *Electrophysiological basis of MEG signals* (pp. 1-2). London,
1067 UK: Oxford Univ. Press.
1068
1069 Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B., &
1070 Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious
1071 repetition priming. *Nature neuroscience*, 4(7), 752-758.
1072
1073 Dehaene, S., Le Clec'H, G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The
1074 visual word form area: a prelexical representation of visual words in the fusiform
1075 gyrus. *Neuroreport*, 13(3), 321-325.
1076
1077 Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for
1078 written words: a proposal. *Trends in cognitive sciences*, 9(7), 335-341.
1079
1080 Fruchter, J., Stockall, L., & Marantz, A. (2013). MEG masked priming evidence for
1081 form-based decomposition of irregular verbs. *Frontiers in human neuroscience*, 7,
1082 798.
1083
1084 Fruchter, J., & Marantz, A. (2015). Decomposition, lookup, and recombination: MEG
1085 evidence for the Full Decomposition model of complex visual word recognition. *Brain*
1086 and language, 143, 81-96.
1087
1088 Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C.,
1089 Parkkonen, L. and Hämäläinen, M. S. (2014). MNE software for processing MEG
1090 and EEG data. *Neuroimage*, 86, 446-460.

- 1091
1092 Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993).
1093 Magnetoencephalography—theory, instrumentation, and applications to noninvasive
1094 studies of the working human brain. *Reviews of modern Physics*, 65(2), 413.
1095
1096 Hauk, O. (2004). Keep it simple: a case for using classical minimum norm estimation
1097 in the analysis of EEG and MEG data. *Neuroimage*, 21(4), 1612-1621.
1098
1099 Helenius, P., Tarkiainen, A., Cornelissen, P., Hansen, P. C., & Salmelin, R. (1999).
1100 Dissociation of normal feature analysis and deficient processing of letter-strings in
1101 dyslexic adults. *Cerebral Cortex*, 9(5), 476-483.
1102
1103 Lewis, G., Solomyak, O., & Marantz, A. (2011). The neural basis of obligatory
1104 decomposition of suffixed words. *Brain and language*, 118(3), 118-127.
1105
1106 Lin, F. H., Belliveau, J. W., Dale, A. M., & Hämäläinen, M. S. (2006). Distributed
1107 current estimates using cortical orientation constraints. *Human brain mapping*, 27(1),
1108 1-13.
1109
1110 Liu, A. K., Belliveau, J. W., & Dale, A. M. (1998). Spatiotemporal imaging of human
1111 brain activity using functional MRI constrained magnetoencephalography data:
1112 Monte Carlo simulations. *Proceedings of the National Academy of Sciences*, 95(15),
1113 8945-8950.
1114
1115 Liu, A. K., Dale, A. M., & Belliveau, J. W. (2002). Monte Carlo simulation studies of
1116 EEG and MEG localization accuracy. *Human brain mapping*, 16(1), 47-62.
1117
1118 Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and
1119 MEG-data. *Journal of neuroscience methods*, 164(1), 177-190.
1120
1121 Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social cognitive and
1122 affective neuroscience*, 2(1), 67-70.
1123
1124 Rolls, E. T. (2000). Functions of the primate temporal lobe cortical visual areas in
1125 invariant visual object and face recognition. *Neuron*, 27(2), 205-218.
1126
1127 Simon, D. A., Lewis, G., & Marantz, A. (2012). Disambiguating form and lexical
1128 frequency effects in MEG responses using homonyms. *Language and Cognitive
1129 Processes*, 27(2), 275-287.
1130
1131 Solomyak, O., & Marantz, A. (2009). Lexical access in early stages of visual word
1132 processing: A single-trial correlational MEG study of heteronym recognition. *Brain
1133 and language*, 108(3), 191-196.
1134
1135 Solomyak, O., & Marantz, A. (2010). Evidence for early morphological decomposition
1136 in visual word recognition. *Journal of Cognitive Neuroscience*, 22(9), 2042-2057.
1137
1138 Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R.
1139 (1999). Dynamics of letter string perception in the human occipitotemporal
1140 cortex. *Brain*, 122(11), 2119-2132.

- 1141
1142 Thesen, T., McDonald, C. R., Carlson, C., Doyle, W., Cash, S., Sherfey, J.,
1143 Felsővalyi, O., Girard, H., Barr, W., Devinsky, O., Kuzniecky, R. and Halgren, E.
1144 (2012). Sequential then interactive processing of letters and words in the left fusiform
1145 gyrus. *Nature communications*, 3, 1284.
1146
1147 Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007).
1148 Hierarchical coding of letter strings in the ventral stream: dissecting the inner
1149 organization of the visual word-form system. *Neuron*, 55(1), 143-156.
1150
1151 Warrington, E. K., & Shallice, T. I. M. (1980). Word-form dyslexia. *Brain*, 103(1), 99-
1152 112.
1153
1154 Zweig, E., & Pylkkänen, L. (2009). A visual M170 effect of morphological
1155 complexity. *Language and Cognitive Processes*, 24(3), 412-439.
1156 112.