

# Representations of conceptual information during automatic and active semantic access

Antonietta Gabriella Liuzzi<sup>a\*</sup>, Silvia Ubaldi<sup>b</sup>, Scott Laurence Fairhall<sup>b</sup>

<sup>a</sup> Laboratory for Cognitive Neurology, Department of Neurosciences, KU Leuven, Belgium

<sup>b</sup> Center for Mind/Brain Sciences, University of Trento, Trento, 38068, Italy

\*Corresponding author:

Antonietta Gabriella Liuzzi

[antoniettagabriella.liuzzi@kuleuven.be](mailto:antoniettagabriella.liuzzi@kuleuven.be)

Laboratory for Cognitive Neurology

Department of Neurosciences

KU Leuven

Belgium

## **Abstract**

When we read a word or see an object, conceptual meaning is automatically accessed. However, previous research investigating non-perceptual sensitivity to semantic class has employed active tasks. In this fMRI study, we tested whether conceptual representations in regions constituting the semantic network are invoked during passive semantic access and whether these representations are modulated by the need to access deeper knowledge. Seventeen healthy subjects performed a semantically active typicality judgment task and a semantically passive phonetic decision task, in both the written and the spoken input-modalities. Stimuli consisted of one hundred forty-four concepts drawn from six semantic categories. Multivariate Pattern Analysis (MVPA) revealed that the left posterior middle temporal gyrus (pMTG), posterior ventral temporal cortex (pVTC) and pars triangularis of the left inferior frontal gyrus (IFG) showed a stronger sensitivity to semantic category when active rather than passive semantic access is required. Using a cross-task training/testing classifier, we determined that conceptual representations were not only active in these regions during passive semantic access but that the neural representation of these categories was common to both active and passive access. Collectively, these results show that while representations in the pMTG, pVTC and IFG are strongly modulated by active conceptual access, consistent representational patterns are present during active and passive conceptual access in these same regions.

Keywords: semantic representations, words, fMRI, pMTG, pVTC, BA45

## **Introduction**

When we see an object, we automatically and effortlessly understand its significance and meaning. Likewise, as evidenced by the Stroop task (Stroop, 1935), the meaning of a word is automatically and obligatorily retrieved when we read or hear that word. This reflects the passive and automatic access to conceptual representation that allows us to effectively interact with the world. At the same time, specific tasks and behavioural goals can trigger the retrieval of richer information about an object. Owls are not just birds; they are nocturnal, hunt mice and can rotate their neck 270 degrees in either direction. Forms of active semantic access underlie much of the human capacity for thought and higher understanding. The degree to which automatic conceptual access and active conceptual access to deeper associated meaning are served by the same or different neural representations remains an open question.

Conceptual representations can be accessed regardless of the input-modality: conceptual representation derived by a concept presented as spoken word is generally similar to the one derived by the same concept presented as written word. Accordingly, brain areas that represent such conceptual knowledge are supramodal in nature. In the human brain, conceptual knowledge is represented by a distributed semantic system. This includes regions that respond more strongly to semantically richer stimuli: the angular gyrus, lateral and ventral temporal cortex, ventromedial prefrontal cortex, inferior frontal gyrus, dorsal medial prefrontal cortex and the precuneus/posterior cingulate gyrus (Binder et al., 2009). Studies employing multivariate pattern analysis (MVPA) have determined the non-perceptual sensitivity of elements of the semantic system to semantic content (Fairhall & Caramazza, 2013; Devereux et al., 2013, Clark & Tyler 2014, Simanova et al. 2014, Bruffaerts et al., 2013, Liuzzi et al., 2015, 2017, 2019, 2020, Borghesani et al, 2016, Martin et al., 2018). Nevertheless, all these studies adopted active semantic tasks: naming task (Devereux et al., 2013, Clark & Tyler 2014), judgment of semantic consistency (Simanova et al., 2014), property verification task (Bruffaerts et al., 2013, Liuzzi et al., 2015, 2017, 2019, Martin et al, 2018), semantic decision (Borghesani et al, 2016), or a typicality task (Fairhall & Caramazza 2013, Liuzzi et al, 2020). Representations of word-meaning have been studied during the naturalistic presentation of narratives (Huth et al., 2016, Deniz et al., 2020). These results suggest that representations may be widespread under these conditions. However, the relationship

between these representations to single word processing and the relationship between active and passive conceptual access remain uncertain.

Magnetoencephalographic studies employing word stimuli indicate that access to semantic content during active semantic access is a multistage process. Using representational similarity analysis and a semantic typicality task, Giari and colleagues (2020) demonstrated that access to the conceptual content of word stimuli occurs in two stages, the first ranging from 230 to 335 msec, the second from 360 to 585 msec. This finding supports the possibility that semantic access proceeds in an initial, rapid, automatic phase, followed by a latter phase in which semantic representations are actively accessed. Furthermore, the initial peak of a representation of conceptual content in all modalities was observed at 360ms, which coincides with the N400 potential. The N400 is an index of semantic processing and its amplitude is influenced by predictability of a stimulus (Lau et al., 2008) and integration of semantic information with the working context (Hagoort 2013).

The mechanism by which active semantic access is instantiated is thought to be mediated through control circuitry that guides access to task- and goal-relevant semantic representation. Due to its involvement in selecting between multiple competing semantic responses and in making infrequent semantic associations, the left inferior frontal gyrus (IFG) has been attributed a key role in semantic control (Thompson-Schill et al., 1997, Martin and Chao 2001; Wagner et al. 2001; Thompson-Schill 2003, Lambon-Ralph et al., 2017). Recent models have also posited that the control circuitry may extend to include additional regions, such as the posterior middle temporal gyrus (pMTG; Lambon-Ralph et al., 2017).

In the current study, we address the question of whether the semantic system codes for conceptual information in the same way when we interact with the world as when we internally think about meaning. We do so by adopting a phonetic decision task and a typicality task, which require automatic and active semantic access, respectively. By means of a whole-brain decoding Multivariate Pattern Analysis (MVPA), as well as Region-Of-Interest (ROI) analysis, we determined whether—and to what degree—representations of semantic category during passive or active semantic access rely on shared or distinct neural patterns. Results reveal that the semantic system is sensitive to semantic class during active tasks and

that, among those regions composing this system, the left pMTG, pVTC and IFG share common semantic representations between active and passive semantic access.

## **1. Materials and Methods**

### **2.1 Participants**

Seventeen participants took part in this study. All participants were native Italian speakers, right-handed and free of neurological or psychiatric disorders. All procedures were approved by the Human Research Ethics Committee on the Use of Human Subjects in Research of the University of Trento and the experiments were performed in accordance with the approved guidelines. Participants confirmed that they understood the experimental procedure and gave their written informed consent.

### **2.2 Stimulus dataset**

One hundred forty-four concepts derived from the Paisà corpus (<https://www.corpusitaliano.it/>) were divided into six semantic categories (Mammals, Birds, Plants, Food, Furniture and Manipulable Objects). Stimuli were further subdivided into “conceptual triplets,” which consisted of three tightly related concepts (e.g. rosemary, oregano, thyme or gorilla, baboon, gibbon; eight conceptual triplets per category).

Stimuli were selected to ensure there were no significant differences between semantic categories for log Word Frequency (WF) or Word Length (WL, range: 4 and 12 characters) and no significant difference between semantic categories in within-category semantic similarity. Specifically, semantic similarity was calculated from the Paisà corpus by correlating the log co-occurrence vectors of the 144 concepts with 1,978 features. Next, for each semantic category, we averaged the Fisher-transformed correlation values of each conceptual triplet. In this way, the similarity between different triplets within each category was determined. To determine that the similarity between triplets comprising a category did not differ between categories, the off-diagonal elements of each semantic category were entered in a two-sample t-test for each pair of semantic categories (all p-values > .05). Finally, categories were also matched such that the number of words starting with vowel/consonant was consistent across semantic categories (see phonetic task).

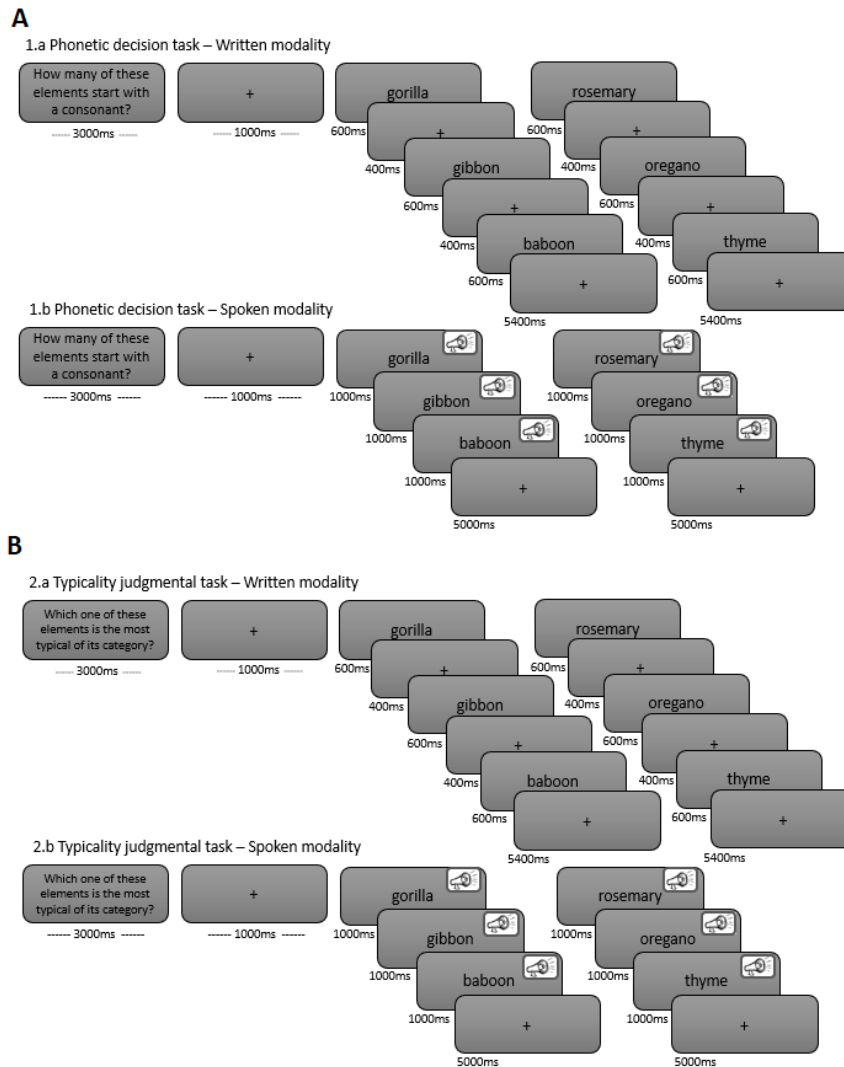
Audio stimuli were recorded through the QuickTime and MacOSX ('Mojave') speech-to-text using the Italian male voice. Written words were presented in black on a gray background, subtending .41 (SD .06) degrees of visual angle vertically and 1.65 (SD .45) horizontally.

### **2.3 Task and experiment details**

Seventeen healthy subjects participated in a block-event-related fMRI experiment. In each block, subjects performed either a phonetic decision task or a typicality judgment task on either written or spoken words. Each block was composed of 8 events, in which conceptual triplets belonging to all semantic categories were presented in a pseudo-randomized order. In the spoken modality, each concept was presented for 1000ms. In the visual modality, three words were each presented for 600ms followed by a fixation cross for 400ms. After the final word of each trial, the fixation cross was displayed for an additional 5000ms. Each block was preceded by a task instruction for 3000ms, followed by a fixation cross for 1000ms (Fig 1A)

In phonetic decision blocks, subjects were asked to count how many words started with a consonant. In typicality judgment blocks, they were required to judge the most typical concept for its category among the three concepts presented. Subjects held a response box in their right hand and pressed a button from 1 to 3 according to their answer.

The experiment consisted of 48 blocks divided into 6 runs: 12 blocks each for the phonetic decision task in written modality, the phonetic decision task in spoken modality, the typicality task in written modality and the typicality task in spoken modality. For each task and for each modality, each conceptual triplet was repeated so that each semantic category (which was composed of 8 conceptual triplets) was presented 16 times (Fig 1B).



**Figure 1.** Tasks. (A) Phonetic decision task, (B) Typicality judgmental task.

## 2.4 Image acquisition

A SIEMENS Prisma scanner (field strength: 3 Tesla) with a 64-channel head coil provided structural and functional images. Structural imaging sequences consisted of a T1-weighted MPRAGE (repetition time (TR) = 2290ms, echo time (TE) = 2.74ms, flip angle = 12, slice thickness = 1mm). Functional data were acquired using a multiband EPI sequence consisting of 78 slices (slice thickness = 1.5mm, TR = 2000ms, TE = 28ms, flip angle = 75, voxel size 2x2x1.5 mm<sup>3</sup>). A total of 1632 volumes were acquired during the experiment. The experiment was split into 6 runs, each lasting 272 volumes. All volumes were AC/PC aligned. Stimuli were presented on a 42", MR-compatible Nordic NeuroLab LCD monitor positioned at the back of

the magnet bore. Participants saw the monitor through a mirror located in front of them. Stimuli were presented using a custom PsychToolBox 3 script running on top of Matlab R2018.

### 3. Data analysis

#### 3.1 Pre-processing

Data were pre-processed with Statistical Parametric Mapping - SPM12 (Wellcome Trust Centre for Neuroimaging, University College London, UK). Functional images were realigned and resliced, and a mean functional image was created. Next, the structural image was co-registered with the mean function image and segmented. Functional images were normalized to the Montreal Neurological Institute (MNI) T1 space, resampled to a voxel size of 2x2x2 mm<sup>3</sup> and spatially smoothed with 8mm FWHM kernel.

#### 3.2 Univariate analysis

A univariate analysis was run as a sanity check of the data. Subject-specific response estimates (beta weights) were derived by fitting a general linear model (GLM). For each subject, and for each run, 24 regressors (6 categories x 2 tasks x 2 input-modalities) were created, starting from the onset of the first concept of each conceptual triplet. The six motion parameters from the realignment procedure were included in the model as regressors of no interest. The main effect of task was computed: [All regressors referring to the typicality task] minus [All regressors referring to the phonetic task] and inverse.

#### 3.3 Whole-brain decoding MVPA

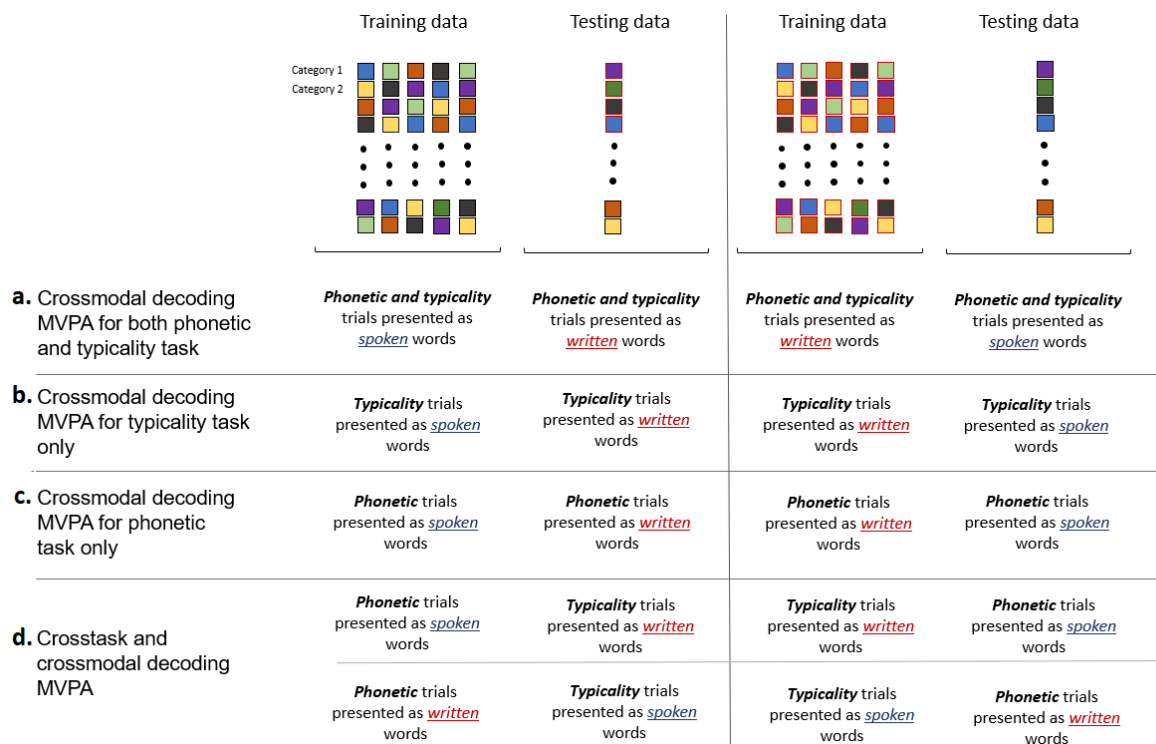
Searchlight **crossmodal** decoding MVPA was performed to investigate regions sensitive to category, by means of CoSMoMVPA ([www.cosmomvpa.org](http://www.cosmomvpa.org)) (Oosterhof et al., 2016). A latent discriminant (LDA) **multiclass** classifier was used to classify **the six object-categories from the** patterns of responses extracted from subject-specific  $\beta$  weights derived from the GLM. For each voxel in the brain, a spherical neighbourhood of  $\beta$  values was defined, with a variable radius set to include the 200 voxels nearest to the centre voxel. **For each searchlight, the classifier was required to classify the 6 semantic categories. As the analysis was cross-modal, the classifier was trained on trials presented as spoken words and tested on trials presented as written words, and vice-versa. Subject-specific accuracy maps were obtained**



by means of a leave-one-out cross— validation procedure: for each input-modality tested, each run (six in total) was left out from the training dataset and used as the testing dataset. Accordingly, each modality was tested six times (6 iterations for each input modality), for a total of 12 iterations. For each searchlight, accuracy scores from different iterations were averaged and the resulting classification accuracy value was summarized at the centre voxel of the sphere. Finally, **chance level accuracy, 1 divided by the number of semantic categories (i.e. 0.166̄), was subtracted from each classification map.** Then subject-specific accuracy maps were smoothed with 6 mm FWHM kernel (Fairhall & Caramazza 2013, Devereux et al., 2013, Simanova et al., 2014, Clarke et al., 2014, Liuzzi et al., 2020) and entered into a one-sample t test.

Several whole-brain **crossmodal** decoding analyses were performed, depending on the type of information we wanted the classifier to decode. To identify brain regions able to discriminate semantic categories regardless of the task, we performed a whole-brain crossmodal decoding analysis, irrespective of task (Fig. 2a). Specifically, for each subject, the classifier was built on 60  $\beta$  weights (corresponding to (6 category\*2 tasks)\*5runs) and its validity was tested on 12  $\beta$  weights (corresponding to 1  $\beta$ \*6 category\*2 tasks), by means of a **leave-one-out** cross-validation measure. A total of 12 iterations were performed. For 6 iterations, the classifier was trained on the spoken modality and tested on the written modality. For the remaining 6 iterations, the input modalities were switched. Secondly, to detect differences in sensitivity to category between tasks, we computed a whole-brain crossmodal decoding analysis for each task separately (Fig. 2b,c). **For each analysis, the classifier was trained and tested on trials derived by one of the two tasks.** Parameters were identical to the previous analysis. Finally, in order to test the consistency of the semantic representation across automatic and active access, an additional whole-brain decoding analysis was performed with training and testing iterations taking place across task as well as across modality (Fig. 2d). Here, the classifier was built on a specific task and modality (e.g. Phonetic task in spoken modality) and its validity was tested on the other task and modality (e.g. Typicality task in written modality) by means of a cross-validation measure. Importantly, the same task was never included in both training and testing folds, insuring that positive results cannot be driven by the presence of category information in only one of the tasks. For each subject, a total of 24 iterations were computed: 6 iterations, including a training phase on Phonetic task in written modality and a test phase on Typicality task in spoken modality, 6

iterations in which data for training and testing were reversed, 6 iterations for a training phase on Phonetic task in spoken modality and a testing phase on Typicality task in written modality, and 6 iterations for the reverse type of data for training and testing.



**Figure 2.** Schematic overview of all whole-brain decoding analysis performed. For each decoding MVP analysis, the type of trials adopted during training and testing phase are provided.

### 3.4 ROI definition and analysis

Each region of interest (ROI) was extracted from a whole-brain decoding analysis. Only local maxima of significant clusters at  $P_{FWE-CORR}$  (cluster level)  $<0.05$  were taken into account. For each local maximum, a sphere with variable radius of 15mm was created. Next, the sphere was intersected with the whole-brain decoding accuracy map. ROI-based analysis was performed by first extracting subject-specific decoding accuracy patterns from subject-specific un-smoothed decoding maps and then averaging the pattern within subjects. Significance was determined by means of a one-sample t-test.

## 4. Results

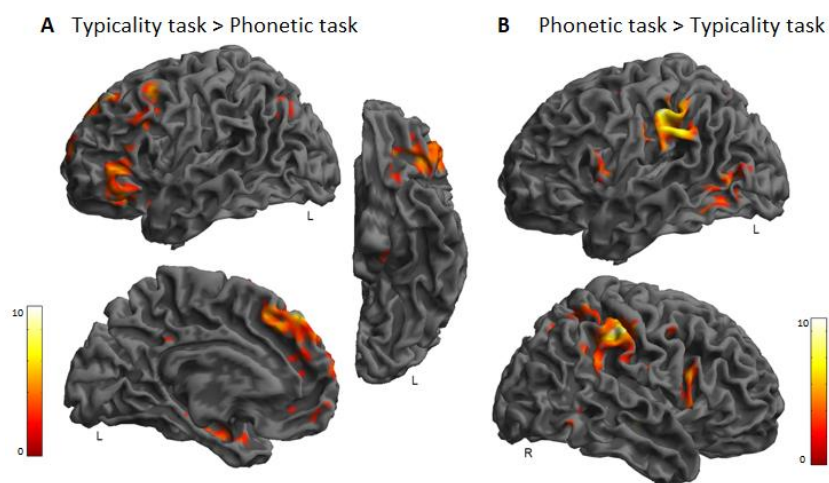
### 4.1 Behavioural Results.

A 3-way Repeated Measures ANOVA with reaction times as output was computed. As expected, there was a strong main effect of task ( $F(1,16) = 28,2, p < .00001$ ) and main effect of input-modality ( $F(1,16) = 24,2, p < .00001$ ). Subjects were faster during the phonetic task (mean = 1.19s, SEM = .019) than during the typicality task (mean = 1.32s, SEM = .017) and they were faster for the written modality (mean = 1.18s, SEM = .017) compared to the spoken modality (mean = 1.33, SEM = .018). A less pronounced effect of category was also observed ( $F(2,8,45.6) = 4,2, p = .01$  (Greenhouse-Geisser)), which did not interact with other factors.

## 4.2 fMRI Results

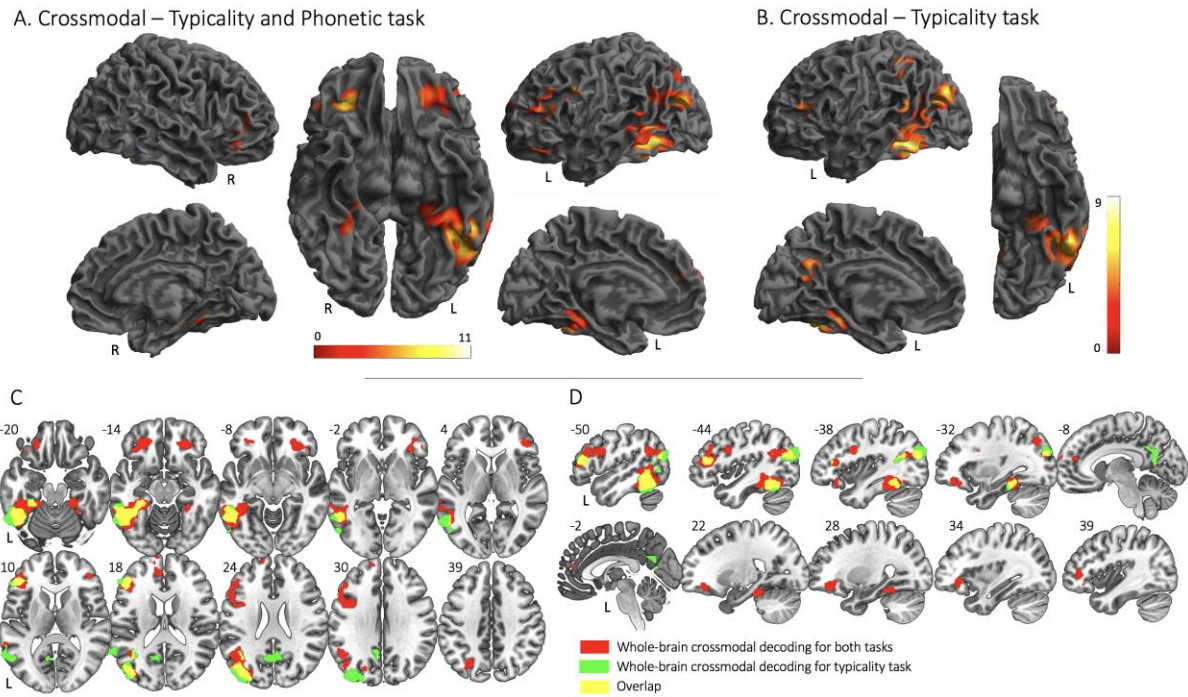
### 4.2.1 Univariate analysis

Compared to the phonetic task, the typicality task produced larger responses in the left middle and inferior frontal gyrus (pars triangularis), dorsomedial prefrontal cortex, the posterior cingulate, the hippocampus, right orbitofrontal cortex and bilateral intraparietal sulcus. Compared to the typicality task, the phonetic task yielded bilaterally the supramarginal gyrus and the pars opercularis of the IFG. (Figure 3).



**Figure 3.** Univariate analysis. 3D rendering of the main effect of task: (A) Typicality > Phonetic task and (B) Phonetic > Typicality task. The significance level was set at voxel-level threshold of uncorrected  $p < 0.001$ , combined with cluster-level inference of  $p < 0.05$  corrected for the whole brain volume.

## 4.2.2 Crossmodal Category Decoding



**Figure 4.** Amodal representations of category. 3D rendering of the whole-brain crossmodal category decoding for (A) typicality and phonetic tasks and (B) typicality task only. No significant clusters for the phonetic task only were detected. (C) Axial and (D) Sagittal slices showing binary map of the whole-brain crossmodal decoding for both tasks (red), typicality task (green) and the overlap between the two maps (yellow). The significance level was set at voxel-level threshold of uncorrected  $p < 0.001$ , combined with cluster-level inference of  $p < 0.05$  corrected for the whole brain volume.

<i>Whole-brain crossmodal category decoding for typicality and phonetic task</i>					
<i>VOI</i>	<i>MNI Coordinates</i>			<i>Extent</i>	<i>PFWE-Corr (Cluster level)</i>
	<i>x</i>	<i>y</i>	<i>z</i>		
Left BA45	-38	6	28	968	<0.001
BA44	-44	28	16		
Left pMTG	-50	-56	-10	2277	<0.001
pVTC	-34	-38	-14		
Right OFC	26	38	-10	502	<0.001
BA45	44	38	4		

Left Occ-Parietal Junction	-40	-80	20	1140	<0.001
IPS	-26	-68	44		
Left OFC	-36	30	-16	335	0.001
Right pVTC	20	-38	-20	199	0.010
Medial PFC	-8	66	20	173	0.020
<i>Whole-brain crossmodal category decoding for typicality task</i>					
Left pMTG	-54	-50	-18	2205	<0.001
pSTG	-68	-54	16		
pVTC	-28	-36	-16		
Left Occ-Parietal Junction	-42	-82	18	909	<0.001
Left BA45	-52	34	14	325	0.009
Precuneus	2	-60	26	367	0.005

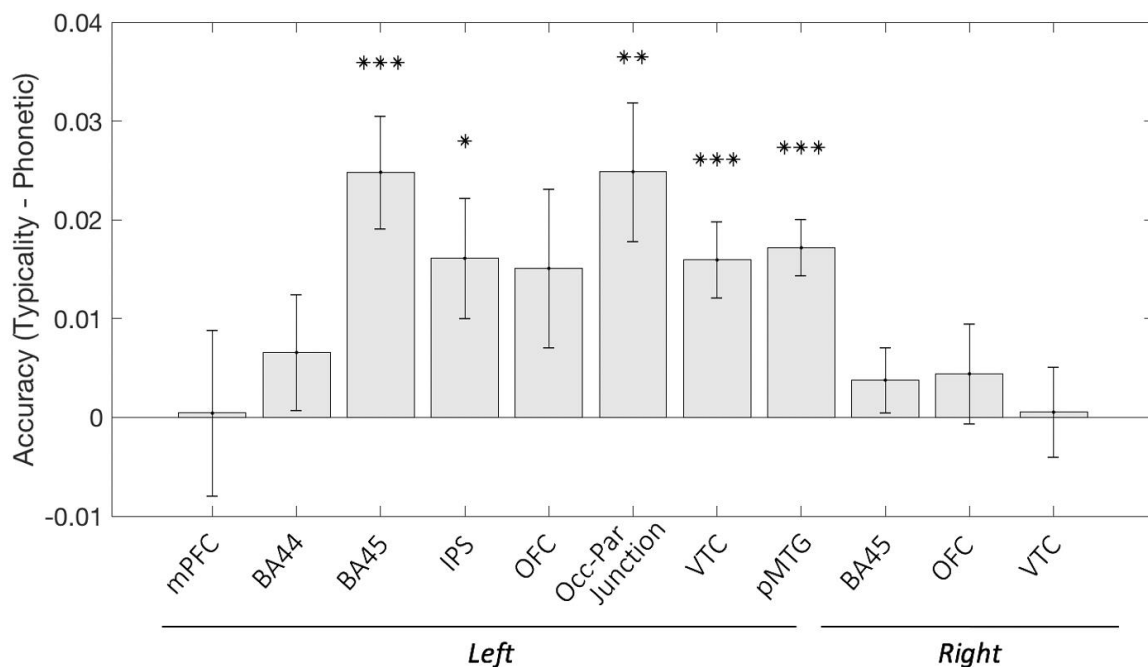
**Table 1:** Clusters showing significant accuracy for whole-brain crossmodal category decoding for both tasks (typicality and phonetic) and for the typicality task only (whole-brain crossmodal decoding for phonetic task only did not yield any significant clusters). The significance level was set at voxel-level threshold of uncorrected  $p < 0.001$  combined with cluster-level inference at  $p < 0.05$  corrected for the whole brain volume. Extent refers to the number of  $2 \times 2 \times 2$  mm<sup>3</sup> voxels. *Abbreviations:* pMTG: posterior middle temporal gyrus; VTC: ventral temporal cortex; IPS: intraparietal sulcus; OFC: orbitofrontal cortex; PFC: prefrontal cortex; pSTG: posterior superior temporal gyrus.

A whole-brain crossmodal decoding for both typicality and phonetic tasks revealed seven clusters sensitive to category. On the lateral surface, sensitivity to category was detected in left and right BA 45, left BA 44, left pMTG, left occipito-parietal junction and left IPS. Ventrally, sensitivity was present in the left and right VTC, left and right OFC and in the medial prefrontal cortex (Fig. 4A; Table 1). Whole-brain crossmodal decoding for the typicality task revealed a left-lateralized network composed of the left pMTG, VTC, pSTG, left occipito-parietal junction, BA45 and the precuneus (Fig. 4B; Table 1). No significant clusters were detected for the whole-brain crossmodal decoding for the phonetic task only.

To assess whether the representation of semantic category differed between spoken and written input modalities, we performed the equivalent whole-brain category-decoding analysis separately for both typicality and phonetic tasks. No significant differences were evident between spoken and written input modalities ( $p > .05$ ).

#### 4.2.3 Differential conceptual representation during active and passive semantic access.

Searchlight crossmodal decoding analysis revealed that several brain regions are sensitive to category for the typicality task (Fig 4B; Table 1) but none for the phonetic task. In order to formally interrogate this difference, we selected ROIs from the whole-brain crossmodal decoding for both tasks (Fig 4A; Table 1) and interrogated them for typicality/phonetic differences. As during crossmodal decoding analysis for both tasks the classifier was blind to the task, ROI definition and interrogation is independent. Raw decoding accuracies for each task are not presented, as values would be biased by the ROI selection. In the active typicality task, we saw significantly stronger representations of category in left BA45 ( $t(16)=4.36$ ,  $p=.0005$ ), VTC ( $t(16)=4.14$ ,  $p=.0008$ ), pMTG ( $t(16)= 6.04$ ,  $p<.0001$ ), occipito-parietal junction ( $t(16)=3.54$ ,  $p=.003$ ), and, at an uncorrected threshold, in the left IPS ( $t(16)=2.65$ ,  $p=.017$ ) (Fig. 5).



**Figure 5:** Enhanced category-sensitive for active access compared to passive access. Decoding accuracy difference (above chance) between typicality and phonetic task for each ROI derived from the searchlight decoding for both tasks. Error bars indicate standard error of the mean. \* corresponds to  $p < .05$ , \*\* corresponds to  $p < .01$ , \*\*\* corresponds to  $p < .001$ . *Abbreviations:* PFC: prefrontal cortex; IPS: intraparietal sulcus; OFC: orbitofrontal cortex; VTC: ventral temporal cortex; pMTG: posterior middle temporal gyrus.

Repeated measures ANOVA revealed that the enhancement of categorical representations in the active semantic task were not equivalent across these regions ( $F(5.4,86)=2.552$ ,  $p=.030$ ; Greenhouse-Geisser corrected). Notably, right hemispheric categorical representations appear not to be modulated by active semantic access: Right BA45, pVTC and OFC showed no significant differences between active and passive semantic access, and a significantly reduced enhancement was seen compared to their left hemispheric counterparts in BA45 ( $p = .007$ ) and pVTC ( $p = .004$ ).

#### **4.2.4 Task-general conceptual representations**

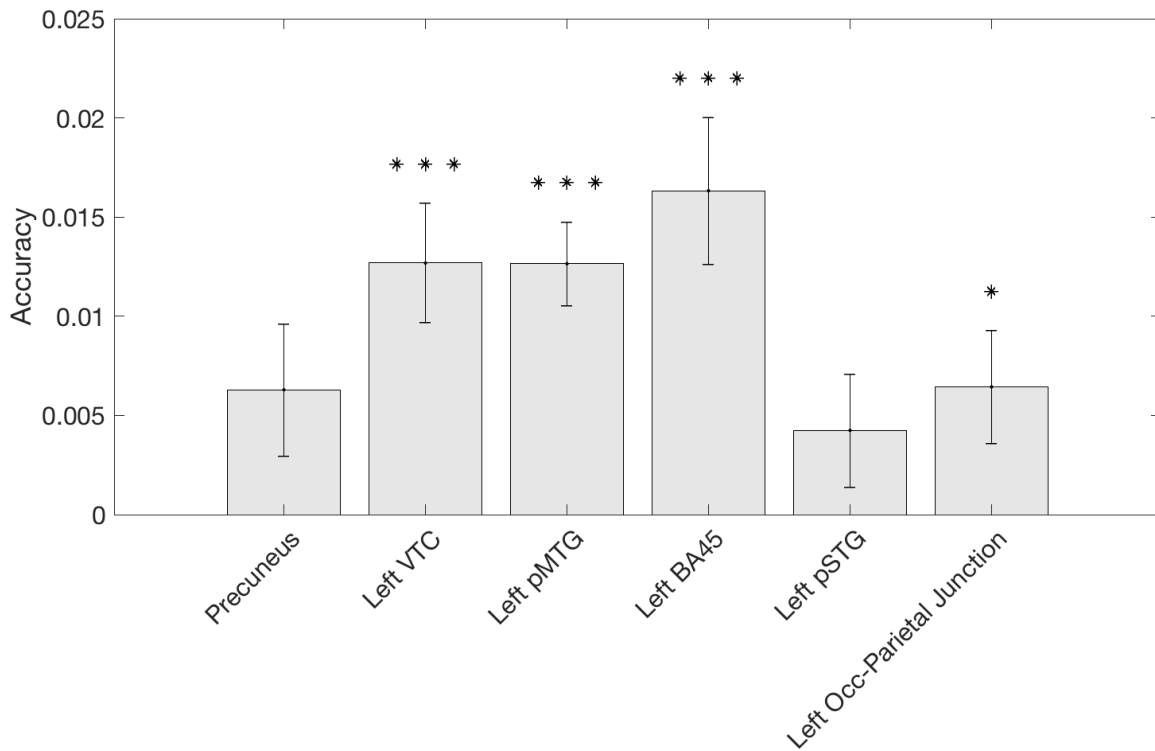
The absence of searchlight results in the phonetic task and the stronger representations evident during active access insinuates that automatic access to meaning does not elicit conceptual representation in the semantic system. However, this may also be a question of sensitivity, with weaker, less-detectable, categorical representation during passive access. In this next analysis we address two questions: are subtle conceptual representations present during passive semantic access? If so, are conceptual representations consistent over active and passive access?

The presence of task-generalised semantic representations was assessed by training and testing across both task and modality. In this way, the classifier is trained on one task (e.g. typicality) and tested on the other (e.g. phonetic) and will detect only conceptual representations that are consistent across both modes of access (active/passive), as well as modality of input (spoken/written). This whole-brain crossmodal and crosstask decoding analysis yielded three significant clusters at the significant voxel-level threshold of uncorrected  $p < 0.001$  combined with cluster-level inference at  $p < 0.05$  corrected for the whole brain volume: One cluster encompassing the left posterior ITG/MTG and the posterior portion of the left fusiform gyrus (local maxima -36 -38 -20, KE 741, Cluster level  $P_{FWE-CORR} < .0001$ ) and another in the left pars triangularis of the left inferior frontal gyrus (local maxima -42 34 10, KE 183, Cluster level  $P_{FWE-CORR} = .031$ ), overlapped with the ROIs used in this analysis. Additionally, a region of the left OFC (local maxima -30 48 -18, KE 261, Cluster level  $P_{FWE-CORR} = .005$ ), showed significant crossmodal and crosstask decoding MVPA. This provides evidence that left OFC, which was also identified in the crossmodal analysis (figure 4A), represents semantic categories in a task-general manner.

Furthermore, to ensure independence while increasing sensitivity, we examined phonetic decoding accuracy within ROIs derived from the searchlight decoding in the typicality task (Fig 4B, Table 1, see methods). Common representations were evident in the same regions showing the greatest enhancement in conceptual representation during active semantic access: the left VTC ( $t(16) = 4.21$ ,  $p = .0007$ ), left pMTG ( $t(16) = 5.98$ ,  $p < .0001$ ) and left BA45 ( $t(16) = 4.42$ ,  $p = .0004$ ) (Fig. 5). At an uncorrected threshold, task-generalised conceptual representations were also present in the left occipito-parietal junction ( $t(16) = 2.27$ ,  $p = .04$ ) (Fig. 6). **To assess whether effects were driven by the direction of training and testing across modalities, we generated subject-specific accuracy maps for two types of iterations: (1) iterations where the classifier was trained on patterns derived by the phonetic task (for written and spoken words) and tested on patterns derived by the typicality task (for written and spoken words) and (2) iterations where the classifier was trained on patterns derived by the typicality task (for written and spoken words) and tested on patterns derived by the phonetic task (for written and spoken words). Overall (averaging across regions), the direction of training/testing did not significantly influence classifier performance ( $t(16) < 1$ ) and follow-up paired-sample t-test did not reveal any significant difference ( $p > 0.05$ ) for any of the regions tested.**

Repeated measures ANOVA revealed that task-generalised categorical representations were not equivalent across ROIs ( $F(5,80) = 3.31$ ,  $p = .009$ ).





**Figure 6: *Task-general conceptual representations*** within ROIs derived from the searchlight crossmodal decoding for the typicality task only. Above chance mean decoding accuracy of a cross-task and cross-modal decoding analysis. Error bars indicate standard error of the mean. \* corresponds to  $p < .05$ , \*\* corresponds to  $p < .01$ , \*\*\* corresponds to  $p < .001$ . *Abbreviations:* VTC: ventral temporal cortex; pMTG: posterior middle temporal gyrus; pSTG: posterior superior temporal gyrus.

## 5. Discussion

In this study, we investigated the representation of conceptual information during automatic and active semantic access to address whether overlapping or distinct neural populations underlie these processes. We observed that active access enhanced representation in pMTG, pVTC and BA45 (Fig 5). At the same time, we found that neural representations in these same cortical regions show a subtle pattern of category representation that is common to both active and passive conceptual access (Fig 6).

### 5.1 pMTG, pVTC and BA45: Enhanced representation during active access and conceptual representations that generalise across active and passive retrieval.

Crossmodal decoding irrespective of task revealed amodal representations of semantic categories in a left-lateralized network, consistent with the general semantic network (Fig 4A) (Binder et al, 2009). When we investigated the role of task in these regions, we found that active access resulted in an enhanced representation of object-category in the left BA45, pVTC and pMTG (Fig 5). While a whole-brain MVPA for the phonetic task did not yield any significant clusters, a more sensitive analysis in which a classifier was built on a specific task and modality and tested on the other task and modality, showed that in left BA45, pVTC and pMTG, conceptual categories were represented in the same way during both forms of access (Fig 6). Collectively, these results indicate that the same regions that show the most pronounced enhancement in conceptual representations resulting from active access are the very same regions that show the strongest reactivation of these category-selective representations during passive semantic access. This underscores the importance of these regions in both active and passive conceptual access, as well as demonstrating that the nature of the neuro-conceptual representation is common across forms of access.

Enhanced amodal representations for active semantic access in the pMTG is consistent with previous findings. In addition to being activated in response to active semantic tasks for several input-modalities (Vandenberghe et al., 1996, Hickok et al., 2004, Stoeckel et al., 2003, Binder et al, 2009), pMTG is sensitive to semantic categories (Fairhall & Caramazza 2013). Importantly, our current results show that while categorical representations are enhanced in the left pMTG during active semantic access, representation in this region is characterised by a common neural response for active and passive semantic tasks. This provides support for a central role of this region in coding semantic information, regardless of the type of access. However, a parallel literature supports the role of this region in the controlled retrieval of semantic information (Davey et al. 2016, Lambon-Ralph et al., 2017). In line with this latter account, it has been recently suggested (Jefferies et al., 2020) that the pMTG would allow a tailored retrieval of conceptual knowledge by dynamically uploading conceptual context. Although our results support the role of pMTG in semantic representation, this does not preclude the possibility that the pMTG has a dual role in semantic control.

Similarly to the pMTG, left BA45 showed consistent semantic representations across types of task and types of input-modality. Involvement of the left BA45 in semantic processing has been extensively demonstrated by means of active semantic tasks (semantic judgmental tasks

for written words and pictures (Vandenberghe et al., 1996); semantic decision tasks for written words (Wagner et al., 2001)). The role of the IFG in semantic access is most often attributed to semantic control and the guided retrieval of goal-relevant semantic information (Martin and Chao 2001; Wagner et al. 2001; Thompson-Schill 2003, Lambon-Ralph et al., 2017, Jefferies et al., 2020). However, this region is also sensitive to semantic content: similarity in the neural patterns produced by concepts conforms to the similarity in meaning between concepts (Liuzzi et al., 2017, Deniz et al., 2019).

In the current study, we found that neural conceptual representations present in BA45 were enhanced during active access and persisted during passive, incidental access where no semantic control demand was present. This provided tentative support for a role of semantic representation within BA45. At the same time, the results of the current study, where participants were asked to compare three concepts to one another, are also in line with a role of the left IFG in domain-specific semantic working memory (Gabrieli et al., 1998).

## **5.2 Evidence for regional dissociation in the role of active semantic access**

While left BA45, pVTC and pMTG showed involvement in both active and passive conceptual access, there was some evidence for regional dissociation in their relative role in these processes. Enhancements in conceptual representation were not equally evident across the semantic system. Notably right hemisphere regions appear less involved in active semantic access, with BA45 and pVTC showing no significant enhancement in representation during active semantic access compared to passive access, a pattern that was significantly different from their left hemispheric counterparts. This implies that right-hemisphere conceptual representations are equally involved in passive and active conceptual access and, by extension, are not part of the network involved in effortful semantic access.

Results also indicate that the presence of task-generalised categorical representation was not equivalent across the six regions within which this effect was analysed. Here, task-generalised representations did not reach significance in the precuneus or left pSTG. Although this provides tentative evidence for a functional dissociation between these regions, future work is needed to verify the nature of these differences.

## **6. Conclusions**

In the current study we investigated whether active and passive semantic access rely on shared or distinct neural substrates. We found that active access enhanced representation in left pMTG, pVTC and BA45, and through a cross-task decoding analysis, we showed that these same regions exhibit a common neural representation for both types of access, active and passive. Collectively, these results show that the same cortical regions code for conceptual information both when we interact with the world and when we internally think about its meaning.

#### **DATA AVAILABILITY**

Data, stimuli and analysis scripts are available upon request from the corresponding author.

#### **ACKNOWLEDGEMENTS**

The project was funded by the European Research Council (ERC) Starting Grant CRASK - Cortical Representation of Abstract Semantic Knowledge, awarded to Scott Fairhall under the European Union's Horizon 2020 research and innovation program (grant agreement no. 640594).

#### **BIBLIOGRAPHY**

1. 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.
2. Borghesani, V., Pedregosa, F., Buiatti, M., Amadon, A., Eger, E., & Piazza, M. (2016). Word meaning in the ventral visual path: a perceptual to conceptual gradient of semantic coding. *NeuroImage*, *143*, 128-140.
3. Bruffaerts, R., Dupont, P., Peeters, R., De Deyne, S., Storms, G., & Vandenberghe, R. (2013). Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. *Journal of Neuroscience*, *33*(47), 18597-18607.
4. Clarke, A., & Tyler, L. K. (2014). Object-specific semantic coding in human perirhinal cortex. *Journal of Neuroscience*, *34*(14), 4766-4775.
5. Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., ... & Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *NeuroImage*, *137*, 165-177.
6. Deniz, F., Nunez-Elizalde, A. O., Huth, A. G., & Gallant, J. L. (2019). The representation of semantic information across human cerebral cortex during listening versus reading is invariant to stimulus modality. *Journal of Neuroscience*, *39*(39), 7722-7736.

7. Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *Journal of Neuroscience*, 33(48), 18906-18916.
8. Fairhall, S. L., & Caramazza, A. (2013). Brain regions that represent amodal conceptual knowledge. *Journal of Neuroscience*, 33(25), 10552-10558.
9. Gabrieli, J. D., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the national Academy of Sciences*, 95(3), 906-913.
10. Giari, G., Leonardelli, E., Tao, Y., Machado, M., & Fairhall, S. (2020). Spatiotemporal properties of the neural representation of conceptual content for words and pictures—an MEG study. *NeuroImage*, 116913.
11. Hagoort, P. (2007). 11 The memory, unification, and control (MUC) model of language. *Automaticity and control in language processing*, 1, 243.
12. Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67-99.
13. Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600), 453-458.
14. Jefferies, E., Thompson, H., Cornelissen, P., & Smallwood, J. (2020). The neurocognitive basis of knowledge about object identity and events: dissociations reflect opposing effects of semantic coherence and control. *Philosophical Transactions of the Royal Society B*, 375(1791), 20190300.
15. Lambon-Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42.
16. Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920-933.
17. Liuzzi, A. G., Bruffaerts, R., Dupont, P., Adamczuk, K., Peeters, R., De Deyne, S., ... & Vandenberghe, R. (2015). Left perirhinal cortex codes for similarity in meaning between written words: Comparison with auditory word input. *Neuropsychologia*, 76, 4-16.
18. Liuzzi, A. G., Bruffaerts, R., Peeters, R., Adamczuk, K., Keuleers, E., De Deyne, S., ... & Vandenberghe, R. (2017). Cross-modal representation of spoken and written word meaning in left pars triangularis. *NeuroImage*, 150, 292-307
19. Liuzzi, A. G., Dupont, P., Peeters, R., Bruffaerts, R., De Deyne, S., Storms, G., & Vandenberghe, R. (2019). Left perirhinal cortex codes for semantic similarity between written words defined from cued word association. *NeuroImage*, 191, 127-139.
20. Liuzzi, A. G., Aglinskas, A., & Fairhall, S. L. (2020). General and feature-based semantic representations in the semantic network. *Scientific Reports*, 10(1), 1-12.
21. Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current opinion in neurobiology*, 11(2), 194-201.
22. Martin, C. B., Douglas, D., Newsome, R. N., Man, L. L., & Barense, M. D. (2018). Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. *Elife*, 7, e31873.
23. Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in neuroinformatics*, 10, 27.
24. Simanova, I., Hagoort, P., Oostenveld, R., & Van Gerven, M. A. (2014). Modality-independent decoding of semantic information from the human brain. *Cerebral cortex*, 24(2), 426-434.
25. Stoeckel, M. C., Weder, B., Binkofski, F., Buccino, G., Shah, N. J., & Seitz, R. J. (2003). A fronto-parietal circuit for tactile object discrimination:: an event-related fmri study. *Neuroimage*, 19(3), 1103-1114.

26. Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6), 643.
27. Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences*, 94(26), 14792-14797.
28. Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring "how" from "where". *Neuropsychologia*, 41(3), 280-292.
29. Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, 383(6597), 254-256.
30. Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron*, 31(2), 329-338.