Neuropsychologia 137 (2020) 107275

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: http://www.elsevier.com/locate/neuropsychologia

Imageability effect on the functional brain activity during a naming to definition task

Francesca Garbarini[a,b,c,h], Fabrizio Calzavarini[b,i], Matteo Diano[g,h,*], Monica Biggio[e], Carola Barbero[b,d], Daniele P. Radicioni[b,f], Giuliano Geminiani[a,b,g,h], Katiuscia Sacco[a,b,h,1], Diego Marconi[b,d,1]

[a] Imaging and Plasticity Research Group at Department of Psychology, and Neuroscience Institute of Turin, University of Turin, Torino, 10100, Italy
[b] Center for Logic, Language and Cognition (LLC), Turin, Italy
[c] MANIBUS Lab, Department of Psychology, University of Turin, Turin, Italy
[d] Department of Philosophy and Educational Sciences, University of Turin, 10123, Torino, Italy
[e] Department of Experimental Medicine, Section of Human Physiology and Centro Polifunzionale di Scienze Motorie, University of Genoa, 16132, Genoa, Italy
[f] Department of Computer Science, University of Turin, 10123, Torino, Italy
[g] GCS-fMRI, Koelliker Hospital, Turin, Italy
[h] Department of Psychology, University of Turin, 10123, Torino, Italy
[i] Department of Letter, Philosophy, Communication. University of Bergamo, Italy

ARTICLE INFO

Keywords:
Functional magnetic resonance imaging (fMRI)
Lexical inferential competence
Visual imagery
Ventral temporal cortex
Posterior parietal lobe

ABSTRACT

Lexical competence includes both the ability to relate words to the external world as accessed through (mainly) visual perception (referential competence) and the ability to relate words to other words (inferential competence). We investigated the role of visual imagery in lexical inferential competence by using an auditory version of an inferential naming-to-definition task, in which visual imageability of both definitions and target words was manipulated. A visual imageability-related brain activity (bilateral posterior-parietal lobe and ventrotemporal cortex, including fusiform gyrus) was found during a "pure" inferential performance. The definition effect in high vs. low imageability contrast suggests that a visual-imagery strategy is spontaneously activated during the retrieval of a word from a high imageable definition; such an effect appears to be independent of whether the target word is high or low imageable. This contributes to the understanding of the neural correlates of semantic processing and the differential role of spontaneous visual imagery, depending on the semantic properties of the processed stimuli.

1. Introduction

Lexical semantic competence, i.e. knowledge of word meaning, is traditionally assessed by means of a variety of verbal semantic tasks, both in production and in comprehension. Some reflections in the philosophy of language (Marconi, 1997) have suggested a distinction between two different classes of verbal semantic tasks, i.e. referential and inferential tasks (Marconi et al., 2013; see also Calzavarini, 2017; Calzavarini, 2019). Referential tasks involve the language–world relation as mediated by perception, particularly by vision (i.e., picture naming, word-picture matching), while inferential tasks require the ability to deal with semantic relations among lexical items. Examples of inferential tasks are word-word matching, which requires selecting which word among various alternatives is best related in meaning to a probe word, and naming to definition, in which subjects are asked to recover a target word from a verbal definition. Property/sentence verification, sentence completion, and semantic relatedness decision can likewise be considered as inferential tasks (Calzavarini, 2017; 2019).

Inferential tasks have been found to correlate with increased activation in language-related areas, such as left superior and middle

---

[*] Corresponding author.

E-mail addresses: francesca.garbarini@unito.it (F. Garbarini), fabrizio.calzavarini@unibg.it (F. Calzavarini), matteo.diano@unito.it (M. Diano), monica.biggio@edu.unige.it (M. Biggio), carola.barbero@unito.it (C. Barbero), daniele.radicioni@unito.it (D.P. Radicioni), giulianocarlo.geminiani@unito.i (G. Geminiani), katiuscia.sacco@unito.it (K. Sacco), diego.marconi@unito.it (D. Marconi).

URL: http://www.llc.unito.it (M. Diano).

[1] Co-last authors.

https://doi.org/10.1016/j.neuropsychologia.2019.107275
Received 4 August 2018; Received in revised form 20 November 2019; Accepted 20 November 2019

temporal gyri (Marconi et al., 2013). In a previous study (Marconi et al., 2013) some typical visual related areas were found to be also engaged by purely inferential tasks, not involving visual recognition of objects or pictures. Such activation might reflect visual imagery/visual semantic processes triggered by "visually loaded" sentences and words used in the tasks (in spite of the fact that visual load was not experimentally controlled in that study). In psycholinguistic research, it is common to use the term « visual imageability » to indicate the property of a particular word or sentence to produce an experience of visual imagery (e.g., Jonides et al., 1975). Intuitively, words such as *cat* or *apple* refer to concrete entities and are associated with high visual imageability. On the contrary, low imageable words such as *democracy* or *truth* are not clearly associated with the production of visual mental images.

Many previous neuroimaging studies have investigated the effect of visual imageability, and multimodal imageability more generally, on brain measures (for a review, see Binder, 2007). Some among such studies have compared concrete, high imageable material vs abstract, low imageable material during inferential tasks such as word-word matching (Bedny and Thompson-Schill, 2006; Hoffman et al., 2015; Sabsevitz et al., 2005), property verification (Pexman et al., 2007), sentence verification (Just et al., 2004) and sentence completion (Mestres-Missé et al., 2009). Most such studies have shown that inferential performances with concrete, high imageable words correlate with selective activation of visual-related areas. At a general level, this finding is consistent with the results of several neuroimaging studies conducted within the so-called Simulation Framework that have reported activation of visual cortex during language tasks with concrete, high imageable words (for a review, see Kemmerer, 2010). According to the Simulation Framework, sensorimotor cortex is critical for semantic competence. As restricted to visual modality, the Simulation Framework predicts that the visual cortex, particularly the visual association cortex in ventral temporal lobes (e.g., Martin, 2007), stores visual conceptual knowledge. For instance, according to Simmons and Barsalou (2003), "visual feature maps" in ventral temporal stream represent visual properties of concrete objects that are routinely mobilized not only in visual recognition, but also in more general conceptual and linguistic processing including inferential tasks.

Although the above mentioned data seem to support the imageability hypothesis, caution should be taken for several reasons. The first is that not all neuroimaging studies investigating the role of imageability in inferential tasks reported selective activation of visual related areas during the processing of concrete, high imageable words (e.g., Giesbrecht et al., 2004; Pexman et al., 2007). Secondly, some inferential studies that did report an imageability effect in visual related regions were potentially confounded by the explicit request to imagine the semantic content of words or sentences (Just et al., 2004; Mellet et al., 1998). Thirdly, few studies have manipulated imageability during language-related tasks with auditory stimuli such as words (Wise et al., 2000) or sentences (Desai et al., 2010; Just et al., 2004). Finally, few studies manipulated imageability during language tasks involving complete sentences as stimuli (e.g., Desai et al., 2010, 2011; Mellet et al., 1998); more importantly, no previous studies manipulated imageability during a naming to definition task.

In the present fMRI study, we further investigate the role of visual imageability in inferential performances by collecting behavioural and neuroimaging data during a naming to definition task. Note that participants were not explicitly instructed to generate a visual mental image of the definition's content. The visual imageability of definitions and target words was manipulated throughout the experiment by comparing high with low visual imageable stimuli. Thus for the same target, whether high imageable (e.g. 'cat')or low imageable (e.g. 'intelligence'), there were two alternative definitions, high imageable (e.g. for 'cat': 'The mice hunter with whiskers and a long tail is … '; for 'intelligence': 'People believed that a big brain was a symptom of … ') and low imageable (e.g. for 'cat': 'The domestic feline with nine lives is … '; for 'intelligence': 'The quality of people who easily solve hard problems is called … '). See details in Methods.

Our expectation here was that the manipulated variable (i.e. the visual load of both targets and definitions) would modulate the imageability-related brain activity during the naming to definition task. In our 2 × 2 factorial design, a significant interaction would suggest that the visual loads of both target and definition contribute to activate visual imageability-related areas; thus, we expected that high imageable definitions of high imageable targets would show greater imeageability-related brain activity with respect to the other conditions. Alternatively, a specific effect of either factor (target/definition) might allow us to assess whether the imageability-related brain activity is mainly associated with visual imageability of the verbal definitions (irrespective of whether the target word is high or low imageable) or with visual imageability of the recovered target words (irrespective of whether the definition contains high or low imageable words).

2. Materials and methods

2.1. Participants

Twenty-five healthy volunteers, native Italian speakers, were recruited for the experiment. Due to technical issues during fMRI acquisition (excess of head movement while speaking in six subjects and/or failure of voice recording during scan in two subjects), eight subjects were excluded from the analysis. Thus, 17 subjects (9 females and 8 males), 21–30 years of age (mean ± standard deviation = 23.5 ± 2.5) were considered for the main analysis (an additional behavioral analysis including 23 subjects is presented in Supplementary Online Materials SOM). All participants gave their written informed consent. The procedure was approved by the ethical committee of the University of Turin (N: 134681), in accordance with the Declaration of Helsinki (World Medical Association, 1991). Participants were all naïve to both the experimental procedure and the aims of the study.

2.2. Experimental design

Participants were asked to perform an inferential 'naming to definition' task inside the MRI scanner. During the task, a sentence was pronounced and the participants were instructed to listen to the stimulus given in the headphones (listening phase) and to overtly name, as accurately and as fast as possible, the target word corresponding to the definition (answering phase). During the experiment, accuracy and reaction time were collected. At the end of the experimental session, subjects were administered a visual imageability questionnaire: they had to rate on a 1–7 Likert scale the intensity of the visual imagery experience they perceived as related to each target and each definition. For each sentence, the level of visual imageability (high vs low) was manipulated either in the "definition factor" (i.e. the statement that describes the target could have either a high level or a low level of visual imageability) or in the "target factor" (the word naming the object described in the definition could have either a high level or a low level of visual imageability). This design gave rise to four experimental conditions: High Imageable Target \ High Imageable Definition (HIT\HID; e.g., "*The bird of prey with large wings flying over the mountains is the … eagle*"); High Imageable Target \ Low Imageable Definition (HIT\LID; e.g., "*The hottest of the four elements of the ancients was … fire*"); Low Imageable Target \ High Imageable Definition (LIT\HID; e.g., "*Pinocchio's nose stretched when he told a … lie*"); Low Imageable Target \ Low Imageable Definition (LIT\LID; e.g., "*The quality of people that easily solve hard problems is called … intelligence*").

Thus, for each condition there were 24 sentences, 96 sentences overall. Number of words (nouns and adjectives), word frequency, and (syntactic dependency) structure of the included sentences were uniform across conditions (see Supplementary analysis of the set of stimuli in SOM).

2.3. fMRI paradigm

We designed a fast event related paradigm according to the following timeline (Supplementary Fig. 1): 1) 5 s of listening phase (participant hears the definition); 2) 2 s of answering phase (subject's verbal response); 3) random jitter from 1 to 2 s (mean 1.5 s). For each condition (HIT\HID, HIT\LID, LIT\HID, LIT\LID), 24 stimuli were presented in pseudo-random order. Note that there were 48 definitions and 48 targets, 24 targets repeated twice; for each target we had both a visual (HI) and a non-visual (LI) definition. Each trial lasted 7 s (5 s for the question and 2 s for the answer) plus random jittering (1–2 s). We performed 1 run per participant; each run was composed of 24 trials for each condition (96 in all). The total functional protocol lasted about 17 min (see details in Supplementary Fig. 1). In a training period prior to fMRI, participants familiarized with the task in the control room, using a short version of the naming to definition task on sentences not included in the main experiment.

2.4. Behavioral data analysis

Details of behavioral data acquisition within the MRI scanner are reported in SOM. In the main analysis, including 17 subjects, the participants' performance was evaluated by recording, for each response, the reaction time (*RT)* and the accuracy (*AC)*. Then, *RT* and *AC* were combined into the *Inverse Efficiency Score* (IES), a metrics commonly used to aggregate reaction time and accuracy and to summarize them by using the formula $IE=(RT/AC)*100$ (Townsend and Ashby, 1978). The mean IES value was used as the dependent variable and entered in a 2 × 2 ANOVA with "Target" ("High imageable"; "Low imageable") and "Definition" ("High imageable"; "Low imageable") as within-subjects' factors. *Post hoc* comparisons were performed using the Duncan test. Additionally, we performed separate analyses for RT and AC (see SOM) and we performed the IES analysis on 23 subjects (including the six participants previously excluded due to head movements). These additional analyses confirmed the results of the main analysis (see SOM).

Finally, to analyze the visual imageability questionnaire, we computed for each item (definition and target) a mean imageability score, obtained by averaging the participants' ratings. The resulting data were analyzed using non-parametric Wilcoxon test, contrasting High vs. Low imageable Targets and Definitions. Furthermore, to perform correlations with the beta values extracting from visual-related clusters highlighted by the fMRI analysis (see below), we averaged the imageability score for each definition item across subjects, obtaining a normative value for each element in the questionnaire. All behavioral analyses, included correlations with fMRI data (see below), are performed by using *STATISTICA* software (https://www.tibco.com/products/data-science).

2.5. FMRI data acquisition, preprocessing and analysis

The experiment was performed at the Koelliker Hospital of Turin (Italy), by means of a 1.5-T magnetic resonance (MR) scanner (Ingenia, Philips Healthcare, Best, The Netherlands) with a Philips dStream HeadSpine coil 15 channels head coil optimized for functional imaging. A MRI compatible headphones system (NordicNeuroLab AS, Bergen, Norway, http://www.nordicneurolab.com/) was used to present auditory stimulation via E-Prime software (Psychology Software Tools, Inc., Pittsburgh, PA), which also ensured synchronization with the MR scanner and the behavioral data collection.

For fMRI acquisition: echo planar imaging (EPI) with matrix size 64 × 64, Time Repetition 2.5 s, Time Echo 41 ms, 25 slices with interleaved scan order, isotropic voxel size 4 mm, slice gap 0.5 mm, phase encode directions (Posterior-Anterior), functional run of 418 volumes. For T1 weighted acquisition: 3-dimensional high resolution T1w, fast field echo sequence, Time Repetition 25 ms, ultra-short Time Echo, 30° flip angle, in-plane voxel resolution 0.98 × 0.98 mm with 1.5 mm slice thickness, 107 sagittal slices. The functional EPI was preprocessed using FSL 5.0.8 (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/, Jenkinson et al., 2012) following a standard pipeline: I) motion correction using mcflirt to align the data to the middle volume; II) EPI distortion correction using TOPUP method (FSL - 5 volumes phase encode direction Anterior-Posterior); III) the functional data were coregistered into native structural T1w image using Boundary Based Registration (BBR); IV) then, the subject-specific T1w was coregistered into standard MNI152 2 mm (isotropic voxel size) template using linear affine transformation with 12 Degrees of Freedom (FLIRT - FSL); V) the intensity in the EPI images was normalized using a single scaling factor ("grand mean scaling") for each volume; VI) the functional image was 7 mm spatial smoothed and high pass temporal filtered at 90 s. To quantify the cerebral activity in response to the task (first level analysis), a double-gamma Hemodynamic Response Function (HRF) was fitted in the design matrix in FILM (FMRIB's Improved Linear Model) approach adding temporal derivatives to shift the convolved design in time to compensate for regional change in the HRF.

In the fMRI analysis, to achieve the group result, we adopted the two level analysis approach in FSL. At the first level, we estimated T contrast of parameter estimates (cope) within the functional run for each subject. At the second level, we performed a group analysis controlling the between-subject's variance using mixed-effect approach (FLAME1 option in FSL) using a cluster forming threshold of $z = 2.3$ and FWE cluster corrected at $p < 0.05$.

fMRI analyses were performed using an ANCOVA-like approach. In the GLM design, we modelled the paradigm phase named "listening" (see Supplementary Fig. 1) because our focus was on the understanding-and-search process (inferential performance). Note that we used T-contrasts, instead of F-contrasts, as we were directly interested in the directionality of the High vs. Low effect. The T-contrasts were defined as: "Definition" with 'High imageable" [(HIT\HID + LIT\HID)-(HIT\LID + LIT\LID)] and "Low imageable" [(HIT\LID + LIT\LID)-(HIT\HID + LIT\HID)]; "Target" with "High imageable" [(HIT\HID + HIT\LID)-(LIT\HID + LIT\LID)] and "Low imageable" [(LIT\HID + LIT\LID)-(HIT\HID + HIT\LID)]; "Interaction" [(HIT\HID + LIT\LID)-(LIT\HID + HIT\LID)] and [(LIT\HID + HIT\LID)-(HIT\HID + LIT\LID)]. "Interaction" derives from the contrasts between conditions in which definitions and target words were either both high imageable or both low imageable ("congruent conditions"), and conditions in which this was not the case ("incongruent conditions"). Our main goal was to investigate cerebral areas that are involved during the task, excluding the difficulty of the task itself, mirrored by the subject's performance. Therefore, IES was included in the design as a covariate of no interest (see SOM). Afterwards, we created a single mask for each cluster of activation observed in a group level whole brain analysis of trials containing highly imageable definitions versus trials with low imageability definitions. From within these masks, and at the single subject level, we extracted an average beta value from across all voxels and for each individual trial. Then these values were averaged across subjects for each item. To compute the single trial beta for each stimulus, we fitted into GLM analysis the single trial regressor plus a regressor with all other trials (Mumford et al., 2012). Subsequently, these group average values were correlated with average subjective ratings of the imageability scores (see section "Behavioral data analysis" above). To explore the presence of such a correlation, we performed two ANCOVA models in which subjective ratings at the visual imageability questionnaire were used to predict beta values in visual areas showing a Definition effect in the High vs Low T-contrast (i.e. Right and Left Fusiform Gyrus), also controlling for the condition effect (e.g. Burin et al., 2017; Garbarini et al., 2014). In two separated ANCOVA models, we used as dependent variables the beta values extracted from either the right or left fusiform areas. In both models, the four experimental conditions (HIT\HID, HIT\LID, LIT\HID, LIT\LID) were used as categorical predictor and the subjective ratings at the questionnaire were used as continuous predictor.

All the anatomical localization and tables were derived from Xjview (http://www.alivelearn.net/xjview), which allows to describe, for each cluster, the location of the maximum peak and how many voxels are comprised in the *WFU_PickAtlas database (*Maldjian et al. 2003*).*

3. Results

3.1. Behavioral results

The ANOVA results showed a significant effect of the within-subject factor “target” [F (1,16) = 12.385; p = 0.003), with the IES values significantly lower in the high (mean ± sem = 696.877 ± 69.387) than in the low (mean ± sem = 845.848 ± 97.163) visual imageable targets. No significant effect was found for the “definition” factor. We also found a significant “target*definition” interaction [F(1,16) = 5.016; $p <$ 0.041), showing that high imageable definitions improved performance with high imageable targets and worsened performance with low imageable targets (Duncan *post hoc* comparison*:* p = 0.007). On the contrary, with low imageable definitions no difference between high and low imageable targets was found (p = 0.238). See Fig. 1. See additional behavioral analysis in Supplementary Online Materials (SOM).

Comparing ratings for high imageable (mean ± sem = 5.95 ± 0.08) and low imageable (mean ± sem = 4.64 ± 0.10) definitions in the visual imageability questionnaire results, we found a significant difference (p < 0.001; Wilcoxon test). Comparing ratings for high imageable (mean ± sem = 5.66 ± 0.11) and low imageable (mean ± sem = 4.93 ± 0.13) targets we likewise found a significant difference (p < 0.001).

3.2. fMRI results

The main result of our fMRI analysis show a visual imageability effect in the Definition (High vs Low) contrast, including several bilateral clusters in visual imagery related areas (such as bilateral Fusiform Gyrus, Middle Temporal Gyrus, Precuneus, Inferior and Superior Parietal Lobule (see details in Table 1, Supplementary Table 1 and Fig. 2A)). In the Definition reverse contrast (Low vs. High) a greater brain activity was found in the left Angular Gyrus (Supplementary Table 2; Supplementary Fig. 2). In the Target contrast, only a very specific activity in the left mouth sensorimotor areas was found in the High vs. Low contrast (see Supplementary Table 3 and Supplementary Fig. 3). No Target effect was found in the reverse contrast. In the Interaction contrast, when HIT\LID + LIT\HID and HIT\HID + LIT\LID are compared, a significantly greater activity was found in the left Angular Gyrus, Left Inferior and Middle Frontal Gyrus (see Supplementary Table 4 and Supplementary Fig. 4). No Interaction effect was found in the reverse contrast.

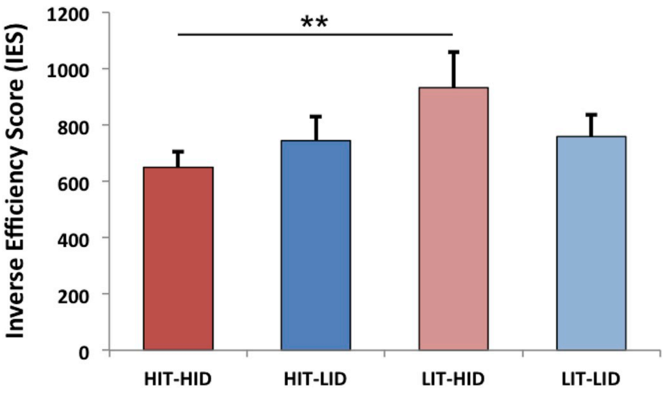

Fig. 1. Behavioral results. The graph shows, for each condition, the mean IES with standard error of the mean (sem). Significant differences between conditions were marked with asterisks (**p < 0.005).

Table 1
The table shows the imageability-related brain areas and the coordinates at maximum peak of the cluster in the Definition contrast (High vs Low). **Sub-regions** indicate the extension of the activation blobs through some areas. The anatomical location of each single cluster was defined using “cluster report” in xjview MATLAB script (http://www.alivelearn.net/xjview, for more details see Supplementary Tables 1–4). The coordinates X-Y-Z are relativized to MNI standard space.

Area	z-score	X	Y	Z
LOW IMAGEABLE DEFINITION				
Left Angular Gyrus	3.8639	−44	−74	42
HIGH IMAGEABLE DEFINITION				
Left Fusiform Gyrus	4.7461 *Sub-regions* Left Inferior Temporal Gyrus Left Middle Temporal Gyrus Left Inferior Occipital Gyrus	−42	−52	−16
Right Inferior Temporal Gyrus	3.9769 *Sub-regions* Right Inferior Parietal Lobule Right Superior Parietal Lobule Right Postcentral Gyrus Precuneus Right Fusiform Gyrus	48	−54	−6
Left Superior Parietal Lobule	4.5263 *Sub-regions* Left Inferior Parietal Lobule Precuneus Left Postcentral Gyrus Left Superior Occipital Gyrus	−24	−70	40

3.3. Correlations between visual imageability questionnaire and beta values of the brain areas showing a definition effect

The ANCOVA model for the Left Fusiform Gyrus revealed that the subjective ratings at the visual imageability questionnaire significantly predicts the beta values (F(1,91) = 8.007, p = 0.005), even when the effect of condition was controlled for. On the contrary, the ANCOVA model for the Right Fusiform Gyrus did not show a significant effect of the visual imageability questionnaire in predicting the beta values (F (1,91) = 0.166, p = 0.683). See Fig. 2B.

4. Discussion

The primary aim of our study was to further investigate the role of visual imageability during inferential performances, particularly during an inferential naming to definition task. To this aim, we manipulated the visual imageability of both definitions and target words, so that both high and low visual imageable targets and high and low visual imageable definitions were included in the experimental design. Visual imageability related effects were found at both the behavioral and the neuroimaging level.

Behavioral results show that, overall, the participants’ performance was significantly faster and more accurate when they had to name high than low imageable targets. This confirms the advantage for words quickly and spontaneously associated with mental images, as previously described by using both similar (Kiran and Tuchtenhagen, 2004) and different (e.g., Hoffman et al., 2015) semantic inferential tasks. Furthermore, our results show that high imageable definitions improve performance with high imageable targets and worsen it with low imageable targets (see Fig. 1). This result somehow mirrors the stimuli list construction phase, where it was easier to define high imageable targets (e.g., *eagle*) using high imageable words, and low imageable targets (e.g., *intelligence*) using low imageable words, than *vice versa* (defining high imageable. Targets by low imageable words, etc.). We can

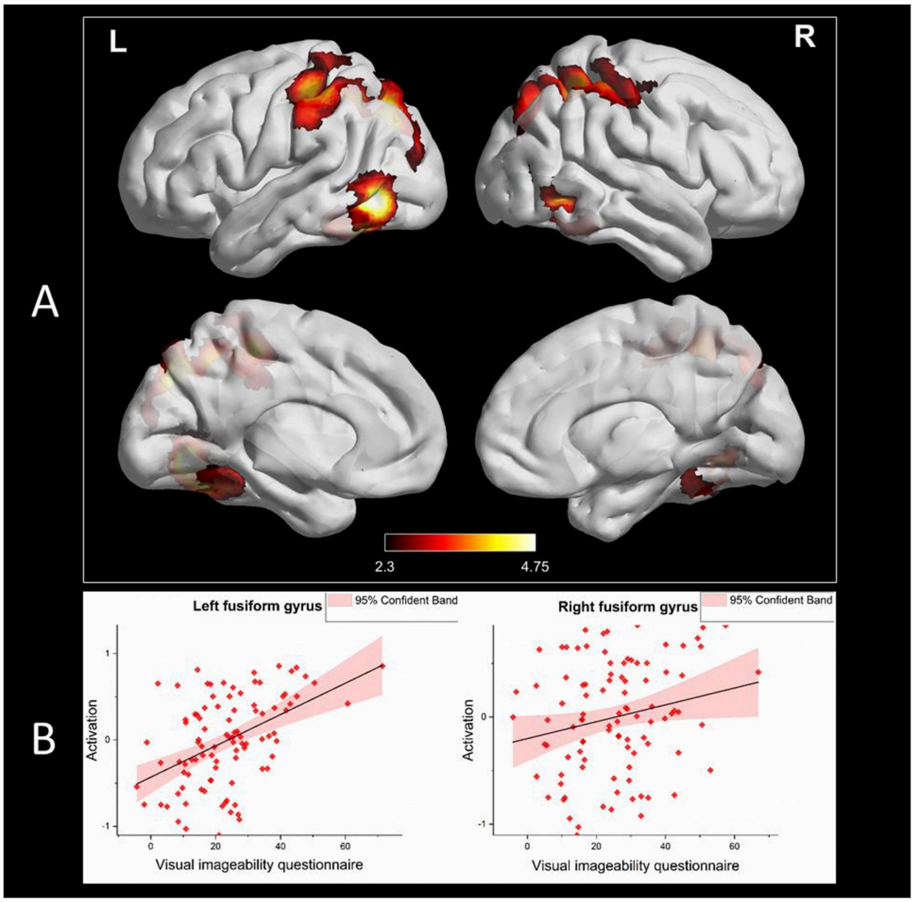

Fig. 2. A: fMRI results. The figure shows imageability-related brain activity in the Definition contrast (High vs. Low). The results are thresholded at $z = 2.3$ and cluster corrected at $p < 0.05$. L = Left, R = Right. **B**: **Correlations results.** The graph shows correlations between subjective ratings on the visual imageability questionnaire and beta values of the high imageable definition related area. Left Fusiform Gyrus $F(1,91) = 8.007$ $p = 0.005$, Right Fusiform Gyrus $F(1,91) = 0.166$ $p = 0.683$.

speculate that for low imageable words (or at least those included in our list) it is harder both to produce (stimuli construction) and to associate (task) high imageable definitions, while for high imageable words the opposite is true. Further research is needed to investigate this aspect.

As far as the fMRI results are concerned, the main findings of the present study are that a visual imageability related brain activity (i.e. bilateral ventral temporal cortex, including fusiform gyri) was found in a specific inferential task (i.e. naming to definition) and that such activity was selectively associated with high visual imageability of the definitions, irrespective of whether the target word to be retrieved was high or low visual imageable (see Fig. 2A).

The first result supports the view that visual imagery may play a role in inferential performances with high imageable stimuli. Previous neuroimaging studies, using different inferential tasks or different methods, showed similar imageability effects in ventral temporal (e.g., Hoffman et al., 2015; Sabsevitz et al., 2005) and/or posterior parietal cortex (e.g., Sabsevitz et al., 2005). In the study of Bookheimer et al. (1998), subjects underwent PET while performing a naming to definition task with concrete, high imageable (three-word) definitions and target words. In addition to areas associated to language processing, listening to high imageable definitions elicited increased activation in the visual cortices (i.e., bilateral occipital gyri, left fusiform gyrus). However, no low imageable definitions or target words were used in that study, so that it was impossible to investigate neural activation elicited by high vs. low imageable stimuli. In the study of Mellet et al. (1998), subjects were asked to listen to 15 words and their definitions. In the first ("concrete") condition, both definitions and target words were high imageable. In the second ("abstract") condition, the words were ordinary words for 'abstract' concepts. A strong visual imagery effect was found in bilateral inferior temporal/fusiform gyrus, as well as in the left inferior parietal lobule (among other areas). However, critically, subjects were not asked to retrieve and say aloud the word corresponding to a given definition. Finally, in contrast with our study, subjects in that experiment were explicitly requested to imagine the semantic content of the linguistic stimuli used in the high imageable condition, thereby making it impossible to verify if visual activation was spontaneously triggered in inferential performances. On the contrary, our study shows the selective activation of visual-imagery related areas in a "pure" naming to definition task (with no concurrent visual imagery task), with stimuli specifically designed to oppose high and low visual imageability.

Importantly, neither an interaction effect nor a target effect was found in visual imageability related brain areas. This suggests that a visual imagery strategy may be unconsciously activated by a high imageable definition, irrespective of whether the target word is high or low imageable. Nevertheless, a potential limitation of this study is that, during the listening phase (on which we focused our analysis), it is not possible to precisely disentangle the BOLD response relative to the definition comprehension from the BOLD response relative to the search of the target word (though it is plausible to conjecture that search for the target word started during the comprehension phase, even before comprehension was completed). Therefore, more research is needed to further investigate this point.

Furthermore, our fMRI results show a significant Interaction effect in the left Angular Gyrus and the Left Middle and Inferior Frontal Gyrus (see Supplementary Figure and Table 4). Even if the left Angular Gyrus also showed a Definition (Low vs High) effect, the Interaction effect makes clear that the Definition effect was led by the HIT\LID condition, including low imageable definitions of high imageable targets. A possible explanation of our findings is that defining a high imageable word by means of sentences consisting of low imageable terms, and viceversa, may be perceived by the subjects as a violation of their semantic expectations (*semantic violation*). This interpretation is consistent with previous reports of frontal lobe activations in the processing of semantic anomalous sentences (Maess et al. 2006), and with studies highlighting the role of the prefrontal cortex and posterior parietal cortex in the recognition of semantic violations in action representations (Balconi et al., 2014). Finally, this interpretation is consistent with the role of the angular gyrus in semantic processing (e.g., Binder 2007; Schwartz et al., 2011; Price et al., 2015) and with the role of the inferior frontal gyrus in semantic control (e.g., Whitney et al., 2011, 2012). Indeed, semantic control has been regarded as "necessary to specifically retrieve context-relevant and task-appropriate semantic information from the representational system especially when unusual, uncharacteristic or anomalous meanings need to be accessed" (Wawrzyniak et al., 2017, p.1).

Finally, a Target (High vs Low) effect was found in left Precentral/Postcentral Gyrus (see Supplementary Figure and Table 3). According to sensorimotor functional localization of lips and tongue (for greater detail on lip localization see (Makin et al., 2015), our finding overlaps with the left mouth sensorimotor area. This suggests an involvement of cerebral areas deputized to oral speech production. To test this hypothesis we used Neurosynth database (July 2017); and, indeed, this cluster shows a maximum correlation with the term "speech production" ($r = 0.256$). This activity is likely explained by the fact that, in some trials, the subjects did not wait for the end of the definition phase in order to pronounce the target word. These cases of anticipated response were more frequent with high imageable targets, which, coherently, showed lower RT with respect to low imageable targets. We acknowledge that this result might represent a limitation of the present study. In principle, during the listening phase, no vocalizations had to be produced by the subjects (as they were explicitly instructed to produce the target word during the response phase). However, we cannot identify these anticipated responses to discharge them from the analysis because vocal responses were not recorded during the listening phase but only during the response phase.

It is worth noting that the imageability-related effect we found in the fMRI results seems to be independent of the participants' performance during the task. Indeed, the visual load manipulation acts in different ways on the behavioral parameter (IES) and on brain activity (hemodynamic response). IES values exhibited a Target effect (performance was faster and more accurate with high than with low imageable targets), while hemodynamic response exhibited a Definition effect (greater imageability-related brain activity with high than with low imageable definitions). It should be noted that in the fMRI analysis, the IES values were used as a covariate, in order to control for the effect of the subjects' performance on the brain activity results.

Finally, for both targets and definitions, subjective ratings on the visual imageability questionnaire suggest that our intuitive categorization of each sentence within the four experimental conditions (HIT\HID, HIT\LID, LIT\HID, LIT\LID) is generally in agreement with the subjects' own intuitive judgements. Furthermore, we found that subjective ratings at the visual imageability questionnaire significantly predict the beta values extracted from the left Fusiform Gyrus (but not those extracted from the homologous areas in the right hemisphere), also controlling for the condition effect (see Fig. 2B). This result supports the view that, for the activity of the Fusiform Gyrus in the left hemisphere (dominant for language), a linear increase was predicted by the imageability level of each definition, irrespective of the definition effect shown by our whole brain analysis.

Taken together, our findings improve our understanding of visual imagery related processing during semantic processing, providing neural correlates of (what can be regarded as) the visual imageability effect in an inferential task.

Funding

This work has been funded by San Paolo Foundation (2012) grant (D15E13000360003) to D.M., by MIUR-SIR 2014 grant (RBSI146V1D) and by the San Paolo Foundation 2016 grant (CSTO165140) to F.G. M.D. is supported by an ERC Consolidator Grant 2017 (LIGHTUP, project # 772953, P.I. Marco Tamietto).

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

The authors are grateful to all volunteers involved in the study. The authors heartily thank the Reviewers for the time devoted and effort regarding this work which has been considerably improved.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2019.107275.

References

Balconi, M., Canavesio, Y., Vitaloni, S., 2014. Activation of the prefrontal cortex and posterior parietal cortex increases the recognition of semantic violations in action representation. Brain Stimul 7 (3), 435–442. https://doi.org/10.1016/j.brs.2014.01.011. May-Jun, Epub 2014 Jan 29.

Bedny, M., Thompson-Schill, S.L., 2006. Neuroanatomically separable effects of imageability and grammatical class during single-word comprehension. Brain Lang. 98 (2), 127–139. https://doi.org/10.1016/j.bandl.2006.04.008.

Binder, J.R., 2007. Effects of word imageability on semantic access: neuroimaging studies. In: Kraut, M.A., Hart, J. (Eds.), Neural Basis of Semantic Memory. Cambridge University Press, pp. 149–181. https://doi.org/10.1017/CBO9780511544965.007.

Bookheimer, S.Y., Zeffiro, T.A., Blaxton, T.A., Gaillard, W.D., Malow, B., Theodore, W.H., 1998. Regional cerebral blood flow during auditory responsive naming: evidence for cross-modality neural activation. Neuroreport 9 (10), 2409–2413.

Burin, D., Garbarini, F., Bruno, V., Fossataro, C., Destefanis, C., Berti, A., Pia, L., 2017. Movements and body ownership: evidence from the rubber hand illusion after mechanical limb immobilization. Neuropsychologia 107, 41–47.

Calzavarini, F., 2017. Inferential and referential lexical semantic competence: a critical review of the supporting evidence. J. Neurolinguistics 44, 163–189.

Calzavarini, F., 2019. Brain and the Lexicon. The Neural Basis of Inferential and Referential Competence. Springer.

Desai, et al., 2010. Activation of sensory-motor areas in sentence comprehension. Cerebr. Cortex 20, 468–478.

Desai, et al., 2011. The neural career of sensory-motor metaphors. J. Cogn. Neurosci. 23, 2376–2386.

Garbarini, F., Boero, R., D'Agata, F., Bravo, G., Mosso, C., Cauda, F., Duca, S., Geminiani, G., Sacco, K., 2014. Neural correlates of gender differences in reputation building. PLoS One 9, e106285.

Giesbrecht, B., Camblin, C.C., Swaab, T.Y., 2004. Separable effects of semantic priming and imageability on word processing in human cortex. Cerebr. Cortex 14, 521–529.

Hoffman, P., Binney, R.J., Lambon Ralph, M.A., 2015. Differing contributions of inferior prefrontal and anterior temporal cortex to concrete and abstract conceptual knowledge. Cortex; a Journal Devoted to the Study of the Nervous System and Behavior 63, 250–266. https://doi.org/10.1016/j.cortex.2014.09.001.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. FSL. *NeuroImage* 62, 782–790.

Jonides, J., Kahn, R., Rozin, P., 1975. Imagery instructions improve memory in blind subjects. Bull. Psychon. Soc. 5 (5), 424–426, 1975.

Just, M.A., Newman, S.D., Keller, T.A., McEleney, A., Carpenter, P.A., 2004. Imagery in sentence comprehension: an fMRI study. Neuroimage 21, 112–124.

Kemmerer, D., 2010. How words capture visual experience: the perspective from cognitive neuroscience. In: Malt, B., Wolff, P. (Eds.), Words and the Mind: How Words Capture Human Experience. Oxford University Press, Oxford, UK, pp. 289–329.

Kiran, S., Tuchtenhagen, J., 2004, January 1. Imageability effects in normal bilingual adults: evidence from naming to definition and semantic priming tasks [Clinical Aphasiology Abstract], Retrieved January 23, 2016, from. http://aphasiology.pitt.edu/archive/00001615/.

Makin, T.R., Scholz, J., Henderson Slater, D., Johansen-Berg, H., Tracey, I., 2015. Reassessing cortical reorganization in the primary sensorimotor cortex following arm amputation. Brain: J. Neurol. 138 (8), 2140–2146. https://doi.org/10.1093/brain/awv161.

Maldjian, J.A., Laurienti, P.J., Burdette, J.B., Kraft, R.A., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19, 1233–1239.

Marconi, D., 1997. Lexical Competence. MIT Press.

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.

Martin, A., 2007. The representation of object concepts in the brain. Annu. Rev. Psychol. 58, 25–45. https://doi.org/10.1146/annurev.psych.57.102904.190143.

Mellet, E., Tzourio, N., Denis, M., Mazoyer, B., 1998. Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition. Neuroreport 9 (5), 803–808.

Mestres-Missé, A., Münte, T.F., Rodriguez-Fornells, A., 2009. Functional neuroanatomy of contextual acquisition of concrete and abstract words. J. Cogn. Neurosci. 21 (11), 2154–2171. https://doi.org/10.1162/jocn.2008.21171.

Mumford, J.A., Turner, B.O., Ashby, F.G., Poldrack, R.A., 2012. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. Neuroimage 59 (3), 2636–2643. https://doi.org/10.1016/j.neuroimage.2011.08.076.

Pexman, P.M., Hargreaves, I.S., Edwards, J.D., Henry, L.C., Goodyear, B.G., 2007. Neural correlates of concreteness in semantic categorization. J. Cogn. Neurosci. 19 (8), 1407–1419. https://doi.org/10.1162/jocn.2007.19.8.1407.

Price, A.M., Bonner, M.F., Peelle, J.E., Grossman, M., 2015. Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. J. Neurosci. 35 (7), 3276–3284. https://doi.org/10.1523/JNEUROSCI.3446-14.2015.

Sabsevitz, D.S., Medler, D.A., Seidenberg, M., Binder, J.R., 2005. Modulation of the semantic system by word imageability. Neuroimage 27 (1), 188–200. https://doi.org/10.1016/j.neuroimage.2005.04.012.

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. Proc. Natl. Acad. Sci. 108 (20), 8520–8524. https://doi.org/10.1073/pnas.1014935108.

Simmons, W.K., Barsalou, L.W., 2003. The similarity-in-topography principle: reconciling theories of conceptual deficits. Cogn. Neuropsychol. 20 (3), 451–486. https://doi.org/10.1080/02643290342000032.

Townsend, J.T., Ashby, F.G., 1978. Methods of modeling capacity in simple processing systems. In: Castellan, J, Restle, F (Eds.), Cognitive theory. Erlbaum Associates, Hillsdale, NJ, pp. 200–239. Vol. III.

Wawrzyniak, M., Hoffstaedter, F., Klingbeil, J., Stockert, A., Wrede, K., Hartwigsen, G., et al., 2017. Fronto-temporal interactions are functionally relevant for semantic control in language processing. PLoS ONE 12 (5), e0177753. https://doi.org/10.1371/journal.pone.0177753.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M.A., Jefferies, E., 2011. The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. Cerebral Cortex (New York, N. Y.: 1991) 21 (5), 1066–1075. https://doi.org/10.1093/cercor/bhq180.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M.A., Jefferies, E., 2012. Executive semantic processing is underpinned by a large-scale neural network: revealing the contribution of left prefrontal, posterior temporal, and parietal cortex to controlled retrieval and selection using TMS. J. Cogn. Neurosci. 24 (1), 133–147. https://doi.org/10.1162/jocn_a_00123.

Wise, R.J., Howard, D., Mummery, C.J., Fletcher, P., Leff, A., Büchel, C., Scott, S.K., 2000. Noun imageability and the temporal lobes. Neuropsychologia 38 (7), 985–994.

World Medical Association, 1991. World Medical Association declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 310 (20), 2191–2194. https://doi.org/10.1001/jama.2013.281053, 2013 Nov 27.