On the neural basis of word reading:

a meta-analysis of fMRI evidence using Activation Likelihood Estimation.

Murphy, K.A., Jogia, J. & Talcott, J.B.

Aston Brain Centre, School of Life and Health Sciences, Aston University, Birmingham, UK.

Corresponding author

Kelly Murphy, email: murphyk5@aston.ac.uk

Running Title: ALE analysis of single word reading in fMRI

1. *Introduction*

Reading is a complex, yet widespread phenomenon, which relies on a vast and strategically coordinated network of brain regions. Expert readers can fluently decipher once arbitrary contours on a page into the meaning and sound constituents of thousands of words at rapid speeds, as if it was an innate and natural ability. It is the extent and complexity of the linguistic and neural development needed to become an expert reader and why it fails in otherwise healthy individuals that are revealing of the nature of the human brain.

Recent progressions in neuroimaging techniques have allowed researchers to make more subtle manipulations and valuable observations and of the reading network. Studies using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have reasonably established the spatial extent of the expert reading network. However, due to factors such as technical constraints, researcher interests and linguistic complexities, the range of tasks used alongside these modalities vary widely. Because of a wide spectrum of task demands, the cost to the collective outcomes of research is that there is only partial agreement on the acute roles of particular cortical regions. For example, single regions can be associated with multiple, reading-relevant processes (e.g. left AG is associated with semantic retrieval, attention and memory in language paradigms (see Seghier, 2013 for a review)). Likewise, one process can be associated with multiple, spatially distinct regions (e.g. word form recognition is associated with left occipitotemporal region (Dehaene & Cohen, 2011) and fusiform gyrus (McCandliss et al, 2003). While it may truly be the case that one region has multiple functions, or one cognitive process is spread over multiple regions, it is also difficult to quantify the extent that neural activity is affected by the niche requirements of tasks used to probe the reading response.

The general consensus from meta analytic data is that the reading network is distributed across the left occipitotemporal area, fusiform gyrus and left cerebellum for orthographic processing, phonological processing is distributed across the left superior temporal cortex, inferior frontal gyrus

(IFG); and left parietal cortex; and the left IFG is heavily implicated in semantics (Turkletaub et al, 2002; Jobard et al, 2003; Vigneau et al, 2006, 2011; Cattinelli et al, 2013; McNorgan et al, 2015). According to Reichle (2015), an overwhelming majority of the reading literature agrees that these regions combine to form two routes for word and novel/nonword reading as was originally proposed by Coltheart et al (2001). However, meta analytic work by Cattinelli et al (2013) proposes a third attentional route that overlaps the word and nonword routes when mediated by lexical difficulty. More recent findings suggest that word and nonword reading across orthographies and imaging modalities manifest differently depending upon the demands of the cognitive paradigm (McNorgan et al, 2015); but the extent of task demands remains uncertain if extraneous variables such as orthographies and imaging modalities are not also controlled. Without establishing the extent of this variance across tasks, there is heightened potential for unwitting bias and misinterpretations of the neural signature for real word reading in future experiments.

To investigate the extent of variance introduced by differing methodological practices we focused on the two most prevalent paradigms used to assess the typical reading network in neuroimaging: single word reading (SWR) and the lexical decision task (LDT; Carrieras et al, 2007). SWR paradigms can be conducted silently or aloud, to mimic typical reading processes. However, due to extra scanning considerations such as breaking the scanning sequence for articulation, reading aloud paradigms have the potential to incur excess noise in the data (Poeppel et al, 2001). On the flip side, it is difficult to check a participant's engagement in the task, or distinguish accurate from inaccurate responses with passive silent reading. Successful completion of LDT paradigms involve differentiating words from nonwords normally via a finger-tap method or articulation, so despite not being as naturalistic of typical reading processes LDT is often favoured because participant engagement, head movement; and accuracy can all be controlled (Perea et al, 2002; Carrieras et al, 2007). It is for these reasons that LDT is such a widely adopted paradigm; but only SWR paradigms have been meta analysed in isolation due to their naturalistic relevance (Turkletaub et al, 2002). It would be sensible to infer that LDT studies therefore warrant being quantitatively summated in their own right. Interestingly, it has been implied that because of the added decisional mechanism used to generate the measured behavioural response in LDT relative to SWR, the two tasks are in fact measuring overlapping behaviours (Carrieras et al, 2007).   If this is the case, it would not be unreasonable to suggest that the neural response from fMRI reading tasks are difficult to interpret when aggregated for review purposes, for as long as the extent of the variance between tasks remains unknown.

As is the nature of reviews, experimental techniques and the body of published work will continue to advance, so regular calls for quantitative analyses to be conducted in dynamic fields such as the neural basis of reading persist. In this case, the most recent meta analysis of reading to our knowledge does not cover papers published 2011 (McNorgan et al, 2015), so opportunity for inclusion of studies published post-2011 in a quantitative review exists. To compensate for the inevitable fate of reviews eventually becoming outdated, meta analytic findings should aim to enhance their relevance by predicting the outcomes of future experimentation. Therefore, due to the increase in popularity of fMRI for neuroimaging research relative to PET, we have decided focus on an all-fMRI data-set.

As with all meta analyses there is a trade-off between a refined research objective and the scope of data up for inclusion large enough to gain statistical power. Herein so far, it is more commonplace to prioritise power and summate data across languages, task type and imaging methods. An unfortunate by-product of loose inclusion criteria is the accumulation of variance within the data, which diminishes the clarity of the results (Turkletaub et al, 2002; Jobard et al, 2003; Vigneau et al, 2006 & 2011; Cattinelli et al, 2013). Although as previously stated, the ongoing cycle of publishing means that more studies continue to be available for review. This means that there is now greater opportunity to more tightly restrict inclusion criteria without much cost to statistical power, since a greater volume of studies within a body of literature increases the likelihood of shared traits as well as task differences. Thus, here we focus solely on studies conducted in the English language so task related differences can be assessed without the additional uncertainties introduced through nuances between alphabetic orthographies.

To quantitatively summarise and compare the neural activation for each method, we use the activation likelihood estimation (ALE) method. Unlike other cluster-based methods such as kernel density analysis (KDA), the ALE algorithm employs a kernel permutation technique to test the null hypothesis and assess effect size, whereby fMRI foci are used as novel data points plotted on a template brain as Gaussian kernels, weighted by the number of participants that contributed towards them. Since foci are weighted and permutations occur at the voxel-level, a high foci resolution can be maintained with ALE and both whole brain and region of interest (ROI) studies remain for inclusion. In turn, the availability of data is bolstered and statistical power of the findings is relatively heightened (Eickhoff et al, 2012). However, because foci need to be overlaid on a template brain, the ability of ALE to handle negative blood oxygen level dependent (BOLD) values is poor. It is a particular criticism of functional neuroimaging that negative BOLD values are poorly understood in the context of human cognition and so are rarely reported (see Pasley, Inglis & Freeman, 2007). Thus overly stringent algorithms which account for negative BOLD values such as signed difference mapping (SDM) are deemed unnecessary for addressing our question at hand. For these reasons, the ALE method is particularly advantageous for the current purpose.

A second and equally serious motive for conducting a meta analysis of reading lies in the fact that the ALE algorithm has been used for several meta analyses published before 2016 (e.g. Turkletaub et al, 2002 and McNorgan et al, 2015). Prior to 2016, the GingerALE software contained errors which hindered the ability of the false discovery rate (FDR) to adequately correct results for multiple comparisons (Eickhoff et al, 2016). In light of using newly-corrected FDR algorithms, we use the ALE algorithm to investigate the spatial extent of task-specific activation differences between SWR and LDT fMRI studies conducted in the English language published from 1990 to August-2014.

As SWR is a deliberate and typical method of normal reading, we hypothesise that both the silent and aloud tasks will elicit activity across the entirety of the word-specific and phonological routes in the network. Alternatively, as LDT judgements include a decisional element emphasising visual word forms, a distinguishable reduction in semantic and phonological activations is expected relative to SWR, with activations related to decisional and motor planning mechanisms becoming apparent (Gold & Shadlen, 2007; Vigneau et al, 2011). Since SWR and LDT tasks are the two most prolifically used paradigms in clinical and cognitive research (Carrieras et al, 2007; McNorgan et al, 2015), the potential impact of any task related demands cannot be underestimated.

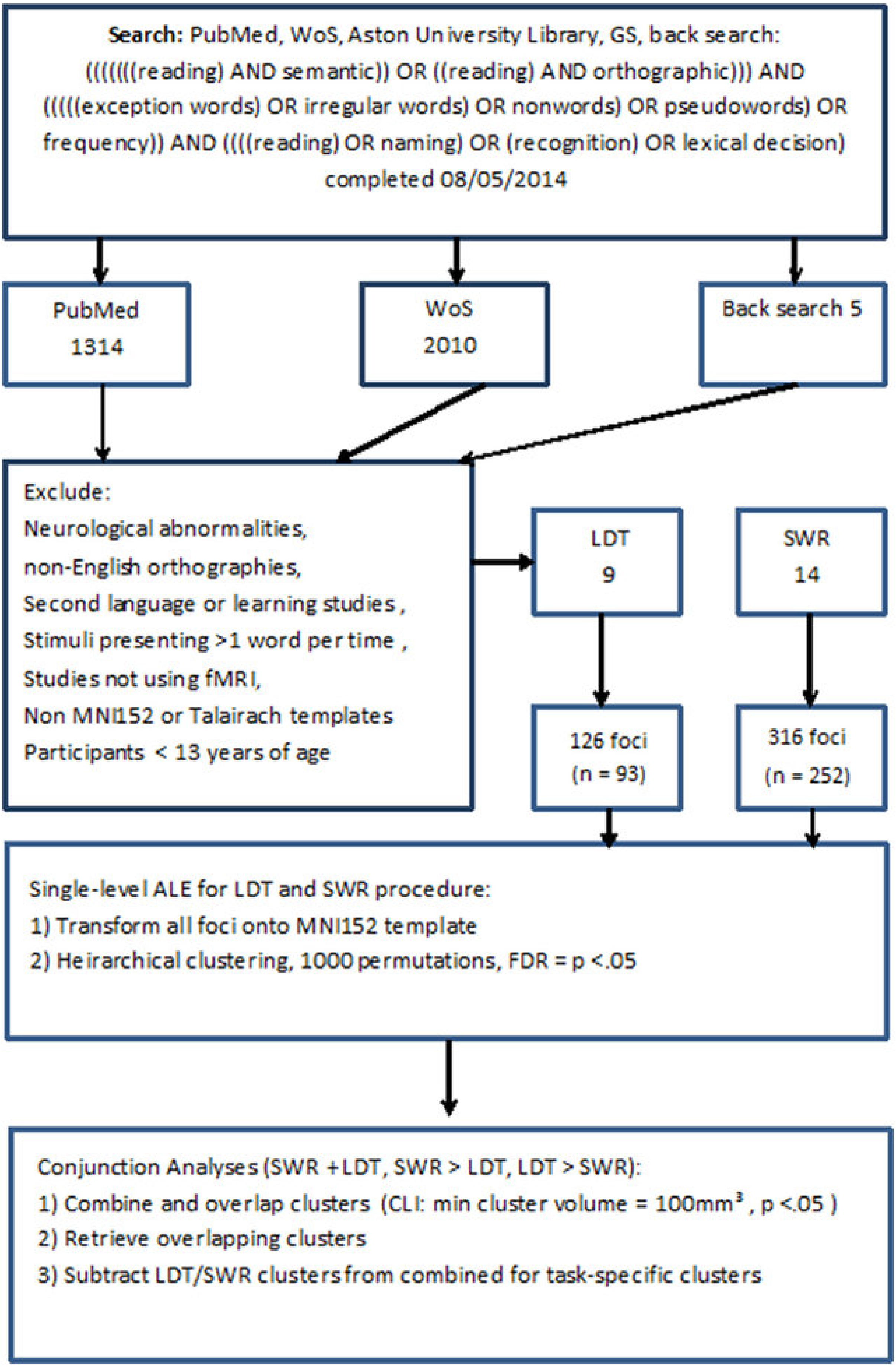
2. Methods

2.1 Literature Search

Since a major objective of this meta analysis was to limit the variance in the findings, the ALE method was seen as more favourable over other label and foci derived approaches for the quantitative analysis. As opposed to label-driven methods, the ALE algorithm plots maxima voxel co-ordinates in 3D stereotactic space; and cortical regions are labelled secondarily. In 3D space foci are plotted on a template brain as Gaussian distributions, the height and full width at half maxima (FWHM) of which are determined by the group size from which the reported co-ordinates were found. The larger the group size, the taller the kernel, the narrower the FWHM, the greater the spatial-specificity of the plot and the better the statistical certainty that the voxel is associated with a function as opposed to random noise. Kernel height at every voxel co-ordinate is plotted on a histogram, which is then randomly permuted to create synthetic data-sets, by which the original data compared via a series of t-tests. Regions of overlap surpassing a chosen critical volume are used to create clusters and retrieve cortical labels (see Turkeltaub et al, 2002 for more information).

This method of quantitative analysis is highly sensitive to specific cortical regions, thus allowing a more concise insight into the reading network; and key-differences due to task-demands.

A illustrated by figure 1, two Boolean medical subject headings (MeSH) searches were conducted to identify peer-reviewed studies of BOLD-fMRI studies available online. It was found that including task-specific keywords (i.e. [AND/OR] “lexical”, “decision”, “task”, “LDT”, “single”) did not generate any novel finds; and were only explicitly included in a second search in order to check whether the initial search had been thorough.



Search: PubMed, WoS, Aston University Library, GS, back search:
((((((reading) AND semantic)) OR ((reading) AND orthographic))) AND
(((((exception words) OR irregular words) OR nonwords) OR pseudowords) OR
frequency)) AND ((((reading) OR naming) OR (recognition) OR lexical decision)
completed 08/05/2014

PubMed
1314

WoS
2010

Back search 5

Exclude:
Neurological abnormalities,
non-English orthographies,
Second language or learning studies,
Stimuli presenting >1 word per time,
Studies not using fMRI,
Non MNI152 or Talairach templates
Participants < 13 years of age

LDT
9

SWR
14

126 foci
(n = 93)

316 foci
(n = 252)

Single-level ALE for LDT and SWR procedure:
1) Transform all foci onto MNI152 template
2) Heirarchical clustering, 1000 permutations, FDR = p <.05

Conjunction Analyses (SWR + LDT, SWR > LDT, LDT > SWR):
1) Combine and overlap clusters (CLI: min cluster volume = $100mm^3$, p <.05)
2) Retrieve overlapping clusters
3) Subtract LDT/SWR clusters from combined for task-specific clusters

Fig. 1. PRISMA flow diagram showing the literature search pipeline. PubMed and Web of Science (WoS) databases were searched independently and cross-overs of relevant papers were examined before conducting back searches. Once exclusion criteria had been applied 126 foci from 9 studies contributed to the LDT data set, whilst 316 foci from 14 studies contributed to the SWR data set. All foci were mapped onto the MNI152 template following a transformation algorithm for foci provided in Talaraich coordinates.

Inclusion criteria were drawn so that the fundamental aspects of the tasks used single words, non-words or letter-strings; and decisions based around the question: “Is it a word?” constituted eligibility for LDT. Similar paradigms such as “does it sound like or mean the same [as a prime word]?” were not up for inclusion since these paradigms incorporate a priming or memorisation element. Go/no-go studies were included within the LDT dataset, on the basis that both paradigms require a response to stimulus based upon a judgement of lexicality. The subtle difference between LDT and go/no-go tasks is that in go/no-go, participants are required to read real word stimuli aloud and remain unresponsive to nonwords, whereas judgments in LDT are ‘yes/no’ responses (Perea et al, 2002). It could be argued that go/no-go tasks are therefore qualitatively similar to SWR aloud paradigms as well as LDT; but the shared traits between SWR aloud and go/no-go tasks involve articulation – which is often modelled out of the data. Therefore, the decisional mechanism is of interest in this meta analysis; and so go/no-go tasks are grouped with LDT.

Both region of interest (ROI) studies were of interest in addition to whole-brain data, as clusters retrieved by the ALE algorithm are independent of each other, i.e. the likelihood that a given cluster reaches statistical significance does not depend on that of any other cluster, regardless of whether it contains foci either exclusive to, or a combination of whole-brain and ROI data. Rather, including ROI data in addition to whole-brain data will add power to a cluster. It is statistically improbable that a BOLD activation in a given region would reach significance in an ROI study but not a whole-brain analysis

Our exclusion criteria covered studies of learning, left-handedness, non-English dialects, memory and sentence reading, in order to further restrict potential task-related differences. Foci irretrievable in either MNI or Talairach space were also excluded since it is uncommon practice to use any other type of stereotactic co-ordinate system for group-level data; and the ALE software cannot account for individual head-spaces when modelling the foci onto a 3D template. Foci that were not significant or reported as deactivations were excluded. Negative BOLD responses are not commonly reported and their neural underpinnings are poorly understood, therefore, any negative BOLD values in our data would be difficult to interpret and lack practical value. Reporting negative BOLD changes is rare; but their inclusion would be likely to incur power issues; and thus have a dampening effect in the pooled data (Fox et al, 1998).

When extracting BOLD foci, if single participant and group-level data were reported, only the group data was included. However, if only statistically significant single participant data were reported, it was still included on the basis that the ALE algorithm accounts for differences in group size by weighting foci modelled on the 3D template (see Turkletaub et al, 2002 for further information). When multiple contrasts were reported, task versus baseline results were prioritised to avoid repeatedly sampling activations included in multiple contrasts. Contrasts not specific to word processing (e.g. faces > words, in Woodhead et al; 2011) were ignored; but in the absence of task > baseline foci, word-specific contrasts (e.g. words > faces, in Woodhead et al; 2011) were included for because activations in these contrasts are based on lower-level task > baseline comparisons. However, baseline comparisons were considered preferable, as it is likely that some reading-specific activations could be dampened by a task > task comparison relative to a task > baseline. In the same light as ROI foci, non-baseline comparisons still include voxel activations statistically independent from those elsewhere in the cortex, so can still be considered useful by contributions to the analysis as a whole.

In all, 1314 papers were found in PubMed; and the Web of Science database produced a further 2010 search results; but no relevant findings and after duplicates were removed. Abstracts from articles of interest were read to assess for methodological practices in concordance with the inclusion criteria, if this need was met, the method and results sectioned were read in their entirety to ensure there were no grounds for exclusion. Papers that used SWR or LDT paradigms alongside an additional task were included but only task-specific contrasts were extracted.

2.2 ALE Methods

The ALE method averages statistically significant focal activations using 3D labels given in stereotactic space, weighted by group size (Turkletaub et al, 2002). This solution overcomes spatial uncertainty issues associated with using fixed foci; and better allowing for false-effects and multiple-comparisons to be compensated for in the ALE algorithm (Laird et al, 2005). Initially, a single-level analysis is used to describe all the pooled data within a given group, secondarily to this, a subtraction analysis retrieves both shared and unique clusters of activation between 2 data-sets.

The GingerALE 2.3.6 software package (Eickenhoff et al, 2009, 2011, 2012 & 2016; Turkeltaub et al, 2012) was used to conduct this meta-analysis, whereby foci in Talairach space were transformed into MNI(SPM) space using a reverse Lancaster transform (Lancaster et al, 2007) and applied to the less conservative MNI mask in order to account for potential spatial differences between mask-type (Eickenhoff et al, 2009). The false discovery rate (FDR pIN) was used to control for type 2 error whilst accounting for multiple comparisons (Chumbley & Friston, 2009; Laird et al, 2005) by finding ALE scores across 1000 permutations of the original data-set, recording the most statistically significant; and repeating the process so that only a given percentage of false discoveries are found, which was set to p<.05 in this study, with no a priori assumptions of the data before anatomical labels are assigned to common foci. GingerALE software versions predating the 2016 2.3.6 package were reported to inaccurately rank P-values during the permutation process, causing data to be 'right-shifted' and increasing the likelihood of type 1 error (Eickhoff et al, 2016). Therefore, we argue that the findings in this study are more dependable than similar meta analyses conducted on earlier GingerALE softwares.

In order to be able to conduct a thorough review of the available data, contrast-specific foci used in all the analyses were divided by task-type and are provided in supplementary text B. When foci were not provided in the paper, authors were contacted where possible.

This resulted in 14 SWR and 9 LDT studies being included in the analysis, including 6 papers published after Cattinelli et al (2013). All of the data was used for the initial ALE analyses concerned with SWR, LDT task-related differences (LDT, SWR, LDT > SWR, SWR > LDT). Methodological differences within SWR studies were investigated (ALOUD, SILENT, ALOUD > SILENT, SILENT > ALOUD, WORDS, NONWORDS, WORDS > NONWORDS, NONWORDS > WORDS) but only contrasts where we could be certain of these characteristics were included for this level of investigation. Since the focus of this meta analysis was to establish the effects of task demands on the reading network as a whole, foci from words and nonwords were pooled, with nonwords referring to both pseudowords and consonant strings.

Of the 23 articles that met the inclusion criteria, 14 contained 305 foci from 417 fluent English participants undergoing SWR tasks and 9 LDT studies detailing 144 foci from 226 fluent English participants (study characteristics are outlined in supplementary text A). Table I shows demographics for foci from contrasts of interest used in the ALE analyses but tables showing contrast-specific foci can be found in supplementary text A. In all, 43% of SWR contrasts were single-level data, contributing 16 foci coordinates. All 15 contrasts used in the LDT analysis represented group-level data. A chi-square test found SWR studies were more likely to adopt block design over event-related (EV); but LDT studies exclusively used EV designs ($\chi^2(1)= 15.7$, $p<.001$, $\phi = 0.83$). For LDT studies, control motor/response tasks were noted to assess where decisional/finger-tapping movements were subtracted from the active task BOLD in addition to a comparative condition in the reported contrasts. Interestingly, no LDT study described the use of an additional decision-making task; and only 2 LDT papers described the use of a motor-control task to account for finger-tapping responses (Rossell et al, 2001; Bruno et al, 2008). In terms of whole brain or ROI analyses, there was no difference between SWR and LDT studies.

Within the SWR studies, a chi-square goodness-of-fit test did not find any significant differences between the use of silent, or aloud reading paradigms ($\chi^2(1) = .23$, $p > .05$, $\phi = .021$). The elimination of some foci from the original SWR group due to combined aloud and silent contrasts showed a reduction in power, thus clusters were thresholded so CLI FDRpN = .05. Within the reading aloud studies, only one study included a control-task to accommodate articulatory artefacts (Seghier et al, 2008). All clusters found in the single task estimates survived a significance threshold of $p < .05$ when the minimum cluster volume threshold was set at 100mm^3 and the number of threshold permutations was set to 1000.

Table I. Studies used in the meta analysis

Authors (year)	**Task**	**Design**	**No. of participants**	**Tesla**	**Template**	**Control/ Contrast tasks**	**No. of foci**
Authors	Task	Design	n	Tesla	Template	Task Contrast	N foci
Mano et al, 2013	SWR aloud	Block	8	3	Talairach	Pseudowords, letter strings	16
Cummine et al, 2013	SWR aloud	Block	20	1.5	Talairach	Pseudowords	20
Church et al, 2011	SWR	EV	24	1.5	Talairach	Rest	13
Purcell et al, 2011	SWR silent	Block	17	3	Talairach	Rest	9
Newman & Jonnisse, 2011	LDT	EV	13	4	Talairach	Nonwords & congruency	21
Woodhead et al 2011	SWR silent	Block	12	1.5	MNI	Faces, grey noise images	3

Graves et al, 2010	SWR aloud	EV	20	3	Talairach	Rest	21
Bruno et al, 2008	LDT	EV	28	3	Talairach	Rest	3
Hauk et al, 2008	SWR silent	Block	21	3	MNI	Rest	11
Seghier et al, 2008	SWR aloud	Block	43	1.5	MNI	Rest	25
Mechelli et al, 2006	SWR silent	Block	6	2	Talairach	False fonts	10
Binder et al, 2005	SWR aloud	Block	24	1.5	Talairach	Pseudo & irregular words	152
Dietz et al, 2005	SWR aloud & silent	Block	16	1.5	MNI	Pseudowords	4
Edwards et al, 2005	LDT	Block	18	3	Talairach	Rest	10
Mechelli et al, 2005	SWR silent	Block	22	1.5	MNI	False fonts	6
Binder et al, 2003							

	LDT	EV	24	1.5	Talairach	Nonwords	26
Mechelli et al, 2003	SWR silent	Block	20	1.5	MNI	Pseudowords	10
Rossell et al, 2003	LDT	EV	8	2	MNI	Relatedness	19
Henso et al, 2002	LDT	Block	12	2	Talairach	Nonwords	7
Polk et al, 2002	SWR silent	Block	8	1.5	Talairach	Consonant strings	9
Rossell et al, 2001	LDT	Block	12	1.5	Talaiarach	Rest	8
Kiehl et al, 1999	LDT	Block	6	1.5	Talairach	Rest	19

3. Results

3.1 Single data set ALE

Information regarding retrieved ALE values are described in table 2 and shown in Figure 2. In total, 12 clusters were found for SWR, with the largest in volume and extrema weighted probability score was formed from 7 foci in the left fusiform gyrus, whilst the second largest cluster was formed from 2 foci but was a very close overlap to cluster 1 in the left fusiform gyrus (Brodmann areas 37). Across the left cerebral cortex, 5 additional ventral clusters cover the medial, transverse and superior temporal gyrus (STG) and IFG (Brodmann areas 9, 6, 41 respectively), whilst 5 dorsal clusters were found along the precuneus, superior and inferior parietal lobule, precentral and medial frontal gyri (Brodmann areas 19, 17, 7 and 40; Figure 1). The two right-lateralised clusters are also the smallest, with the lowest extrema values and are homologous to leftward clusters in the inferior parietal lobule and IFG (Brodmann areas 7 and 9).

For LDT foci, only 3 clusters were found, with the largest cluster containing 9 foci covering the right IFG (Brodmann areas 9). Another right-lateralised cluster was found in the fusiform and lingual gyri (Brodmann areas 19 and 18), whilst the sole leftwards clusters were found in the fusiform gyrus (Brodmann area 37). Clusters common to both single-set analyses include Brodmann area 37 (fusiform gyrus) in the left hemisphere. Further subtraction analyses were used to examine the extent of cluster overlap and uniqueness for each paradigm type.

Table 2. Central voxel co-ordinates in MNI152 space and labelled using the AAL atlas in MRIcron per contrast. (L = left, R = right).

Contrast	Extrema Value (10^{-2})	x	y	z	AAL Label	Brodmann Area
SWR	2.27	-50	-48	-8	L Fusiform Gyrus	37
	2.26	-50	-8	44	L Precentral Gyrus	4
	2.24	-40	28	24	L Middle Frontal Gyrus	9
	2.22	-54	-16	8	L Superior Temporal Gyrus	41
	2.09	-42	-54	-18	L Fusifrm Gyrus	37
	2.03	-22	-68	48	L Precuneus	7
	2.01	-44	6	26	L Inferior Frontal Gyrus	9
	1.95	48	12	24	R Inferior Frontal Gyrus	9
	1.94	-24	-98	-4	L Inferior Occipital Gyrus	17
	1.81	-4	-2	56	L Medial Frontal Gyrus	6
	1.73	34	-56	50	R Superior Parietal Lobule	7
LDT	2.51	41	9	28	R Inferior frontal gyrus	9

	2.13	-38	-60	-14	L Fusiform Gyrus	37
	1.75	20	-78	-10	R Lingual Gyrus	18
LDT + SWR	1.84	-41	-53	-15	L Fusiform gyrus	37
	1.73	46	12	24	R Inferior Frontal Gyrus	9
LDT > SWR	3.29	43	6	32	Right precentral gyrus	6
SWR > LDT	..	..	..	..	..	..

3.2 Subtraction Analyses

A subtraction analysis was conducted on the single-level SWR and LDT data to identify clusters that are shared and unique to each paradigm. Unless otherwise stated, all clusters in the subtraction analyses survived a corrected significance threshold of p<.05 at a permutation threshold of 1000 and minimum cluster volume of 100mm^3. Images are provided in figure 2. A conjunction analysis showing the correspondence across both SWR and LDT foci confirmed common findings in the single data set analyses by retrieving 2 shared clusters, the largest of which covers the left fusiform gyrus (Brodmann area 37), whilst the remaining cluster can be found over the right precentral gyrus (Brodmann area 6).

A subtraction analysis of SWR > LDT foci found no significant clusters. The only retrievable cluster in this analysis was located over the left fusiform gyrus when the significance value was uncorrected for multiple comparisons (Brodman area 37). For the LDT > SWR contrast, only one cluster was found over the right precentral gyri (Brodmann area 6).

A within SWR subtraction analysis was conducted to explore potential differences between aloud and silent conditions; and found only very subtle differences in cortical activation patterns that related to speech production and perception in reading. More details on this analysis can be found in supplementary texts B & C.

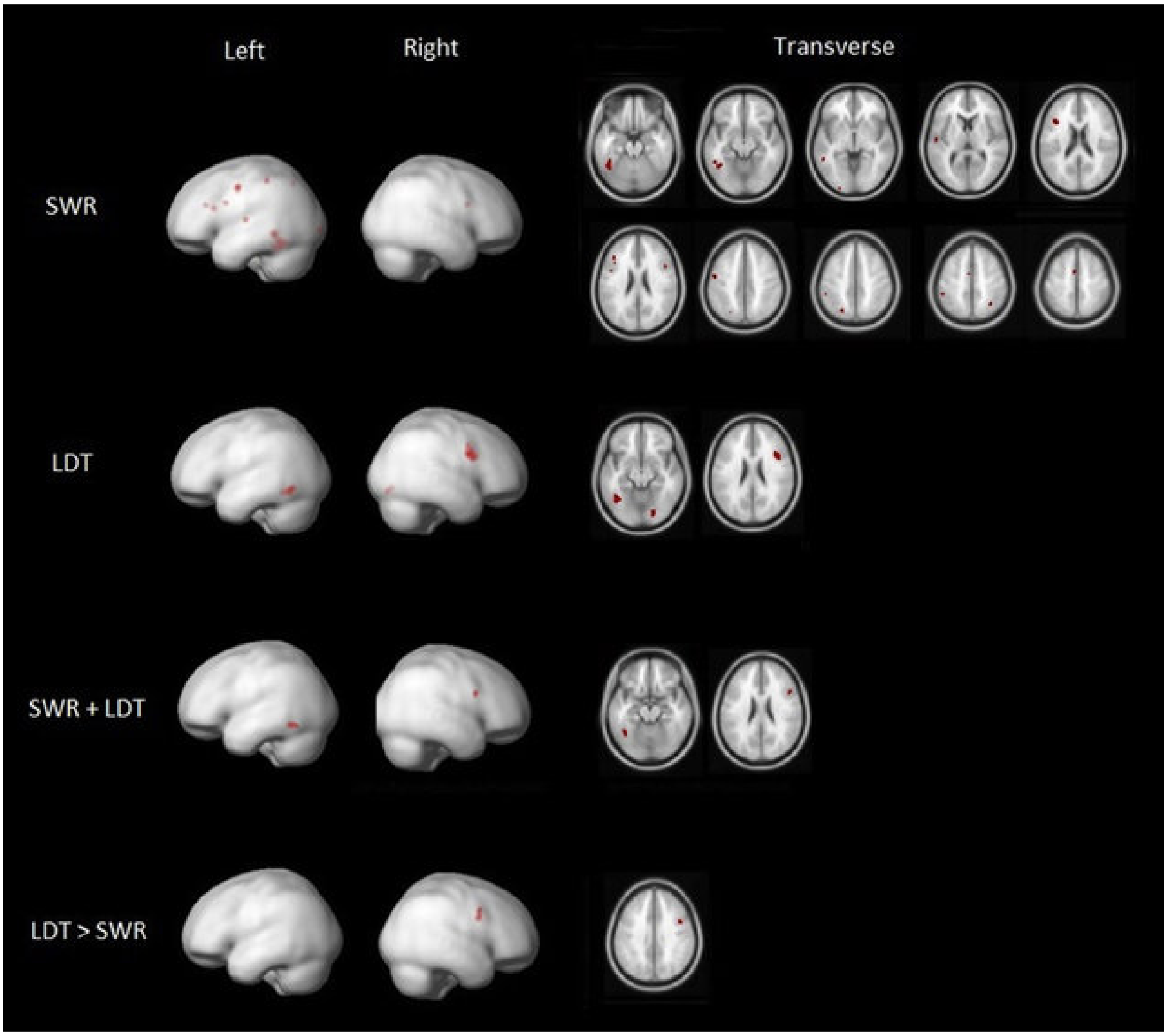


Figure 2. Clusters retrieved from SWR, LDT and conjunction analyses rendered on inflated cortical surfaces, accompanied by a view of all clusters in the transverse plane (left = left, right = right). All clusters survived FDR multiple comparison correction when p<.05. SWR + LDT images only show clusters shared between tasks, LDT > SWR shows clusters specific to LDT in comparison to SWR. SWR > LDT did not retrieve clusters after FDR correction.

4. Discussion

The aim of this meta analysis was to provide quantitative summations of the two most prevalent experimental tasks used to investigate English single word reading in fMRI (SWR & LDT), in order to assess their differences and therefore establish the extent to which task-demands can affect the data. Our results show that both tasks retrieve common left lateralised clusters in regions associated with orthographic processing; and to some smaller extent, semantic processing, which are assumed to reflect common cognitive activity. Interestingly, the LDT clusters were not found to be associated with phonological processing; but instead executive language, decision-making and motor-planning functions. We found a heavily predictable clear left-lateralized reading network in SWR, which was only partially evident in LDT, so we suggest that SWR and LDT are measuring different psychological phenomena that share functions associated with reading sub processes.

4.1 Single Word Reading

The single dataset analysis for SWR clusters distributed across the left cortical hemisphere in regions that have previously been heavily associated with orthographic, phonological and semantic processing, which closely resemble previous meta analytic results (Turkletaub et al, 2002; Vigneau et al, 2011; Cattinelli et al, 2013). The cluster with the highest weighted-centre was located over the left posterior fusiform gyrus, which is responsible for word and letter storage (Moore & Price, 1999; Dehaene et al, 2002; Deheane et al, 2005; Wang et al, 2011); and processing contours and simple patterns, like alphabetic letters (Puce et al, 1996; Ishai et al, 2000; Carreiras et al, 2007; Wang et al, 2011). Other clusters were found in the left inferior and middle temporal gyri which are both involved with word retrieval (Nobre et al, 1994; Price et al, 1996; Paulesu et al, 2001) and spelling (Rapcsak & Beeson, 2004), indicating that orthographic processes may be the most critical components of SWR reading.

The remaining clusters are distributed ventrally along the left superior temporal lobe toward the inferior frontal lobe; and along left dorsal parietal and frontal areas. Ventral clusters lie on the STG approximate to Wernicke's area and the primary auditory cortex, for language-specific comprehension, semantic processing; and orthographic-phoneme conversion (Paulesu et al, 2001; Poldrack et al, 2001; Bigler et al, 2007; Visser & Ralph, 2011). The insular cluster corresponds to previous findings linking the region to visual-auditory integration and phonological processing and language comprehension (Engelien et al, 1995; Habib et al, 1995; Hyman & Tranel, 1989; Mutschler et al, 2009). Interestingly, IFG activation was found bilaterally for SWR tasks which is indicative of semantic association (Vingerhoets et al, 2003). Whilst the left IFG is prolifically associated with semantic comprehension (Costafreda et al, 2006), the right IFG is better associated with executive processes that govern semantic comprehension (Vigneau et al, 2011). As the clusters for SWR are spatially distributed along regions associated with semantic and phonological processing, our findings are in keeping with ventral and dorsal visual and reading routes predicted by the dual-route theory of reading and connectionist models (Plaut et al, 1996; Coltheart et al, 2001; Vigneau et al, 2006, 2011)) and add weight to findings from functional connectivity studies (Hampson et al, 2005)

The remaining clusters include the precuneus, precentral and medial frontal gyri which have been related to attention and visuospatial processes (Simon et al, 2002; Knauff et al, 2003; Cattinelli et al, 2013). Meanwhile, activity found within the inferior parietal lobule (IPL) remains difficult to specify, although it is generally accepted that activity in this region is involved in language comprehension and orthography-phonology conversion (Clark et al, 2000; Moore & Price, 1999), articulation (Takayama et al, 1993; Brown et al, 2009) and semantic-phonological mediation (Jindal & Liu et al, 2006). However, there is much debate in the literature over the precise roles of the inner regions within the IPL such as the angular gyrus (AG), which itself is sub-dividable into areas related to both default mode network and language-related activities amongst others (Long et al, 2008; Greicius et al, 2009; Seghier et al, 2010 & 2013). To illustrate, the left AG is understood to be involved with left – right discrimination (Hirnstein et al, 2011); and semantic integration (Seghier et al, 2010), whereas the right AG plays an essential role in guiding visuospatial attention (Cattaneo et al, 2009), orthography – phonology conversion (Clark et al, 2000), pseudoword reading (Horwitz et al, 1998) and phonological assessments (Hartwigsen et al, 2010). Bilateral AG activity could be indicative of a left-to-right stream of letter processing; and would be an interesting avenue of investigation as it may contribute to serial or parallel theories of reading. As this is a meta-analysis and not first-level experimental data, it would be too far to conjecture the role(s) of the IPL in SWR; but from looking cross the body of fMRI literature, it is clear that the precise nature of activity occurring within these sub-regions require further investigation to clarify.

4.2 Lexical Decision Task

The most weighted and perhaps controversial cluster in LDT covered the right IFG and precentral gyrus, which are attributed to having key executive and working memory functions, not specific to language or meaning (Vigneau et al, 2011), including inhibiting responses (Menon et al, 2001; Rubia et al, 2003; Aron et al, 2004; Nakic et al, 2006; Hampshire et al, 2010) and planning finger and eye movements (Yousry et al, 1997; Muggleton et al, 2003). Importantly, nonwords are a vital component of LDT and go/no-go tasks, whereby the typical response method is to button press or verbalise a real word and refrain from responding to a nonword. Since the right IFG has been linked to response inhibition and words and nonwords were pooled in this meta analysis, it would not be overly presumptuous to assume that response inhibitions could be infiltrating the measured reading activation for nonwords.

Similarly, right IFG activity can further represent nonword conditions by having been associated with facilitating word and syllable segregation and silent reading of pseudowords when spelling-sound consistency is not readily accessed (Poldrack et al, 1999). This notion corroborates with the co-presence of the right-lateralised lingual gyrus cluster, which has been shown to be recruited in LDT studies for assisting orthographic processes when alphabetic readers process word-like patterns with similar contour characteristics as orthographic symbols or random letter strings (Changizi et al, 2006; Vogel et al, 2013).

Since LDT paradigms require hasty judgements of lexicality, it could be argued that the lack of left IFG activation relative to SWR indicates that semantic input may not always be a necessary component and some responses can be made superficially. The presence of the left fusiform gyri cluster in the absence of other well-defined reading regions suggests that rapid lexicality judgements can be based mostly on orthographic access (Vingerhoets et al, 2003). New research supports this notion by suggesting that lexicality can be judged within first instances (Gwiliams et al, 2017), which would mean that the emphasis on speedily and accurate responses in LDT could promote pattern recognition styles of processing in place of reading.

4.3 Subtraction analysis: SWR vs LDT

Prior to subtraction, a conjunction analysis was conducted to assess the extent of the overlap between SWR and LDT. Out of the 3 all-LDT clusters, 2 were shared with SWR: left fusiform and right precentral gyri. Since orthographic access is a shared necessity for completing either task, the left fusiform cluster is not surprising. As has been discussed, the right IFG has been found to be facilitate word and syllable segregation and phonological processing when spelling-sound consistency is not readily accessed (Poldrack et al, 1999). Nonwords were pooled with real words in the SWR and LDT datasets in order to investigate the extent of task demands on the reading network as a whole which could explain the presence of this shared cluster.

For the LDT > SWR analysis, the most weighted cluster was again found over the right inferior frontal and precentral gyrus. On top of extra semantic and phonological processing, the right IFG has also been associated with response inhibition as previously discussed; and like the all-LDT findings this cluster may have even more relevance to LDT tasks when the precentral gyrus is considered. The right frontal eye field is again implicated and is strongly associated with planning saccadic eye movements and guiding visual attention (Muggleton et al, 2003; Buschman & Miller, 2009); and the precentral gyrus aids planning and executing motor movements of the hands (Yousry et al, 1997; Connolly et al, 2002).

Despite hand-movements being an integral part of the response modality, the right hand was used throughout all studies, which would be expected to promote left-lateralised activity. However, a meta analysis of finger tapping studies found that in response to target visual cues activity in the right inferior frontal region can co-occur with bilateral occipital activity, similar to the findings in this meta analysis (Witt et al, 2008). Perhaps an assumption when conducting LDT studies is that finger and eye movements will be subtracted from the BOLD response when comparing correct responses to target stimuli from across two conditions. In this review, 4 studies provided foci from active versus rest comparisons (see Supplementary text A) and only one of these studies explicitly stated the use of a subtractive finger-tapping task (Rossell et al, 2001). Furthermore, it cannot be ruled out that anticipation of an upcoming target trial may also be affecting saccadic movements in ways that cannot be subtracted when compared to a rest condition. Without experimental manipulation it is difficult to ascertain the relevance of this cluster. Yet it seems clear that a large part of this activation is likely to result from the task demands of LDT on top of extra semantic and phonological processing. Moreover, a decisional mechanism coupled with an emphasis on a speedy and accurate behavioural response in LDT could be affecting how participants anticipate upcoming trials when responding correctly/quickly/slowly and to words/nonwords etc. For certain better caution could be applied in future LDT studies to avoid motor response contamination.

In the SWR > LDT analysis, no statistically significant clusters were found which was somewhat unexpected. SWR is a slower process than LDT, more deliberate consideration of the stimulus is allowed which should bolster the BOLD response (Cohen et al, 2002; Gould et al, 2016). In this meta analysis we aimed to constrain as much variance as possible in attempts to enhance consistency across tasks so differences are a direct result of task demands. One could argue that stringent inclusion criteria could ultimately lead to a relative paucity of data; but we do not expect this to be the ultimate case, nor do we expect that there are no quantitative differences between SWR and LDT since there are obvious qualitative ones. Instead, two similar explanations for a lack of statistically significant clusters in SWR > LDT comparisons come to light.

Firstly, changing semantic, phonological and frequency characteristics have been shown to interact for successful known word reading - which leads to fluctuating recruitment of regions along the reading network in SWR (Graves et al, 2010; Cummine et al, 2013); but this has yet to be shown at the whole-brain level using LDT. Thus as the rapid response element of LDT can encourage superficial judgements of lexicality (Vingerhoets et al, 2003; Gwilliams et al, 2017), then it would be assumed that the neural response is more consistent across words compared to fluctuating SWR activity (Graves et al, 2010; Cummine et al, 2013). Interestingly, experimental work suggests that differences between LDT and SWR only small differences in the neural responses to changes in lexicality exist across tasks independent of task demands (Carrieras et al, 2007); but potential differences will certainly be exaggerated in meta analyses.

Secondly, the SWR dataset used for this meta analysis was composed of aloud and silent reading data, so even though it is standard practice to remove articulation activations from reading aloud measurements (Perea et al, 2001), variance as a result of differing cognitive engagement and difficulty leading up to the articulation process of words and nonwords can still affect the BOLD response and further increase the variance (Taylor, Rastle & Davis, 2014). Therefore, voxel coordinates reported across SWR tasks could be more likely to be spread more diffusely across the cortex, which would diminish the odds of clusters surviving statistical thresholds relative to a more consistent LDT dataset in an ALE meta analysis. Greater consistency across LDT BOLD measurements can provide reasonable explanation for an absence of SWR > LDT clusters after multiple comparison correction. Although further work investigating differences between SWR and LDT for words and nonwords is required simply to better quantify the differences between tasks, it is accepted that varying lexical traits are also representative of typical reading in English.

4.4 Additional Analyses: Single word reading – silent and aloud reading

Since the aim of this meta-analysis was to investigate the influence of task-type on the reading network, within-task SWR silent vs aloud conditions were investigated. The participant demographics for the sub SWR analyses are listed in supplementary text B; and a table of results are provided in supplementary text C. Three clusters remained when subtracting silent from aloud activations, with the largest spanning the left STG and precentral gyrus. Previous studies have shown bilateral precentral gyrus activity is implicated in frontal-motor regions heavily associated with articulation, particularly control of the tongue and larynx (Braun et al, 1997; Bookheimer et al, 2000; Riecker et al, 2000; Brown et al, 2008; Brown et al, 2009). The remaining clusters encompass the left fusiform gyrus and culmen of the cerebellum – which has previously been implicated to attention allocation (Kübler et al, 2005; Cattinelli et al, 2013). Only one cluster remained when subtracting aloud from silent clusters, which was in the left middle temporal gyrus, previously noted to be associated with word retrieval (Nobre et al, 1994; Price et al, 1996; Paulesu et al, 2001). As with LDT studies, perhaps reading aloud paradigms should include a control speech task so to ensure reading-specific neural activations are not flawed by task-specific noise.

4.5 Methodological Implications

Compared to silent reading paradigms, LDT and aloud SWR tasks have the advantage of responses being used as checking mechanisms to assure the participant can indeed read the stimuli; and that they are paying attention to the task at hand (Carreiras et al, 2007). Although precautionary, this conjecture may be undervaluing the integrity of the participants since there is no evidence to suggest that the results of reading paradigms are flawed in anyway by reading silently. In fact, silent reading elicits full lexical, phonological and semantic activity without being tainted by finger/articulatory motor command activations and minute in-scanner body-movements associated with a response modality (Hajnal et al, 1994; Yetkin et al, 1995), nor by attentional mechanisms to the same extent as reading aloud, or LDT. Because of this, there is less opportunity for misinterpretations of the data. However, we accept that sub vocal activations are highly intertwined with silent reading as a normal by-product, in fact, readers must learn to inhibit vocalisations in reading through instruction (Daneman & Newson, 1992), so it should just be accepted as part of the fundamental reading network despite not contributing to the deciphering of word stimuli.

In LDT however, there is an additional risk of measuring false-positives, as there is heightened opportunity for incorrectly recognising unknown words (and incorrectly rejecting highly plausible pseudowords, which will then be incorrectly removed from subsequent analyses) because of experimental emphasis on the speed of the behavioural response. We do accepted that a definitive response to target words, such as a binary button-press option, may be essential for addressing some research questions. Therefore, we suggest a finger tap be used in conjunction with dummy decision making trials in LDT; and that the speed of the response not be emphasised during the task if it is not essential to addressing the research question at hand, as it is a prime candidate for what hinders full lexical and phonological access in LDT.

Finally, aside from the issues between paradigm choices, another constraint to be considered is the inclusion of English speakers only. It is understood that a large body of data of alphabetic reading was by-passed in this meta-analysis but orthographic depth has potential to be a large source of variance; and since there is a large body of literature on native English readers (Share, 2008). Anyway, a deep orthography like English is likely to strain the reading network and fully utilise sound-spelling and semantically-driven routes, as there are still many words that follow regular spelling-sound rules that are read in a similar style to words in shallow alphabetic orthographies, so novel substrates should not become apparent if this meta analysis were to be repeated using a range of orthographic transparencies. However, it would be of interest to conduct a large-scale study of the normal reading network using shallow and non-alphabetic orthographies so that differences in activation patterns across the reading network can be clearly disseminated.

5. *Conclusion*

This meta analysis found that for fMRI studies of fluent monolingual English word reading, SWR tasks are more than sufficient for eliciting activity across the full network than LDT, which may not depend on full utilisation of the reading network for successful completion and is far more susceptible to being flawed by task demands. SWR tasks retrieve highly predictable orthographic, phonological and semantic related clusters, whereas LDT elicits bilateral orthographic clusters, a relatively smaller semantic response; and right hemispheric activity more closely associated with decision making and movement planning. Contrary to conventional practice in neuroimaging, we suggest that validation mechanisms are not strictly essential for studying typical reading processes and that if a validation mechanism is required, the full implications of choosing a finger tapping mechanism over single word reading should be carefully considered.

Acknowledgements:

I would like to thank Klaus Kessler and Paul Furlong for their useful and constructive feedback on this project. Together, Joel, Jigar and myself would like to declare that there are no conflicts of interests associated with this meta analysis.

References

*(References marked with a * were included in the meta analysis)*

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in cognitive sciences*, *8*(4), 170-177.

Bigler, E. D., Mortensen, S., Neeley, E. S., Ozonoff, S., Krasny, L., Johnson, M., ... & Lainhart, J. E. (2007). Superior temporal gyrus, language function, and autism. *Developmental neuropsychology*, *31*(2), 217-238.

*Binder, J., McKiernan, K., Parsons, M., Westbury, C., Possing, E., Kaufman, J., & Buchanan, L. (2003). Neural correlates of lexical access during visual word recognition. *Cognitive Neuroscience, Journal of*, *15*(3), 372-393.

*Binder, J. R., Medler, D. A., Desai, R., Conant, L. L., & Liebenthal, E. (2005). Some neurophysiological constraints on models of word naming. *Neuroimage*,*27*(3), 677-693

Bookheimer, S.Y., Zeffiro, T.A., Blaxton, T.A., Gaillard,W., Theodore, W.H., (2000). Activation of language cortex with automatic speech tasks. Neurology 55, 1151–1157

Braun, A.R., Varga, M., Stager, S., Shulz, G., Selbie, S., Maisog, J.M., Ludlow, C.L., (1997). Altered patterns of cerebral activity during speech and language production in developmental stuttering. An H215O positron emission tomography study. Brain, 120, 761–784.

Brown, S., Ngan, E., Liotti, M., (2008). A larynx area in the human motor cortex. Cereb. Cortex 18, 837–845.

Brown, S., Laird, A. R., Pfordresher, P. Q., Thelen, S. M., Turkeltaub, P., & Liotti, M. (2009). The somatotopy of speech: phonation and articulation in the human motor cortex. *Brain and cognition*, *70*(1), 31-41.

*Bruno, J. L., Zumberge, A., Manis, F. R., Lu, Z. L., & Goldman, J. G. (2008). Sensitivity to orthographic familiarity in the occipito-temporal region.*Neuroimage*, *39*(4), 1988-2001.

Buschman, T. J., & Miller, E. K. (2009). Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron*,*63*(3), 386-396.

Caminiti R, Genovesio A, Marconi B, Mayer AB, Onorati P, Ferraina S, et al. (1999) Early coding of reaching: frontal and parietal association connections of parieto-occipital cortex. Eur J Neuroscience; 11: 3339–45.

Carreiras, M., Mechelli, A., Estévez, A., & Price, C. J. (2007). Brain activation for lexical decision and reading aloud: two sides of the same coin?. Journal of cognitive neuroscience, 19(3), 433-444.

Cattaneo, Z., Silvanto, J., Pascual-Leone, A., & Battelli, L. (2009). The role of the angular gyrus in the modulation of visuospatial attention by the mental number line. *Neuroimage*, *44*(2), 563-568.

Cattinelli, I., Borghese, N. A., Gallucci, M., & Paulesu, E. (2013). Reading the reading brain: a new meta-analysis of functional imaging data on reading.*Journal of Neurolinguistics*, *26*(1), 214-238.

Changizi, M.A. et al. (2006) The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. Am. Nat. 167, 117–139

Chumbley, J. R., & Friston, K. J. (2009). False discovery rate revisited: FDR and topological inference using Gaussian random fields. *Neuroimage*, *44*, 62-70.

*Church, J. A., Balota, D. A., Petersen, S. E., & Schlaggar, B. L. (2011). Manipulation of length and lexicality localizes the functional neuroanatomy of phonological processing in adult readers. *Journal of cognitive neuroscience*,*23*(6), 1475-1493

Clarke, C. (2000). Perceptual adjustments to foreign accented English.*Research on Spoken Language Processing Progress Report*, (24), 312-335.

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, *125*, 1054-1069

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud.*Psychological review*, *108*(1), 204.

Connolly, J. D., Goodale, M. A., Menon, R. S., & Munoz, D. P. (2002). Human fMRI evidence for the neural correlates of preparatory set. *Nature neuroscience*, *5*(12), 1345.

Costafreda, S. G., Fu, C. H., Lee, L., Everitt, B., Brammer, M. J., & David, A. S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. *Human brain mapping*,*27*(10), 799-810.

*Cummine, J., Gould, L., Zhou, C., Hrybouski, S., Siddiqi, Z., Chouinard, B., & Borowsky, R. (2013). Manipulating instructions strategically affects reliance on the ventral-lexical reading stream: converging evidence from neuroimaging and reaction time. *Brain and language*, *125*(2), 203-214

Daneman, M., & Newson, M. (1992). Assessing the Importance of Subvocalization in Normal Silent Reading. Reading and Writing: An Interdisciplinary Journal, 4, 55-77.

Dehaene, S., Le Clec'H, G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport*, *13*(3), 321-325.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in cognitive sciences*, *9*(7), 335-341.

*Dietz, N. A., Jones, K. M., Gareau, L., Zeffiro, T. A., & Eden, G. F. (2005). Phonological decoding involves left posterior fusiform gyrus. *Human brain mapping*, *26*(2), 81-93

*Edwards, J. D., Pexman, P. M., Goodyear, B. G., & Chambers, C. G. (2005). An fMRI investigation of strategies for word recognition. *Cognitive brain research*, *24*(3), 648-662.

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human brain mapping*, *30*(9), 2907-2926.

Eickhoff, S. B., Bzdok, D., Laird, A. R., Roski, C., Caspers, S., Zilles, K., & Fox, P. T. (2011). Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. Neuroimage, 57(3), 938-949.

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *Neuroimage*, *59*(3), 2349-2361.

Eickhoff, S. B., Laird, A. R., Fox, P. M., Lancaster, J. L., & Fox, P. T. (2016). Implementation errors in the GingerALE Software: Description and recommendations. *Human Brain Mapping*.

Engelien, A., Sibersweig, D., Stern, E., Huber, W., Frith, C., & Frackowiak, R. S. J. (1995). The functional anatomy of recovery from auditory agnosia A PET study of sound categorization in a neurological patient and normal controls.*Brain*, *118*(6), 1395-1409.

Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading.*Proceedings of the National Academy of Sciences*, *95*(3), 914-921.

Fox, P. T., Parsons, L. M., & Lancaster, J. L. (1998). Beyond the single study: function/location metanalysis in cognitive neuroimaging. *Current opinion in neurobiology*, *8*(2), 178-187.

Gould, L., McKibben, T., Ekstrand, C., Lorentz, E., & Borowsky, R. (2016). The beat goes on: the effect of rhythm on reading aloud. *Language, Cognition and Neuroscience*,*31*(2), 236-250.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.*, *30*, 535-574.

*Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2010). Neural systems for reading aloud: a multiparametric approach. *Cerebral Cortex*, *20*(8), 1799-1815.

Greicius, M. D., Supekar, K., Menon, V., & Dougherty, R. F. (2009). Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cerebral cortex*, *19*(1), 72-78.

Gwilliams, L., Poeppel, D., Marantz, A., & Linzen, T. (2017). Phonological (un) certainty weights lexical activation. *arXiv preprint arXiv:1711.06729*.

Habib, M., Daquin, G., Milandre, L., Royere, M. L., Rey, M., Lanteri, A., ... & Khalil, R. (1995). Mutism and auditory agnosia due to bilateral insular damage—role of the insula in human communication. *Neuropsychologia*, *33*(3), 327-339.

Hajnal, J. V., Myers, R., Oatridge, A., Schwieso, J. E., Young, I. R., & Bydder, G. M. (1994). Artifacts due to stimulus correlated motion in functional imaging of the brain. Magnetic resonance in medicine, 31(3), 283-291.

Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. Neuroimage, 50(3), 1313-1319.

Hampson, M., Tokoglu, F., Sun, Z., Schafer, R. J., Skudlarski, P., Gore, J. C., & Constable, R. T. (2006). Connectivity–behavior analysis reveals that functional connectivity between left BA39 and Broca's area varies with reading ability.*Neuroimage*, *31*(2), 513-519.

Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *Proceedings of the National Academy of Sciences*, *107*(38), 16494-16499.

*Hauk, O., Davis, M. H., & Pulvermüller, F. (2008). Modulation of brain activity by multiple lexical and word form variables in visual word recognition: A parametric fMRI study. *Neuroimage*, *42*(3), 1185-1195.

*Henson, R. N. A., Price, C. J., Rugg, M. D., Turner, R., & Friston, K. J. (2002). Detecting latency differences in event-related BOLD responses: application to words versus nonwords and initial versus repeated face presentations.*Neuroimage*, *15*(1), 83-97.

Hirnstein, M., Bayer, U., Ellison, A., & Hausmann, M. (2011). TMS over the left angular gyrus impairs the ability to discriminate left from right.*Neuropsychologia*, *49*(1), 29-33.

Horwitz, B., Rumsey, J. M., & Donohue, B. C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proceedings of the National Academy of Sciences*, *95*(15), 8939-8944.

Hyman, B. T., & Tranel, D. (1989). Hemianesthesia and aphasia: an anatomical and behavioral study. *Archives of neurology*, *46*(7), 816-819.

Ishai, A., Ungerleider, L. G. L. G., Martin, A., & Haxby, J. V. J. V. (2000). The representation of objects in the human occipital and temporal cortex. *Cognitive Neuroscience, Journal of*, *12*(Supplement 2), 35-51.

Jensen, E. J., Hargreaves, I., Bass, A., Pexman, P., Goodyear, B. G., & Federico, P. (2011). Cortical reorganization and reduced efficiency of visual word recognition in right temporal lobe epilepsy: A functional MRI study.*Epilepsy research*, *93*(2), 155-163.

Jindal, N., & Liu, B. (2006). Mining comparative sentences and relations. In *AAAI* (Vol. 22, pp. 1331-1336).

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *Neuroimage*,*20*(2), 693-712.

Katz, L., Lee, C. H., Tabor, W., Frost, S. J., Mencl, W. E., Sandak, R., ... & Pugh, K. R. (2005). Behavioral and neurobiological effects of printed word repetition in lexical decision and naming. Neuropsychologia, 43(14), 2068-2083.

*Kiehl, K. A., Liddle, P. F., Smith, A. M., Mendrek, A., Forster, B. B., & Hare, R. D. (1999). Neural pathways involved in the processing of concrete and abstract words. *Human brain mapping*, *7*(4), 225-233.

Knauff, M., Fangmeier, T., Ruff, C. C., & Johnson-Laird, P. N. (2003). Reasoning, models, and images: Behavioral measures and cortical activity.*Journal of cognitive neuroscience*, *15*(4), 559-573.

Kübler, A., Murphy, K., & Garavan, H. (2005). Cocaine dependence and attention switching within and between verbal and visuospatial working memory. European Journal of Neuroscience, 21(7), 1984-1992.

Laird, A. R., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., ... & Fox, P. T. (2005). ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human brain mapping*, *25*(1), 155-164.

Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., ... & Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human brain mapping*, *28*(11), 1194-1205

Long, X. Y., Zuo, X. N., Kiviniemi, V., Yang, Y., Zou, Q. H., Zhu, C. Z., ... & Zang, Y. F. (2008). Default mode network as revealed with multiple methods for resting-state functional MRI analysis. *Journal of neuroscience methods*, *171*(2), 349-355.

*Mano, Q. R., Humphries, C., Desai, R. H., Seidenberg, M. S., Osmon, D. C., Stengel, B. C., & Binder, J. R. (2013). The role of left occipitotemporal cortex in reading: reconciling stimulus, task, and lexicality effects. *Cerebral Cortex*,*23*(4), 988-1001.

*Mechelli, A., Price, C. J., Noppeney, U., & Friston, K. J. (2003). A dynamic causal modeling study on category effects: bottom–up or top–down mediation?*Journal of cognitive neuroscience*, *15*(7), 925-934.

*Mechelli, A., Crinion, J. T., Long, S., Friston, K. J., Ralph, M. A. L., Patterson, K., ... & Price, C. J. (2005). Dissociating reading processes on the basis of neuronal interactions. *Journal of cognitive neuroscience*, *17*(11), 1753-1765.

*Mechelli, A. Friston, K.J. Price, C. (2006) The effects of presentation rate during word and pseudoword reading: a comparison of PET and Fmri, *Journal of Cognitive Neuroscience,* 12, 145 - 156

Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. Human brain mapping, 12(3), 131-143.

Moore, C. J., & Price, C. J. (1999). Three distinct ventral occipitotemporal regions for reading and object

Muggleton, N. G., Juan, C. H., Cowey, A., & Walsh, V. (2003). Human frontal eye fields and visual search. *Journal of neurophysiology*, *89*(6), 3340-3343.

Mutschler, I., Wieckhorst, B., Kowalevski, S., Derix, J., Wentlandt, J., Schulze-Bonhage, A., & Ball, T. (2009). Functional organization of the human anterior insular cortex. *Neuroscience letters*, *457*(2), 66-70.

Nakic, M., Smith, B. W., Busis, S., Vythilingam, M., & Blair, R. J. R. (2006). The impact of affect and frequency on lexical decision: the role of the amygdala and inferior frontal cortex. *Neuroimage*, *31*(4), 1752-1761.

*Newman, R. L., & Joanisse, M. F. (2011). Modulation of brain regions involved in word recognition by homophonous stimuli: An fMRI study. *Brain research*,*1367*, 250-264

Nobre, A. C., Allison, T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature*, *372*(6503), 260-263.

Pasley, B. N., Inglis, B. A., & Freeman, R. D. (2007). Analysis of oxygen metabolism implies a neural origin for the negative BOLD response in human visual cortex. *Neuroimage*, *36*(2), 269-276.

Paulesu, E., Démonet, J. F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., ... & Frith, U. (2001). Dyslexia: cultural diversity and biological unity.*Science*, *291*, 2165-2167.

Perea, M., Rosa, E., & Gómez, C. (2002). Is the go/no-go lexical decision task an alternative to the yes/no lexical decision task?. *Memory & Cognition*, *30*(1), 34-45.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological review*, *103*(1), 56.

Poeppel, D. (2001). New approaches to the neural basis of speech sound processing: introduction to special section on brain and speech. *Cognitive Science*, *25*(5), 659-661.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, *10*(1), 15-35.

Poldrack, R., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., & Gabrieli, J. (2001). Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. *Cognitive Neuroscience, Journal of*, *13*(5), 687-697.

*Polk, T. A., & Farah, M. J. (2002). Functional MRI evidence for an abstract, not perceptual, word-form area. *Journal of Experimental Psychology: General*,*131*(1), 65.

Price, C. J., Moore, C. J., Humphreys, G. W., Frackowiak, R. S. J., & Friston, K. J. (1996). The neural regions sustaining object recognition and naming.*Proceedings of the Royal Society of London. Series B: Biological Sciences*,*263*(1376), 1501-1507.

Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, *62*(2), 816-847.

*Purcell, J. J., Napoliello, E. M., & Eden, G. F. (2011). A combined fMRI study of typed spelling and reading. *Neuroimage*, *55*, 750-762.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *The Journal of Neuroscience*, *16*(16), 5205-5215.

Rapcsak, S. Z., & Beeson, P. M. (2004). The role of left posterior inferior temporal cortex in spelling. *Neurology*, *62*(12), 2221-2229.

Reichle, E. D. (2015). Computational models of reading: A primer. *Language and Linguistics Compass*, *9*(7), 271-284.

Riecker, A., Ackermann, H., Wildgruber, D., Meyer, J., Dogil, G., Haider, H., Grodd, W., 2000. Articulatory/phonetic sequencing at the level of the anterior perisylvian cortex: a functional magnetic resonance imaging (fMRI) study. Brain Lang. 75, 259–276

*Rossell, S. L., Bullmore, E. T., Williams, S. C., & David, A. S. (2001). Brain activation during automatic and controlled processing of semantic relations: a priming experiment using lexical-decision. *Neuropsychologia*, *39*(11), 1167-1176.

*Rossell, S. L., Price, C. J., & Nobre, A. C. (2003). The anatomy and time course of semantic priming investigated by fMRI and ERPs. *Neuropsychologia*,*41*(5), 550-564.

Rosso, C., Valabregue, R., Arbizu, C., Ferrieux, S., Vargas, P., Humbert, F., ... & Lehéricy, S. (2014). Connectivity between Right Inferior Frontal Gyrus and Supplementary Motor Area Predicts After-Effects of Right Frontal Cathodal tDCS on Picture Naming Speed. *Brain stimulation*, *7*(1), 122-129.

Rubia, K., Smith, A. B., Brammer, M. J., & Taylor, E. (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. Neuroimage, 20(1), 351-358.

*Seghier, M. L., Lee, H. L., Schofield, T., Ellis, C. L., & Price, C. J. (2008). Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. *Neuroimage*, *42*(3), 1226-1236.

Seghier, M. L., Fagan, E., & Price, C. J. (2010). Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. *The Journal of Neuroscience*, *30*(50), 16809-16817.

Seghier, M. L. (2013). The angular gyrus multiple functions and multiple subdivisions. *The Neuroscientist*, *19*(1), 43-61.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: the perils of overreliance on an" outlier" orthography. *Psychological bulletin*, *134*(4), 584.

Simon, O., Mangin, J. F., Cohen, L., Le Bihan, D., & Dehaene, S. (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron*, *33*(3), 475-487.

Takayama, Y., Sugishita, M., Kido, T., Ogawa, M., & Akiguchi, I. (1993). A case of foreign accent syndrome without aphasia caused by a lesion of the left precentral gyrus. *Neurology*, *43*(7), 1361-1361.

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage*, *16*(3), 765-780.

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Human brain mapping*, *33*(1), 1-13.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., ... & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage*, *30*(4), 1414-1432.

Vigneau, M., Beaucousin, V., Hervé, P. Y., Jobard, G., Petit, L., Crivello, F., ... & Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. *Neuroimage*, *54*(1), 577-593.

Vingerhoets, G., Van Borsel, J., Tesink, C., van den, N.M., Deblaere, K., Seurinck, R., Vandemaele, P., Achten, E., 2003. Multilingualism: an fMRI study. Neuroimage 20, 2181–2196.

Vogel, A. C., Church, J. A., Power, J. D., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2013). Functional network architecture of reading-related regions across development. *Brain and language*, *125*(2), 231-243.

Visser, M., & Ralph, M. L. (2011). Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *Journal of Cognitive Neuroscience*, *23*(10), 3121-3131.

Wang, X., Yang, J., Shu, H., & Zevin, J. D. (2011). Left fusiform BOLD responses are inversely related to word-likeness in a one-back task.*Neuroimage*, *55*(3), 1346-1356.

Witt, S.T., Laird, A.R., & Meyerand, E. (2008) Functional neuroimaging correlates of finger-tapping task variations: an ALE meta-analysis. *Neuroimage* 42, 343-356.

*Woodhead, Z. V. J., Wise, R. J. S., Sereno, M., & Leech, R. (2011). Dissociation of sensitivity to spatial frequency in word and face preferential areas of the fusiform gyrus. *Cerebral Cortex*, bhr008.

Yetkin, F. Z., Hammeke, T. A., Swanson, S. J., Morris, G. L., Mueller, W. M., McAuliffe, T. L., & Haughton, V. M. (1995). A comparison of functional MR activation patterns during silent and audible language tasks. *American Journal of Neuroradiology*, *16*(5), 1087-1092.

Yousry, T. A., Schmid, U. D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., & Winkler, P. (1997). Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. *Brain: a journal of neurology*, *120*(1), 141-157.