Inferior parietal lobule is sensitive to different semantic similarity relations for concrete and abstract words

Maria Montefinese[1,2], Paola Pinti[3,4,#], Ettore Ambrosini[2,5,6,#], Ilias Tachtsidis[3], David Vinson[1]

[1]*Department of Experimental Psychology, University College London, London, United Kingdom*

[2]*Department of General Psychology, University of Padova, Padova, Italy*

[3]*Department of Medical Physics and Biomedical Engineering, Malet Place Engineering Building, University College London, London, United Kingdom*

[4]*Institute of Cognitive Neuroscience, Alexandra House, University College London, London, United Kingdom*

[5]*Department of Neuroscience, University of Padova, Padova, Italy*

[6]*Padova Neuroscience Center, University of Padova, Padova, Italy*

#These authors have contributed equally.

Corresponding author: Maria Montefinese

Department of General Psychology, University of Padova,

Via Venezia 12, 35131 Padova, Italy

Email: maria.montefinese@gmail.com

Abstract

Similarity measures, the extent to which two concepts have similar meanings, are the key to understand how concepts are represented, with different theoretical perspectives relying on very different sources of data from which similarity can be calculated. While there is some commonality in similarity measures, the extent of their correlation is limited. Previous studies also suggested that the relative performance of different similarity measures may also vary depending on concept concreteness and that the inferior parietal lobule (IPL) may be involved in the integration of conceptual features in a multimodal system for the semantic categorization.

Here, we tested for the first time whether theory-based similarity measures predict the pattern of brain activity in the IPL differently for abstract and concrete concepts. English speakers performed a semantic decision task, while we recorded their brain activity in IPL through fNIRS. Using Representational Similarity Analysis, results indicated that the neural representational similarity in IPL conformed to the lexical co-occurrence among concrete concepts (regardless of the hemisphere) and to the affective similarity among abstract concepts in the left hemisphere only, implying that semantic representations of abstract and concrete concepts are characterized along different organizational principles in the inferior parietal lobule. We observed null results for the decoding accuracy. Our study suggests that the use of the Representational Similarity Analysis as a complementary analysis to the decoding accuracy is a promising tool to reveal similarity patterns between theoretical models and brain activity recorded through fNIRS.

1. Introduction

Similarity/relatedness measures, the extent to which two concepts have similar meaning, can reveal the nature of semantic representation: the knowledge we have of the world (Montefinese, 2019; Vigliocco, Meteyard, Andrews, & Kousta, 2009). Numerous efforts have been made to understand the nature of semantic representation, assessing the extent to which a given theoretically based similarity measure predicts patterns in data. Some perspectives focus upon similarity in our affective and sensorimotor experience as inferred from verbal features (e.g., featural similarity; Montefinese, Vinson, & Ambrosini, 2018; Montefinese, Zannino, & Ambrosini, 2015; Vigliocco, Vinson, Lewis, & Garrett, 2004) or property ratings (e.g., affective content; Fairfield, Ambrosini, Mammarella, & Montefinese, 2017; Montefinese, Ambrosini, Fairfield, & Mammarella, 2014; Warriner, Kuperman, & Brysbaert, 2013), others upon regularities in spoken and written language (e.g., lexical co-occurrence; Andrews, Vigliocco, & Vinson, 2009; Griffiths, Steyvers, & Tenenbaum, 2007; Landauer & Dumais, 1997; Lund & Burgess, 1996). The similarity relation between two concepts may also be modelled as a measure of associative strength that reflects the probability of one concept to evoke another one in a free word association task (De Deyne, Navarro, Perfors, Brysbaert, & Storms, 2019; Nelson, McEvoy, & Schreiber, 2004). Although these similarity measures are correlated to each other to some extent in characterizing semantic representation, they are not entirely overlapping and seem to target different aspects of word meaning (Montefinese & Vinson, 2017). This allows us to investigate whether the different similarity measures also relate differently to the patterns of brain activity, and to identify which theoretical approach relates the most to observed effects of meaning similarity.

Relative effects of different theory-based similarity measures in predicting the participants' performance in different semantic tasks (such as, for example, lexical decision, Heyman et al., 2015; Brunelliere et al., 2017; Vigliocco et al., 2004; Montefinese et al., 2018) may also vary depending on concept concreteness: the degree to which a concept denoted by a word refers to an entity that can be perceived through the senses (Brysbaert, Stevens, De Deyne, Voorspoels, & Storms, 2014). This dimension is usually assessed by participants on Likert scales, in which concrete concepts lie on one side of the scale, referring to single, bounded, identifiable referents that can be perceived through the senses (Borghi et al., 2017). Abstract concepts lie on the opposite side of the scale and lack clearly

perceivable referents (even if they might evoke scenes and emotional experiences) and are more strongly reliant on interoception (i.e., sensations inside the body; Connell, Lynott, & Banks, 2018; Montefinese, Ambrosini, Visalli, & Vinson, 2020). Indeed, the latter are acquired later and mostly through language and social interaction compared to concrete concepts (see recent review by Dove, Barca, Tummolini & Borghi, 2020). By contrast, concrete concepts are more imageable (Paivio, 1990) and have greater availability of contextual information (Schwanenflugel et al., 1991).

Different organisational principles have been argued to govern semantic representations of concrete and abstract concepts: concrete concepts are predominantly organized by featural similarity and abstract concepts by associative relations (Crutch & Warrington, 2005; Hill, Kiela, & Korhonen, 2013). It has been proposed that abstract conceptual representation could be based more on linguistic information arising through patterns of co-occurrence and syntactic information (Gleitman & Papafragou, 2005; Vigliocco, Kousta, Vinson, Andrews, & Del Campo, 2013); however, some studies suggest in literature either lexical co-occurrence-based models behave similarly for concrete and abstract words concepts (Rotaru, Vigliocco, & Frank, 2018) or account better for concrete than abstract concepts (Hill et al., 2013).

Concrete concepts can be also inscribed into definite domains, such as natural kinds vs. artefacts, and they are organized into hierarchical categories, while abstract concepts are considerably more variable and not organized into well-defined categories (Borghi et al., 2017). Moreover, participants agree more when they produce properties and associations for concrete words like "dog" compared with abstract words like "justice" or "freedom" (De Mornay Davies & Funnell, 2000; Tyler, Moss, Galpin, & Voice, 2002). Together, these studies suggest indirectly that concepts' concreteness should be considered when comparing theory-based similarity measures.

In terms of the neural circuitry underlying the semantic system, two meta-analyses of neuroimaging studies showed a left-lateralized brain network (Binder, Desai, Graves, & Conant, 2009; Wang et al., 2010). Wang et al.'s meta-analysis identified neural differences in abstract and concrete concept representations in two main systems, with abstract concepts relying on the verbal-language system and concrete concepts relying on the imagery and perceptual systems (Wang et al., 2010). In addition, Binder et al.'s meta-analysis showed that the semantic system includes seven brain regions:

1) lateral temporal cortex, 2) ventral temporal cortex, 3) dorsomedial prefrontal cortex, 4) inferior frontal gyrus, 6) ventromedial prefrontal cortex, 6) posterior cingulate gyrus and 7) inferior parietal lobule (IPL) (i.e.,, the angular gyrus and the adjacent supramarginal gyrus; Pulvermüller, 2013). The common denominator of these regions is their role in high-level integrative processes. Indeed, they are known to receive extensively processed, multimodal and supramodal inputs. In particular, the IPL is assumed to have the role of a convergence hub sustaining integration in a multimodal system (Binder & Desai, 2011). It has been suggested indeed that IPL may integrate features for semantic categorization (Koenig et al., 2005) and enable increasingly abstract, supramodal representations of perceptual experience (Binder & Desai, 2011). In other words, this high-level convergence zone binds representations from two or more modalities, and the resulting supramodal representations capture similarity structures that define categories (Binder & Desai, 2011). The information for encoding abstract versus concrete concept representations in this region may thus reflect the abstract/concrete distinction on a semantic comparison level, as a consequence of the differences in either sensorimotor information from mental imagery or associated verbal contexts (Binder, Westbury, McKiernan, Possing, & Medler, 2005; Dhond, Witzel, Dale, & Halgren, 2007; Pexman, Hargreaves, Edwards, Henry, & Goodyear, 2007; Sabsevitz, Medler, Seidenberg, & Binder, 2005; J. Wang, Baucom, & Shinkareva, 2013; Wang, Conder, Blitzer, & Shinkareva, 2010). For these reasons, here we aim to further investigate which kind of information IPL integrates depending on the concept concreteness.

In recent years, the advancement in analytical approaches for fMRI data, like the development of multivariate pattern analysis (MVPA) techniques, has allowed a more direct investigation of abstract and concrete conceptual representations by examining, for example, whether the pattern of functional brain responses (e.g., fMRI voxels or functional Near Infrared Spectroscopy (fNIRS) channels) can discriminate between two stimuli conditions. Being a data-driven analysis, MVPA does not require any hypothesis and exhibits a higher sensitivity than traditional univariate statistical methods (Liu et al., 2013). Recently, this analytical approach has been successfully extended to the analysis of fNIRS data (Emberson, Zinszer, Raizada, & Aslin, 2017; Zinszer, Bayet, Emberson, Raizada, & Aslin, 2017). In particular, MVPA has been used in fNIRS studies for different purposes, such as to discriminate children with attention-deficit/hyperactivity disorder from healthy controls during a working memory

task (Gu et al., 2018), to decode visual and auditory stimuli in infants (Emberson et al., 2017), or to classify activation patterns associated with spoken and signed language in monolinguals (Mercure et al., 2020). To date, Zinszer and colleagues' study (2017) is the only one that investigated whether semantic representations are encoded in fNIRS neuroimaging data in a semantic task. The authors examined if participants' neural response to spoken words and their corresponding pictures predicts the pattern of similarity between concepts computed as a lexical co-occurrence measure. Results showed that neural activity pattern in the occipital cortex was predicted by the lexical co-occurrence across concepts. This work represents a substantial step forward in the investigation of semantic representation with fNIRS, moving from the conventional univariate analyses, typically used to localize which brain region is involved in semantic tasks by analysing one measurement location or channel at a time (e.g., Amiri et al., 2014), to multivariate methods, that consider the full spatial pattern of brain activity and channels simultaneously (Haynes and Rees, 2006).

Within this framework, Representational Similarity Analysis (RSA) (Kriegeskorte & Mur, M., & Bandettini, 2008), may represent a promising tool for fNIRS-based decoding of brain activity, enabling the comparison between data from different sources. For instance, the semantic similarity structure of a word set (that can be modelled in various ways) can be correlated with the similarity structure of the regional hemodynamic response pattern elicited by the same words. RSA was demonstrated to be particularly fruitful for fMRI research in investigating the relation between the functional response pattern and higher-level semantic representations between concepts, with a heavy focus on concrete knowledge (Devereux, Clarke, Marouchos, & Tyler, 2013; Fairhall & Caramazza, 2013; Kriegeskorte & Mur, M., & Bandettini, 2008; but see e.g., X. Wang et al., 2017 for abstract concepts). Here, we aim to further expand and show the potential of MVPA-based methods to decode brain activity patterns from fNIRS data.

Whilst there are limitations to fNIRS, such as the lack of anatomical information, a lower spatial resolution, and the ability to record only from the surface of the cortex, it presents some advantages that allow novel cognitive neuroscience investigations (Emberson et al., 2017; Pinti, Scholkmann, Hamilton, Burgess, & Tachtsidis, 2019; Zinszer et al., 2017). For example, unlike fMRI, fNIRS does not require participants to lie in a confined scanner environment, and it is more robust to movements,

making it suitable for a wide range of participant populations and tasks (Pinti et al., 2019). However, being a relatively recent technology, fNIRS still lacks standardized analysis procedures (Hocke et al., 2018; Pinti et al., 2019) and sophisticated analytical approaches comparable to those used for fMRI.

The specific question that we seek to answer is whether theory-based similarity measures predict the pattern of brain activity in IPL, and whether this differs for abstract and concrete words as recorded by fNIRS. By taking the finding from Zinszer et al. (2017) as a starting point, i.e., that fNIRS can be used successfully to predict neural patterns from concept similarity, we address this question by advancing in several ways by: 1) using a larger sample of stimuli, 2) comparing directly abstract and concrete concepts and alternative semantic models, 3) employing RSA which is particularly appropriate for comparing similarity patterns between data sets which can be from very different sources.

We hypothesize that MVPA can be used to classify brain activation in response to abstract and concrete words. We also predict that the IPL encodes different kinds of similarity measure depending on the word concreteness. Moreover, we hypothesize that information in the left hemisphere will be more critical than the one in the right hemisphere to this classification and similarity pattern.

2. Method

2.1. Participants

Thirteen healthy adults (four females; mean age = 26.7 years; range: 20 – 40 years) took part in the fNIRS study, which included two experimental sessions. A sensitivity power analysis (G*Power 3 software; Faul, Erdfelder, Lang, & Buchner, 2007) revealed that our sample size was large enough to have a statistical power of .80 to detect the significant (α = .05, two-tailed test) within-subject differences of interest (i.e., a main effect of hemisphere, word Category, or their interaction; see section 2.5.2) in either representation similarities or decoding accuracies with a medium effect size (Cohen's f = 0.3) assuming a correlation between repeated measures of .75.

All participants were right-handed, healthy and native English speakers with no history of neurological or psychiatric disorders and normal or corrected-to-normal vision. They were compensated for their participation and gave written informed consent to the experimental protocol approved by the University College London local research ethics committee. Due to technical issues, one participant

completed 4 out of 6 runs (1 run in session 1, 3 runs in session 2) and another participant 3 out of 6 runs (3 runs of session 1).

2.2. Stimuli

The experimental stimuli consisted of 160 English words denoting 80 abstract and 80 concrete concepts derived from the English semantic feature norms (Buchanan, Holmes, Teasley, & Hutchison, 2013; Buchanan, Valentine, & Maxwell, 2019; McRae, Cree, Seidenberg, & Mcnorgan, 2005; Vinson & Vigliocco, 2008). The words were assigned to abstract/concrete type based on a cut-off value estimated from the distribution of the 5-point concreteness ratings for English language (Brysbaert et al., 2014) for all words in the English semantic norms (Buchanan et al., 2014). In particular, the visual inspection of the distribution of concreteness ratings (estimated using a gaussian kernel smoothing function with an optimized bandwidth) indicated bimodality, which was confirmed by the Hartigan's dip test of unimodality ($p < .001$), suggesting the existence of (at least) two sub-distributions reflecting concreteness ratings for abstract and concrete words. The distribution was thus submitted to a gaussian mixture analysis, which confirmed that two sub-distributions were required to fit the data ($R^2 = .97$; a model including a third sub-distribution did not significantly increase the fit, $p > .05$). The cut-off value was then estimated as the concreteness value corresponding to the intersection of these sub-distributions.

The selection of experimental stimuli was restricted to the nouns for which the following English norms and measures were available that were necessary to compute the similarity measures for the RSA analysis (see below): semantic norms (Buchanan et al., 2013), lexical British National Corpus (Leech, Garside, & Bryant, 1994), association norms (De Deyne et al., 2019); affective norms (Warriner et al., 2013). In order to obtain a more representative set of abstract and concrete concepts while maximizing their variability (and, thus, RSA efficiency; see e.g., Meersmans et al., 2020), abstract and concrete concepts were selected based on a series of clustering analyses performed on the measures mentioned above, so that they could be grouped into four categories each, composed by a variable number of words (for the abstract concepts: social constructs, social attributes, cognitive events/states, and other abstract constructs; for the concrete concepts: professions, animals, vehicles, and buildings). The small number of words for each category, and their uneven number across categories, prevented us to further

investigate between-categories differences. Moreover, the experimental stimuli were selected so to maximize the range in semantic similarity within abstract and concrete stimuli while keeping correlations among the semantic, lexical, and association similarities as low as possible (see below). A complete list of the selected stimuli is provided in the Supplementary Material.

A two-tailed independent *t*-test confirmed the significant difference between abstract and concrete words for the concreteness measure (abstract: $M = 2.72$, $SD = 0.73$, concrete: $M = 4.61$, $SD = 0.35$; $t_{(78)} = 26.305$, $p < .0001$, Cohen's $d = 3.296$). Abstract and concrete stimuli were naturally balanced for word length, word frequency, valence, and dominance (Brysbaert et al., 2014; Warriner et al., 2013); as shown by two-tailed independent *t*-test comparisons ($t_{(78)} = 1.524$, $p > .132$, Cohen's $d < .241$). However, we refrained to balance them for further semantic-lexical variables. Indeed, as already noted, it is important to select variable stimuli for RSA (e.g., Meersmans et al., 2020) and matching abstract and concrete concepts on all semantic-lexical variables would have resulted in a stimuli set including highly specific concepts, which would have lowered RSA efficiency.

We obtained five types of representational dissimilarity matrices (RDMs) for abstract and concrete words, one for each of the similarity measures considered here, i.e. featural similarity, association strength, lexical co-occurrence, affective ratings, and orthographic similarity (as a control model). Each RDM is a symmetric $n \times n$ matrix, where n is the number of experimental conditions (i.e., $n = 160$ words in this study) and each off-diagonal element indicates the distance for each pair of words in a given measure. We maximised the range in semantic similarity within the concrete and abstract stimulus group as we assumed that a wider range of semantic similarities might increase the sensitivity of the representational similarity analysis of the fNIRS patterns (e.g., Meersmans et al., 2020).

The five similarity measures were computed as follows. The featural similarity measure was derived from English semantic norms in which participants were asked to list the properties of each word, such as its physical, functional, and categorical properties (Buchanan et al., 2013); it was calculated as the cosine angle between vectors for each pair of words (Buchanan et al., 2013). The association strength values were gathered by a continuous association task on large scale for English words (i.e., to produce multiple associations for each cue word) (De Deyne et al., 2019). Association strength was computed as the proportion of participants that gave the target response to a given cue

word. The lexical co-occurrence was computed as the natural logarithm of co-occurrence frequency plus one of each word pair in a symmetrical 5-word window in the British National Corpus (Leech, Garside, & Bryant, 1994). Since the results may be severely compromised whether these measures of similarity were too highly correlated, we tried to keep these correlations as low as possible (the within-domain correlations between the three measures of similarity were all lower than .27); we also selected materials that, while still representing high levels of similarity overall, varied substantially within each domain (i.e., abstract, concrete) in relative similarity according to the three measures. The affective similarities were calculated as the Euclidean distance in the three-dimensional space characterized by valence (the way an individual judges a situation, from pleasant to unpleasant), arousal (the degree of activation an individual feels toward a given stimulus, from calm to exciting) and dominance (the degree of control an individual feels over a specific stimulus, from out of control to in control) measured on a 9-point scale (Warriner et al., 2013). Finally, the orthographic distance values were calculated as the Levenshtein distance (OLD20, Yarkoni, Balota, & Yap, 2008) between each pair of words, reflecting the number of deletions, insertions and substitutions necessary to turn a word of the pair into the other one (for example, the orthographic distance between "career" and "deer" is three, reflecting two deletions ("c" and "a") and a substitution (from "r" to "d").

2.3. Procedure

The experimental task consisted of a semantic decision task in which the participants were asked to assess the concreteness (abstract vs. concrete) of the word stimuli. The words were presented at the centre of the screen in black helvetica font against a grey background.

Each subject participated to six 12-min fNIRS runs divided in two sessions, which were performed in separate days, one week apart. In each run, all the 160 experimental stimuli (word trials) were presented exactly once along with 40 interspersed null trials (consisting in the presentation of a fixation cross), whose orders were pseudorandomized so that there were no more than three concepts in a row from the same category (abstract, concrete) and no null event repetitions. By repeating experimental stimuli six times, but in different runs, we were able to maximize design efficiency (and the precision of the estimates) while avoiding temporal clustering, which is undesirable in this kind of design (Kriegeskorte et al., 2008). Different trial orders were used per subject.

The trial duration was 3 s, during which a fixation cross (200 ms) preceded the stimulus (i.e., a word or a null event, presented for 500 ms) which in turn was followed by the blank screen (ITI: 2300 ms) in which participants could make their response by pressing the right or left arrow keys on a computer keyboard. The input key mapping was counterbalanced across participants. Participants were required to perform a semantic decision task, in which they judged whether the word denoted an abstract or concrete concept.

The fact that each word was repeated six times during the experiment introduces a possible source of noise because of novelty effects on the first presentation. We therefore asked participants to perform a paper-and-pencil version of the semantic decision task on the same stimuli of the computerized version at the beginning of the first experimental session to reduce any novelty signals during the fNIRS session.

2.4. fNIRS data acquisition

fNIRS, like fMRI, is a neuroimaging technique based on neurovascular coupling that measures the changes in brain hemodynamics (oxyhemoglobin -HbO_2- and deoxyhemoglobin -HbR) following neuronal activations using near-infrared light (Pinti et al., 2019). brain hemodynamic and oxygenation changes were recorded over the inferior parietal cortex bilaterally using a wearable fNIRS device (LIGHTNIRS, Shimadzu Corp., Kyoto, Japan). The fNIRS system is equipped with 8 light sources, emitting light at 780, 805, and 830 nm, and 8 light detectors that were split and arranged into two 4×2 arrays as shown in Figure 1. The source-detector separation was set at 3 cm. Raw intensity signals were sampled at 13.33 Hz from 20 measurement channels, 10 per hemisphere.

For each participant, the coordinates of the fNIRS optodes and of the 10-20 standard anatomical landmarks (Nasion, Inion, right and left pre-auricular points, Cz) were recorded using a 3D magnetic digitizer (Liberty, Polhemus, Vermont). These were used to co-register the fNIRS channels locations onto a standard brain template using the SPM for fNIRS toolbox (https://www.nitrc.org/projects/spm_fnirs). The MNI coordinates and the anatomical locations of each channel were then estimated. The median group MNI coordinates, the corresponding anatomical locations, and the atlas-based probabilities are listed in Supplementary Table 1.
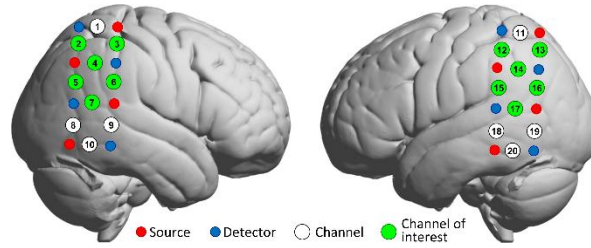
Figure 1. Schematic representation of the fNIRS channel configuration. Sources (red circles) and detectors (blue circles) are arranged in a 4 x 2 configuration on each hemisphere, creating 20 measurement channels (white circles) in total. The channels of interest, covering the left and right IPL, are indicated in green.

2.5. fNIRS data pre-processing and analysis

The Homer2 software package (Huppert, Diamond, Franceschini, & Boas, 2009) was used to preprocess the fNIRS signals. Raw intensity data were first visually inspected and the presence of the heart beat frequency (~1 Hz) in the signal power spectral density was assessed in order to identify the channels showing poor signal-to-noise due to detectors saturation or poor optical coupling (Pinti et al., 2019). No channels were excluded from further analysis due to poor signal quality. Raw intensity signals were converted into optical density changes (Homer2 function: *hmrInteensity2OD*) and motion artifacts were corrected using the wavelet-based approach (Homer2 function: *hmrInteensity2OD*; *iqr* = 1.5; Molavi & Dumont, 2012). In order to reduce high-frequency (e.g., heart rate) and very low frequency noise, a band-pass filter was applied (Homer2 function: *hmrBandpassFilt*; order: 3rd; band-pass frequency range [0.01 0.6] Hz) and the concentration changes of HbO_2 and HbR were then calculated with the modified Beer-Lambert law (Homer2 function: *hmrOD2Conc*; DPF = 6). The correlation-based signal improvement (CBSI, Cui, Bray, & Reiss, 2010) was used to combine HbO_2 and HbR into the so-called 'activation signal' (Scholkmann et al., 2014). This was done to infer functional brain activity on one signal including both HbO_2 and HbR at the same time that can help in reducing false positives at the statistical analysis stage (Tachtsidis & Scholkmann, 2016).

2.5.1. Contrast effects analysis

A General Linear model (GLM) approach (Friston et al., 1994) was adopted to estimate the first-level (or single-subject) *β*-values on the CBSI signals. This was done using the *SPM for fNIRS* toolbox on the fNIRS activation signals, downsampled to 3 Hz, and for each of the 6 runs individually. Specifically, our design matrix included 160 regressors, one for each word, computed through the convolution of the

events timeline (modelled as stick functions) with the canonical hemodynamic response function. The regressors were used to fit to the fNIRS activations signals and the single-subject β-values were estimated. Then, for each word we computed the t-statistics on the β estimates testing for the hypothesis that that word was significantly related to the CBSI signal. This was carried out on the channels covering the IPL bilaterally (green-filled circles in Figure 1).

2.5.2. Representational Similarity Analysis and multivariate classification analysis

We performed a Representational Similarity (RS) Analysis based on Spearman partial correlations computed between, on the one side, a similarity matrix based on the neural activity patterns (brain) and, on the other side, the similarity measures across abstract and concrete concepts (models). The brain similarity matrix was computed as the correlation between the neural activations for each pair of words. The five theoretical models were based on featural similarity, word association, lexical co-occurrence, affective content, and orthographic similarity (control model), as detailed in the Stimuli section. We also performed a leave-one-out item-level multivariate classification analysis, carried out using the procedure in Emberson et al. (2017; see also, Anderson, Zinszer, & Raizada, 2016) to decode the concreteness category of single words (trial-level decoding).

Subject-wise RSs and mean trial-level decoding accuracies were compared between hemispheres and abstract and concrete categories with within-subjects GLM analysis. Post-hoc one-tailed one-sample t-tests were performed to test for simple effects (i.e., RSs and decoding accuracies significantly higher than 0 and .5, respectively). The same analysis was carried out on HbO_2 and HbR separately and results are included in the Supplementary Material.

3. Results

3.1. Representational Similarity Analysis

We found different brain-model RSs for co-occurrence between abstract and concrete words, regardless of the hemisphere [$F_{(1,12)} = 16.00$, $p = .002$, Cohen's $f = 0.55$; Figure 2]. This was due to a greater (and significant) RS for concrete [$t_{(12)} = 2.13$, $p = .027$] as compared to abstract words [$t_{(12)} = -2.76$, $p = .955$]. The hemisphere by concreteness interaction was not significant [$F_{(1,12)} = 0.20$, $p = .662$, $f = 0.06$].

We also found different brain-model RSs for affective content between hemispheres and word concreteness [hemisphere by concreteness interaction: $F_{(1,12)} = 6.62$, $p = .024$, $f = 0.36$]. This interaction

was explained by a significantly greater (and significant) RS for abstract [$t_{(12)} = 2.02$, $p = .033$] as compared to concrete [$t_{(12)} = -1.20$, $p = .873$] words for the left hemisphere; the RSs for the abstract [$t_{(12)} = -1.44$, $p = .913$] and concrete [$t_{(12)} = 0.83$, $p = .210$] words were not significant for the right hemisphere.

No significant concreteness effect or hemisphere by concreteness interaction was found for the brain-model RSs for featural similarity [respectively, $F_{(1,12)} = 1.00$ and 0.19, $p = .337$ and $.067$, $f = 0.14$ and 0.06], word association [respectively, $F_{(1,12)} = 1.81$ and 0.07, $p = .203$ and $.792$, $f = 0.19$ and 0.04], and orthographic similarity [respectively, $F_{(1,12)} = 1.51$ and 2.31, $p = .241$ and $.155$, $f = 0.17$ and 0.21].
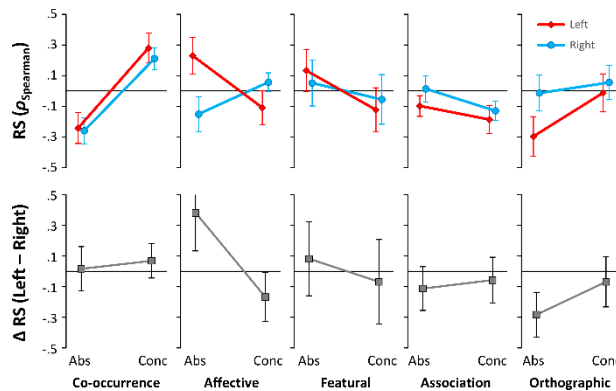

Figure 2. RSA results. The plots in the upper row show the mean representational similarity (RS, Spearman partial correlations) as a function of word concreteness (Abs, abstract; Conc, concrete) and hemisphere (left, in red; right, in blue) for each of the five theoretical models. The plots in the lower row show the corresponding differences between RSs in the left and right hemisphere as a function of word concreteness. Error bars represent within-subjects standard errors of the mean (Morey, 2008).

3.2. Decoding Analysis

No significant concreteness effect or hemisphere by concreteness interaction was found for the brain-model RSs for featural similarity [respectively, $F_{(1,12)} = 4.09$ and 0.01, $p = .066$ and $.958$, $f = 0.28$ and 0.01], co-occurrence [respectively, $F_{(1,12)} = 0.14$ and 1.02, $p = .717$ and $.333$, $f = 0.05$ and 0.14], and orthographic similarity [respectively, $F_{(1,12)} = 3.95$ and 0.92, $p = .070$ and $.356$, $f = 0.28$ and 0.13]. A significant concreteness effect was found for word association [$F_{(1,12)} = 17.40$, $p = .001$, $f = .58$], but the decoding accuracy was not significantly above chance level for both abstract and concrete words (all ps > .621). Moreover, a significant hemisphere by concreteness interaction was found for affective content [$F_{(1,12)} = 6.14$, $p = .029$, $f = .34$], but again the decoding accuracy was not significantly above chance level in any case (all ps > .104).

4. Discussion

In this study we tested for the first time whether theory-based similarity measures predicted the pattern of brain activity in the IPL differently for abstract and concrete words. To this aim, we asked native English speakers to perform a semantic decision task requiring an explicit coding of the concreteness dimension of the words, while we recorded their brain activity in IPL through fNIRS. Using RSA on fNIRS data, we found that the neural representational similarity in IPL conformed to the lexical co-occurrence among concrete concepts (regardless of the hemisphere) and to the affective similarity among abstract concepts in the left hemisphere only. This concordance between neural and semantic-affective relationships within supramarginal and angular gyri suggests that these regions encode semantic-affective information depending on word concreteness.

The role of the IPL in semantic processing has been supported by convergent evidence from human functional imaging studies (Binder & Desai, 2011; Binder et al., 2009; J. Wang, Conder, Blitzer, & Shinkareva, 2010), implicating this region in semantic representational aspects. Although this evidence would suggest that IPL may encode feature similarity among concepts, our results do not confirm this assumption. Indeed, we did not find a significant similarity pattern between feature similarity between concepts and neural activation pattern in IPL. Rather we found a similarity pattern in this region with lexical co-occurrence between concrete concepts in IPL for both hemispheres. These results appear consistent with a meta-analysis of 120 fMRI and PET studies on semantic processing that identified the bilateral angular gyrus to be consistently engaged more in concrete than in abstract concept processing (Binder et al., 2009), but this does not explain why we found a reliable similarity pattern only for lexical co-occurrence and not for feature similarity. The IPL also responds to statistical regularity of meaningful events (words/pictures sequences) (Hoenig & Scheef, 2009; Kuperberg et al., 2003; Tinaz, Schendan, Schon, & Stern, 2006; Tinaz, Schendan, & Stern, 2008) and it is part of a context-related processing network (Bar, Aminoff, & Schacter, 2008; Fornito, Harrison, Zalesky, & Simons, 2012). As lexical co-occurrence captures statistical regularities in language and reflects co-occurrence of words in similar contexts, IPL may be sensitive to regularities in written and spoken language as well as to the spatial and temporal regularities of events. Support for this view comes from a body of evidence revealing that the text-based model predicted activity in a distributed network extending to the bilateral inferior parietal lobule during language and semantic processing (Anderson,

Bruni, Lopopolo, Poesio, & Baroni, 2015; Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016) Indeed, like in our study, Huth and colleagues (2016) found that semantic representation of concepts modelled as a corpus-based space recruits both hemispheres for comprehension of natural speech. But this does not explain why we only observed a reliable similarity pattern only for concrete, and not also for abstract words. This result was unexpected and at odds with the view that sensorimotor properties would be more relevant in IPL for concrete concepts, and linguistic (or at least relational) properties would be more relevant for abstract (Crutch & Warrington, 2005). If confirmed in future studies, this result would highlight the importance of the linguistic and context information for concrete concepts as well.

It is also worth noting that single word processing in a semantic task (as in our study) elicits more context-related information (Price, 2010) compared to the processing of a word embedded in a sentence. Concrete concepts are strongly associated to a limited number of contexts compared to abstract concepts (i.e., abstract concepts have greater contextual diversity; see Brysbaert et al., 2014; Schwanenflugel, 1991) and are thus processed more efficiently, particularly when little or no context is provided with the concept presentation. This characteristic is also reflected in our set of stimuli (two-tailed independent *t* test on contextual diversity values shows a difference between abstract and concrete concepts: $t_{(78)} = 3.80$, $p < .001$, $d = .589$). On this view, in our study the greater availability of context information for concrete concepts conveyed by lexical co-occurrence may have determined the similarity pattern between corpus-based space and activity pattern of IPL for concrete concepts only. It remains to be seen whether this is an important representational difference, related to this broad difference between concrete and abstract concepts overall, or a product of this characteristic of our stimulus set.

We also found a similarity pattern between affective content-based model and activity pattern in IPL for abstract concepts. While we did not have a strong hypothesis about the IPL's role in abstract representation, this pattern of results was not surprising: while concrete concepts usually have direct sensory referents (Crutch & Warrington, 2005; Montefinese, Ambrosini, Fairfield, & Mammarella, 2013b, 2013a), abstract concepts tend to be more emotionally valenced (Crutch, Troche, Reilly, & Ridgway, 2013; Kousta et al., 2011; Montefinese et al., 2020; Vigliocco et al., 2014) and have low

sensorimotor grounding (for a concise review, Montefinese, 2019). Indeed, several authors have emphasized the peculiar role of affective and social experiences in semantic representation of abstract concepts (Binder et al., 2016; Borghi, Flumini, Cimatti, Marocco, & Scorolli, 2011; Katja Wiemer-Hastings & Xu, 2005; Kousta et al., 2011; Vigliocco et al., 2009; Zdrazilova & Pexman, 2013). brain activations in the left (or bilateral) IPL have been observed in response to emotional salience of words (Kensinger & Schacter, 2006; Kensinger & Corkin, 2004; Skipper & Olson, 2014). brain activations in the left IPL have also been shown to be modulated by the presence of both valence and arousal dimensions of the word (Kensinger & Corkin, 2004).

Together, the results of our study indicate that the IPL represents distributional and affective information depending on the word concreteness, implying that semantic representations of abstract and concrete concepts are characterized along different organizational principles in the IPL. Although some fMRI studies found a similarity pattern between the word association based-model and other brain regions (Liuzzi et al., 2019; Meersmans et al., 2020), we did not find a significant similarity pattern between this model and activity pattern in the IPL. These results suggest that the association strength for written words could represent less similarity between concepts compared to the other measures in the IPL. However, future research is needed to confirm this conclusion.

We also observed unexpected, null results for the decoding accuracy: i.e., the decoding accuracy was not significantly above chance level with any semantic model. This could be related to the intrinsic limitations of the fNIRS technology and the low number of encoding channels we used. In fact, fNIRS instruments are typically equipped with a limited number of channels and, in general, brain activity is sampled at sparse and discrete locations of the cortex, with optodes placed 2-3 cm apart from each other. fNIRS has also a spatial resolution of 2-3 cm and recordings are much less finely grained than fMRI (Emberson et al., 2017). Therefore, MVPA is applied on a lower number of encoding channels than fMRI and these include information from a bigger brain volume than the fMRI voxels, which can impact the performance of the decoding procedure. Nonetheless, our study showed that the use of the RSA as a complementary analysis to the decoding accuracy is a promising tool to reveal similarity pattern between theoretical models and brain activity recorded through fNIRS. Future technological advances can help in improving the performance of MVPA-based analyses on fNIRS data

and fully establish multivariate analyses as a solid method for fNIRS-derived brain decoding. For instance, the development of whole-head diffuse optical tomography systems can increase the data quality, depth and spatial resolution of fNIRS recordings by using dense grids of sources and detectors with different separations and overlapping measurements at different depths (Zhao & Cooper, 2017).

CONFLICT OF INTEREST
All authors declare that they have no financial or other conflicts of interest.

ORCID
Maria Montefinese https://orcid.org/0000-0002-7685-1034
Paola Pinti https://orcid.org/0000-0002-6667-3481
Ettore Ambrosini https://orcid.org/0000-0002-1439-5491
Ilias Tachtsidis https://orcid.org/0000-0002-8125-0313
David Vinson https://orcid.org/0000-0002-6226-3258

References

Amiri, M., Pouliot, P., Bonnéry, C., Leclerc, P. O., Desjardins, M., Lesage, F., & Joanette, Y. (2014). An exploration of the effect of hemodynamic changes due to normal aging on the fNIRS response to semantic processing of words. *Frontiers in neurology, 5,* 249. https://doi.org/10.3389/fneur.2014.00249

Anderson, A. J., Bruni, E., Lopopolo, A., Poesio, M., & Baroni, M. (2015). Reading visually embodied meaning from the brain: Visually grounded computational models decode visual-object mental imagery induced by written text. *NeuroImage*, *120*, 309–322. https://doi.org/10.1016/j.neuroimage.2015.06.093

Anderson, A. J., Zinszer, B. D., & Raizada, R. D. S. (2016). Representational similarity encoding for fMRI: Pattern-based synthesis to predict brain activity using stimulus-model-similarities. *NeuroImage*, *128*, 44–53. https://doi.org/10.1016/j.neuroimage.2015.12.035

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, *116*(3), 463–498. https://doi.org/10.1037/a0016261

Bar, M., Aminoff, E., & Schacter, D. L. (2008). Scenes unseen: The parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *Journal of Neuroscience*, *28*(34), 8539–8544. https://doi.org/10.1523/JNEUROSCI.0987-08.2008

Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., &

Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology*, *33(3-4)*, 130–174. https://doi.org/10.1080/02643294.2016.1147426

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527–536. https://doi.org/10.1016/j.tics.2011.10.001

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*(December), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, *17*(6), 905–917. https://doi.org/10.1162/0898929054021102

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, *143*(3), 263–292. https://doi.org/10.1037/bul0000089

Borghi, A. M., Flumini, A., Cimatti, F., Marocco, D., & Scorolli, C. (2011). Manipulating objects and telling words: a study on concrete and abstract words acquisition. *Frontiers in Psychology*, *2*(February), 15. https://doi.org/10.3389/fpsyg.2011.00015

Brunellière, A., Perre, L., Tran, T., & Bonnotte, I. (2017). Co-occurrence frequency evaluated with large language corpora boosts semantic priming effects. *Quarterly Journal of Experimental Psychology*, *70*(9), 1922–1934. https://doi.org/10.1080/17470218.2016.1215479

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods, 46*(3), 904-911. https://doi.org/10.3758/s13428-013-0403-5

Buchanan, E. M., Holmes, J. L., Teasley, M. L., & Hutchison, K. A. (2013). English semantic word-pair norms and a searchable Web portal for experimental stimulus creation. *Behavior Research Methods*, *45*(3), 746–757. https://doi.org/10.3758/s13428-012-0284-z

Buchanan, E. M., Valentine, K. D., & Maxwell, N. P. (2019). English semantic feature production norms: An extended database of 4436 concepts. *Behavior Research Methods*, *51*(4), 1849–1863. https://doi.org/10.3758/s13428-019-01243-z

Connell, L., Lynott, D., & Banks, B. (2018). Interoception: the forgotten modality in perceptual grounding of abstract and concrete concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), 20170143. https://doi.org/10.1098/rstb.2017.0143

Crutch, S. J., Troche, J., Reilly, J., & Ridgway, G. R. (2013). Abstract conceptual feature ratings: the role of emotion, magnitude, and other cognitive domains in the organization of abstract conceptual knowledge. *Frontiers in Human Neuroscience*, *7*(May), 1–14. https://doi.org/10.3389/fnhum.2013.00186

Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *brain*, *128*(3), 615–627. https://doi.org/10.1093/brain/awh349

Cui, X., Bray, S., & Reiss, A. L. (2010). Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *NeuroImage*, *49*(4), 3039–3046. https://doi.org/10.1016/j.neuroimage.2009.11.050

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987–1006. https://doi.org/10.3758/s13428-018-1115-7

De Mornay Davies, P., & Funnell, E. (2000). Semantic representation and ease of predication. *brain and Language*, *73*(1), 92–119. https://doi.org/10.1006/brln.2000.2299

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. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(48), 18906–18916. https://doi.org/10.1523/JNEUROSCI.3809-13.2013

Dhond, R. P., Witzel, T., Dale, A. M., & Halgren, E. (2007). Spatiotemporal cortical dynamics underlying abstract and concrete word reading. *Human brain Mapping*, *28*(4), 355–362. https://doi.org/10.1002/hbm.20282

Dove, G., Barca, L., Tummolini, L., & Borghi, A. M. (2020 in press, Psychological Research). Words have a weight: language as a source of inner grounding and flexibility in abstract concepts. *Psychological Research* PsychArxiv. doi: https://doi.org/10.31234/osf.io/j6xhe

Emberson, L. L., Zinszer, B. D., Raizada, R. D. S., & Aslin, R. N. (2017). Decoding the infant mind: Multivariate pattern analysis (MVPA) using fNIRS. *PLOS ONE*, *12*(4), e0172500. https://doi.org/10.1371/journal.pone.0172500

Fairfield, B., Ambrosini, E., Mammarella, N., & Montefinese, M. (2017). Affective Norms for Italian Words in Older Adults: Age Differences in Ratings of Valence, Arousal and Dominance. *PloS ONE*, *12*(1). https://doi.org/10.1371/journal.pone.0169472

Fairhall, S. L., & Caramazza, A. (2013). brain Regions That Represent Amodal Conceptual Knowledge. *Journal of Neuroscience*, *33*(25), 10552–10558. https://doi.org/10.1523/JNEUROSCI.0051-13.2013

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fornito, A., Harrison, B. J., Zalesky, A., & Simons, J. S. (2012). Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(31), 12788–12793. https://doi.org/10.1073/pnas.1204185109

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. -P, Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human brain Mapping*, *2*(4), 189–210. https://doi.org/10.1002/hbm.460020402

Gleitman, L., & Papafragou, A. (2005). Language and Thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (p. 633–661). Cambridge, UK: Cambridge University Press.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*(2), 211–244. https://doi.org/10.1037/0033-295X.114.2.211

Gu, Y., Miao, S., Han, J., Liang, Z., Ouyang, G., Yang, J., & Li, X. (2018). Identifying ADHD children using hemodynamic responses during a working memory task measured by functional near-infrared spectroscopy. *Journal of Neural Engineering*, *15*(3), 035005. https://doi.org/10.1088/1741-2552/aa9ee9

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience, 7*(7), 523-534. https://doi.org/10.1038/nrn1931

Heyman, T., Van Rensbergen, B., Storms, G., Hutchison, K. a, & De Deyne, S. (2015). The influence of working memory load on semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 911–920. https://doi.org/10.1037/xlm0000050

Hill, F., Kiela, D., & Korhonen, A. (2013). Concreteness and corpora: A theoretical and practical analysis. CMCL 2013, page 75, Sofia.

Hocke, L., Oni, I., Duszynski, C., Corrigan, A., Frederick, B., & Dunn, J. (2018). Automated Processing of fNIRS Data—A Visual Guide to the Pitfalls and Consequences. *Algorithms*, *11*(5), 67. https://doi.org/10.3390/a11050067

Hoenig, K., & Scheef, L. (2009). Neural correlates of semantic ambiguity processing during context verification. *NeuroImage*, *45*(3), 1009–1019. https://doi.org/10.1016/j.neuroimage.2008.12.044

Huppert, T. J., Diamond, S. G., Franceschini, M. A., & Boas, D. A. (2009). HomER: A review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied Optics*, *48*(10), D280–D298. https://doi.org/10.1364/AO.48.00D280

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. https://doi.org/10.1038/nature17637

Katja Wiemer-Hastings, K., & Xu, X. (2005). Content Differences for Abstract and Concrete Concepts. *Cognitive Science*, *29*(5), 719–736. https://doi.org/10.1207/s15516709cog0000_33

Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective and Behavioral Neuroscience*, *6*(2), 110–126. https://doi.org/10.3758/CABN.6.2.110

Kensinger, E. a, & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences*, *101*(9), 3310–3315. https://doi.org/10.1073/pnas.0306408101

Koenig, P., Smith, E. E., Glosser, G., DeVita, C., Moore, P., McMillan, C., … Grossman, M. (2005). The neural basis for novel semantic categorization. *NeuroImage*, *24*, 369–383. https://doi.org/10.1016/j.neuroimage.2004.08.045

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, *140*(1), 14–34. https://doi.org/10.1037/a0021446

Kriegeskorte, N., & Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4. https://doi.org/10.3389/neuro.06.004.2008

Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience*, *15*(2), 272–293. https://doi.org/10.1162/089892903321208204

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*. https://doi.org/10.1037/0033-295X.104.2.211

Leech, G., Garside, R., & Bryant, M. (1994). CLAWS4. In *Proceedings of the 15th conference on Computational linguistics* - (Vol. 1, p. 622). Morristown, NJ, USA: Association for Computational Linguistics. https://doi.org/10.3115/991886.991996

Levenshtein, Vladimir I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Cybernetics and Control Theory, 10*(8), 707–710.

Liu, P., Qin, W., Wang, J., Zeng, F., Zhou, G., Wen, H., … Tian, J. (2013). Identifying Neural Patterns of Functional Dyspepsia Using Multivariate Pattern Analysis: A Resting-State fMRI Study. *PLoS ONE*, *8*(7), e68205. https://doi.org/10.1371/journal.pone.0068205

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. https://doi.org/10.1016/j.neuroimage.2019.02.011

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203–208. https://doi.org/10.3758/BF03204766

McRae, K., Cree, G. S., Seidenberg, M. S., & Mcnorgan, C. (2005). Semantic feature

production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559. https://doi.org/10.3758/BF03192726

Meersmans, K., Bruffaerts, R., Jamoulle, T., Liuzzi, A. G., De Deyne, S., Storms, G., … Vandenberghe, R. (2020). Representation of associative and affective semantic similarity of abstract words in the lateral temporal perisylvian language regions. *NeuroImage*, *217*, 116892. https://doi.org/10.1016/j.neuroimage.2020.116892

Mercure, E., Evans, S., Pirazzoli, L., Goldberg, L., Bowden-Howl, H., Coulson-Thaker, K., … MacSweeney, M. (2020). Language Experience Impacts brain Activation for Spoken and Signed Language in Infancy: Insights From Unimodal and Bimodal Bilinguals. *Neurobiology of Language*, *1*(1), 9–32. https://doi.org/10.1162/nol_a_00001

Molavi, B., & Dumont, G. A. (2012). Wavelet-based motion artifact removal for functional near-infrared spectroscopy. *Physiological Measurement*, *33*(2), 259–270. https://doi.org/10.1088/0967-3334/33/2/259

Montefinese, M. (2019). Semantic representation of abstract and concrete words: A minireview of neural evidence. *Journal of Neurophysiology*, *121*(5), 1585–1587. https://doi.org/10.1152/jn.00065.2019

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2013a). Semantic memory: A feature-based analysis and new norms for Italian. *Behavior Research Methods*, *45*(2), 440–461. https://doi.org/10.3758/s13428-012-0263-4

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2013b). The "subjective" pupil old/new effect: Is the truth plain to see? *International Journal of Psychophysiology*, *89*(1), 48–56. https://doi.org/10.1016/j.ijpsycho.2013.05.001

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behavior Research Methods*, *46*(3), 887–903. https://doi.org/10.3758/s13428-013-0405-3

Montefinese, M., Ambrosini, E., Visalli, A., & Vinson, D. (2020). Catching the intangible: A role for emotion? *Behavioral and brain Sciences*, *43*. https://doi.org/10.1017/S0140525X19002978

Montefinese, M., Buchanan, E. M., & Vinson, D. (2018, June 20). How well do similarity measures predict priming in abstract and concrete concepts?. https://doi.org/10.31234/osf.io/ypvgw

Montefinese, M., & Vinson, D. (2017). Resemblance among similarity measures in semantic representation. Retrieved August 27, 2020, from https://cogsci.mindmodeling.org/2017/papers/0809/

Montefinese, M., Vinson, D., & Ambrosini, E. (2018). Recognition memory and featural similarity between concepts: The pupil's point of view. *Biological Psychology*, *135*, 159–169. https://doi.org/10.1016/j.biopsycho.2018.04.004

Montefinese, M., Zannino, G. D., & Ambrosini, E. (2015). Semantic similarity between old and new items produces false alarms in recognition memory. *Psychological Research*, *79*(5), 785–794. https://doi.org/10.1007/s00426-014-0615-z

Morey R.D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology, 4*, 61-64. https://doi.org/10.20982/tqmp.04.2.p061

Nelson, D. L., McEvoy, C. L., & Schreiber, T. a. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Paivio, A. (1990). Dual Coding Theory. *Mental Representations,* 53–83. https://doi.org/10.1093/acprof:oso/9780195066661.003.0004

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). Neural correlates of concreteness in semantic categorization. *Journal of Cognitive*

Neuroscience, *19*(8), 1407–1419. https://doi.org/10.1162/jocn.2007.19.8.1407

Pinti, P., Scholkmann, F., Hamilton, A., Burgess, P., & Tachtsidis, I. (2019). Current Status and Issues Regarding Pre-processing of fNIRS Neuroimaging Data: An Investigation of Diverse Signal Filtering Methods Within a General Linear model Framework. *Frontiers in Human Neuroscience*, *12*, 505. https://doi.org/10.3389/fnhum.2018.00505

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, *1191*(1), 62–88. https://doi.org/10.1111/j.1749-6632.2010.05444.x

Pulvermüller, F. (2013, September 1). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*. Elsevier Current Trends. https://doi.org/10.1016/j.tics.2013.06.004

Rotaru, A. S., Vigliocco, G., & Frank, S. L. (2018). modeling the Structure and Dynamics of Semantic Processing. *Cognitive Science*, *42*(8), 2890–2917. https://doi.org/10.1111/cogs.12690

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

Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Mata Pavia, J., Wolf, U., & Wolf, M. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage*, *85*, 6–27. https://doi.org/10.1016/j.neuroimage.2013.05.004

Schwanenflugel, P. J. (1991). Capter 2 Contextual Constraint and Lexical Processing. *Advances in Psychology*, *77*(C), 23–45. https://doi.org/10.1016/S0166-4115(08)61528-9

Skipper, L. M., & Olson, I. R. (2014). Semantic memory: Distinct neural representations for abstractness and valence. *brain and Language*, *130*, 1–10. https://doi.org/10.1016/j.bandl.2014.01.001

Tachtsidis, I., & Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. *Neurophotonics*, *3*(3), 031405. https://doi.org/10.1117/1.nph.3.3.031405

Tinaz, S., Schendan, H. E., Schon, K., & Stern, C. E. (2006). Evidence for the importance of basal ganglia output nuclei in semantic event sequencing: An fMRI study. *brain Research*, *1067*(1), 239–249. https://doi.org/10.1016/j.brainres.2005.10.057

Tinaz, S., Schendan, H. E., & Stern, C. E. (2008). Fronto-striatal deficit in Parkinson's disease during semantic event sequencing. *Neurobiology of Aging*, *29*(3), 397–407. https://doi.org/10.1016/j.neurobiolaging.2006.10.025

Tyler, L. K., Moss, H. E., Galpin, A., & Voice, J. K. (2002). Activating meaning in time: The role of imageability and form-class. *Language and Cognitive Processes*, *17*(5), 471–502. https://doi.org/10.1080/01690960143000290

Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The Neural Representation of Abstract Words: The Role of Emotion. *Cerebral Cortex*, *24*(7), 1767–1777. https://doi.org/10.1093/cercor/bht025

Vigliocco, G., Kousta, S., Vinson, D., Andrews, M., & Del Campo, E. (2013). The representation of abstract words: What matters? Reply to Paivio's (2013) comment on Kousta et al. (2011). *Journal of Experimental Psychology: General*, *142*(1), 288–291. https://doi.org/10.1037/a0028749

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, *1*(02), 219–247. https://doi.org/10.1515/LANGCOG.2009.011

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive*

Psychology, *48*(4), 422–488. https://doi.org/10.1016/j.cogpsych.2003.09.001

Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*(1), 183–190. https://doi.org/10.3758/BRM.40.1.183

Wang, J., Baucom, L. B., & Shinkareva, S. V. (2013). Decoding abstract and concrete concept representations based on single-trial fMRI data. *Human brain Mapping*, *34*(5), 1133–1147. https://doi.org/10.1002/hbm.21498

Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human brain Mapping*, *31*(10), 1459–1468. https://doi.org/10.1002/hbm.20950

Wang, X., Wu, W., Ling, Z., Xu, Y., Fang, Y., Wang, X., … Bi, Y. (2017). Organizational Principles of Abstract Words in the Human brain. *Cerebral Cortex*, (November), 1–14. https://doi.org/10.1093/cercor/bhx283

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Zdrazilova, L., & Pexman, P. M. (2013). Grasping the invisible: semantic processing of abstract words. *Psychonomic Bulletin & Review*, *20*(6), 1312–1318. https://doi.org/10.3758/s13423-013-0452-x

Zhao, H., & Cooper, R. J. (2017). Review of recent progress toward a fiberless, whole-scalp diffuse optical tomography system. *Neurophotonics*, *5*(01), 1. https://doi.org/10.1117/1.nph.5.1.011012

Zinszer, B. D., Bayet, L., Emberson, L. L., Raizada, R. D. S., & Aslin, R. N. (2017). Decoding semantic representations from functional near-infrared spectroscopy signals. *Neurophotonics, 5(*01), 1. https://doi.org/10.1117/1.NPh.5.1.011003