



Research Report

Functional neuroanatomy of lexical access in contextually and visually guided spoken word production



Natascha Marie Roos ^{a,*}, Atsuko Takashima ^b and Vitória Piai ^{b,c}

^a Radboud University, Donders Centre for Cognition, Nijmegen, the Netherlands

^b Radboud University, Donders Institute for Brain, Cognition, and Behavior, Nijmegen, the Netherlands

^c Radboudumc, Donders Institute for Brain, Cognition, and Behavior, Department of Medical Psychology, Nijmegen, the Netherlands

ARTICLE INFO

Article history:

Received 28 April 2022

Reviewed 15 July 2022

Revised 9 September 2022

Accepted 21 October 2022

Action editor Cynthia Thompson

Published online 19 December 2022

Keywords:

Referential/inferential naming

Taxonomic/thematic semantics

Prediction

Visual word form area

Conceptual access

ABSTRACT

Lexical access is commonly studied using bare picture naming, which is visually guided, but in real-life conversation, lexical access is more commonly contextually guided. In this fMRI study, we examined the underlying functional neuroanatomy of contextually and visually guided lexical access, and its consistency across sessions. We employed a context-driven picture naming task with fifteen healthy speakers reading incomplete sentences (word-by-word) and subsequently naming the picture depicting the final word. Sentences provided either a constrained or unconstrained lead-in setting for the picture to be named, thereby approximating lexical access in natural language use. The picture name could be planned either through sentence context (constrained) or picture appearance (unconstrained). This procedure was repeated in an equivalent second session two to four weeks later with the same sample to test for test-retest consistency. Picture naming times showed a strong context effect, confirming that constrained sentences speed up production of the final word depicted as an image. fMRI results showed that the areas common to contextually and visually guided lexical access were left fusiform and left inferior frontal gyrus (both consistently active across-sessions), and middle temporal gyrus. However, non-overlapping patterns were also found, notably in the left temporal and parietal cortices, suggesting a different neural circuit for contextually versus visually guided lexical access.

© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

* Corresponding author. Postbus 9104, 6500 HE Nijmegen, the Netherlands.

E-mail addresses: natascha.roos@donders.ru.nl (N.M. Roos), atsuko.takashima@donders.ru.nl (A. Takashima), vitoria.piai@donders.ru.nl (V. Piai).

<https://doi.org/10.1016/j.cortex.2022.10.014>

0010-9452/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One crucial thing we do while speaking is retrieving information from memory. The access to the mental lexicon, where we store information about words, is commonly studied using picture naming (Levelt et al., 1991). However, simple object naming is not a big part of our daily language use in conversation. Yet it remains difficult to study lexical access in spontaneous speech, and especially its underlying neural mechanisms. This issue is at the core of the present study.

So far, spoken word production has been studied extensively using bare picture naming paradigms, in combination with electrophysiological as well as hemodynamic imaging methods. A commonly employed method is functional magnetic resonance imaging (fMRI), which measures neuronal activity based on the blood oxygen level dependent (BOLD) signal. Most of such picture naming studies using fMRI have found activity in bilateral language networks including the occipitotemporal and parietal cortex, as well as the left inferior frontal and dorsal premotor areas (Liljeström et al., 2008; Price et al., 2005). Picture naming represents word production based on a concept and therefore always starts with perceiving and recognizing the depicted object. Next, the speaker selects a concept for the object and accesses the associated lexical representation and then phonological information of the target word (Levelt et al., 1999). The retrieval of the conceptual, lexical, and phonological information from the mental lexicon is what we consider ‘word retrieval’ for the purpose of the present study. However, bare picture naming does not closely resemble how we access word information in conversation, where sentence context induces word retrieval. It has been argued that understanding a sentence context via reading can be considered similar to saying the sentence by oneself (Griffin & Bock, 1998). This further allows one to study the access and retrieval of words from the lexicon based on the context they appear in (Federmeier, 2007).

Trying to stay closer to natural language production in the present fMRI study, we use a picture naming paradigm based on sentence context: pictures appear in place of the last word of a lead-in sentence. As fMRI is dependent on the BOLD signal, which has a relatively low temporal resolution in the range of several seconds (Veldsman et al., 2015), dynamic changes associated with language processes could be challenging to capture. But with this paradigm, we can dissociate the timing of the word retrieval process. The context of the sentence is either uninformative and unconstrained, or informative and thus constrained towards the last word of the sentence (i.e., the target word/picture). In unconstrained sentences, people can only retrieve the target word depicted by the picture when this appears. Hence, in unconstrained sentences, the identification and selection of the target word is visually guided by the picture. In constrained sentences, however, people can retrieve the target word based on the information given in the sentence context (e.g., Griffin & Bock, 1998). Thus, processes of word retrieval can happen before picture appearance (Hustá et al., 2021; Piai et al., 2014, 2020). This allows one to study word production in a way that approximates a more naturalistic setting, with the potential to demonstrate the commonalities and differences between

conceptually driven word planning either based on sentence context or on visual cues such as the picture. This was the primary goal of the present study.

Admittedly, this picture-naming task is based on sentence reading of the lead-in sentences (word by word). Previous fMRI studies investigating the effects of contextual constraint during sentence reading have found involvement of the left angular and supramarginal gyri for sentences with high contextual constraint compared to low constraint, and this was interpreted as encoding of semantic sentence processing (Schuster et al., 2021). Another study showed that semantic unification during sentence reading modulates BOLD signal changes in the left and right inferior frontal gyrus, as well as left anterior cingulate cortex and superior and middle temporal gyrus (Zhu et al., 2019). Note, however, that our study is based on a picture naming task, which requires participants to engage in word production processes, that is retrieving conceptual, lexical, and phonological information in order to produce a response.

The employed context-driven picture naming paradigm has been extensively studied with electrophysiology. Together with faster picture naming times following constrained contexts, these studies have shown alpha-beta power decreases before picture appearance, linking the power decreases to processes of conceptual and lexical retrieval (Hustá et al., 2021; Piai et al., 2014, 2015, 2020). In a magnetoencephalography (MEG) study testing the across-session consistency of these patterns, the behavioral facilitation from constrained contexts and the clusters of alpha-beta power decreased in the left temporal and inferior parietal lobule replicated over sessions within the same individuals (Roos & Piai, 2020). Knowing how reliable and valid fMRI maps cognitive processes for a specific paradigm is especially crucial for clinical purposes, as it is still the most common method used for preoperative mapping in neurosurgical populations. This has previously been investigated for fMRI, using a range of different language paradigms, with the conclusion that the balance between reliability and validity is best when using sentence completion tasks (Wilson et al., 2017). The fact that the spatial resolution of MEG is lower than that of fMRI motivated the realization of a two-session fMRI study. Thus, as a secondary goal, we tested the suitability and across-session consistency of context-driven word production with fMRI, targeting time-dynamic word retrieval processes that happen at a timescale of milliseconds.

Based on the findings of our previous studies using this paradigm and stimulus materials (Hustá et al., 2021; Klaus et al., 2020; Piai et al., 2014, 2015, 2017, 2018; Roos & Piai, 2020), we expected shorter naming times for pictures in constrained compared to unconstrained contexts. To examine the activity profiles of contextually and visually guided word retrieval, we included BOLD contrasts for three different moments within a trial (i.e., first word, pre-picture, and picture appearance). We expected no difference between BOLD increases at the beginning of the sentence since sentential constraint has not yet been established at this point. Further, we were interested in the ‘cross-overlap’ of BOLD increases at the (assumed) moment of word retrieval per context type. This concerns the pre-picture interval of constrained sentences for contextually guided word retrieval

(constrained > unconstrained) and the picture interval of unconstrained sentences for visually guided word retrieval (unconstrained > constrained). As these are the moments at which we assume participants to retrieve conceptual, lexical, and phonological information for the target word, both contrasts should yield a profile of brain activity patterns during word retrieval, which may occur at different time points in the sentence for the two conditions. Power decreases of electrophysiological data in the alpha-beta frequency range have been shown to correlate with BOLD signal increases for picture naming (Conner et al., 2011; Liljeström et al., 2015). Based on these findings, we hypothesized BOLD signal increases for constrained relative to unconstrained sentences prior to picture appearance (i.e., contextually guided) in the left temporal cortex more broadly and inferior parietal lobule, which are the areas in which the electrophysiology studies revealed alpha-beta power decreases linked to contextually guided word retrieval. Regarding visually guided word retrieval, we expected to find BOLD increases in the inferior temporal gyrus, as an area associated with picture naming and the access of lexical concepts in word production (Indefrey & Levelt, 2004; Price, 2012; Roelofs, 2014).

2. Methods

The present study falls under the blanket approval for standard studies of the accredited ethical reviewing committee, CMO Arnhem-Nijmegen, following the declaration of Helsinki. We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. No part of the study procedures or analyses was preregistered prior to the research being conducted. Data was collected at the Donders Institute for Cognitive Neuroimaging in Nijmegen in the Netherlands. All raw data and code are available via the Donders Repository (<https://doi.org/10.34973/72sn-vb83>).

2.1. Participants

Fifteen native speakers of Dutch (11 females) participated in the study, ages ranging from 18 to 26 years ($Mdn = 20$). All were MRI compatible, healthy, right-handed subjects with normal or corrected-to-normal vision and hearing and received monetary compensation for participation. We excluded the dataset of one female participant due to no corrected-to-normal vision leading to a large amount of invalid trials, and missing field maps for session 1. Therefore, an additional participant was recruited to enable an analysis of 15 complete datasets. This number was based on sample sizes of previous studies and the known effect size of the employed task with electrophysiology (Hustá et al., 2021; Piai et al., 2014, 2015; Roos & Piai, 2020).

2.2. Materials

The experiment contained 224 nouns which were presented as pictures (targets) on white background, or a full frame in

case of sceneries. Each target word had two corresponding sentences, one constrained and one unconstrained sentence that preceded the picture, yielding a total of 448 stimulus sentences. Linguistic materials (all Dutch) were taken from previous studies (based on Piai et al., 2014, 2015; identical materials to Roos & Piai, 2020). Target pictures were taken from the BOSS database (Brodeur et al., 2010) and via online search. Target word length varied between 2 and 11 phonemes (mean = 5) and sentence length from 4 to 13 words (mean = 7), including the target word.

2.3. Design

The 224 target words were shuffled three times to form three different lists. Each list was split in half, controlled for frequency, word length, and initial letter between the two halves. We also alternated the order between the two halves, resulting in six different main lists. We randomly assigned participants to one of the six main lists and presented the first half in session 1, and the second half in session 2. Each session included 112 target pictures, which always appeared twice per session (once in each condition), yielding 224 trials in total (112 per condition). The items were arranged in pseudo-randomized order, unique per participant, using Mix (van Casteren & Davis, 2006). We constrained items to a maximum of five consecutive trials of the same condition, as well as a minimum distance of 20 trials between the appearance of the constrained and unconstrained sentences of the same target picture. Test and retest sessions were scheduled between 13 and 28 days apart ($Mdn = 20$). For an example of both trials for one target word and the corresponding sentences see Fig. 1. This figure also highlights the sentence parts that were analyzed as well as the presentation durations and jittered intervals during a trial. Trials always started with a fixation cross of 500 ms, followed by the experimental sentence presented word-by-word in the center of the screen. Words were presented for 300 ms in black on a gray background with 200 ms blank intervals in between. The intervals before and after picture presentation were randomly jittered, to capture the BOLD response at different stages. The pre-picture interval thus ranged from 1250 to 3000 ms duration, and the fixation cross between trials from 3000 to 6500 ms. Target pictures were presented for 1000 ms in place of the last word of each sentence.

2.4. Procedure

We instructed participants about the task and scanning session and clarified open questions before they signed the informed consent form. Then, we screened them for MRI compatibility and familiarized them with the pictures of the experiment and the corresponding target words in a slideshow. Picture familiarization is a common procedure in picture naming studies to control for naming variability and increase accuracy, as well as to lower potential repetition effects at the second appearance of the pictures in the experiment. Once in the scanner, the session started with four practice trials and another opportunity to resolve remaining uncertainties. Sentences and target pictures were projected to a screen behind the scanner that can be seen with a mirror

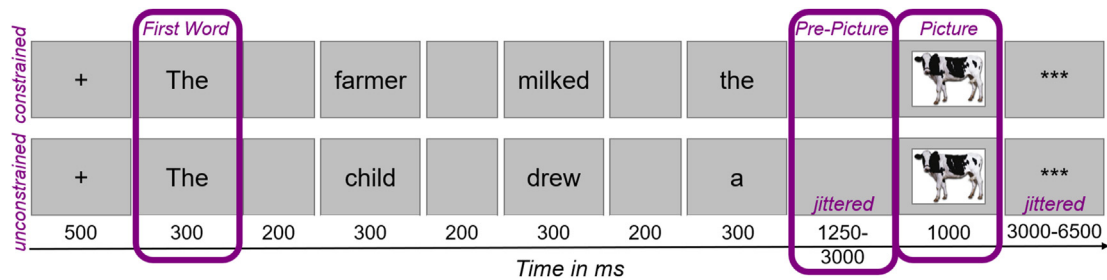


Fig. 1 – Overview and timeline of constrained (top) and unconstrained (bottom) trial events for target word cow. Analyzed sentence parts circled in purple. Boxes represent the screen participants saw at each timepoint. Reproduced with permission from the authors from <https://doi.org/10.17605/OSF.IO/2FVGB>.

mounted on the head coil. Stimuli were presented using Presentation software (Neurobehavioral Systems). Participants had to silently read the words presented on screen and name the pictures overtly with the previously familiarized (target) word for it. They were asked to keep their fixation to the center of the screen and move their jaw and head as little as possible. One scanning session lasted approximately 45 min.

2.5. Response time analyses

We audio recorded participants' responses to monitor their picture naming times and accuracy. Recordings lasted for 2500 ms, starting with the picture appearance. Responses starting before or after the trial recordings were not considered for response time analyses. Further, we excluded hesitations, stutters, and multiple responses from the response time as well as the BOLD analyses. Cases in which participants responded with plausible synonyms of the target word were considered as correct responses for all analyses. Response times were calculated manually using the speech editor Praat (Boersma & Weenink, 2017), with the coder blinded for condition, and statistically analyzed in R (R Core Team, 2017). The response times (in milliseconds) were evaluated using a linear mixed-effects regression (fixed effects of condition, session, and their interaction) with by-participant and by-item random intercepts and slopes for condition. Additionally, as a measure of effect size, we calculated Cohen's *d* for paired samples for each session, using the average variance (J. Cohen, 1988, p. 2).

2.6. MRI acquisition

Participants had to wear metal-free clothing on their upper body and change into scanner clothing if necessary. Then they were taken into the scanner room and positioned on the scanner bed with cushions underneath their knees and elbows. Their forehead was taped to the lower part of the head coil with crepe tape to minimize their head movement (Krause et al., 2019), and an emergency button was placed on their belly. All scans were acquired on a 3T Siemens PrismaFit scanner with a 32-channel head coil using echo-planar imaging (EPI), employing a multiband sequence (multiband acceleration factor 6, 2 mm isotropic voxels, 66 slices, TR = 1000 ms, TE = 34 ms, FoV = 210 × 210 × 132 mm, flip angle = 60°). The experiment started with a pulse countdown

from nine to zero followed by all 224 trials consecutively. Field maps were acquired at the end of the run (TR = 620 ms, TE 1 = 4.7 ms, 64 slices, voxel size 2.4 × 2.4 × 2 mm, FoV = 210 × 210 × 132 mm, flip angle = 60°) to calculate voxel displacement maps (VDM) for each session. Structural T1-weighted MPRAGE images (TR = 2300 ms, TE = 3.03 ms, 192 slices, FoV = 256 × 256 × 192 mm, 1 mm isotropic voxels) for anatomical reference were acquired after session 1.

2.7. fMRI preprocessing

fMRI preprocessing was performed session-individually using MATLAB and SPM12 (SPM12 Software—Statistical Parametric Mapping). The first nine volumes of each session were discarded as dummy scans to allow the magnetization to reach a steady state. All other images were realigned with reference to the 10th volume and unwarped by applying the calculated session-specific voxel displacement map (VDM). T1 image was segmented into different tissue types based on probability maps such as gray matter, white matter, cerebrospinal fluid, bone, and soft tissue and normalization parameter to the template MNI space was estimated. This normalization parameter was applied to the functional images, and a smoothing was applied with a Gaussian kernel of twice the voxel size (FWHM = 4 mm).

2.8. fMRI BOLD analysis

The fMRI analysis was done on the whole-brain level by means of a general linear model (GLM) per participant. The model included eight task-specific regressors per session as well as the six motion parameters, to account for further movement artifacts of participants in the scanner. The three condition-specific regressors were time-locked to the onsets of 1) first word, 2) pre-picture interval, and 3) picture appearance for the two conditions separately. Two extra regressors were modeled: one for the onset of the first word and another for the onset of the pre-picture interval, both for the incorrect trials (*M* = 3, *SD* = 4 per session). Excluding errors, each session consisted of an average of 110 trials per condition (*SD* = 2). The onsets were modeled as stick functions (duration = 0) and convolved with the canonical hemodynamic response function. A high-pass filter of 128 s was implemented to remove slow signal drifts. The two sessions were modeled as one GLM. Contrast images of interest were

computed for each participant separately per session, as well as averaged over both sessions. Furthermore, difference contrasts between constrained and unconstrained contexts were computed for pre-picture interval and picture appearance, as well as their interaction averaged over sessions. These should yield the expected ‘cross-overlap’, i.e., the profile of brain activity during word retrieval per context, and the interaction between them. To explore the consistency of this cross-overlap across sessions, we compared the respective contrasts (i.e., the differences between contexts at both assumed moments of word retrieval) of both sessions. All participant-specific contrast images were tested using One-sample T-tests on the group level. To investigate across-session consistency pooled over the two conditions, we used a full-factorial model in SPM12. For all comparisons, we used a threshold of $p = .001$ uncorrected on the voxel-level, and then used a cluster-size statistics thresholded at $p = .05$ Family-Wise Error (FWE) corrected (Hayasaka & Nichols, 2003). All tables list significant FWE-corrected clusters with the corresponding brain areas. All reported anatomical labels are based on the MNI space coordinates obtained with SPM12 from the Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002), using the MATLAB function `cuixufindstructure.m` (Cui, 2012).

3. Results

3.1. Response times

For both sessions, all participants were faster in naming the picture if the sentence context was constrained, compared to unconstrained, with an across-session response time effect of 202 ms ($Mdn = 213$, $SD = 60$). This is also displayed in Fig. 2, which shows the mean response times of all participants per condition and session, where each participant's data is connected by a line across conditions and color-coded across sessions. In session 1, the mean response time for unconstrained sentences was 896 ms ($Mdn = 860$, $SD = 225$) and for constrained sentences 694 ms ($Mdn = 600$, $SD = 269$). In session 2, this was 926 ms ($Mdn = 898$, $SD = 230$) for unconstrained sentences and 723 ms ($Mdn = 666$, $SD = 269$) for constrained sentences. This yielded a mean effect of response time differences between constrained and unconstrained sentences of 201 ms in session 1 ($Mdn = 220$, $SD = 59$) and 203 ms in session 2 ($Mdn = 186$, $SD = 64$) (Fig. 2). The results of the linear mixed-effects regression analysis revealed a significant effect of condition, which confirms that participants were faster in naming the picture when the sentence context was constrained (Estimate = $-.20$, $SE = .02$, $t = -13.26$, $P < .001$). The regression also yielded a significant effect of session (Estimate = $.03$, $SE = .01$, $t = 4.21$, $P < .001$), with participants being slower overall in session 2. The context effect across sessions yielded a Pearson's correlation coefficient of $r = .47$. Session specific linear mixed-effects regression also yielded a strong effect of condition for session 1 (Estimate = $-.20$, $SE = .02$, $t = -12.79$, $P < .001$, Cohen's $d = .81$) and session 2 (Estimate = $-.20$, $SE = .02$, $t = -11.84$, $P < .001$, Cohen's $d = .81$). The overall mean error frequency was 1.5% ($SD = 1.8$), also per session individually, ranging from 0 to 7% per participant in

total. Per condition this was 1.8% following constrained contexts and 1.2% after unconstrained contexts.

3.2. Whole-brain BOLD fMRI

3.2.1. Session average: trial events over rest

We first looked at the session-average of BOLD increases at the three trial events (i.e., first word, pre-picture interval, picture appearance) per condition over rest. These results are shown per trial event in the two upper rows of Fig. 3 (1st row: constrained context, 2nd row: unconstrained context). As expected, no significant differences were found between the two conditions at the beginning of the sentence (‘First Word’, left column in Fig. 3). Here, we see BOLD increases over rest mostly in the visual network. For a list of significant clusters of BOLD increases at the first word see Supplementary Table 1.1.

At the pre-picture interval and picture appearance (middle and right columns in Fig. 3), we obtained distinct results per context condition over rest. At the pre-picture interval, areas showing BOLD increases over rest that were common to both contexts were, amongst others, bilateral middle temporal gyrus and the orbital part of the inferior frontal gyrus, as well as the left middle temporal pole. In constrained contexts, additional clusters of BOLD increases over rest appeared in the left superior parietal lobule, the triangular part of the left inferior frontal gyrus, the left precentral gyrus, amongst others (cf., Supplementary Table 1.2.1). At the pre-picture interval in unconstrained contexts, we also found clusters of BOLD increases over rest in the left inferior and middle occipital gyri and the left angular gyrus (cf., Supplementary Table 1.2.2). For a complete list of clusters for the pre-picture interval see Supplementary Table 1.2.

After picture appearance, common areas of BOLD increases over rest for both conditions were the bilateral post-central gyrus and the right fusiform gyrus, but we again saw different areas exhibiting BOLD increases per context. For constrained contexts, the results yielded additional clusters in the opercular part of the left inferior frontal gyrus, left caudate and left occipital areas, amongst others (cf., Supplementary Table 1.3.1). After picture appearance in unconstrained contexts, additional clusters appeared in the left middle cingulate gyrus and parts of the left cerebellum (cf., Supplementary Table 1.3.2). For a list of all significant clusters after picture appearance see Supplementary Table 1.3.

3.2.2. Session average: differences between contexts

We expected word retrieval to happen during the pre-picture interval in constrained contexts, and after picture appearance in unconstrained contexts. To investigate this expected ‘cross-overlap’ of activity profiles for word retrieval per context type, we looked at the interaction between context condition and trial event (i.e., pre-picture vs picture appearance). The results of this interaction are shown in Table 1 and Fig. 4 (top right). Following the significant interaction, we looked at the BOLD increases for constrained over unconstrained contexts at the pre-picture interval (i.e., contextually guided word retrieval), and for unconstrained over constrained contexts at picture appearance (i.e., visually guided word retrieval). These results are shown in the top row, left and middle columns of Fig. 4 and reported in Tables 2.1 and 2.2 respectively.

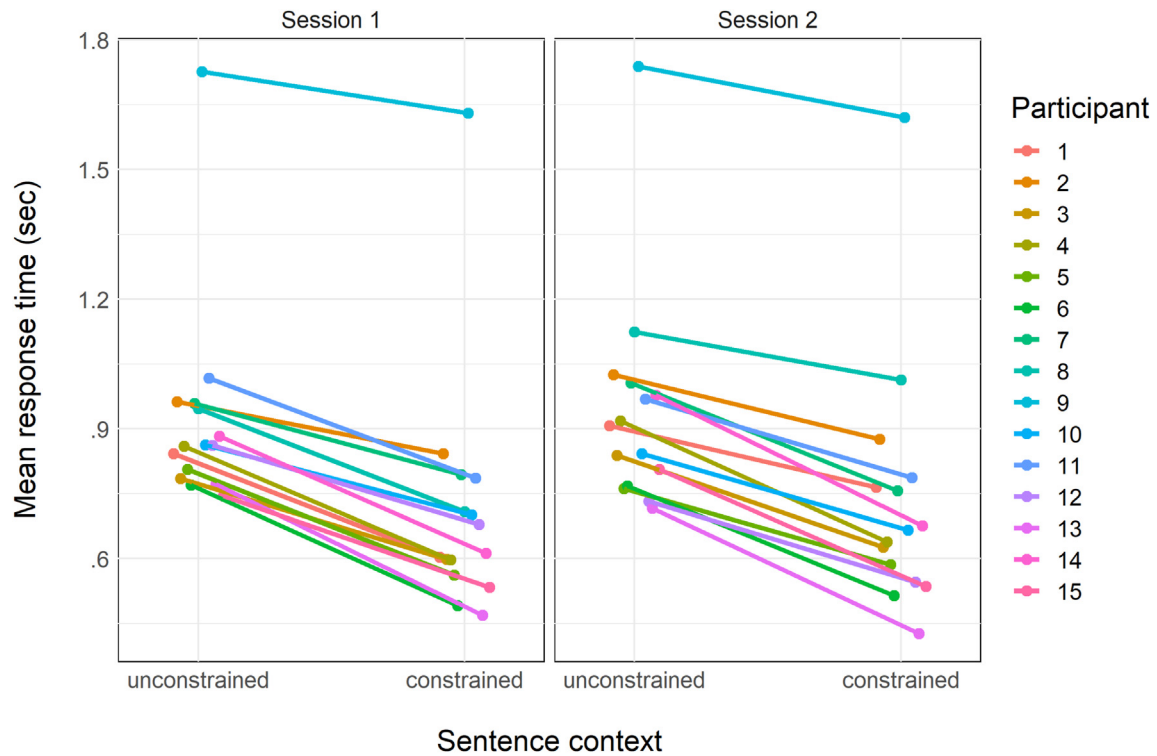


Fig. 2 – Mean picture naming times for each participant (color) per condition and session. Lines connect the mean response times per context for each participant. Reproduced with permission from the authors from <https://doi.org/10.17605/OSF.IO/2FVGB>.

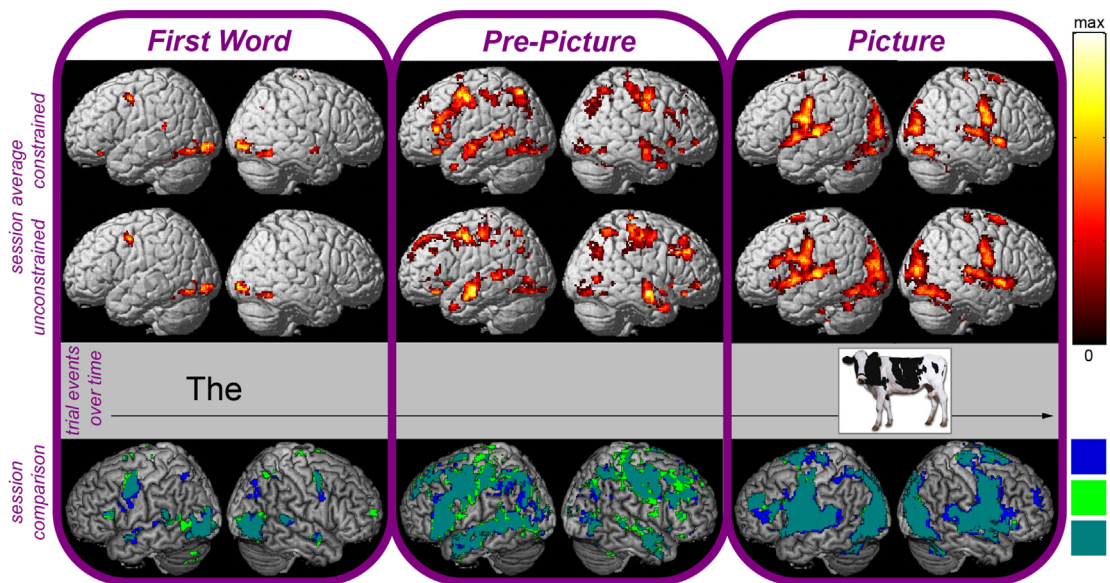


Fig. 3 – Top and middle row: T-contrasts of BOLD increases per trial event, averaged across sessions. Left column = first word, middle column = pre-picture interval, right column = picture appearance. Bottom row: session consistency of effects pooled over conditions by means of a full-factorial model. Blue = Session 1, green = Session 2, teal = common to both sessions. All clusters are significant on cluster level, FWE-corrected $P < .05$. For max T-values, see [Supplementary Table 1](https://doi.org/10.17605/OSF.IO/2FVGB). Reproduced with permission from the authors from <https://doi.org/10.17605/OSF.IO/2FVGB>.

Common areas that exhibited BOLD increases during the assumed moment of word retrieval in both contexts (purple, ‘overlap’ in lower right column of Fig. 4) were the mid to

posterior portion of the left middle temporal gyrus, the left fusiform gyrus, and the triangular part of the left inferior frontal gyrus. Additional BOLD increases at the pre-picture

Table 1 – Session Average. Interaction between Context and Trial Event: Pre-Picture interval constrained > unconstrained & Picture onset unconstrained > constrained.

| Cluster | | Peak Voxel | (MNI Coordinates) | | | Anatomical Location (AAL) |
|-----------------|------|----------------|-------------------|----------|----------|---|
| <i>p</i> (FWEc) | Size | <i>T</i> value | <i>x</i> | <i>y</i> | <i>z</i> | |
| .000 | 220 | 11.2* | 30 | 26 | –2 | Right Insula |
| .000 | 1608 | 9.01 | –50 | 16 | 0 | Left Inferior Frontal Gyrus (triangularis) |
| .000 | 612 | 8.45 | –50 | –60 | –14 | Left Inferior Occipital Gyrus |
| | | 6.89 | –32 | –40 | –24 | Left Fusiform Gyrus |
| .000 | 937 | 8.25 | 0 | 0 | 68 | Left Supplementary Motor Area |
| .000 | 481 | 7.51 | –50 | –56 | 12 | Left Middle Temporal Gyrus (posterior) |
| | | 6.72 | –66 | –30 | 6 | Left Middle Temporal Gyrus (mid) |
| | | 6.32 | –54 | –24 | 0 | Left Middle Temporal Gyrus (mid) |
| .000 | 124 | 7.14 | 36 | –40 | –16 | Right Fusiform Gyrus |
| .004 | 69 | 6.67 | 38 | 12 | 26 | Right Inferior Frontal Gyrus (opercularis) |
| | | 5.02 | 42 | 26 | 22 | Right Inferior Frontal Gyrus (triangularis) |
| .017 | 54 | 6.63 | 16 | –28 | –10 | Right Lingual Gyrus |
| | | 5.18 | 20 | –22 | –16 | Right Parahippocampal Gyrus |
| .000 | 130 | 6.52 | –6 | –8 | 8 | Left Thalamus |
| | | 5.16 | 4 | –14 | 8 | Right Thalamus |
| .000 | 186 | 6.18 | 44 | –12 | 32 | Right Postcentral Gyrus |
| | | 4.64 | 54 | –4 | 36 | Right Precentral Gyrus |
| .000 | 103 | 5.96 | –58 | 0 | 24 | Left Precentral Gyrus |
| | | 4.56 | –56 | –10 | 32 | Left Postcentral Gyrus |
| .039 | 46 | 5.79 | 54 | –32 | –2 | Right Middle Temporal Gyrus |
| .01 | 59 | 5.77 | –10 | –34 | –4 | Left Lingual Gyrus |
| | | 4.42 | –20 | –28 | –8 | Left Hippocampus |
| .000 | 96 | 5.67 | –34 | –86 | 24 | Left Middle Occipital Gyrus |
| .000 | 154 | 5.62 | –30 | –66 | 52 | Left Superior Parietal Gyrus |
| | | 5.58 | –30 | –54 | 52 | Left Inferior Parietal Gyrus |
| .000 | 97 | 5.52 | 16 | –84 | –36 | Right Cerebellum |
| .000 | 105 | 5.51 | –38 | 2 | 48 | Left Precentral Gyrus |
| .048 | 44 | 4.94 | 14 | –64 | –28 | Right Cerebellum |
| .001 | 86 | 4.87 | 48 | –66 | –14 | Right Inferior Occipital Gyrus |
| | | 4.62 | 42 | –58 | –12 | Right Inferior Temporal Gyrus |

Note. Peak voxels with an undefined anatomical label are not reported. Asterisk indicates the max *T*-value. FWEc = Family-Wise Error corrected.

interval specific to contextually guided word retrieval (pink, ‘context’ in lower right column of Fig. 4) occurred in the left inferior temporal and occipital gyrus, as well as the left inferior parietal lobule (cf., Table 2.1). Significant clusters for the picture appearance specific to visually guided word retrieval (cyan, ‘visual’ in lower right column of Fig. 4) additionally occurred in the anterior to posterior left middle temporal gyrus, as well as in the right fusiform gyrus (cf., Table 2.2). For completeness we also tested unconstrained versus constrained at the pre-picture interval, and constrained versus unconstrained at picture onset (i.e., the reverse contrasts of the cross-overlap contrasts). For the contrast constrained > unconstrained at picture onset, no significant clusters survived FWE-correction. The results for pre-picture unconstrained > constrained are shown in Table 2.3 and Supplementary Figure 1. Significant clusters of BOLD increases were found in right and left middle frontal gyrus and left anterior cingulate cortex, amongst other areas. Note that there is little to no activation in the left hemisphere perisylvian areas for this contrast. Given that this contrast is of less relevance, it will not be discussed further.

3.2.3. Session consistency: differences between contexts

To investigate the across-session consistency of contextually and visually guided word retrieval, we compared the respective activity profiles of both sessions, meaning the difference in BOLD increases between contexts at both moments of word retrieval. These results are shown in the bottom row of Fig. 4, color coding the session-specific as well as common areas. Here we discuss areas that consistently exhibited BOLD increases across the two sessions for both moments of word retrieval. For all other areas, we refer to Supplementary Table 2. In general, the session-specific results show a tendency of more additional clusters in session 2 than in session 1.

During word retrieval at the pre-picture interval (‘contextually guided’, left column of Fig. 4), consistent clusters of BOLD increases were in the left and right insula, left precentral gyrus, left inferior occipital and temporal lobes, left fusiform gyrus, the opercular part of the left inferior frontal gyrus, and left inferior parietal lobule. Areas that consistently exhibited BOLD increases during word retrieval at the picture appearance (‘visually guided’, middle column of Fig. 4), were the right insula, left middle temporal gyrus, the triangular part

Table 2 – Session average. Cross-overlap of pre-picture interval and picture onset and reverse contrasts.

| Cluster | | Peak Voxel | (MNI Coordinates) | | | Anatomical Location (AAL) |
|---|------|------------|-------------------|-----|-----|--|
| p (FWEc) | Size | T value | x | y | z | |
| 1. Cross-Overlap: constrained > unconstrained at Pre-Picture interval | | | | | | |
| .000 | 1188 | 9.59* | −30 | 24 | 2 | Left Insula |
| | | 9.42 | −50 | 30 | 24 | Left Inferior Frontal Gyrus (triangularis) |
| | | 7.93 | −52 | 16 | 0 | Left Inferior Frontal Gyrus (triangularis) |
| .000 | 129 | 9.56 | 32 | 24 | −4 | Right Insula |
| .000 | 436 | 8.85 | −54 | −2 | 28 | Left Precentral Gyrus |
| | | 6.95 | −48 | −10 | 38 | Left Postcentral Gyrus |
| .000 | 804 | 8.42 | 6 | 4 | 70 | Right Supplementary Motor Area |
| | | 8.25 | −14 | 2 | 72 | Left Superior Frontal Gyrus |
| | | 6.98 | −4 | 6 | 60 | Left Supplementary Motor Area |
| .000 | 364 | 8.21 | −50 | −60 | −14 | Left Inferior Occipital Gyrus |
| | | 7.19 | −44 | −50 | −14 | Left Inferior Temporal Gyrus (posterior) |
| .000 | 339 | 7.34 | 44 | −12 | 32 | Right Postcentral Gyrus |
| | | 6.36 | 50 | −6 | 34 | Right Precentral Gyrus |
| .000 | 93 | 7.09 | −32 | −40 | −24 | Left Fusiform Gyrus |
| .000 | 453 | 6.84 | −32 | −54 | 52 | Left Inferior Parietal Lobule |
| | | 6.05 | −32 | −70 | 50 | Left Inferior Parietal Lobule |
| .000 | 171 | 6.44 | −20 | 6 | 2 | Left Putamen |
| | | 5.68 | −24 | −4 | −6 | Left Pallidum |
| .000 | 99 | 6.11 | −4 | −24 | 10 | Left Thalamus |
| | | 5.26 | 4 | −14 | 8 | Right Thalamus |
| .000 | 96 | 5.35 | 10 | −78 | −36 | Right Cerebellum |
| .016 | 51 | 4.75 | −64 | −42 | 8 | Left Middle Temporal Gyrus (posterior) |
| | | 4.60 | −64 | −30 | 6 | Left Middle Temporal Gyrus (mid) |
| | | 4.21 | −58 | −40 | 2 | Left Middle Temporal Gyrus (posterior) |
| 2. Cross-Overlap: unconstrained > constrained at Picture onset | | | | | | |
| .000 | 258 | 9.31* | 30 | 24 | −2 | Right Insula |
| .000 | 945 | 8.12 | −66 | −30 | 6 | Left Middle Temporal Gyrus (mid) |
| | | 7.93 | −52 | −54 | 14 | Left Middle Temporal Gyrus (posterior) |
| | | 7.33 | −50 | −10 | −16 | Left Middle Temporal Gyrus (mid) |
| .000 | 1294 | 7.91 | −52 | 32 | 16 | Left Inferior Frontal Gyrus (triangularis) |
| | | 7.20 | −42 | 20 | 24 | Left Inferior Frontal Gyrus (triangularis) |
| | | 6.73 | −42 | 12 | 30 | Left Precentral Gyrus |
| .000 | 676 | 7.60 | −50 | −60 | −14 | Left Inferior Occipital Gyrus |
| | | 7.12 | −42 | −52 | −14 | Left Fusiform Gyrus |
| | | 6.63 | −34 | −40 | −26 | Left Fusiform Gyrus |
| .000 | 802 | 6.98 | −10 | 4 | 70 | Left Supplementary Motor Area |
| | | 6.72 | −4 | 20 | 32 | Left Middle Cingulum |
| | | 6.24 | −4 | 18 | 48 | Left Supplementary Motor Area |
| .000 | 658 | 6.97 | 38 | −32 | −20 | Right Fusiform Gyrus |
| | | 6.20 | 48 | −64 | −6 | Right Inferior Temporal Gyrus |
| .004 | 74 | 6.44 | 16 | −84 | −36 | Right Cerebellum |
| .002 | 82 | 6.23 | 2 | −62 | −38 | Right Cerebellum |
| | | 5.95 | 0 | −54 | −40 | Left Cerebellum |
| .000 | 157 | 6.23 | 16 | −28 | −10 | Right Lingual Gyrus |
| | | 6.20 | 20 | −22 | −16 | Right Parahippocampal Gyrus |
| .040 | 49 | 6.19 | 54 | −32 | −2 | Right Middle Temporal Gyrus |
| .000 | 195 | 6.15 | −8 | −56 | 12 | Left Precuneus |
| | | 5.25 | −16 | −62 | 2 | Left Lingual Gyrus |
| | | 4.92 | −18 | −70 | 14 | Left Calcarine Sulcus |
| .001 | 95 | 6.08 | −10 | −34 | −6 | Left Lingual Gyrus |
| | | 5.27 | −8 | −50 | −2 | Left Cerebellum |
| .008 | 66 | 6.01 | −50 | −68 | 22 | Left Middle Temporal Gyrus (posterior) |
| | | 3.96 | −42 | −66 | 16 | Left Middle Temporal Gyrus (posterior) |
| .001 | 95 | 5.94 | −6 | −6 | 6 | Left Thalamus |
| | | 4.84 | −8 | −24 | 10 | Left Thalamus |
| .000 | 199 | 5.72 | −38 | −88 | −2 | Left Middle Occipital Gyrus |
| | | 5.60 | −40 | −86 | 16 | Left Middle Occipital Gyrus |
| .021 | 56 | 5.58 | 16 | −72 | −26 | Right Cerebellum |
| .001 | 89 | 5.35 | 0 | −70 | 24 | Left Cuneus |
| .006 | 70 | 5.27 | 60 | 0 | −12 | Right Superior Temporal Gyrus |
| | | 5.13 | 52 | −8 | −20 | Right Middle Temporal Gyrus |

(continued on next page)

Table 2 – (continued)

| Cluster p (FWEc) | Size | Peak Voxel | (MNI Coordinates) | | | Anatomical Location (AAL) |
|---|------|------------|-------------------|-----|-----|------------------------------|
| | | T value | x | y | z | |
| .001 | 96 | 5.10 | 28 | –70 | 30 | Right Middle Occipital Gyrus |
| 3. Reverse Cross-Overlap: unconstrained > constrained at Pre-Picture interval | | | | | | |
| .000 | 432 | 8.55* | 32 | 26 | 42 | Right Middle Frontal Gyrus |
| | | 5.83 | 20 | 38 | 36 | Right Superior Frontal Gyrus |
| .028 | 46 | 8.22 | –26 | 34 | 46 | Left Middle Frontal Gyrus |
| | | 4.15 | –18 | 34 | 42 | Left Superior Frontal Gyrus |
| .000 | 578 | 8.17 | –10 | 46 | 4 | Left Anterior Cingulum |
| | | 6.5 | 4 | 16 | –6 | Right Olfactory Bulb |
| .015 | 52 | 7.15 | 16 | –68 | 0 | Right Lingual Gyrus |
| .000 | 155 | 6.83 | 2 | –52 | 24 | Right Precuneus |
| .039 | 43 | 5.84 | –24 | –48 | –42 | Left Cerebellum |
| .000 | 125 | 5.71 | 62 | –26 | 38 | Right Supramarginal Gyrus |
| 4. Reverse Cross-Overlap: constrained > unconstrained at Picture onset | | | | | | |
| No significant clusters | | | | | | |

Note. Peak voxels with an undefined anatomical label are not reported. Asterisk indicates the max T-value corresponding to the max T-value in Figs. 3 and 4. FWEc = Family-Wise Error corrected.

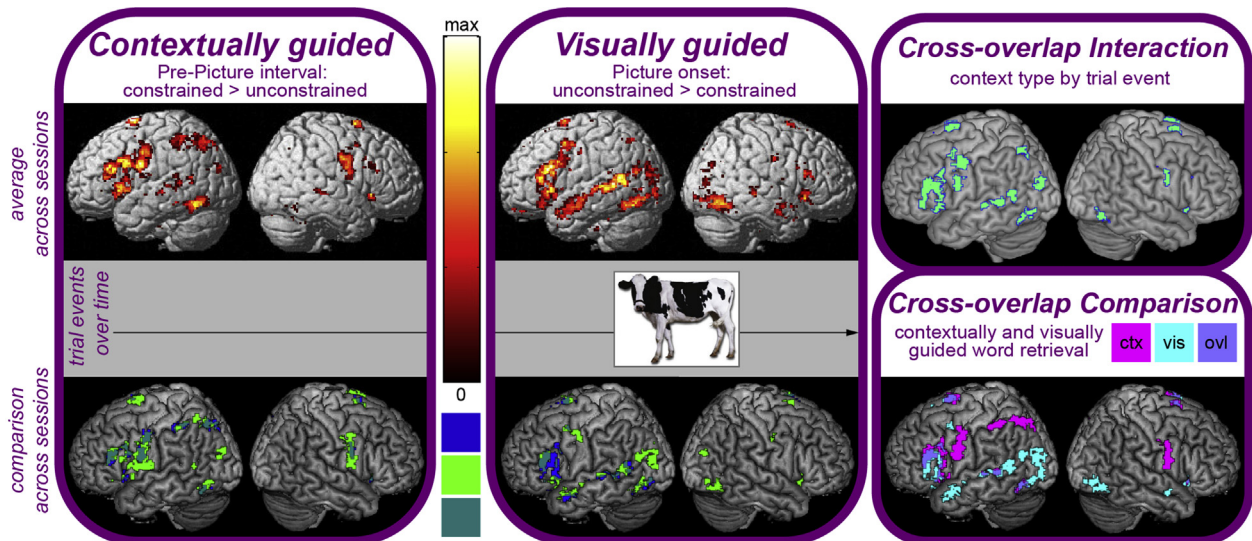


Fig. 4 – Top row: T-contrasts of BOLD increases for constrained over unconstrained contexts at the pre-picture interval (left column), and unconstrained over constrained contexts at the picture appearance (middle column), averaged across sessions. Bottom row: session consistency of differences between contexts (i.e., of contrasts presented in top row). Dark blue = Session 1, green = Session 2, teal = common to both sessions. Right column, top: interaction between context type and trial event (i.e., pre-picture interval vs picture appearance). Right column, bottom: cross-overlap of contextually and visually guided word retrieval. Pink = contextually guided, cyan = visually guided, purple = common to both ways of word retrieval. All clusters are significant on cluster level, FWE-corrected $P < .05$. For max T-values, see Tables 1 and 2. Reproduced with permission from the authors from <https://doi.org/10.17605/OSF.IO/2FVGB>.

of the left inferior frontal gyrus, as well as left and right fusiform gyrus. For completeness, we also show the across session consistency for each of the three trial events pooled over conditions (bottom row of Fig. 3).

4. Discussion

In the present study, we investigated the differences and commonalities in brain areas underlying contextually and visually guided word retrieval. Additionally, we examined the

consistency of this mapping within the same participants across two sessions. In terms of behavioral results, we replicated previous findings of faster picture naming in constrained compared to unconstrained contexts, indicating that people are able to start planning a word based on contextual information in a sentence (e.g., Hústá et al., 2021; Piai et al., 2014; Roos & Piai, 2020), shown in Fig. 2. Our imaging results show that contextually and visually guided word retrieval have distinct and overlapping underlying profiles of brain activity, as can be seen in Figs. 3 and 4. The mapping consistency of BOLD increases across sessions was high when

pooled over conditions (lower row of Fig. 3), but more divergent when looking at the differences between contexts across sessions (lower row of Fig. 4). We first discuss the similarities and differences of the involved brain areas for both types of word retrieval, based on session-averaged results. Then we focus on the session consistency and discuss how far session 1 and 2 yielded overlapping or diverging results in terms of the involved brain areas.

4.1. Brain areas common to both types of word retrieval

One outcome of interest of our study regards the expected ‘cross-overlap’ of activity patterns during word retrieval. For each context type, we assumed that people would retrieve the word at different points in the sentence: After having seen the picture in the unconstrained sentences (i.e., visually guided, middle panel of Fig. 4), and in the constrained sentences during the pre-picture interval, before the picture appeared (i.e., contextually guided, left panel of Fig. 4). We further expected processes of word retrieval to be reflected in BOLD signal increases. We found two different activity profiles for word retrieval depending on the lead-in process, with partial overlap between them (lower right panel of Fig. 4). This overlap is likely to indicate the core process of word retrieval, such as accessing the lexical concept, selecting the correct lemma, and the corresponding phonological form. These steps should not differ depending on whether word retrieval is initiated by sentence context (contextually) or picture appearance (visually; for similar reasoning, see Indefrey & Levelt, 2004). Areas that were common between both ways of word retrieval (‘overlap’, purple clusters in the lower right panel of Fig. 4) were the left fusiform gyrus, including the visual word form area (L. Cohen et al., 2000), the left middle temporal gyrus, and the pars triangularis of left inferior frontal gyrus. The distribution of clusters within the left middle temporal gyrus suggest that the overlap is mostly present in the mid portion (where anterior $y > -7$, mid $-7 \geq y \geq -38$, and posterior $y < -38$, defined in Talairach space, following Indefrey & Levelt, 2004), while the most posterior and anterior extent are found for visually guided word retrieval (see section 4.2.2 below), but not for contextually guided word retrieval. The involvement of the left fusiform, middle temporal, and inferior frontal gyrus at both moments of word retrieval suggests that they are independent of the input modality of the lead-in information (context vs. picture). These areas are thus likely associated with core processes of conceptually driven word production.

The visual word form area is located between the left fusiform and inferior temporal gyrus and has been suggested to be critical for reading, but also for other tasks including visual stimuli such as picture naming (Jobard et al., 2003). Since both contexts included reading as well as the processing of pictorial stimuli, the involvement of the fusiform gyrus at both moments of word retrieval is not surprising. The left fusiform and middle temporal gyri have been associated with conceptual processing, as well as storage and retrieval of semantic knowledge and language use as part of the human semantic network (Binder et al., 2009). In one study, the same areas were identified to disrupt semantic processing when cortically stimulated, independent of modality for visual as

well as auditory naming (Forseth et al., 2018). Activity in the left inferior frontal gyrus has repeatedly been found for tasks involving the retrieval of semantic knowledge (Murtha et al., 1999), as well as semantic processing in object naming and word generation (Binder et al., 2009; Price, 2012; Price et al., 2005). Further, the mid and posterior portion of the left middle temporal gyrus have previously been associated with conceptually driven lexical selection following visual recognition of objects, and phonological code retrieval respectively (Indefrey & Levelt, 2004). This is in line with our present results, as these are all crucial steps for planning a spoken word, in contextually as well as visually guided scenarios.

4.2. Different brain areas for each type of word retrieval

Next to the overlap of brain areas involved in word retrieval depending on the lead-in information, we also found non-overlapping areas (‘context’ in pink and ‘visual’ in cyan, Fig. 4). The main difference between the two ways of word retrieval probably reflects the manner in which the concept is accessed: either through sentence context or picture appearance.

4.2.1. Contextually guided word retrieval

Our results (top row in left panel, Fig. 4) suggest that areas involved in accessing the concept based on sentence context are the left inferior temporal and occipital gyrus, as well as the left inferior parietal lobule. The inferior temporal as well as the occipital gyrus have been repeatedly identified as specific to picture naming in comparison to word generation, suggesting they are not common word production areas, but rather involved in lead-in processes for word retrieval that require specific semantic knowledge (Indefrey & Levelt, 2004). Further, the inferior temporal gyrus has been linked to accessing lexical concepts in word production (Price, 2012; Roelofs, 2014). Finally, the inferior parietal lobule has been shown to be involved in the prediction and integration of semantic knowledge (Binder et al., 2009; Price, 2012). The involvement of the left temporal and inferior parietal lobule in contextually guided word retrieval further corresponds to the results of our previous MEG study, where we found the strongest and most reliable power decreases between constrained and unconstrained sentence contexts in these areas across sessions (Roos & Piai, 2020). Although MEG and fMRI capture different aspects of neuronal activity and the BOLD signal underlies a six to 12 s hemodynamic response curve (Kujala et al., 2014; Liljeström et al., 2009; Vartiainen et al., 2011), it is reassuring to see partial convergence between both methods for the same paradigm, demonstrating the possibility to dissociate processing areas in close temporal proximity using fMRI.

4.2.2. Visually guided word retrieval

We found that accessing the concept based on picture appearance involves more anterior and posterior regions of the left middle temporal gyrus, as well as the right fusiform gyrus (top row in middle panel, Fig. 4). These areas thus seem to be relevant for semantic processes related to visual object recognition. In fact, the anterior temporal lobe has been demonstrated to play a significant role in naming pictures of

faces and houses (Grabowski et al., 2001) and in a study using voxel-based lesion-symptom mapping in aphasia (Schwartz et al., 2009). Further, a review of studies on the anatomy of visual object processing also illustrates the involvement of the anteromedial temporal cortex for processing animate as well as inanimate objects (Bright et al., 2005). As the unconstrained sentence contexts in our study are non-informative with regard to the upcoming picture, they are basically irrelevant for retrieval of the target word depicted by the picture. This makes word retrieval in unconstrained contexts somewhat comparable to bare picture naming tasks.

4.2.3. Organization of the semantic system

The divergence between areas involved in contextually and visually guided word retrieval (Fig. 4) also relates to previously proposed specializations within the semantic system in the brain, such as thematic and taxonomic systems. This regards the relationships between words that are either taxonomically related (apple-pear) or thematically related (dog-leash, similar to constrained contexts in our study). Studies investigating brain organization for taxonomic and thematic relations have found activations in the left anterior temporal lobe, and posterior temporal as well as inferior parietal areas respectively (Mirman et al., 2017; Schwartz et al., 2011). Our findings of the involvement of the left inferior parietal lobule for the constrained condition are in line with this literature.

Another aspect of the semantic system related to our study regards referential and inferential semantic competence (Marconi et al., 2013). Referential naming refers to visual naming of objects, while inferential naming refers to naming to definition. Likewise, theories propose distinct underlying neural processes, while both also rely on the common semantic network of areas that are crucial for accessing semantic knowledge (Binder et al., 2009). Referential naming has been associated with activity in the right fusiform gyrus, which has been linked to visual processing (Marconi et al., 2013). Another study has found activity for referential naming compared to baseline in left middle temporal gyrus and right fusiform gyrus (Farias et al., 2005). This further converges with our results for visually guided word retrieval (middle panel, Fig. 4). The same study has located significant activity for inferential naming compared to baseline to the left inferior temporal and inferior parietal lobule, which overlaps with our results for contextually guided word retrieval (left panel Fig. 4). Thus, visually guided word retrieval is similar to referential naming, and contextually guided word retrieval comparable to inferential naming (although the latter is less similar, as the sentence context in our paradigm plays a slightly different role than in naming to definition, which could again induce differences in the underlying processes of word retrieval).

4.3. Consistency of involved brain areas across sessions

Apart from the brain areas underlying the two different types of word retrieval, which were derived from an average over sessions, we also examined the consistency of this mapping from session 1 to session 2. We approached this by comparing the BOLD increases pooled over both conditions per sentence part (bottom row of Fig. 3). This led to a high consistency across sessions such that clusters of BOLD increases of

sessions 1 and 2 mostly overlapped. However, pooling over conditions is not informative in the scope of the employed paradigm. A more informative across-session consistency regards the activity patterns specific to contextually and visually guided word retrieval. We first focus on the contrasts for both types of word retrieval averaged across sessions (top row of Fig. 4) and to what extent they are based on consistent session-specific contrasts (bottom row of Fig. 4), before we discuss the direct comparison of the specific contrasts of sessions 1 and 2.

4.3.1. Contextually guided word retrieval

For contextually guided word retrieval (left panel, Fig. 4), all but four clusters (the four lowest peak T-values) that were observed in the pre-picture constrained > unconstrained contrast averaged across sessions (Table 2.1) appear in both session-specific contrast (Supplementary Table 2.1 and 2.2) and are thus consistent across sessions. This holds for all left hemisphere areas that we discussed to be specific for contextually guided word retrieval, such as the inferior temporal and occipital gyri and the inferior parietal lobule. However, the results for session 1 do not yield any activity in the left middle temporal gyrus, which we discussed above to be common to both types of word retrieval. Most likely, these clusters did not meet the threshold in session 1, suggesting that the effect size for the left middle temporal gyrus is smaller than for the other regions. Another divergence between the session-specific and session-averaged contrast regards the left inferior frontal gyrus, again suggesting that the effect size in this area is smaller than for other areas.

4.3.2. Visually guided word retrieval

For visually guided word retrieval (middle panel, Fig. 4), the convergence of the session-average and session-specific contrasts was smaller (Table 2.2, and Supplementary Table 2.3 and 2.4 respectively). Here only six clusters with the highest peak values and one additional cluster (out of 20 in total) are common across sessions. This means that, for visually guided word retrieval, less than half of the areas exhibiting BOLD increases are session consistent. Yet, areas that are consistent include the mid and posterior portion of the left middle temporal gyrus, the left inferior frontal gyrus, and bilateral fusiform gyrus. As argued above, the left middle temporal and right fusiform gyrus seem to be at the core of visually guided word retrieval, while the left fusiform and inferior frontal gyrus are common to both types of word retrieval.

4.3.3. Comparison

We see a tendency of more clusters of BOLD increases and a more wide-spread distribution of those for session 2 compared to session 1 (bottom row, Fig. 4). Overall, our results for contextually guided word retrieval seem to be more reliable (replicable across sessions) than those for visually guided word retrieval, as there are more areas consistent across sessions for the former than for the latter. Almost all areas that are session consistent for visually guided word retrieval are, in fact, common to both types of word retrieval. This provides converging evidence on the functional neuroanatomy of lexical access.

4.4. Strengths and limitations

The experimental paradigm we employed has a number of strengths. We provide a setting for picture naming in which each picture appears in the context of a plausible sentence. The moment and way of word retrieval differs between experimental conditions, which is reflected in varying sentence contexts. In other words, this paradigm yields a well-controlled setting for word retrieval based on sentence context or picture appearance within the same modality, while it prevents capturing any activity due to differences between conditions that are not of interest. Another strength of our study is that we tested the same group of people in two sessions to see test-retest reliability.

In terms of validity for language mapping in combination with fMRI, this paradigm seems promising. In a previous fMRI study, sentence completion yielded the highest validity and across-session consistency, while picture naming without a proper control condition did not result in valid left-lateralized activity maps (Wilson et al., 2017). The same study also investigated different analysis parameters and concluded that a priori selected regions of interest improve results for fMRI but are yet a risky approach for research and especially for clinical contexts. In that sense, the paradigm we employed offers a good control condition, yielding left lateralized activity patterns for picture naming, simply by conducting whole brain analyses with standard parameters.

These strengths, however, also come with limitations. The event-related design likely reduces experimental power for this paradigm in combination with fMRI. An adaptation of the current or similar paradigms to tackle this limitation and increase power would be a blocked design (Wilson et al., 2017). As the BOLD response also reflects long-lasting processes, it is quite sensitive to variations from trial to trial (Liljeström et al., 2009). Especially with a sensitive contrast between conditions that only differs in the context of the sentence preceding the picture, an event-related design might be too subtle to capture differences between conditions within one session. Further, we have argued that word retrieval in unconstrained sentences is similar to bare picture naming. Yet, by embedding visually guided word retrieval in the context of a plausible albeit unconstrained sentence, we steer away from settings of bare picture naming, which makes our results less comparable to previous findings in the literature of such studies in the field.

Regarding our sample size, we admit that a sample of 15 participants generally differs from sample sizes used for other fMRI studies, and thereby constitutes a limitation of the current study. However, we included two sessions per participant and mostly looked at data averaged over both sessions. In addition, we also show the test-retest consistency for contextually and visually guided naming, which should ultimately reinforce the confidence in our findings. Further, the small sample size might have prevented detecting effects of smaller magnitude, which we consider to be of smaller relevance. In this study we limited our target words to only nouns. For future research of lexical access, however, it would be interesting to add to the current findings and also include picture naming of action verbs. Another methodological

limitation regards thresholding in fMRI. We only conducted whole-brain analyses which were not restricted to a priori selected regions of interest. Thus, our results yielded activity in a vast amount of brain regions but taking a closer look at the session consistency for the respective contrasts shows us that the average contrasts are not always consistent with the session-specific ones (although many areas that are session consistent for visually guided retrieval are also engaged in conceptually driven word retrieval). More targeted analyses using pre-defined regions of interest would likely result in higher across-session consistency.

5. Conclusion

We investigated the functional neuroanatomy of visually and contextually guided lexical access and the consistency of this mapping across sessions. While contextually guided retrieval of words specifically involves the left inferior temporal and occipital gyrus, as well as the left inferior parietal lobule, visually guided retrieval recruits the anterior to posterior left middle temporal gyrus, as well as the right fusiform gyrus. Regardless of its lead-in processes, lexical access consistently involved the mid to posterior portion of the left middle temporal gyrus, the left fusiform gyrus, and the triangular part of the left inferior frontal gyrus. These findings inform our theories about the neurobiology of lexical access in word production. Combining the BOLD increases of both conditions yielded consistent mappings for session 1 and 2. However, the across-session consistency of BOLD signal differences between conditions was lower. This suggests that more power is needed to detect reliable and meaningful condition-specific activity patterns that are consistent between multiple fMRI sessions.

Author contributions

Natascha M Roos: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration.

Atsuko Takashima: Methodology, Software, Writing - Review & Editing, Supervision.

Vitória Piai: Conceptualization, Methodology, Writing - Review & Editing, Supervision.

Open practices

The study in this article earned Open Data and Open Materials badges for transparent practices. Materials and data for the study are available at <https://osf.io/2fvgb/>.

Acknowledgements

We would like to thank the organizers and attendees of the Functional MRI Analysis Meeting at the Donders Institute for

helpful feedback and discussion, as well as Laura Giglio for much appreciated support throughout the process. This work was supported by Gravitation Grant 024.001.006 from the Dutch Research Council (NWO) to the Language in Interaction Consortium.

Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cortex.2022.10.014>.

REFERENCES

- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767–2796. <https://doi.org/10.1093/cercor/bhp055>
- Boersma, P., & Weenink, D. (2017). *Praat, software for speech analysis and synthesis*. University of Amsterdam.
- Bright, P., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2005). The anatomy of object processing: The role of anteromedial temporal cortex. *The Quarterly Journal of Experimental Psychology Section B*, 58(3–4b), 361–377. <https://doi.org/10.1080/02724990544000013>
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *Plos One*, 5(5). <https://doi.org/10.1371/journal.pone.0010773>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: L. Lawrence Earlbaum Associates, 10.1016.B978-0-12-179060-8.50006-2.
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain: A Journal of Neurology*, 123(2), 291–307. <https://doi.org/10.1093/brain/123.2.291>
- Conner, C. R., Ellmore, T. M., Pieters, T. A., DiSano, M. A., & Tandon, N. (2011). Variability of the relationship between electrophysiology and BOLD-fMRI across cortical regions in humans. *Journal of Neuroscience*, 31(36), 12855–12865. <https://doi.org/10.1523/JNEUROSCI.1457-11.2011>
- Core Team, R. (2017). *R: A language and environment for statistical computing*.
- Cui, X. (2012). Find anatomical location of a list of MNI coordinates – Xu Cui while (alive)[learn;] <https://www.alivelearn.net/?p=1456>.
- Farias, S. T., Harrington, G., Broomand, C., & Seyal, M. (2005). Differences in functional MR imaging activation patterns associated with confrontation naming and responsive naming. *AJNR. American Journal of Neuroradiology*, 26(10), 2492–2499. <http://www.ajnr.org/content/26/10/2492>.
- Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505. <https://doi.org/10.1111/j.1469-8986.2007.00531.x>
- Forseth, K. J., Kadipasaoglu, C. M., Conner, C. R., Hickok, G., Knight, R. T., & Tandon, N. (2018). A lexical semantic hub for heteromodal naming in middle fusiform gyrus. *Brain: A Journal of Neurology*, 141(7), 2112–2126. <https://doi.org/10.1093/brain/awy120>
- Grabowski, T. J., Damasio, H., Tranel, D., Ponto, L. L. B., Hichwa, R. D., & Damasio, A. R. (2001). A role for left temporal pole in the retrieval of words for unique entities. *Human Brain Mapping*, 13(4), 199–212. <https://doi.org/10.1002/hbm.1033>
- Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, 38(3), 313–338. <https://doi.org/10.1006/jmla.1997.2547>
- Hayasaka, S., & Nichols, T. E. (2003). Validating cluster size inference: Random field and permutation methods. *Neuroimage*, 20(4), 2343–2356. <https://doi.org/10.1016/j.neuroimage.2003.08.003>
- Hustá, C., Zheng, X., Papoutsis, C., & Piai, V. (2021). Electrophysiological signatures of conceptual and lexical retrieval from semantic memory. *Neuropsychologia*, 161, Article 107988. <https://doi.org/10.1016/j.neuropsychologia.2021.107988>
- Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92(1), 101–144. <https://doi.org/10.1016/j.cognition.2002.06.001>
- Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A meta-analysis of 35 neuroimaging studies. *Neuroimage*, 20(2), 693–712. [https://doi.org/10.1016/S1053-8119\(03\)00343-4](https://doi.org/10.1016/S1053-8119(03)00343-4)
- Klaus, J., Schutter, D. J. L. G., & Piai, V. (2020). Transient perturbation of the left temporal cortex evokes plasticity-related reconfiguration of the lexical network. *Human Brain Mapping*, 41(4), 1061–1071. <https://doi.org/10.1002/hbm.24860>
- Krause, F., Benjamins, C., Eck, J., Lühns, M., Hoof, R., & Goebel, R. (2019). Active head motion reduction in magnetic resonance imaging using tactile feedback. *Human Brain Mapping*, 40(14), 4026–4037. <https://doi.org/10.1002/hbm.24683>
- Kujala, J., Sudre, G., Vartiainen, J., Liljeström, M., Mitchell, T., & Salmelin, R. (2014). Multivariate analysis of correlation between electrophysiological and hemodynamic responses during cognitive processing. *Neuroimage*, 92, 207–216. <https://doi.org/10.1016/j.neuroimage.2014.01.057>
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences*, 22(1), 1–38. <https://doi.org/10.1017/S0140525X99001776>
- Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, 98(1), 122–142. <https://doi.org/10.1037/0033-295X.98.1.122>
- Liljeström, M., Hultén, A., Parkkonen, L., & Salmelin, R. (2009). Comparing MEG and fMRI views to naming actions and objects. *Human Brain Mapping*, 30(6), 1845–1856. <https://doi.org/10.1002/hbm.20785>
- Liljeström, M., Stevenson, C., Kujala, J., & Salmelin, R. (2015). Task- and stimulus-related cortical networks in language production: Exploring similarity of MEG- and fMRI-derived functional connectivity. *Neuroimage*, 120, 75–87. <https://doi.org/10.1016/j.neuroimage.2015.07.017>
- Liljeström, M., Tarkiainen, A., Parviainen, T., Kujala, J., Numminen, J., Hiltunen, J., Laine, M., & Salmelin, R. (2008). Perceiving and naming actions and objects. *Neuroimage*, 41(3), 1132–1141. <https://doi.org/10.1016/j.neuroimage.2008.03.016>
- Marconi, D., Manenti, R., Catricalà, E., Della Rosa, P. A., Siri, S., & Cappa, S. F. (2013). The neural substrates of inferential and referential semantic processing. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 49(8), 2055–2066. <https://doi.org/10.1016/j.cortex.2012.08.001>
- Mirman, D., Landrigan, J.-F., & Britt, A. E. (2017). Taxonomic and thematic semantic systems. *Psychological Bulletin*, 143(5), 499–520. <https://doi.org/10.1037/bul10000092>
- Murtha, S., Chertkow, H., Beauregard, M., & Evans, A. (1999). The neural substrate of picture naming. *Journal of Cognitive*

- Neuroscience, 11(4), 399–423. <https://doi.org/10.1162/089892999563508>
- Zhu, Z., Bastiaansen, M., Hakun, J. G., Petersson, K., Wang, S., & Hagoort, P. (2019). Semantic unification modulates N400 and BOLD signal change in the brain: A simultaneous EEG–fMRI study. *Journal of Neurolinguistics*, 52(100855). <https://doi.org/10.1016/j.jneuroling.2019.100855>.
- Neurobehavioral Systems. (n.d.). Retrieved March 27, 2020, from https://www.neurobs.com/menu_support/menu_archives/citation_guide.
- Piai, V., Klaus, J., & Rossetto, E. (2020). The lexical nature of alpha-beta oscillations in context-driven word production. *Journal of Neurolinguistics*, 55, Article 100905. <https://doi.org/10.1016/j.jneuroling.2020.100905>
- Piai, V., Meyer, L., Dronkers, N. F., & Knight, R. T. (2017). Neuroplasticity of language in left-hemisphere stroke: Evidence linking subsecond electrophysiology and structural connections. *Human Brain Mapping*, 38(6), 3151–3162. <https://doi.org/10.1002/hbm.23581>
- Piai, V., Roelofs, A., & Maris, E. (2014). Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. *Neuropsychologia*, 53, 146–156. <https://doi.org/10.1016/j.neuropsychologia.2013.11.014>
- Piai, V., Roelofs, A., Rommers, J., & Maris, E. (2015). Beta oscillations reflect memory and motor aspects of spoken word production. *Human Brain Mapping*, 36(7), 2767–2780. <https://doi.org/10.1002/hbm.22806>
- Piai, V., Rommers, J., & Knight, R. T. (2018). Lesion evidence for a critical role of left posterior but not frontal areas in alpha–beta power decreases during context-driven word production. *European Journal of Neuroscience*, 48(7), 2622–2629. <https://doi.org/10.1111/ejn.13695>
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, 62(2), 816–847. <https://doi.org/10.1016/j.neuroimage.2012.04.062>
- Price, C. J., Devlin, J. T., Moore, C. J., Morton, C., & Laird, A. R. (2005). Meta-analyses of object naming: Effect of baseline. *Human Brain Mapping*, 25(1), 70–82. <https://doi.org/10.1002/hbm.20132>
- Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: The WEAVER++/ARC model. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior*, 59, 33–48. <https://doi.org/10.1016/j.cortex.2014.07.001>
- Roos, N. M., & Piai, V. (2020). Across-session consistency of context-driven language processing: A magnetoencephalography study. *European Journal of Neuroscience*, 52(5), 3457–3469. <https://doi.org/10.1111/ejn.14785>
- Schuster, S., Himmelstoss, N. A., Hutzler, F., Richlan, F., Kronbichler, M., & Hawelka, S. (2021). Cloze enough? Hemodynamic effects of predictive processing during natural reading. *NeuroImage*, 228, 117687. <https://doi.org/10.1016/j.neuroimage.2020.117687>.
- Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Brecher, A., Faseyitan, O. K., Dell, G. S., Mirman, D., & Coslett, H. B. (2011). Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proceedings of the National Academy of Sciences*, 108(20), 8520–8524. <https://doi.org/10.1073/pnas.1014935108>
- Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Faseyitan, O., Brecher, A., Dell, G. S., & Coslett, H. B. (2009). Anterior temporal involvement in semantic word retrieval: Voxel-based lesion-symptom mapping evidence from aphasia. *Brain: A Journal of Neurology*, 132(12), 3411–3427. <https://doi.org/10.1093/brain/awp284>
- SPM12 Software—Statistical Parametric Mapping. (n.d.). Retrieved April 20, 2022, from <https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1), 273–289. <https://doi.org/10.1006/nimg.2001.0978>
- van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, 38(4), 584–589. <https://doi.org/10.3758/BF03193889>
- Vartiainen, J., Liljeström, M., Koskinen, M., Renvall, H., & Salmelin, R. (2011). Functional magnetic resonance imaging blood oxygenation level-dependent signal and magnetoencephalography evoked responses yield different neural functionality in reading. *Journal of Neuroscience*, 31(3), 1048–1058. <https://doi.org/10.1523/JNEUROSCI.3113-10.2011>
- Veldsman, M., Cumming, T., & Brodtmann, A. (2015). Beyond BOLD: Optimizing functional imaging in stroke populations. *Human Brain Mapping*, 36(4), 1620–1636. <https://doi.org/10.1002/hbm.22711>
- Wilson, S. M., Bautista, A., Yen, M., Lauderdale, S., & Eriksson, D. K. (2017). Validity and reliability of four language mapping paradigms. *NeuroImage: Clinical*, 16, 399–408. <https://doi.org/10.1016/j.nicl.2016.03.015>