A Study of Representational Similarity: The Emergence of Object Concepts in Rapid Serial Visual Presentation Streams

Authors:

Ivy Zhou [1], Tijl Grootswagers [1, 2, 3], Blake Segula [1], Amanda Robinson [1, 2, 3,], Sophia Shatek [1], Christopher Whyte [1], Thomas Carlson [1, 2]

Institutions:

1 University of Sydney, Sydney, NSW, 2006 Australia, 2 ARC Centre of Excellence in Cognition and its Disorders, Sydney, 2109, NSW, Australia, 3 Department of Cognitive Science, Macquarie University, Sydney, NSW, 2109, Australia.

Abstract

Guided by the observation that similar words in language occur in similar contexts, linguistic computational models trained on statistics of word co-occurrence in texts were shown to be effective in modelling both human performance in psycholinguistic tasks and semantically imbued representations in the brain. But, it remains unclear whether the semantic representation extracted from the distributional behavior of words in the natural language resemble the knowledge that is primarily acquired outside language, through sensory-perceptual experience. Using Representational Similarity Analysis (RSA), the present study endeavours to identify a direct link between the neural representation of object concepts and computational modelling. The broad aim of this study is to examine the extent to which neural representation of object concepts can be modelled by two types of linguistic computational models: distributional word co-occurrence and lexical hierarchical. The more specific aim of the study is to investigate which of the two types of semantic structure, distributional or hierarchical, best explains the time-varying neural representation of object concepts in the brain. Subsequently, this study first applied time-resolved Multivariate Pattern Analysis (MVPA) to neural responses evoked by naturalistic images portraying a broad and large (n = 1854) set of object concepts and decoded the four general concept categories: natural, animal, food and drink, and clothing. Then, using RSA, the study compared the geometric structure in the time-varying neural representation of object concept categories with the structure in semantic representations produced by the two broad types of linguistic models. Contrary to previous research results, this study shows evidence that the structure of time-varying neural representations of object concepts corresponds primarily with the hierarchical structure produce by WordNet models. But it also notes that despite their different conceptualizations of word meaning, all linguistic models showed similar correlation paths with the neural data. This thesis concludes that the temporal synchrony between the models coupled with the potential influence from non-hierarchical relations in WordNet suggest the rapid transition from perception to representation, compatible with language and conceptual thoughts, is underpinned by concept category distinctive features.

A Study of Representational Similarity:

The Emergence of Object Concepts in Rapid Serial Visual Presentation Streams

Introduction

Object concepts are the memory representation of a class or a category of objects (Martin, 2007). Representation of object concepts is crucial for supporting a range of cognitive processes such as identifying an object, manipulating representation in thought and judging the similarity between objects (Basalou et al., 2003; Barsalou, 2008; Chen & Rogers, 2014; Patterson, Nestor & Rogers, 2007). Representation of objects concepts can be derived from two types of data: experiential data such as sensory-perceptual information derived from directly interacting with object themselves, and the distributional data embedded in word usage patterns in natural language (Andrews, Vigliocco & Vinson, 2009; Olney, Dale, D'Mello, 2012; Pereira, Detre, Botvinick, 2001; Schutze, 1992). While the two types of data are qualitatively different, they both serve the purpose of assigning meanings to the referent objects in the physical world (Andrews, Vigliocco & Vinson, 2009; Glenberg & Robertson, 2000; Patterson, Nestor & Rogers, 2007). For example, as a child first learns that the word '*ocean*' refers to the glimmering, blue area, the child might also in parallel, learn about other objects embedded within the scene such as '*sand*', '*beach*', '*seashells*', '*ice cream'* and so on. In this learning scenario, sensory-perceptual information and words associated with the external referent objects are encountered simultaneously, all of which can be integrated into a more abstract concept - '*holiday*.'

A missing link in the conceptual research has been the question of whether the semantic representation of concrete entities, such as object concepts, produced in language resembled conceptual representation acquired from sensory-perceptual information outside language. Philosophy, cognitive science, neuroscience and linguistics have long been gripped by this question (Landauer & Dumais, 1997; Roger & Wolmetz, 2016; Shepard, 1987). For a long time, semantic or conceptual representation instantiated from these inter- and intra-linguistic data have been investigated mostly independently away from each other, if not mutually exclusive (Andrews, Vigliocco & Vinson, 2009). Nevertheless, a paucity of research that investigated the joint role of these two types of data showed that both sources of information are non-trivial for perceiving and representing knowledge about objects (Louwerse, 2008; Tanenhaus et al., 1995; Vigliocco, Vinson, Lewis & Garrett, 2004). Thus, the central question investigated by the present study is: to what extent do the semantic representation of object concepts derived from word meanings, and their distributional patterns in language correspond with the knowledge that was initially learned through precepts outside language.

Conceptual research is deeply grounded in philosophy and can be traced back to early modern empiricism (Vigliocco, Vinson, Lewis & Garrett, 2004). It was initially proposed by Locke (1975, p. 49) that '*all knowledge is ultimately based upon sensible qualities or sensory data derived through various sensory modalities*.' According to Locke (1975), the formation of concepts relies on experiential information derived from sensory-motor properties. Contemporary conceptual theories that honoured this empiricism tradition mostly focused on contributions from varying types of perceptual information associated with the objects, such as their features, attributes, knowledge domains and sensory-motor properties (Clarke, Taylor & Tyler, 2010; Clark et al., 2012; Farah et al., 1989; Funnell, Sheridan, 1992; Mahon & Caramazza, 2009; McRae, Cree, Seidenberg & McNorgan, 2005; Sim & Kiefer, 2005; Warrington & McCarthy, 1987). Each account was supported by compelling evidence in psychology and neuroscience. For example, fMRI studies consistently revealed a greater activity in the occipital cortex for animals than for tools, a dichotomy that reflects a distinction between living and nonliving things (Capitani, Laiacona, Mahon & Caramazza, 2003; Caramazza & Mahon, 2008; Caramazza & Shelton, 1998; Damasio, 1989; Ishibashi, Pobric, Saito & Lambon Ralph, 2016). Similarly, the natural and man-made distinction was also one of the most robust and consistently supported evidence observed from Event-Related Potentials (ERP) studies on healthy subjects (Paz-Caballero, Cuetos, Dobarro, 2005) and patients with Herpes Simplex Encephalitis (Noppeney et al., 2007; Stewart, Parkin, Hunkin, 1992). Evidence from Tranel, Logan, Frank & Damasio's (1997) study on category-related dissociations patients' data also showed that category-like distinctions are influenced by a variety of traits of concrete entities, such as practicality, familiarity, age of acquisition. In particular, for conceptual categories such as animals and tools/ utensils, the profile of impairments differed significantly on the factors such as familiarity, manipulability, characteristics motion and touch. These findings, taken together, suggested that the perceptual and functional features associated with the objects are crucial for the memory representation of objects.

Recently, an alternative approach that emerged from the domain of Natural Language Processing (NLP) has recast the question of the formation of conceptual knowledge. It was initially proposed by Wittgenstein (1953, cited in Andrews et al., p. 45) that '*the human language mirrors the world; the arrangements of words in the language also reflect the structure of their referent objects in the physical world*'. This intricate connection between the structures of words in the language and the knowledge represented outside language was recapitulated in Firth's (1957) *Distributional Hypothesis*. According to Firth (1957, p. 2, cited in Sahlgren, 2008, p.

23), the meaning of a word or a concept could be derived from its '*habitual collocation'* and that '*words appear in similar linguistic context also share similar meanings*.' Guided by this intuition, linguistic models trained on word co-occurrences statistics were shown to capture important semantics derived from distributional patterns of words in text corpora, without necessarily correspond any specific feature or functional properties (Riordan & Jones, 2011). The semantics derived through the means of distributional data in natural language was known as '*distributional semantics'* (Sahlgren, 2008, p. 8).

Distributional semantics produced by computational linguistic algorithms trained on massive text datasets yielded robust results for predicting human performances on a range measures including the Test of English as a Foreign Language (TOEFL) synonym test (Landauer & Dumais, 1997), word association (Griffiths & Steyvers, 2007; Mikolov, Chen, Corrado & Dean, 2013), analogy (Baroni et al., 2010; Goldberg & Levy, 2014; Pennington, Socher & Manning, 2014), semantic similarity (Mikolov, Yih & Zweig, 2013; Murphy, Talukdar & Mitchell, 2012) and semantic relatedness tasks (Bullinaria & Levy, 2007; Pereira et al., 2016). A seminal fMRI study by Mitchell et al. (2008) also showed that different spatial patterns of neural activation could be reliably predicted by distributional co-occurrence models trained. Recently, a MEG study by Sassenhagen & Fiebach (2019) also demonstrated distributional co-occurrence models outperformed taxonomic linguistic models in modelling the patterns in neural responses induced by English and German concrete nouns.

Despite these compelling evidence, some researchers remained sceptical about the pragmatic utility in conceptual research (Durda, Buchnan & Caron, 2009; Glenberg & Robertson, 2000; Roger & Wolmetz, 2016). A fundamental criticism against the distributional approach has been that while semantics produced by means of distributional data in language may have captured the crucial elements in conceptual knowledge, words themselves are nevertheless, abstract symbols that are detached from their real-world referents and all their associated sensory richness (Andrews, Vigliocco & Vinson, 2019; Glenberg & Robertson, 2000). While psycholinguists claimed distributional approaches to the meaning acquisition are based entirely on the data in language, and thus, the verity of semantic representation produced from word co-occurrence statistics are entirely constrained by the data permitted within text corpora (Miller & Charles, 1991; Sahlgren, 2008). Such that if the data changes, the distributed semantic representation produced by computational models would also change (Sahlgren, 2008, p. 10). As such, the connection between semantic representation produced in the discourse of language and knowledge in the physical world remained elusive.

Representational Similarity

Robust evidence from research suggest sensory-perceptual information and distributional data derived from language both are non-trivial for representing conceptual knowledge (Barsalou, 2008; Clarke et al., 2012; Clarke Taylor & Tyler, 2010, Pulvermüller et al., 2011; Richardson, Smeaton & Murphy, 1994). To link these two types of data requires an integrated quantitative approach. Representational Similarity Analysis (RSA) provided the means to relate the three types of data by comparing the dissimilarity in the patterns of neural activity evoked by a pair of experimental conditions, for example, a pair of concepts, to the semantic representation of the same pair of concepts produced by the computational model, resulting in a representational dissimilarity matrix (RDM) (Kriegeskorte, Mur &

Bandettini, 2008). Recently, studies have applied the logic of RSA to investigate the neural representation of visual objects (Grootswagers, Robinson & Carlson, 2019), the shared neural representation of magnitude (Teichmann, Grootswagers, Carlson & Rich, 2018), visual and phonological representation in reading recovery (Fischer-Baum, Jang & Kajander, 2017), and neural representation of object organisation in the inferior temporal cortex (Carlson, Simmons, Kriegeskorte & Slev, 2013). For research areas such as conceptual research where the underlying neural mechanisms that drive a specific cognitive process are largely unspecified or unknown, RSA's abstraction approach is particularly useful for characterising the information represented in the brain and relating it to the information produced by computational modelling (Kriegeskorte, Mur & Bandettini, 2008).

Next, the study examined the four dominant theoretical accounts that were inspired by the empiricist tradition and each approached the representation of object concepts from objects' perceptual and modality-specific properties or characteristics and their domain. Then, the study explored how psychologically plausible semantic representation could be derived from the semantic similarity of words, based on their explicit semantic relations and the distributional patterns of co-occurrences in natural language.

Attribute and Feature-Based Approaches

Attribute-based approaches were motivated by the notion that correlations between attributes and properties distributed across varying objects would reveal important latent structures in conceptual knowledge (Rosch, 1975; Rosch et al., 1976; Collin & Quillian, 1969). This idea was articulated in the works of Quillian (1967) and Collins & Quillian (1969), where they argued the perceived attributes and properties associated with objects were crucial for the memory representation of objects. In this view, knowledge about objects can be conceptualized as a hierarchy where high-level concept categories are described in terms of subcategories or the constituent objects, and the constituent object was described in terms of its perceived attributes and properties (Quillian, 1967). This notion was recapitulated by Smith et al. (1974) in their featural model for semantic decisions. On this account, conceptual knowledge was described as a multidimensional space that corresponds to the objects' properties or features and learning conceptual representation, according to Smith et al. (1974), also corresponded to learning the intrinsic structure of this multidimensional featural space.

The notion of a distributed featural correlation also underpinned Rosch's (1976) *'typicality effect'* theory. On this account, representations of objects were encoded as the distribution over a broad range of explicit attributes and properties (Rosch, 1976). The correlations between the features or properties were postulated to reveal important latent relations among objects. The attributed based accounts were supported by findings from verbal attribute listing studies that found significant correlations for people's similarity ratings on a wide range of objects and the type of properties or attributes that were deemed to be common among them (Rosch et al., 1976). However, attribute-based approaches have also been criticised for its lack of theoretical coherence as it was not explicitly stated as to why specific attributes or properties were deemed to be more critical to the concept than others, or indeed, factors that might influence such weighing (Murphy & Medin, 1985; Roger & Woltmez, 2013). As such, it was unclear how representation produced from features and attributes alone could serve the ultimate goal of producing semantically imbued behaviours to allow humans to successfully interact with the environment.

Sensory Functional Hypothesis

The notion that knowledge about objects can be dichotomised into living and nonliving things was motivated by the intriguing phenomena observed on patients with focal brain injuries who displayed disproportionate semantic impairments for living, and nonliving things. It was initially articulated in Warrington & Shallice's (1984) Sensory Functional Hypothesis that knowledge of living things relied more on their sensory properties, whereas knowledge of nonliving things, for example, manmade objects relied more on their functional properties. Hence, it was postulated that neuropathology that affected perceptual versus functional knowledge was responsible for the observed disproportionate semantic impairments (Warrington & Shallice, 1984). Evidence from early clinical case studies corroborated the central prediction in the Sensory Functional Hypothesis, where it was found patients with selective semantic impairments for living things, also showed relatively more impaired knowledge about objects' visual properties than their nonvisual properties (Farah, Hammond, Mehta & Ratcliff, 1989; Warrington & Shallice, 1984; Warrington & McCarthy, 1987). However, evidence emerged from later studies directly contradicted this prediction. It was found that some patients' selective semantic impairment for living things had equally affected their knowledge of perceptual and functional properties (Caramazza & Shelton, 1998), and some patients with impaired knowledge of nonliving things also showed deficits in identifying body parts that arguably, fit into either living and nonliving categories (Gainotti & Silveri, 1996; Sacchett & Humphreys, 1992). Moore & Price (1999) argued that these inconsistent findings could be due to patchy neuropathies, which were not uncommon in focal brain injuries and degenerative neural disorders. As such, it was evident was that the postulated sensory/ functional and living/ nonliving boundaries were too simplistic and clear-cut to account for the full range of documented selective semantic impairments.

Domain Specificity and Embodied Cognition

Inspired by neuroscience and computational network theories, the Domain Specificity account (Caramazza & Shelton, 1998) and Embodied Cognition (Barsalou et al., 2003; Barsalou, 2008) both, to varying degrees, claimed that representation of conceptual knowledge is shaped during the course of experience and interactions with the environment. Moreover, the conceptual system was seen as a distributed system that spreads across various domain- and modality-specific neural substrates. The Domain Specificity account (Caramazza & Shelton, 1998; Mahon & Caramazza, 2009) was underpinned by the idea that evolutional pressures had resulted in specialised, anatomically and functionally independent neural modules for representing and retrieving perceptually and conceptually distinct object concepts that were crucial for humans to efficiently interact with the environment (Caramazza & Shelton, 1998). On this account, objects from animals, plant life, and artifacts concept categories were deemed to be crucial for solving practical survival problems (Caramazza & Shelton, 1998).

Akin to the Domain Specificity account, the embodied view also deemed that perceptual information derived from experience and interacting with objects was important for perceiving and representing objects. Different from the Domain Specificity's modular conceptualization of knowledge and the conceptual system, the embodied view postulated that both were distributed across various sensory modality cortices responsible for processing specific sensory-motor information (Barsalou et al., 2003; Barsalou, 2008). Specifically, the embodied view postulated that conceptual knowledge and the conceptual category were integrated during the course of the experience, and grounded directly in, or near various perception, action and affect sensory systems (Barsalou et al., 2003). It was postulated that modality-specific information associated with the object was activated by the conjoint neurons in the adjacent memory systems during the course of constructing a mental imagery of an object (Barsalou, 2008). As such, embodied view predicted that retrieving a concept from long-term memory would entail spontaneous imagery simulation of the perception, sensory and action characteristics associated with the given concept. Consistent with this prediction, Barsalou et al. (2003) found participants spontaneously engaged in mental imagery and produced the same complex distribution of features as those who were explicitly instructed to imagine the features associated with the concept during the verification task.

Consistent with the embodied view, evidence from neuroscience demonstrated that much of the conceptual knowledge associated with perception and action was represented and distributed in regions in the ventral and lateral temporal cortex that overlapped with neural substrates responsible for perceiving and acting (Binder et al., 2016; Barsalou, 2008; Martin, 2007; Martin & Chao, 2001; Humphreys & Price, 2002; Pulvermüller et al., 2005; Vigliocco et al., 2004). However, evidence from other neuroimaging studies also corroborated the Domain Specificity account. For example, it was found that patients with domain-specific impairments displayed distinct patterns of evoked activation across various cortical regions that corresponded to abstract concept domains, e.g., animate/ inanimate (Martin, Wiggs, Ungerleider, Haxby, 1996; Martin, Chao, 2001; Martin, 2007; Mahon & Caramazza, 2009). As the profiles of activation across various cortical regions are largely overlapping, neuroimaging evidence alone was insufficient for adjudicating between these competing theories and disentangle the cognitive mechanisms that support perceiving and knowing.

Object concepts representation: from perceptual to conceptual

The core aim of the present study was to elucidate one question: to what extent do semantic representation produced from the discourse of language resemble the knowledge acquired from experience and interacting with the objects beyond the bounds of texts? The literatures reviewed so far have studies the characterization of conceptual knowledge in terms of objects' featural correlation, perceptual information and modality-specific information, which was derived from the largely non-linear mappings in the brain. These information were then used to identify and differentiate between objects, object categories and knowledge domains. In the field of linguistics and computational linguistics, psychologically plausible semantic representation could also constructed from the semantic similarity between words, based on two general methods of semantic similarity measures.

The lexical hierarchical approach to semantic similarity

In the field of linguistics, semantic similarity is measured through hierarchal taxonomic links between words and their explicit semantic relations in the lexical hierarchy (Miller, 1995; Pedersen, Patwardhan & Michelizzi, 2004; Resnik, 1999, 1995; Riesenhuber & Poggio, 1999). Evidence from psycholinguistics and neuroimaging studies found semantic similarity scores for object and action concepts produced by lexical hierarchical models such as WordNet closely matched human performances in similarity judgments of object labels (Carlson et al., 2013), English nouns (Miller and Charles, 1991; Resnik, 1999), English verbs (Yang & Powers, 2005), and English and German nouns (Sassenhagen & Fiebach, 2019).

The systematic lexical database in WordNet (Miller, 1990) encompasses a broad range of linguistic entities from the American English language. The WordNet database contains over 80,000 concrete nouns that are organized into nine taxonomy hierarchies (Miller, 1995). Similarity metrics implemented in WordNet compute semantic similarity based on explicit dictionary definitions and the information encoded in a taxonomy hierarchy (Pedersen, Patwardhan, & Michelizzi, 2004, p.124). Semantic similarity for a given pair of concepts is represented by the path distance between nodes that correspond to the concept pair, represented in a hypernym/ hyponym hierarchy (i.e., X is a Y) (Figure 1). A hypernym refers to a general categorical word and branches into subordinate hyponym words with more specific meanings. For example, '*cutlery*' is a hypernym of '*spoon*.' Inversely, a hyponym is a word of more specific meaning than a general or superordinate word applicable to it. For example, '*foliage*' is the hyponym for '*leaf*'. The hyper/hyponym relationship accounts for nearly 80 per cent of all link types in the English language (Yang & Powers, 2005). However, WordNet also makes complimentary use of other non-hierarchical relations in the quantification of semantic similarity between concepts.

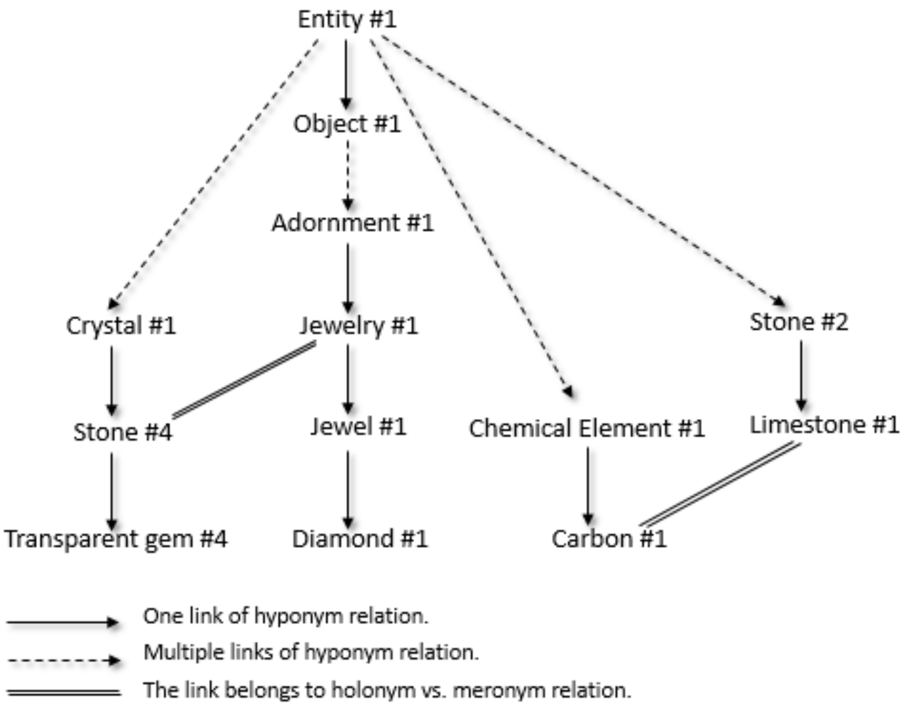


Figure 1. An example of WordNet-style hierarchy. Hierarchical relations: Hyponym/ Hypernym (IS-A/ HAS-A). Non-hierarchical relations: Meronym/ Holonym (Member-of/ Has-Member/ Substance-of / Has-Substance). Adapted from 'Measuring semantic similarity in the taxonomy of WordNet,' by D. Yang, & D. M. Power, 2005, *In Proceedings of the Twenty-eighth Australisian conference on computer Science*, 38, p. 315.

WordNet Wu-Palmer (Wu & Palmer, 1994) and WordNet PATH both are similarity metrics that compute semantic similarity for a pair of concepts based path distance between them in the hierarchy. Both metrics return the similarity score between the range of 0 and 1 that indicate the semantic similarity for the given concepts, such that a higher score indicates a higher degree of semantic similarity. The Wu-Palmer metric identifies the path length to the root node from the Least Common Subsumer (LCS) for the given concepts (Yang & Powers, 2005; Wu & Palmer, 1994). LCS represents the ancestor node deepest in the taxonomy that is shared by the two specific concepts (Pedersen, Patwardhan, & Michelizzi, 2004). In situations where multiple ancestor nodes and multiple paths to the root are available, the longest path will be selected, as for the Wu-Palmer metric, the general class shared by the pair of concepts is most informative for computing semantic similarity (Pedersen, Patwardhan & Michelizzi, 2004). By contrast, the PATH similarity metric takes the shortest path between two nodes (Pedersen, Patwardhan, & Michelizzi, 2004, p.124). As such, similarity scores produced by the Wu-Palmer and the PATH metrics do not always align. For example, under the PATH metric, the concept '*car*' has a moderate similarity correlation of 0.125 to the concept 'boat.' In contrast, in the Wu-Palmer metric, the '*car*' has a significantly higher similarity correlation of 0.695 to '*boat*,' which makes more intuitive sense because while they are different, they also belong to the general concept category of '*machinery*.'

Distributed approaches to semantic similarity

Semantic similarity, according to the distributional co-occurrences approach, is based on the principle that similar words occur in similar contexts (Firth, 1957). This notion is articulated in Sahlgren's (2006, p.3) '*distributional hypothesis*' as if word 1 and word 2 displayed similar distributional properties, such that they consistently occur within the same linguistic context with word 3, this co-occurrence pattern could then be interpreted as word 1, and word 2 belonging to the same linguistic class and thus, shared similar meaning. As such, distributed semantic representation acquired through word co-occurrence patterns can be seen as qualitatively different from WordNet's explicit hypothesis of what semantic knowledge is and how it is organised, as it requires no prior assumption about knowledge or language in order to acquire word meaning (Carlson et al., 2013;

Bullinaria & Levy, 2007; Sahlgren, 2006). Thus, it has been suggested that distributional models captures the emergent structure in word meanings from statistics of word co-occurrence.

In Pereira et al.'s (2016) review of existing state-of-the-art distributional co-occurrence models, Word2Vec (Mikolov, Chen Corrado & Dean, 2013) and Global Context Word-Word Occurrence Count (GloVe) (Pennington, Socher & Manning, 2014) both outperformed other models and produced robust results on a range psycholinguistic tasks. While both models extract distributional semantics based on the word co-occurrences statistics in large text corpora, the quality of the distributed semantic representation is also influenced by factors such as differences in the size of training corpus, vector length and the size of the context window (Pennington, Socher & Manning, 2014). First, Word2Vec is a local context, prediction-based word-embedding model, trained on the Google News dataset with approximately 100 billion tokens. The two log-linear models in Word2Vec: Continuous Bag-of-Words (CBOW) and Skip-Gram, generates a network of semantic representation based on the information available in the local context window, which is typical of five or ten words (Mikolov, Chen Corrado & Dean, 2013). By contrast, the GloVe model is a global context, a word-count model that leverages the global statistics of word co-occurrence in a document and trained on a much larger text corpus that consists a combination of Wikipedia, Gigaword and Common Crawl (Pennington, Socher & Manning, 2014). The specifics of these training sets may limit the type of semantics permitted in each model. Second, while the principle that similar words occur in similar contexts is used by both models to capture the semantic similarity between words, the nature of the context also varies in each model. For the local context, prediction based model Word2vec, this context means words are semantically similar if they occur within the same sentence as only the local information in the sentence is utilised (Mikolov, Chen Corrado & Dean, 2013). For example, The CBOW model predicts the hidden middle word in a five words context window based on the neighbouring words (Mikolov, Chen Corrado & Dean, 2013) (Figure 2). Inversely, the Skip-gram model predicts neighbouring words within a ten words context window based on the given the word in the middle of the window (Mikolov, Yih & Zweig, 2013). It was postulated that this '*many-to-few*' resembles the way humans cognise semantic problems (Mikolov et al., 2013; Turney & Pantel, 2010). By contrast, for the GloVe model, the context refers to the same document as it leverages the global distributional patterns beyond the boundaries of sentence context and hence, some argue the distributed semantic representation generated using this approach captures the more intricate and nuanced relations between words (Baroni et al., 2014; Goldberg & Levy, 2014; Pereira et al., 2016) (Figure 3). In sum, these subtle differences training set and size of the context region limit the type of information from which co-occurrences statistics are derived as well as the subsequent semantic representation acquired by each model.

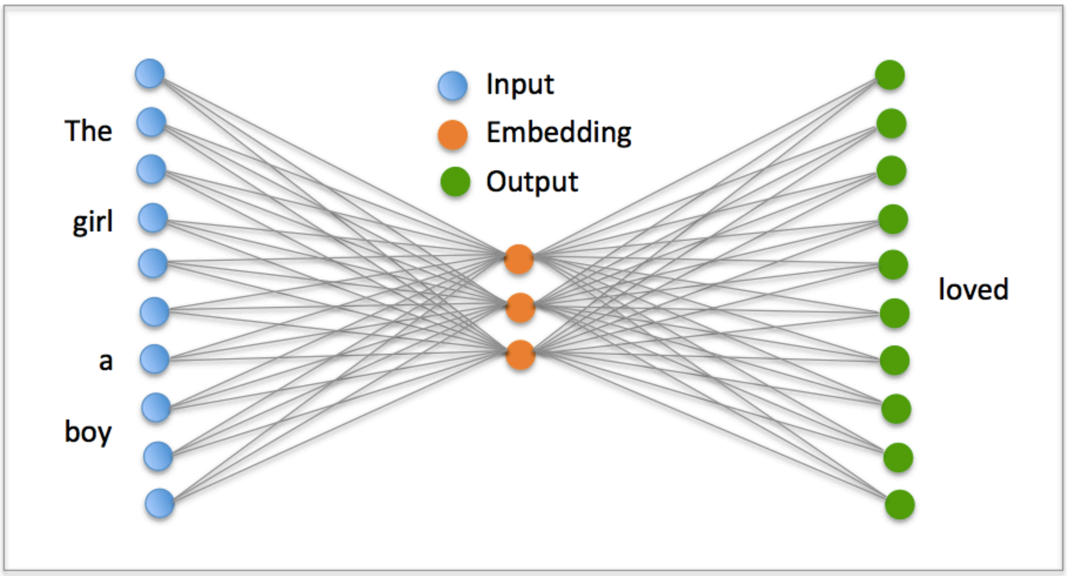


Figure 2. An example of Word2Vec Continuous Bag of Words (CBOW) prediction model for word embedding. An input layer (blue dots) receives a context and learns weights on the internal layer (orange dots), which in turn allow the prediction of the hidden target word (green dots). The Skip-Gram Model does the reverse. Adapted from 'Traces of Meaning Itself: Encoding distributional word vectors in brain activity,' by J. Sassenhagen and C. J. Fiebach, 2005, *bioRxiv,* 603837.

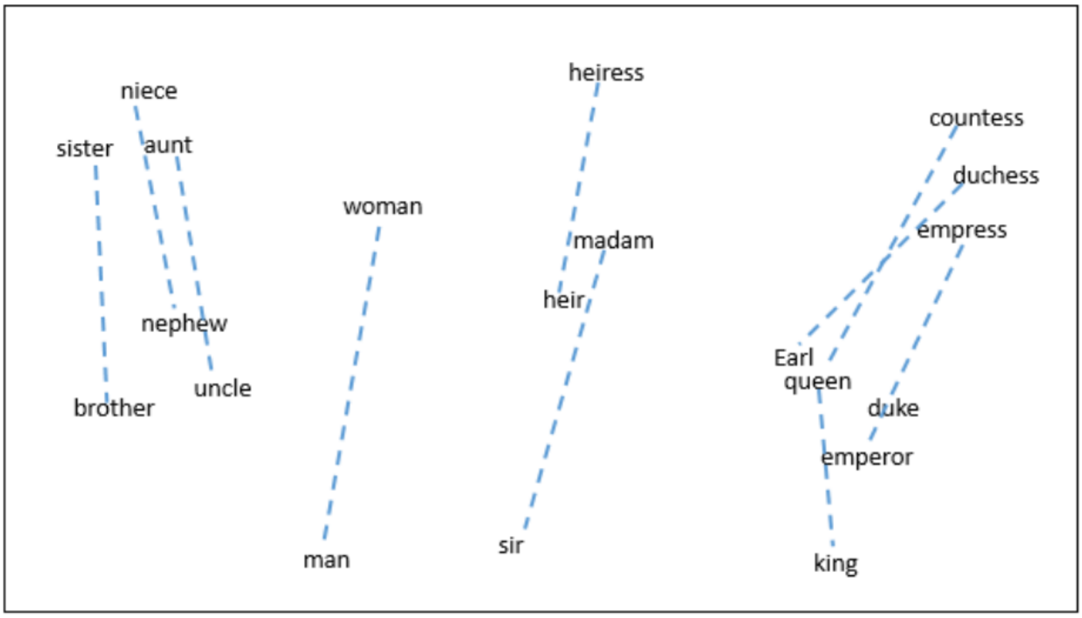


Figure 3. An example of GloVe's word-word occurrences matrix. The GloVe model extracts word co-occurrence statistics within the same document whereby it leverages the global distributional patterns of words. Adapted from 'Glove: Global vectors for word representation,' by J. Pennington, R. Socher and C. Manning, 2014, *In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, p. 1532.

The Present Study

The central aim of this thesis was to examine the extent to which semantic representation produced by linguistic computation models resemble the knowledge acquired from sensory-perceptual information outside language. The more specific question of the thesis was, which type of semantic structure produced by these linguistic models best explained the neural representation of object concepts. In keeping with the literature viewed, the study focused on two types of semantic structures, hierarchical structure of word meaning produced by WordNet's explicit dictionary definitions, and semantic relations of words and emergent structure of word meaning raised from word co-occurrence statistics in text corpora. As the distributed semantic representation acquired by Word2Vec and GloVe is driven primarily by the distributional behaviour of words in text corpora, the nature and the quality of the semantic representation captured by these models, therefore, were postulated to be different. Using RSA, the present study quantitatively related the structure of object concepts captured by these two general methods to the structure of object concept representation in the brain.

Rapid Serial Visual Presentation (RSVP) (Grootswagers et al., 2019) was selected to investigate the neural dynamics of object concept representation in the brain. With its origin traced back to the sixties, RSVP itself is not new to psychology for studying various cognitive processes (Potter, 1976). Past research that used RSVP in studying sentence reading (Potter, Wyble, Hagmann & McCourt, 2014) and picture recognition (Intraub, 1980; Keysers, Xiao, Foldiak & Perrett, 2001; Rousselet, Thorpe & Fabre-Thorpe, 2004) found even at fast presentation rates (~10 ms), people's performances on the behavioural tasks remained efficient. In the current study, the novelty of RSVP was that it was used to study neural responses to each naturalistic thematic image of object concepts in RSVP streams. In RSVP, visual stimuli are briefly flashed on a screen typically at around ten images per second (Marti & Dahaene, 2017). Within this brief moment of onset, previous studies found that people can effortlessly extract high-level concept-specific information such as an object's category (Schenda & Maher, 2009; VanRullen & Thorpe, 2001) and sub-category (Grootswagers et al., 2019) from images of objects. Hence, the present study drew on this core human capacity and adopted RSVP to investigate the neural dynamics of object concept representation in the brain.

Two contradicting hypotheses were of the most interest in the present study.

Hypothesis one: The structure of object concept representation in the brain was hypothesized to correspond to the hierarchical structure of word meaning produced by the two WordNet-based similarity measures (Wu-Palmer, WordNet PATH). As both primarily quantify conceptual similarity as the path distance between concepts in the taxonomy hierarchy, they predict the correlations for concept pairs that share a general concept category will be higher than correlations of concepts pair that belong to distinct concept categories.

Hypothesis two: The structure of object concept representation in the brain was hypothesized to correspond to the emergent structure of word meaning captured by the two distributional co-occurrence models (Word2Vec and GloVe). As both primarily quantify conceptual similarity as the co-occurrence frequency of words, they predict the correlations for concepts pairs that co-occur more frequently in similar linguistic context will be higher than correlations for concepts that seldom appear together in similar linguistic context.

Method

Ethics approval

The experiment was assessed and approved by the University of Sydney's Human Research Ethics Committee (HREC 2016/849).

Participants

32 participants were recruited from the University of Sydney in return for course credit (21 female, mean age = 20 years, SD = 2.3 years, age range = 17 - 28 years). All participants reported normal or corrected-to-normal vision and had no history of psychiatric or neurological disorders. Verbal and written consent was obtained from each participant before the experiment. Two participants were excluded due to uncompleted participation. Hence, the final sample in the present study consisted of 30 participants (19 female, mean age = 18.7 years, SD = 2.6 years, age range = 17 to 28 years). Before the EEG recordings, all participants completed a basic demographic questionnaire that included participants' language backgrounds (Table 1 and 2).

Table 1

Summary of Language Profiles of Participants: Home Language

Home Language	n	%
English	16	53
Chinese	8	27
Vietnamese	2	7
Korean	2	7
Norwegian	1	3

Hindi & Gujarat	1	3

N = 30

Note. Home language refers to the language that is mostly spoken by the participants for everyday interactions at home.

Table 2

Summary of Language Profiles of Participants: Monolingual English and Bilingual English

Language	n	%
English Monolingual	16	53
Bilingual	14	47

N = 30

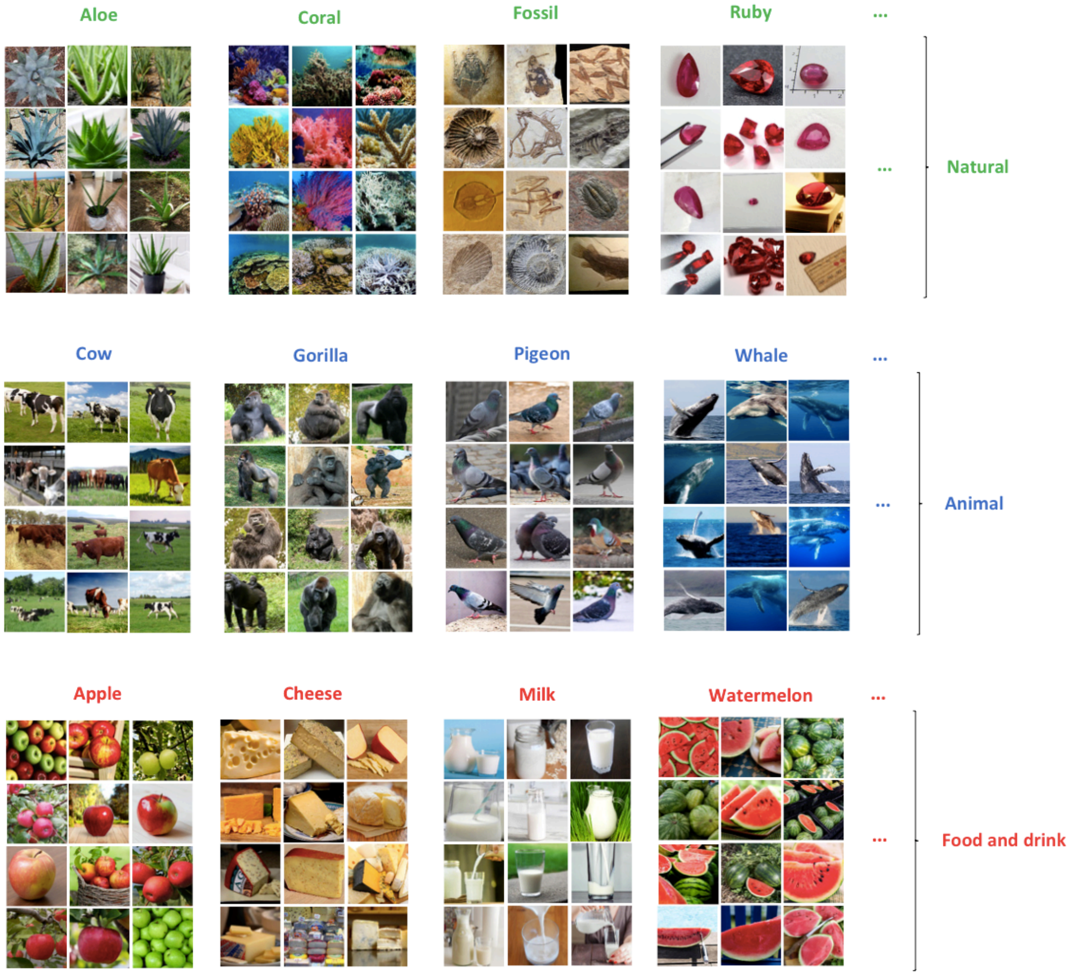
Note: English monolingual refers to participants whose English is both their first language and the language they speak at home. Bilingual refers to participants whose first language is not English but can speak English proficiently and in addition, speak another language for everyday interactions at home.

Apparatus

All stimuli were presented sequentially in random order at approximately 0.5 degrees of visual angle. Stimuli were presented in the centre of a light grey background on a 1920 x 1080 pixel Asus monitor. MATLAB with the Psychtoolbox extension was used for stimulus presentation (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). A BrainVision ActiChamp system with active Ag/AgCI electrodes (GmbH, Herrsching, Germany) was used for EEG recording.

Stimuli

The present study used the THINGS (Hebart et al., 2019) database. The database is comprised of 1,854 of common concrete object concepts that were systematically sampled from the American English language. These object concepts are represented by 26,107 high-quality, naturalistic images sourced from various image databases such as Google, Flickr and ImageNet (Figure 4). The selection process for the candidate concepts and images was rigorous. Concepts and their categories were validated by the workers from Amazon Mechanical Turk and the WordNet word-sense disambiguation (Hebart et al., 2019). These images were also fed through the layers in deep convolutional neural network CorNet-S (Kubilius et al., 2018) to ensure that they retain a sufficient degree of visual variability.



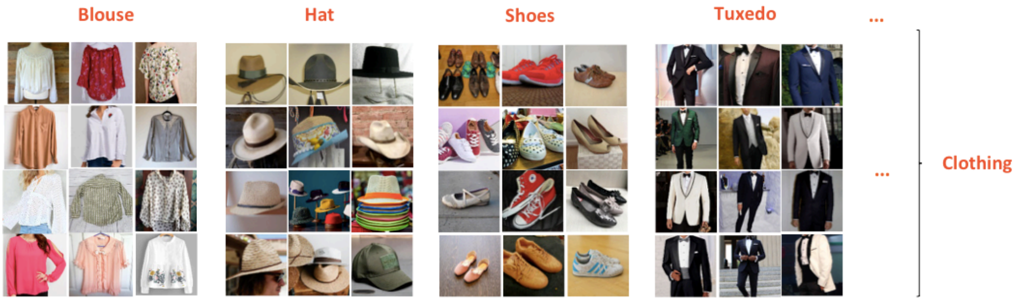


Figure 4. Examples of the object concept images sourced from the THINGS database. Adapted from 'THINGS: A database of 1,854 object concepts and more than 26,000 naturalistic object images' by M. Hebert et al., 2019, *bioRxiv*. 545954.

Procedure

At the start of the trial, each participant was instructed to press a button on a 4-way control box whenever the target superimposed on the centre of a rectangular grey background at approximately 0.5 degrees of visual angle changed to red. The intention of the task was to keep the participants vigilant throughout the trial. Stimulus presentation was controlled by custom-written MATLAB scripts using functions implemented in PsychoPhysics Toolbox (Kleiner et al., 2017; Brainard, 1997; Pelli, 1997). Images were presented at a rate of 10Hz, with the image visible for 50 ms, followed by a 50 ms inter-stimulus interval (ISI) (Figure 5).

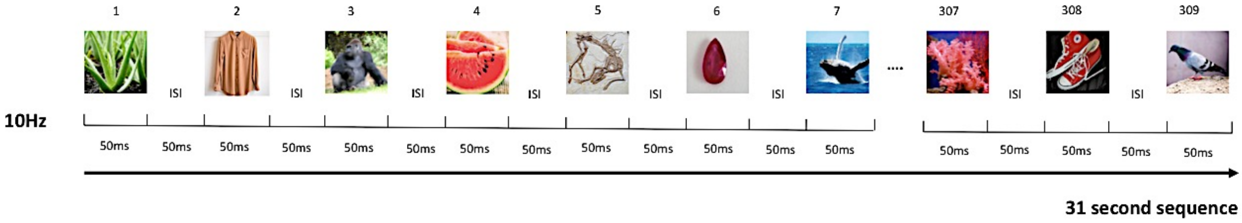
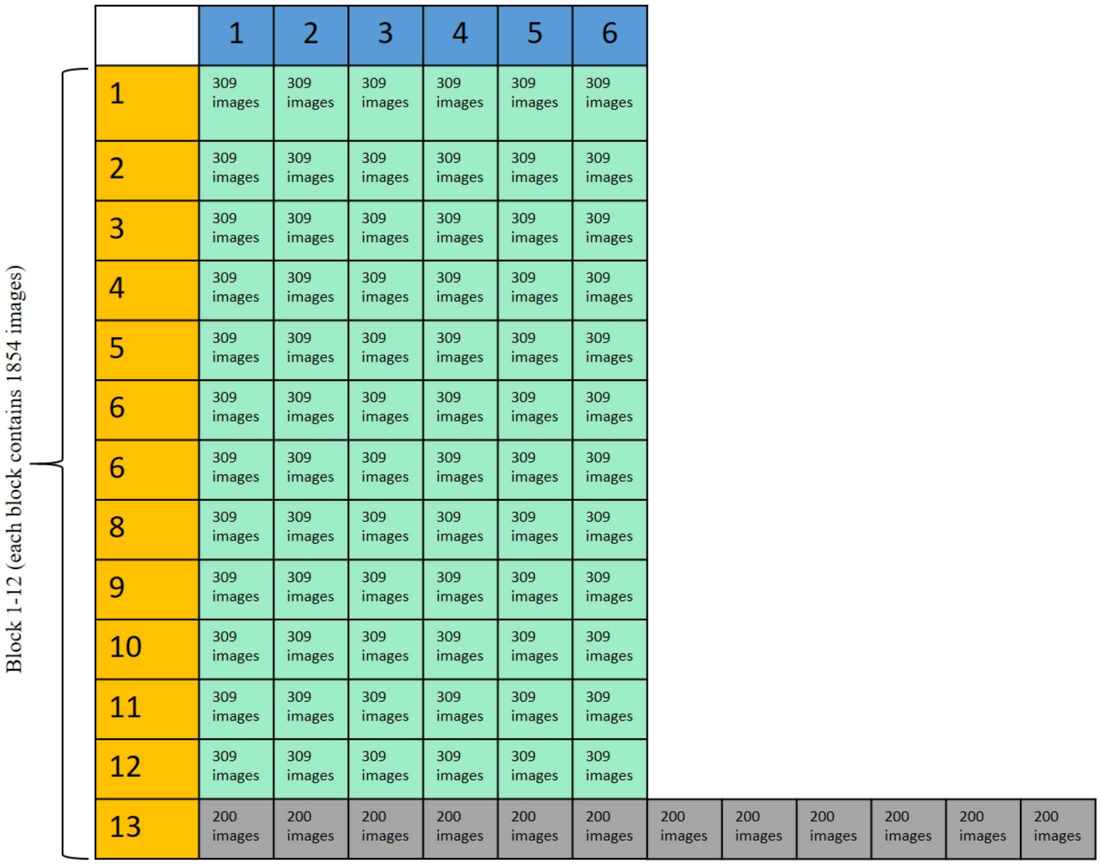


Figure 5. Stimulus presentations in a 10Hz sequence. Each sequence contains 309 images, lasting approximately 31 seconds. Images were presented sequentially in a randomised orders. Each image was presented for 50 ms, followed by a 50 ms ISI.

The experiment was divided into 13 blocks (Figure 6). Blocks 1 to 12 each contained six sequences of rapid serial presentations (10Hz) of 1,854 unique images.

Each sequence lasted approximately 31 seconds. Block 13 contained image from object concepts with more than twelve exemplars. As this difference in exemplar count would have complicated the data analysis and therefore, the data from this block were excluded from the final analysis. 309 images were presented in random order in every sequence, each lasting approximately 31 seconds. Each image was presented only once in the trial over the course of the experiment to control for contributions of low-level features in the decoding performance, resulting in 72 unique sequences.



	1	2	3	4	5	6						
1	309 images	309 images	309 images	309 images	309 images	309 images						
2	309 images	309 images	309 images	309 images	309 images	309 images						
3	309 images	309 images	309 images	309 images	309 images	309 images						
4	309 images	309 images	309 images	309 images	309 images	309 images						
5	309 images	309 images	309 images	309 images	309 images	309 images						
6	309 images	309 images	309 images	309 images	309 images	309 images						
6	309 images	309 images	309 images	309 images	309 images	309 images						
8	309 images	309 images	309 images	309 images	309 images	309 images						
9	309 images	309 images	309 images	309 images	309 images	309 images						
10	309 images	309 images	309 images	309 images	309 images	309 images						
11	309 images	309 images	309 images	309 images	309 images	309 images						
12	309 images	309 images	309 images	309 images	309 images	309 images						
13	200 images	200 images	200 images	200 images	200 images	200 images	200 images	200 images	200 images	200 images	200 images	200 images

Figure 6. Experimental structure. The experiment was divided into 13 blocks. Blocks 1-12 each contained a unique image from each of the 1,854 object concepts. Each block was divided into 6 sequences resulting in 72 unique sequences. Block 13 contained the extra images from some concepts and was not further analyzed in the study.

EEG recording

A BrainVision ActiChamp system with active Ag/AgCI electrodes (GmbH, Herrsching, Germany) digitized at a 1000Hz sample rate was used to record the EEG data continuously. The 64 electrodes corresponded to the 10-10 international standard for electrode placement, and all referenced at Cz during recording (Oostenveld & Paraamstra, 2001). Data/Ground electrodes selected impedance measurement range was 10-50 kΩ. Electrolyte gel was applied before recording to keep the impedances below 10kΩ. Each participant wore an EEG 64-electrode Brain Products cap (standard 64 Channel cap actiCAP snap) (GmbH, Herrshing, Germany) throughout the experiment.

EEG data pre-processing

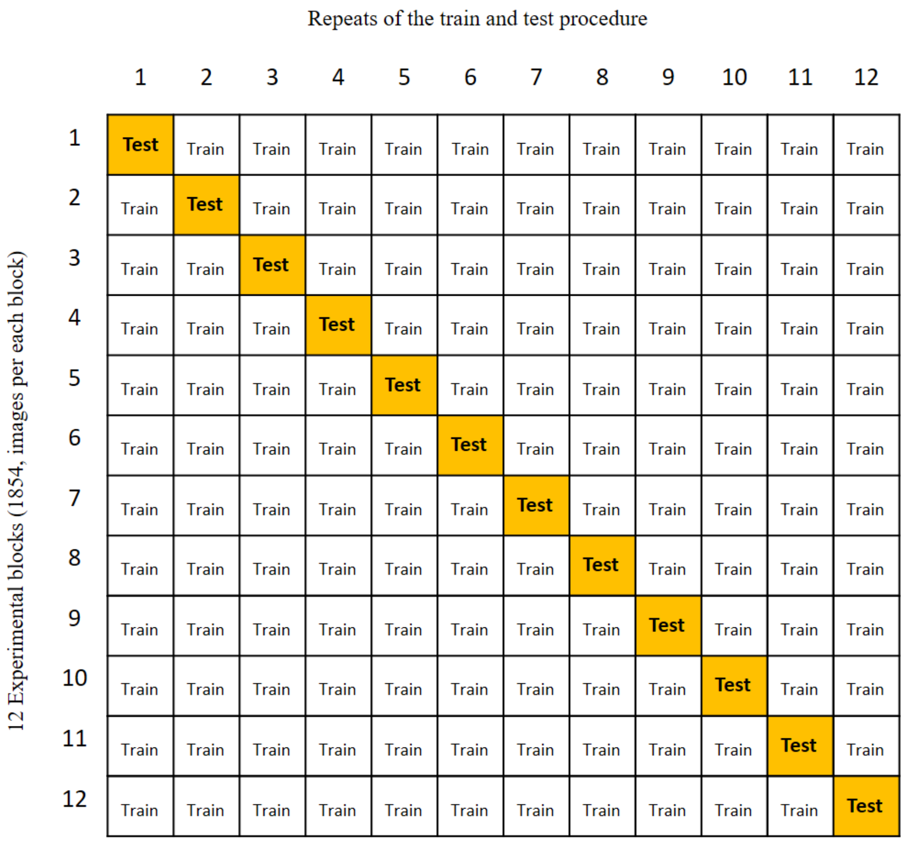
EEG data pre-processing was completed offline using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014). To improve the signal-to-noise ratio in the EEG data, a down-sampling approach was used, which collapses data over time. First the EEG data were filtered using a Hamming windowed FIR filter with 0.1Hz high-pass and 100Hz low-pass filters. Lines at noise level 50Hz were removed using the CleanLine function in EEGlab. The channel voltages at each time point were used for the remainder of the analysis. EEG data from the 72 sequences were epoched from -100 ms to 1000 ms relative to stimulus onset and down sampled to 250Hz. This downsampling approach was used to reduce the computation time for the later decoding analysis (Grootswagers et al., 2017). As the goal of the analysis was to determine whether information relevant to the stimulus was present in the neural signal, EEG data collected 100 ms before the stimulus onset is included as a sanity check (Grootswagers et al., 2017). The assumption is that the neural signal recorded before an image is shown should not contain any specific information about the image. No further pre-processing was done in order to maintain the integrity of the EEG data, in accordance with the advice of Grootswagers et al. (2017).

Pattern Classification

Time-resolved Multivariate Pattern Analysis (MVPA) (Grootswagers et al., 2017) was used to decode the objects from five concept categories: natural, animal, food and drink and clothing.

Time Resolved Multivariate Pattern Analysis (MVPA)

A standard MVPA decoding pipeline (Carlson et al., 2019; Grootswagers et al., 2017) was applied to all 64 EEG channel voltages. The aim of the decoding analyses was to determine whether there exists any information in the EEG signals recorded in response to an observed image that were indicative of the information content in the image. The decoding results serve as validation for the subsequent Representational Similarity Analysis (RSA) where the similarity of evoked neural responses for all possible pairwise concept comparisons were correlated with the similarity of the same concepts computed from two broad types of computational linguistic models. The decoding analysis was implemented in CoSMoMVPA (Oosterhof et al., 2016) and carried out in three steps. First a classification algorithm was trained to find any information in the EEG signal data that allowed it to make a relatively accurate prediction about the concept category of an image, for example, animal. The present study used a Linear Discriminate classifier to find the linear boundary that best labeled the data correctly for category membership. Linear classifiers are commonly used in decoding studies because they do not overfit to the data (Grootswagers et a., 2017). Classifiers were trained separately on EEG data from each 4 ms time window (this is one time point at 250Hz). Then, a set of EEG data that was not included in the training data was used to test whether the classifier could generalize what was learned during the training. For this purpose a 12 fold cross-validation approach was applied by splitting the EEG data into 12 blocks (Figure 7). In this design, the classifier was trained on 11 of these blocks and then tested on the left out block. This process was repeated 12 times to allow each block to serve as the test set. As every image in the stimulus set was unique, images used in the training sets were always different from the images in the test set. This helped to decrease the contribution of semantically uninformative low-level image features on the decoding performance. Finally, the average prediction accuracy of the classifier during the 12 fold cross-validation is taken as an estimate of decoding accuracy for every time point. Classifier performance is compared against a null value of 50%, which is the accuracy that should be observed if the classifier could not learn any distinguishing information. Above-chance classifier performance for any time point indicates that information in the neural signal at that specific time point, to some extent, systematically different based on concept category membership. Decoding performance was calculated for each participant and then averaged together to produce a time-varying measure of decodability for the four concept categories.



Repeats of the train and test procedure

12 Experimental blocks (1854, images per each block)

	1	2	3	4	5	6	7	8	9	10	11	12
1	Test	Train	Train	Train	Train	Train	Train	Train	Train	Train	Train	Train
2	Train	Test	Train	Train	Train	Train	Train	Train	Train	Train	Train	Train
3	Train	Train	Test	Train	Train	Train	Train	Train	Train	Train	Train	Train
4	Train	Train	Train	Test	Train	Train	Train	Train	Train	Train	Train	Train
5	Train	Train	Train	Train	Test	Train	Train	Train	Train	Train	Train	Train
6	Train	Train	Train	Train	Train	Test	Train	Train	Train	Train	Train	Train
7	Train	Train	Train	Train	Train	Train	Test	Train	Train	Train	Train	Train
8	Train	Train	Train	Train	Train	Train	Train	Test	Train	Train	Train	Train
9	Train	Train	Train	Train	Train	Train	Train	Train	Test	Train	Train	Train
10	Train	Train	Train	Train	Train	Train	Train	Train	Train	Test	Train	Train
11	Train	Train	Train	Train	Train	Train	Train	Train	Train	Train	Test	Train
12	Train	Train	Train	Train	Train	Train	Train	Train	Train	Train	Train	Test

Figure 7. 12 fold cross-validation. Over the course of 12 fold cross-validation, the classifier was trained on 11 of the experiment blocks and then tested on the last unseen chunk. The average prediction accuracy of the classifier of the 12 fold cross-validation is taken as an estimate of decoding accuracy for every time point and for every subject.

Statistical inferences for the decoding accuracy

The present study used random effect, non-parametric Monte-Carlo Cluster Statistics (Maris & Oostenveld, 2007) to determine whether the classifier performed above-chance (50%). Threshold-free cluster statistic (TFCE) (Smith & Nichols, 2009) was used to generate a cluster-forming statistic. Based on the assumption that there was some degree of contiguity in signals between time points, TFCE statistic was used to enhance the information from neighbouring time point to facilitate detection of sharp peaks as well as sustained weaker effects (Teichmann et al., 2019, p. 1001). The Monte-Carlo method was chosen to correct the problem of multiple comparisons that are often seen in EEG data analyses (Stelzer, Chen, & Turner, 2013). The Monte-Carlo Cluster Statistics implemented in the CosmoMVPA toolbox (Oosterhof, Connolly & Haxby, 2006) was used to perform a sign permutation test, which involves swapping the signs of the decoding results obtained from all participants at random at every time point, and re-computing the TFCE statistic. (Stelzer, Chen, & Turner, 2013). First, the Monte-Carlo Cluster Statistics function computed the TFCE statistic of the decoding accuracy for every time point. Then, the Monte-Carlo Cluster Statistics function permutated the sign of decoding results, a process that was repeated 10,000 times. Finally, the most extreme value of each null distribution was taken to construct an overall null distribution across the time series (Teichmann et al., 2019, p. 1001). The 95th percentile of this overall null distribution was used to compare the actual observed decoding results and the null hypothesis providing a *p* value ($\alpha = .05$), which is corrected for multiple comparison.

Representational Similarity Analysis (RSA)

The core aim of the present study was to investigate whether there was an overlap between neuronal activity patterns evoked by meaningful images of object concepts and the semantic structures produced by the two types of linguistic models. To do this, RSA was used as it is commonly employed to test hypotheses using classification decoding data (Kriegeskorte and Kievit, 2013; Kriegeskorte et al., 2008).

Using RSA, the semantic similarity between the 1,854 object concepts, as defined by the linguistic models can be seen as a geometrical structure in a high dimensional space, which can be related to a similar representational structure built based on the relationships between the evoked neural responses to images depicting the same concepts. To represent these relationships, representational dissimilarity matrices (RDM) are used (Carlson et al., 2013; Kriegeskorte and Kievit, 2013). RDMs are constructed from every possible pairwise combination of object concepts and assigned a numerical value that quantifies their 'dissimilarity' (Kriegeskorte and Kievit, 2013). These RDM values indicated the degree to which each pair of concepts were distinguishable. The neural EEG RDMs are built based on the dissimilarity between the evoked neuronal activity patterns from object images, based on the averaged decoding accuracies for all subjects. EEG neural RDMs were calculated for each subject individually for each time point and averaged to produce a single RDM for the time series (Figure 8). The study then built four different linguistic models RDMs, two distributional based (Word2Vec, GloVe) and two hierarchically based (WordNet Wu-Palmer, WordNet PATH) (Figure 9). Model RDMs are built based on the semantic dissimilarity between each concept pairs. Similar to the EEG neural RDM, a numerical value was assigned to each cell in the RDM to quantify the dissimilarity between every concept pair, as defined by the algorithms used in each model. Then, we tested whether these models captured the differences in the neural EEG RDMs by correlating model RDMs with the neural RDMs using the Spearman's rank correlation, which resulted in 30 time-varying correlations for the 30 subjects. Finally, the study compared the correlations between the models to test whether the correlations with the neural RDMs were different for hierarchical lexical models (WordNet Wu-Palmer, WordNet PATH) and the distributional co-occurrence models (Word2Vec, GloVe).

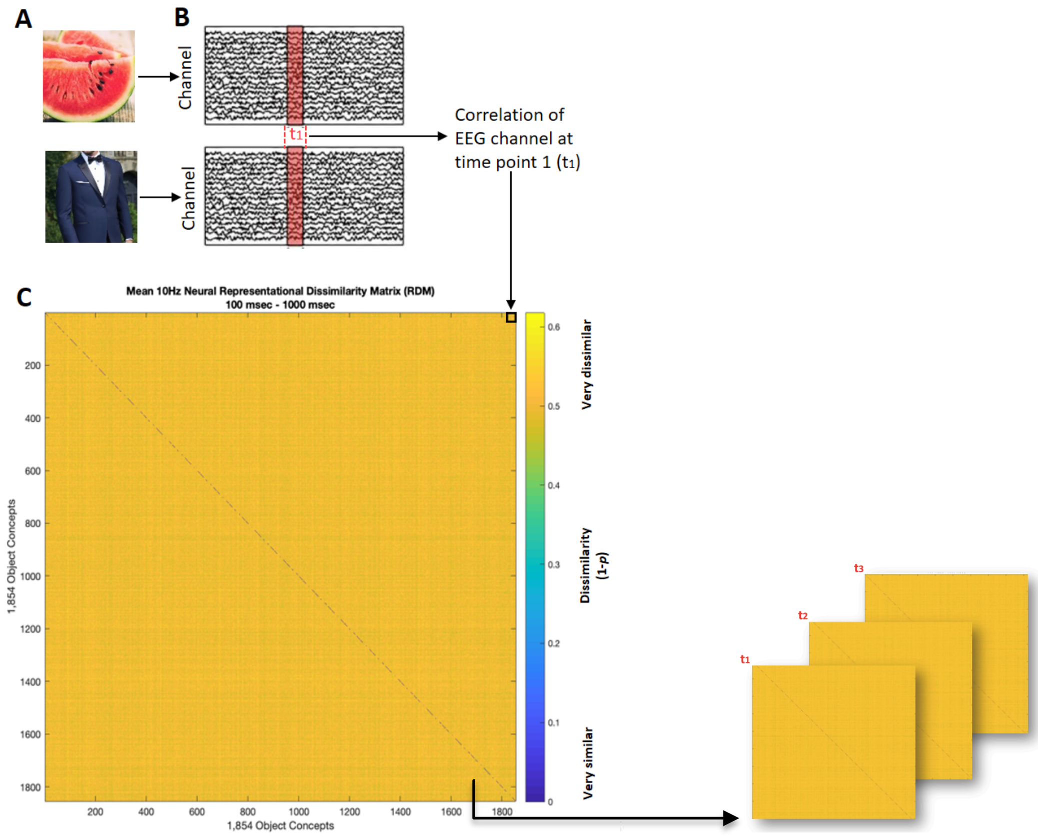


Figure 8. A depicts stimuli seen in two separate trials. B shows the recorded EEG signal in response to these stimuli. The signals from both trials are then correlated at each time window (t1). The correlation values of each stimulus pair are then inserted into the dissimilarity matrix of the corresponding timing window (C). This processing is repeated for all stimulus pairs and at every time window to create a time series of dissimilarity matrices (D) EEG neural RDMs. The full neural RDM comprises all possible object concept comparisons for every sampled time point. Adapted from “Decoding digits and dice with Magnetoencephalography: evidence for a shard representation of magnitude,’ by Teichmann et al., 2019, *Journal of Cognitive Neuroscience*, 30, p. 999.

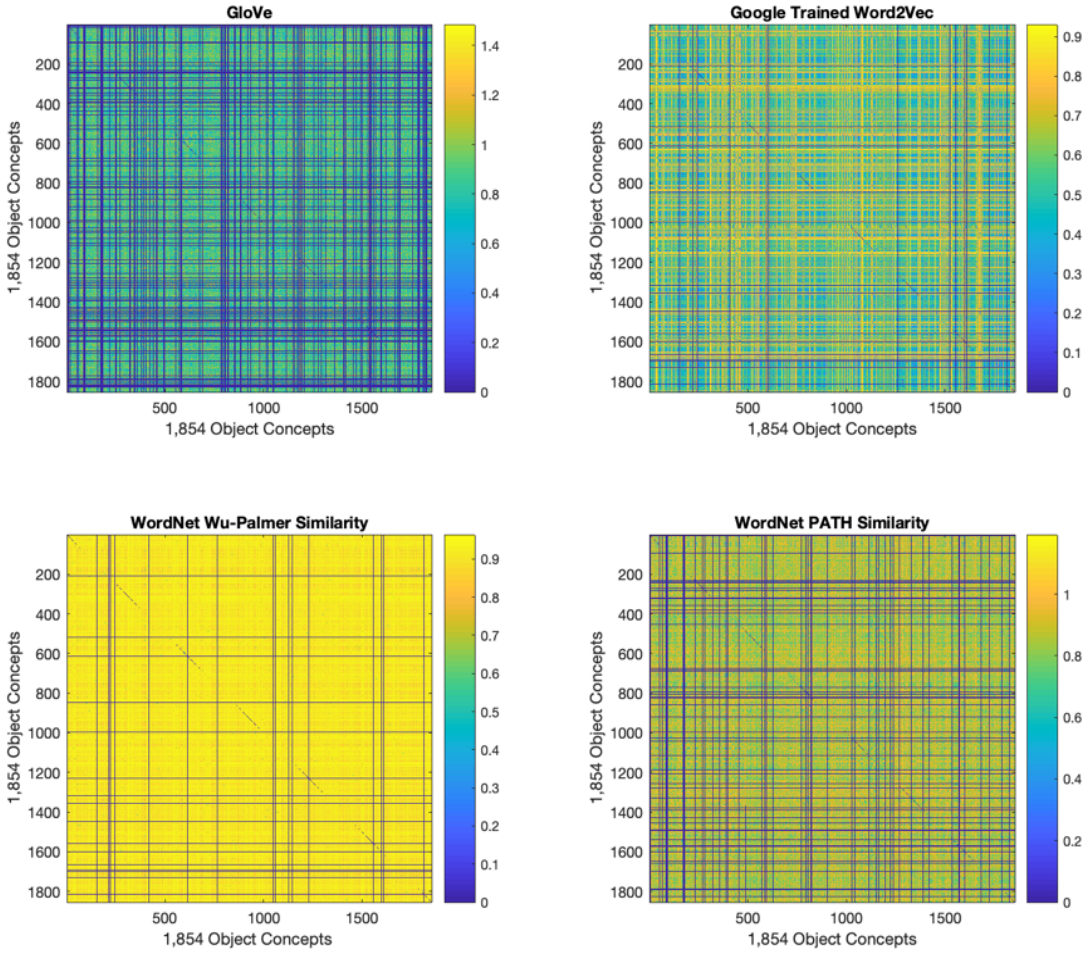


Figure 9. Four linguistic RDMs. Top left: GloVe. Top right: Word2Vec. Bottom left: WordNet Wu-Palmer. Bottom-right: WordNet PATH. Linguistic model RDMs are built based on the semantic dissimilarity between each concept pairs. Similar to the EEG neural RDM, a numerical value was assigned to each cell in the RDM to quantify the dissimilarity between every concept pair, as defined by the algorithms used in each model. Lighter color denotes greater dissimilarity between concept pairs.

Statistical comparisons for the neural RDM and the model RDMs

The study used the same Monte-Carlo Cluster Statistics function (Oosterhof, Connolly & Haxby, 2006) to test whether the correlations were significantly greater than a null hypothesis value of zero. This null value indicates the absence of a systematic linear relationship between the model RDMs and the neural RDMs. For comparing the linguistic models on their correlations with the neural RDMs it indicates no difference, at a given time point, in how well the compared models' semantic relationships reflected the patterns in the neural data. In the same way as the decoding significance test, the Monte-Carlo Cluster Statistics implemented in the CosmoMVPA toolbox (Oosterhof, Connolly & Haxby, 2006) was used to perform a sign permutation, corrected for multiple comparisons. The reported *p* value ($\alpha = .05$) represented the percentage rank of the actual observed correlation value within the null distribution.

Results

In the change detection task participants accurately detected 85% of the targets (*SD* = 2.4%, false alarm rate = 2.1%). The behavioral data was not analysed further as the only goal of the task was to encourage participants maintain vigilance during the rapid visual presentation and to minimize eye movements.

Decoding object concepts categories

Trained classifiers were able to discriminate object concepts categories significantly above chance from the evoked responses (Figure 10, 11, 12, 13). These results served as a validation for the results from the subsequent RSA analysis where the study compares the time-varying neural representation of images of objects to the four linguistic models.

All four concept categories show three distinct decoding peaks approximately at 112 ms, 184 ms, and 300 ms post stimulus onset ($p < .05$). For the natural object concepts, the classifier was able to predict the object category above chance between 84 ms to 476 ms ($p < .05$) (Figure 10). For the animal concepts, above chance decoding was observed for a cluster, stretching between 100 ms to 448 ms ($p < .05$) (Figure 11). For the food and drink concepts, decoding accuracy sustained above chance from 84 ms to 486 ms (*p* < .05) (Figure 12). For the clothing object concepts, above chance decoding was again, observed for a cluster, stretching between 108 ms to 448 ms, displaying two identical peaks at 200 ms and at 300 ms (*p* < .05) (Figure 13). Different from natural and food and drink object concepts where the decoding accuracy reached above chance well before 100 ms after stimulus presentation, above chance decoding for clothing object concepts emerged late in the time series at approximately 108 ms (*p* < .05).

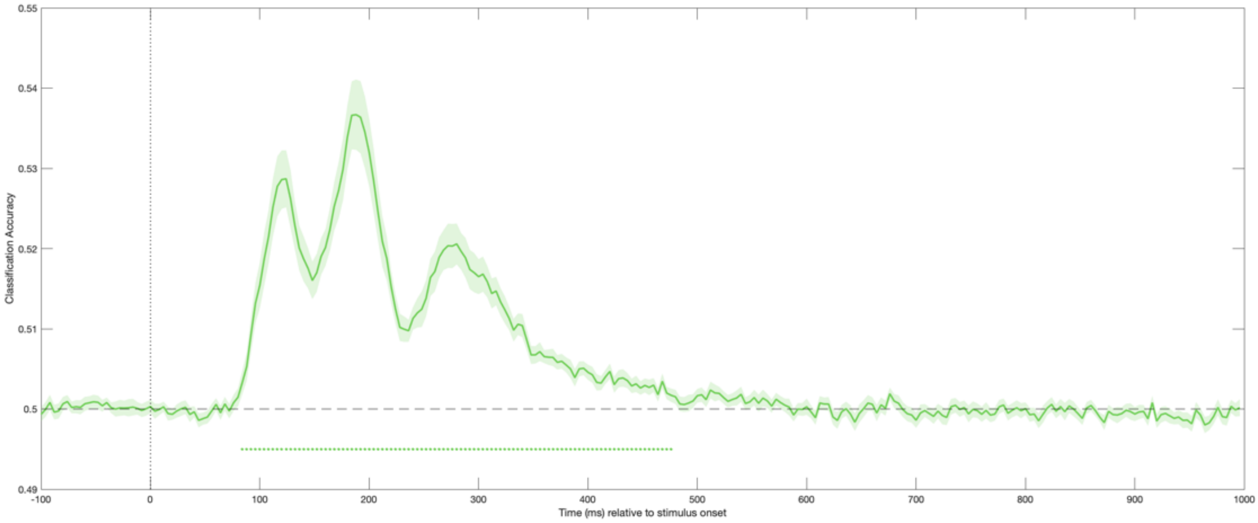


Figure 10. Decoding accuracy for natural object concepts. Decoding accuracy is the average accuracy across all subjects produced by the 12 fold cross-validation. Shades around the line represent the standard error across 30 participants. Vertical line depicts the time of stimulus onset. Colored dots below the x-axis represents the time points where decoding performance is significantly above chance (*p* < .05, corrected for multiple comparisons).

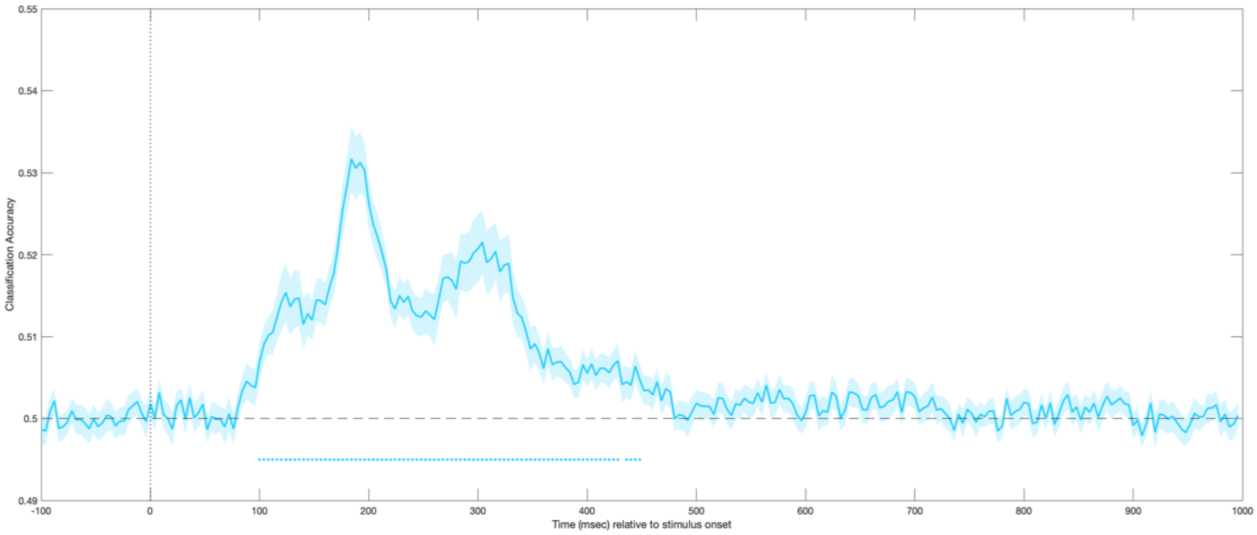


Figure 11. Decoding accuracy for animal concepts. Decoding accuracy is the average accuracy across all subjects produced by the 12 fold cross-validation. Shades around the line represent the standard error across 30 participants. Vertical line depicts the time of stimulus onset. Colored dots below the x-axis represents the time points where decoding performance is significantly above chance (*p* < .05, corrected for multiple comparisons).

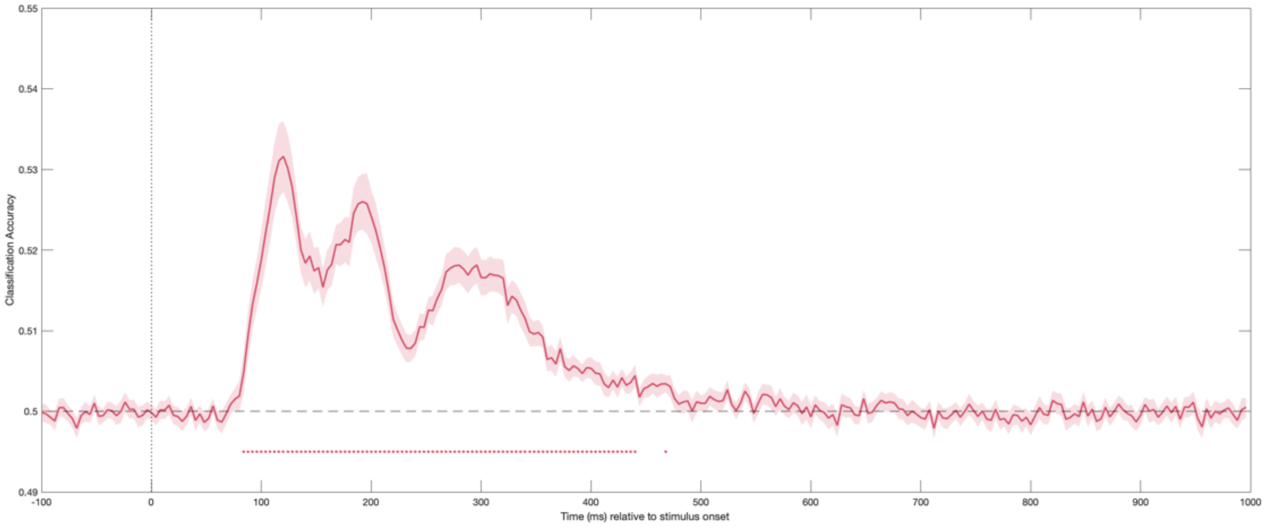


Figure 12. Decoding accuracy for food and drink object concepts. Decoding accuracy is the average accuracy across all subjects produced by the 12 fold cross-validation. Shades around the line represent the standard error across 30 participants. Vertical line depicts the time of stimulus onset. Colored dots below the x-axis represents the time points where decoding performance is significantly above chance (*p* < .05, corrected for multiple comparisons).

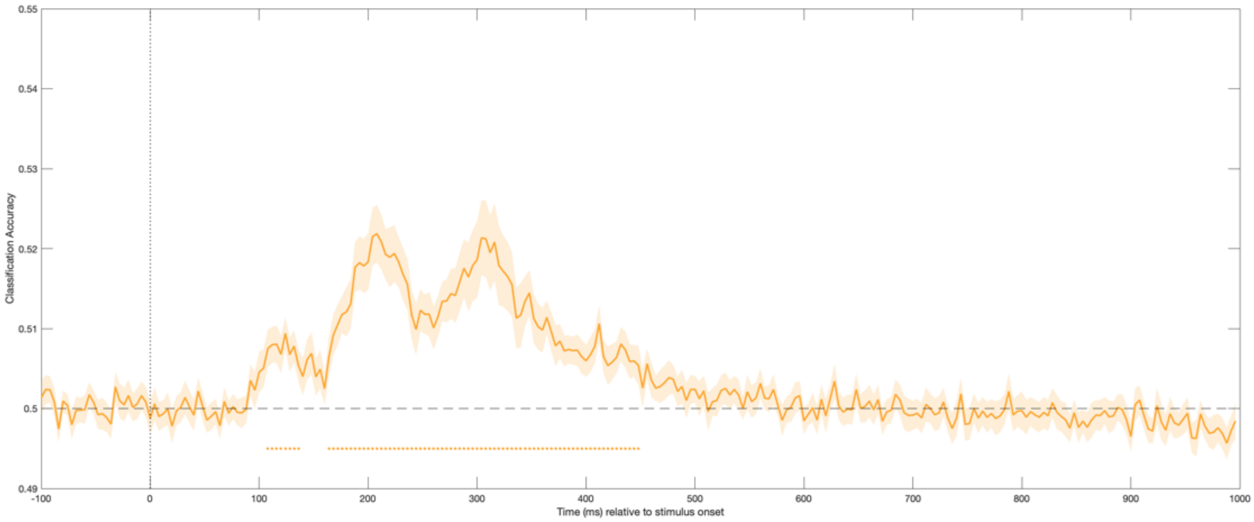


Figure 13. Decoding accuracy for natural object concepts. Decoding accuracy is the average accuracy across all subjects produced by the 12 fold cross-validation. Shades around the line represent the standard error across 30 participants. Vertical line depicts the time of stimulus onset. Colored dots below the x-axis represents the time points where decoding performance is significantly above chance (p < .05, corrected for multiple comparisons).

RSA Analysis

The core aim of the study was to examine the extent to which the structure in semantic representation of object concepts correspond to the neural representation of object images from the four concept categories. The more specific question was, which type of semantic structure, i.e., emergent structure based on word usage patterns (Word2Vec, GloVe) or hierarchical structure, based on explicit dictionary definitions and semantic relations between words (WordNet Wu-Palmer, WordNet PATH), best explained the evoked neural activity patterns by the object concepts.

Results from RSA model testing revealed that evoked neural activity patterns were best captured by the two lexical hierarchical models (WordNet Wu-Palmer, WordNet PATH) (Figure 14). Both lexical hierarchical models' RDMs and the EEG neural RDMs shared a correlation that was significantly above zero between approximately 100 ms to 220 ms post stimulus onset (p < .05). The correlations for the WordNet Wu-Palmer model and neural RDMs peaked at approximately 115 ms, and again at 200 ms ($p < .05$). Similarly, the correlations for the WordNet PATH model and neural RDMs peaked at approximately 115 ms, and again at 190 ms ($p < .05$). For the distributional co-occurrence model Word2Vec, its correlations with neural RDMs peaked at similar time points as the two lexical hierarchical models (115 ms, 200 ms) ($p < .05$), although the correlations were slightly weaker. Interestingly, for the GloVe model, its correlation with the neural RDMs never rose to the level of significance ($p > 0.05$).

We examined whether the correlations for the linguistic model RDMs and the neural RDMs were significantly different as a function of their different measures of semantic similarity of object concepts. Overall, we found the two lexical hierarchical models shared significantly higher correlations with the neural RDMs in comparison to the two distributional co-occurrence models (Figure 15). This difference was greater in contrast tests between the GloVe model and the two hierarchical models. In the WordNet Wu-Palmer and GloVe contrast test, the correlations for the WordNet Wu-Palmer model and neural RDMs were significantly higher than the GloVe model between approximately 100 ms to 200 ms ($p < .05$). In the WordNet PATH and Word2Vec contrasts, the WordNet PATH model's correlations with neural RDMs were again significantly higher than the GloVe model in a cluster stretched between approximately 105 ms to 200 ms ($p < .05$). This differences was lesser in contrasts test for the Word2Vec model and the two hierarchical models. In the WordNet Wu-Palmer and Word2Vec contrast, the Wu-Palmer model's correlation with the neural RDMs were significantly higher than Word2Vec model at 180 ms and 190 ms ($p < .05$). Similarly, the WordNet PATH model's correlation with neural RDMs were also significantly higher than Word2Vec, at approximately 175 ms ($p < .05$). To summarize, the correlations for the two hierarchical models and the neural RDMs were fairly comparable to the Word2Vec model, but both were reliably higher than the GloVe model, which did not reach significance at any point in the time window sampled ($p > 0.05$).

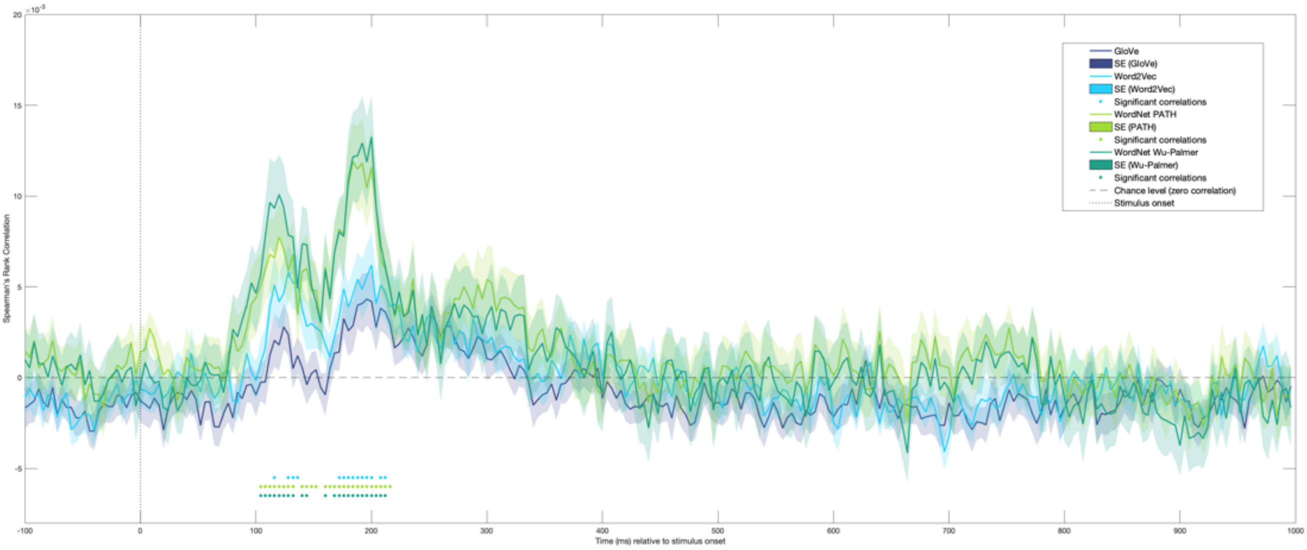


Figure 14. Spearman's rank correlations between the EEG neural RDMs and the four linguistic model RDMs over time. The vertical line indicates the time of stimulus onset. Each colored line depicts the correlations of a linguistic model RDM and the EEG neural RDMs over time. Shades around the colored lines depict standard errors. The colored dotted lines below the x-axis indicate that at the given time point, the correlations of a given model RDM and the EEG neural RDM is significantly above zero ($p < .05$, corrected for multiple comparisons).

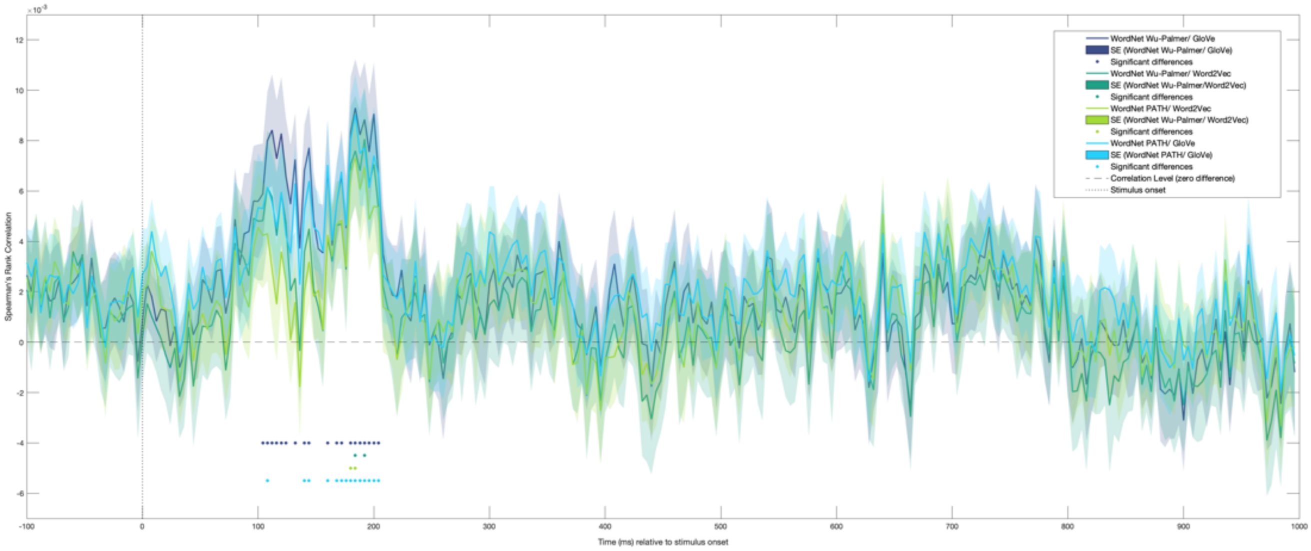


Figure 15. The correlation differences for the four linguistic model RDMs and the over time. The vertical dotted line indicates the time of stimulus onset. The horizontal dashed line indicates the chance level. Each colored line represents the differences in correlations between a pair of model RDMs. Shades around the colored lines depict standard errors. The colored dotted lines below the x-axis depict significant correlation difference for the model pairs. The navy dotted line depicts the comparison of WordNet Wu-Palmer and Glove. The dark green dots depict the comparison of WordNet Wu-Palmer and Word2Vec. The light green dots depict the comparison of WordNet PATH and Word2Vec. The light blue dotted line depicts the comparison of WordNet PATH and GloVe.

As the chosen four linguistic models are based on the English language, the present study suspected these linguistic models, in particular, the two distributional models (Word2Vec, GloVe) that are trained on large English text corpora, might produce different patterns of correlations between the bilingual participants (n = 14) and the English monolingual participants (n = 16). Hence, the study examined whether the linguistic models' correlations with the neural RDMs were different as a function of participants' language profile. The group data was split into two groups (English monolingual, bilingual) based on the language profiles provided by the participant in the demographic questionnaire.

In the monolingual group, the hierarchical models again showed significant correlations with the neural RDMs from approximately 170 ms to 200 ms post stimulus onset (*p* < .05) (Figure 16). Overall, the monolingual group's correlation paths also mirrored the correlation paths in the combined group analysis. For the monolingual group, the correlations for the hierarchical models and the neural RDMs emerged above zero later in the time series, with the PATH model showing more consistent correlations with the neural RDMs than the WordNet Wu-Palmer model during 170 ms to 200 ms post stimulus onset (*p* < .05). Overall, both hierarchical models' correlations with the neural RDMs were visibly more robust in comparison to the distributional models. For the Word2Vec model, its correlations with the neural RDMs were only significantly above zero at 180 ms (*p* < .05). For the GloVe model, its correlations with the neural RDMs was significant at approximately 270 ms after stimulus onset (*p* < .05).

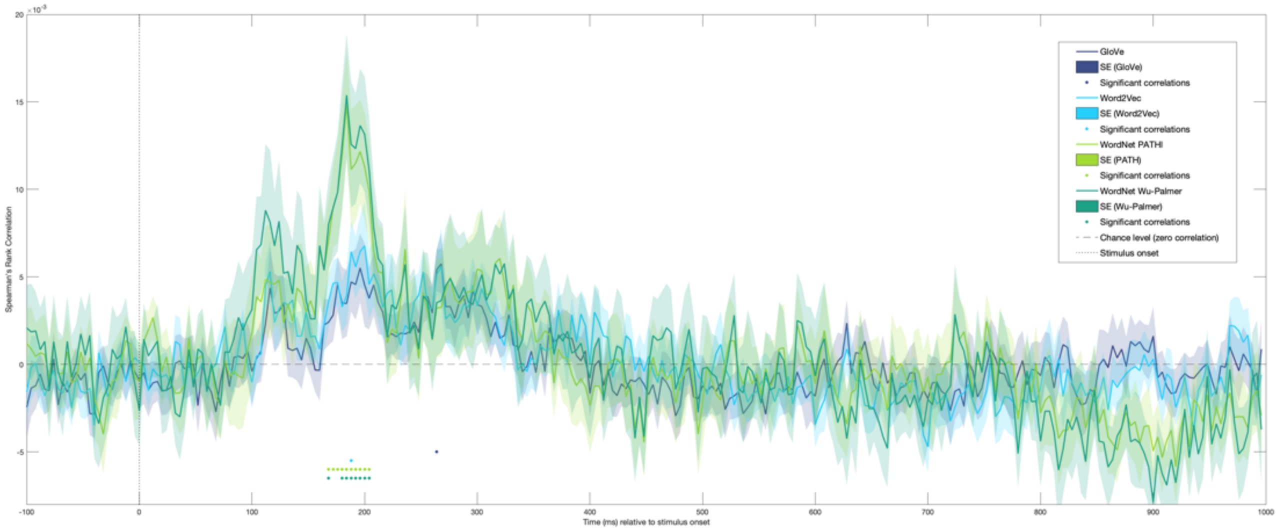


Figure 16. Spearman's rank correlations between the monolingual participants' EEG neural RDMs and the four linguistic model RDMs over time. The vertical line indicates the time of stimulus onset. Each colored line depicts the correlations of a linguistic model RDM and the EEG neural RDM over time. Shades around the colored lines depict standard errors. The colored dotted lines below the x-axis indicate that at the given time point, the correlations of a given model RDM and the EEG neural RDMs were significantly above zero (*p* < .05, corrected for multiple comparisons).

In the bilingual group, the correlations for the WordNet PATH model and the neural RDMs were significantly above zero during 90 ms to 200 ms post stimulus onset (*p* < .05) (Figure 17). Upon visual inspection, the above chance correlation time window for the WordNet PATH and the neural RDMs was noticeably more consistent than the monolingual group, which stretched between 100 ms to 200 ms (*p* < .05). For the WordNet Wu-Palmer model, the correlations with the neural RDMs were significantly above zero from approximately 110 ms to 140 ms (*p* < .05). Unsurprisingly, the correlation with the neural RDMs never rose to significance for neither distributional models during the time window sampled (*p* >.05). Finally, Monte-Carlo significance test found that the four models' correlations with the neural RDMs were not significantly different between the monolingual and the bilingual groups (*p* >.05) (Figure 18).

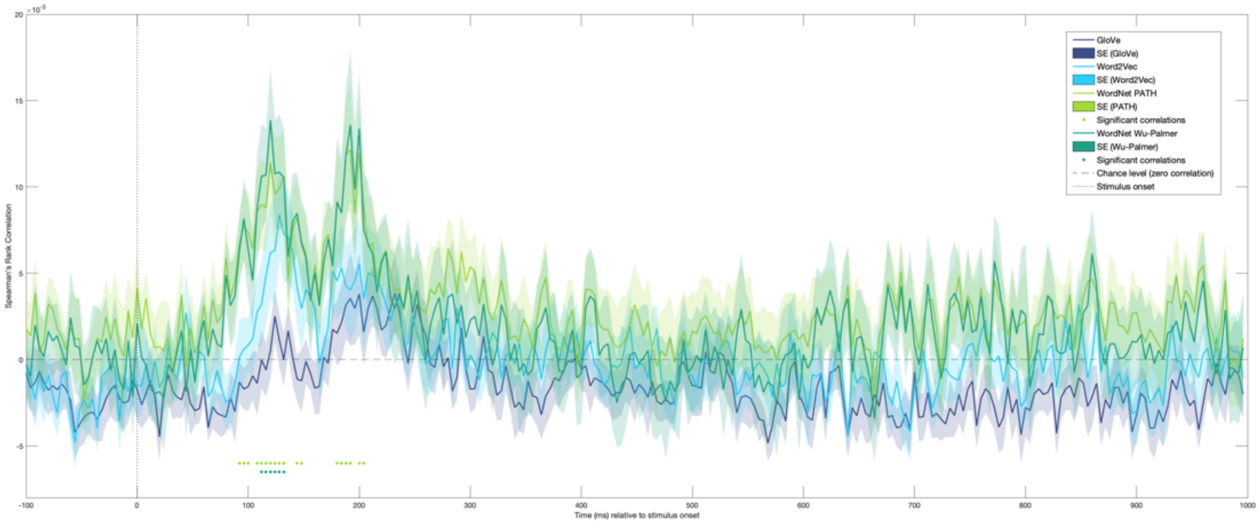


Figure 17. Spearman's rank correlations between the monolingual participants' EEG neural RDMs and the four linguistic model RDMs over time. The vertical line indicates the time of stimulus onset. Each colored line depicts the correlations of a linguistic model RDM and the EEG neural RDM over time. Shades around the colored lines depict standard errors. The colored dotted lines below the x-axis indicate that at the given time point, the correlations of a given model RDM and the EEG neural RDMs were significantly above zero (*p* < .05, corrected for multiple comparisons).

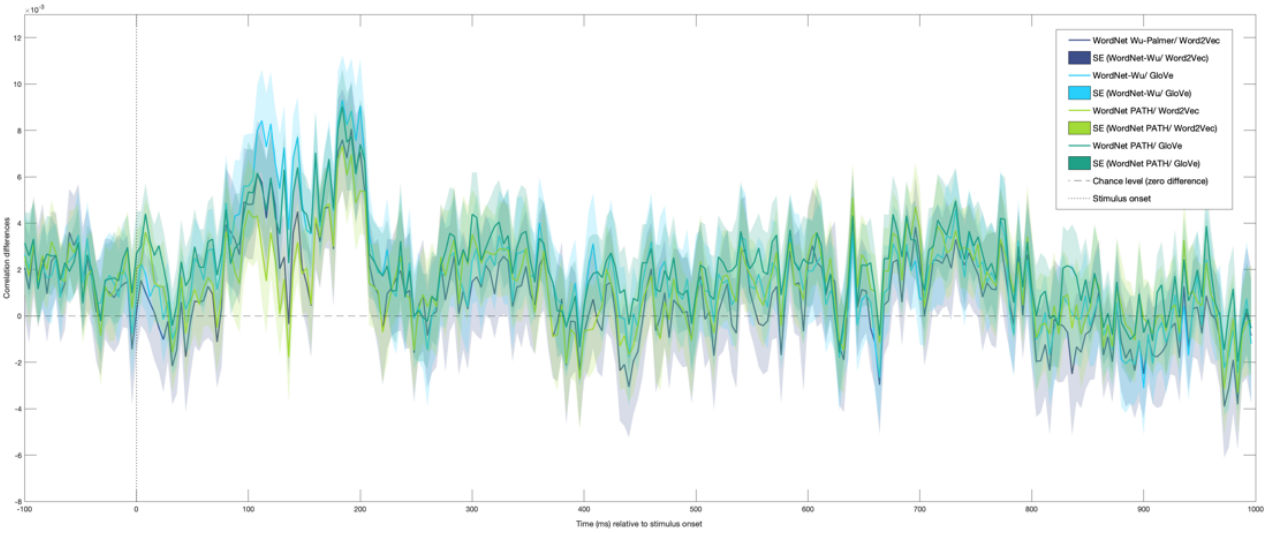


Figure 18. The correlation differences for the four linguistic model RDMs between the monolingual participants and the bilingual participants. The vertical dotted line indicates the time of stimulus onset. The horizontal dashed line indicates the chance level. Each colored line represents the differences in correlations between a pair of model RDMs. Shades around the colored lines depict standard errors. The four linguistic models' correlations with the neural RDMs were not significantly different between the two groups of participants (*p* >.05, corrected for multiple comparisons).

Discussions

Summary

Informed by the intuition in language, computational linguistic models have been incorporated in research studying the relationship between linguistic semantics in language and semantically imbued representation in the brain. Precisely due to their direct and testable predictions, computational models have shown to be effective in modelling human judgments (Pereira, Gershman, Ritter & Botvinick, 2016; Goldberg & Levy, 2014; Landauer & Dumais, 1997) and neural activity involved in processing conceptual and semantic knowledge contents (Howard, Shankar, Jagadisan, 2011; Mitchell et al., 2008; Sassenhagen & Fiebach, 2019). However, the question remains as to what extent do semantic representation extracted from streams of noisy, imprecise, and often incomplete data in natural language resemble conceptual representation acquired outside language. To address this question, the broader aim of the present study was to examine whether the structure in the time-varying neural representation of images representing a multitude of object concept categories corresponds to the structure of semantic representation produced by two general methods of semantic similarity measures. The more specific aim of the study was to test which type of semantic representation, hierarchical or distributional, best explains the neural representation of these object concepts.

Incongruent with previous research results where distributional co-occurrence models are typically better fits to both behavioural data (Turney & Pantel, 2010; Vigliocco, Vinson, Lewis & Garrett, 2014) and neural data (Carlson et al., 2013; Sassenhagen & Fiebach, 2019) than lexical hierarchical WordNet models, the present study found semantic representation produced by WordNet-based similarity measures, based on dictionary word definitions and semantic relationships between words, were the most compatible with time-varying neural representation of images of object concepts. By comparison, the correlations for the two distributional co-occurrence models (Word2Vec, GloVe) and the neural representation were relatively weaker, although they followed a similar correlation path as the two hierarchical models.

As the decoding results serve as a validation for the results in the subsequent RSA analysis, the following section first discusses the time course of object concept processing and the possible contributions from low-level visual properties to the object concept category decoding results. Next, the results are discussed in relation to past works and the role of stimulus characteristics on the decoding analysis. Then, the representational similarity between the neural representation of object concepts and the semantic representation produced by WordNet's lexical hierarchical models is discussed, alongside the potential contribution from WordNet's non-hierarchical relations in capturing information beyond the simple synonym (hyper/hyponym) relations between concepts, such that perceptual and functional properties may also be encoded in the WordNet's hierarchy. The disparity between the currents results and those from past studies and the potential influence from different methodological approaches on the divergent results are also both discussed. Finally, the limitations of using a stimuli set with a strong categorical structure such as the THINGS database are discussed, as are the challenges in using machine learning and neural decodability analyses in areas that currently that drives the decoding classification, such as conceptual research and language.

The temporal profile of object concept category processing

In this study, the goal of the decoding analysis was to differentiate between the neural activity evoked by object images from the four concept categories. The current results show the classifier successfully decoded neural responses evoked by images within the natural, animal, food and drink, and clothing categories. Overall, the concept category decoding was above chance between 84 ms and 400 ms, and at 400 ms, there were already four new images presented. This time window is marked by three distinct decoding peaks at approximately 112 ms, 184 ms, and 300 ms. The decoding results suggest that information that was present in evoked neural activity is unique to the image presented at a specific time point, such that the classifier was able to linearly differentiate between neural activity evoked by concept categories depicted by the stimuli. The more sustained decoding time window was observed for natural object concepts (84 ms to 474 ms) and food and drink object concepts (84 ms to 486 ms) than animal object concepts (100 ms to 448 ms) and clothing object concepts (108 ms to 448 ms).

Varying peak decoding performances also highlighted the temporal profile specific to each concept category. Peak decoding performance first emerged in food and drink concepts at approximately 100 ms, this was followed by animal concepts at 188 ms, natural concepts at 190 ms, and finally, clothing object concepts, with the onset of two comparable peaks at 200 ms and 300 ms. The temporal profiles in the current decoding results suggest thematic images of four concept categories are processed in a cascade, starting with concepts deemed by some researchers as more crucial for solving survival problems (Caramazza & Shelton, 1998; Mahon & Caramazza, 2009), followed by the concomitant activation of natural and animal concepts that are associated more with their perceptual features (Capitani, Laiacona, Mahon & Caramazza, 2003; Martin, Wiggs, Ungerleider, Haxby, 1996; Warrington, 1987; Warington & Shallice,1984), and finally, man-made object concepts that are associated more with their functional features (Tranel et al., 1997; Warrington, 1987; Warrington & Shallice,1984).

Possible contribution from low-level visual properties in decoding accuracy

Low-level visual properties such as color, luminance and contours can be problematic for decoding studies using naturalistic photographs (Cichy, Pantazis & Oliva, 2014). In the past, low-level visual properties have been suspected to enhance the decoding performance in early time series (<100 ms) (Grootswagers et al., 2019). In the RSVP study by Grootswagers et al. (2019), the decoding performance at the exemplar-level emerged slightly earlier (80 ms) in comparison to the category-level (100 ms). Hence, in the present study, it is possible that the decoding performance obtained before 100 ms could be accounted for by decodable low-level visual properties of the stimuli. After 100 ms, however, it becomes unlikely that the decoding accuracy is solely driven by low-level visual properties. This logic is supported by the results from Liu et al.'s (2009) intracranial recording study where category-specific information for objects is distinguished in the brain from as early as 100 ms after stimulus presentation, and this effect remains robust even after extrapolating across various versions of the same image (rotated and scaled). Likewise, Carlson et al. (2011) also found position invariant object information can be extracted from the visual system from as early as 105 ms for objects, and 135 ms for object categories. Importantly, in the final stage of compiling the image database (THINGS - Herbart et al., 2019) all images are validated in CorNet-S (Kubilius et al., 2018), a deep convolutional neural network that controls for spatially localized low-level features (Herbart et al., 2019). This final step increases the degree of visual variability for the exemplars and reduces the likelihood of contributions from low-level visual properties in the classification accuracy, at least from 100 ms onwards. Finally, the backgrounds in the images are preserved to retain an ecologically valid representation of the referent concepts, some images featuring animals, such as whale, cow, also were featured in similar backgrounds, for example, ocean, grass. The background similarity may increase the familiarity associated with certain types of animal concepts. Although, images that potentially may be affected by this issue form a relatively low section of the stimulus set (Table 3).

Table 3

Summary of Images and Concepts in General Object Concept Categories

Low-Level Object Concept Category	Concepts	Images
Natural	610	7320
Animal	177	2214

Food and Drink	314	3768
Clothing	132	1584

Note. General concept categories in THINGS database are generated from human ratings (Amazon Mechanical Turk) and WordNet word-sense disambiguation.

In relation to previous decoding literature

While the distinct temporal profile of emerging category information in the current results suggest that these decoding profiles reflect the processing of object category information, there are two subtle variances between the current results and the existing decoding literature that deserve covering in more detail. Firstly, the concept category decoding peaks in the current study might be seen to emerge slightly earlier in comparison to other decoding studies (Carlson et al., 2011; Carlson, Tovar, Alink, & Kriegeskorte, 2013; Liu, Agam, Madsen & Kreiman, 2009). For example, in Grootswagers et al., (2019), the peak decoding performance in the 5Hz condition for abstract categories, such as animacy, occurred at 150 ms, 200 ms and 400 ms, with peak decoding for object category occurred at 200 ms. By comparison, the three distinct decoding peaks in the current results, under the 10Hz condition, emerged slightly earlier (112 ms, 184 ms, and 300 ms). These differences, however, are likely accounted for by factors such as different visual presentation rates and the characteristics of the stimuli for which there is no consensus in the field (Grootswagers et al., 2019; Carlson et al., 2011; Carlson et al., 2013; Carlson, Tovar, Alink, & Kriegeskorte, 2013; Cichy, Pantazis, & Oliva 2014) as these properties are chosen based on the research questions been asked. Also, the stimuli set used in the present study (26,107 images) is much larger than the stimuli set (200 images) used in Grootswagers et al. (2019), which increases the power in the current analysis.

Secondly, the time spans of above-chance decoding in the current results are also somewhat shorter than previous research where the object category decoding performance sustained until 500 ms post-stimulus onset (Carlson et al. 2011; Grootswagers et al. 2019). However, this also varies with the stimulus set. The stimulus set used by Carlson et al. (2011) for example included human faces, which drive high-level visual cortex activity strongly (Cichy, Pantazis, & Oliva, 2014) while other sets contain cropped images of objects (Grootswagers et al. 2019) which enhance low-level visual properties such as lines and silhouettes and drive a stronger V1 response that can take longer to decay (Cichy, Pantazis, & Oliva, 2014).

The time-varying neural representation of object concepts correspond to WordNet models

Overall, the category decoding results observed here align with prior research, with only small time differences in peak decoding and length of sustained decoding, both of which can vary with presentation rate and stimulus set. Therefore it is unlikely that the differences observed in the RSA comparing neural and linguistic model representational similarity are accounted for by any systematic differences in the neural data or the stimulus set of the present study.

The core aim of the RSA analysis was to compare the structure in the neural representation object categories to the structure in semantic representation produced by linguistic models. The current results suggest that the structure of these neural representations correspond primarily to the semantic representation produced by lexical hierarchical models WordNet Wu-Palmer and WordNet PATH. Both models show above zero correlations with the neural RDMs between 100 ms to 220 ms post-stimulus onset and display similar correlation peaks, with the WordNet Wu-Palmer model's peaks at approximately 115 ms and 200 ms, and the WordNet PATH model's peaks at approximately 115 ms and 190 ms. By contrast, the correlation for the Word2Vec model was relatively weaker. Surprisingly, the correlation for the GloVe model never rose to the level of significance. Between model contrast tests confirm this distinction, revealing that overall, the correlations for the lexical hierarchical models were more robust than the distributional co-occurrence models. As approximately half of the participants are non-native English speakers, it was initially suspected that a different pattern of correlations would be found in bilingual speaker participants, especially as Word2Vec and GloVe both are trained on large English text corpora. However, the individual group analyses indicate that the patterns of correlations are not significantly different between the two groups.

From perceptual to conceptual: modality-specific features encoded in WordNet's taxonomy hierarchy

There is a high likelihood that the semantic representation captured by the two WordNet models also captured featural and functional properties specific to each object concept category. While the two similarity measures primarily quantify conceptual similarity or conceptual distance as the path distance between concepts in the taxonomy hierarchy, this quantification also makes complimentary use of other non –hierarchical relations and short dictionary definitions (gloss) (Miller, 1995). Non-hierarchical relations such as has-part, is-made-of, is-an-attribute-of, mean conceptual relatedness can also be expressed in ways beyond meaning similarity, such as '*a wheel is a part of a car'*, '*a knife is used to cut bread'*, '*night is the opposite of day'*, '*snow is made up of water*' (Pedersen et al., 2004, p. 1024). As such, semantic representation captured using non-hierarchical relations can also be partially construed as modality-specific featural representation. For example, the meaning for the concept '*car*' partly entails its perceptual information, and in WordNet, the concept '*car*' is encoded in the has-part relation as '*car-has-rear window*' (Richardson, 1994) (Figure 19). In line with this argument, the correspondence between the structures in semantic representation produced by WordNet and neural representation could be ascribed to the correspondence between the featural representation of these concrete entities been opaquely encoded in WordNet non-hierarchical relations and the modality-specific neural representation of thematic images of these entities.

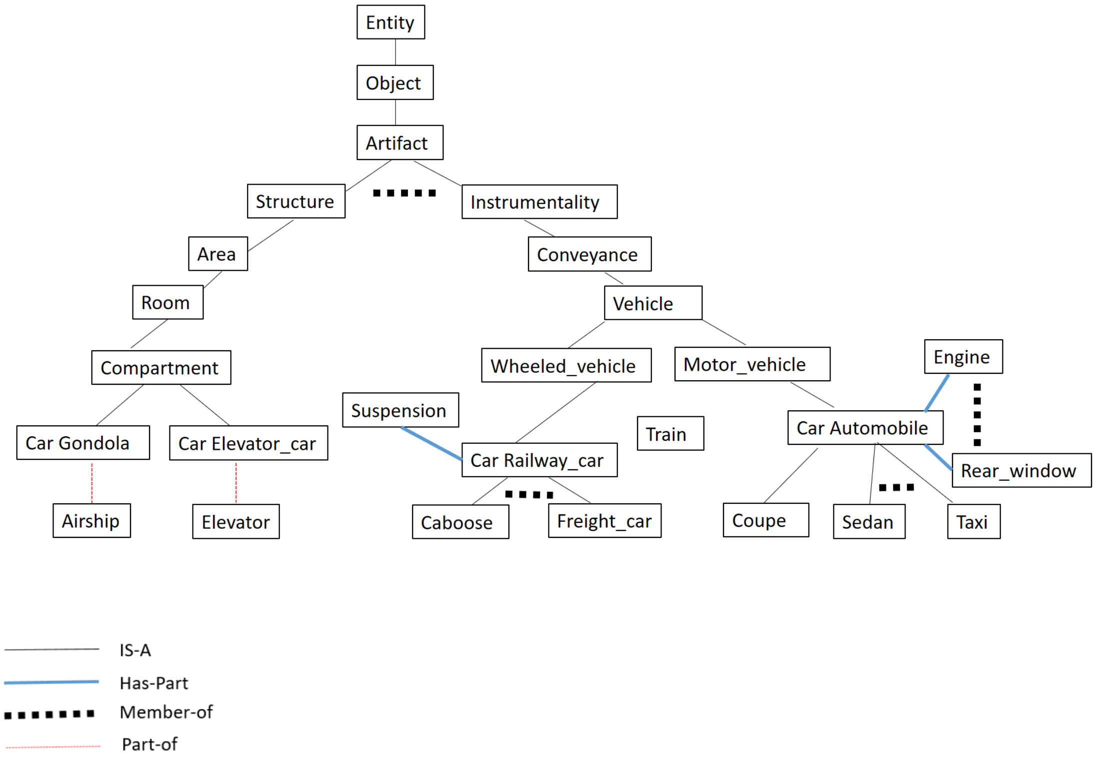


Figure 19. WordNet extract for the concept *'car.'* Adapted from 'Using WordNet as a knowledge base for measuring semantic similarity between words,' by R. Richardson, A.F. Smeaton, J, Murphy, 1994, *Dublin City University*.

This argument leads back to the classic feature-based accounts of semantic representation and semantic memory (Rosch et al., 1976; Smith, Shoben & Rips, 1974). Notably, in work by Smith et al. (1974), the representation of a concept is construed as a binary list of characteristic features in a multidimensional feature space where the features correlate to their external referent object. Similarly, in modern feature-based theories, feature distinctiveness is also deemed to be instrumental for constructing semantic representation and deriving word meanings (Cree, McRae & McNorgan, 1999). According to Cree, McRae & McNorgan (1999), the presence of a distinctive feature distinctiveness allows two concepts to be differentiated on the basis of perceptual dissimilarity and hence, feature distinctiveness is crucial for constructing and representing conceptual knowledge. Strong support for these accounts comes from Cree and McRae's (2003) study on category-specific impairment patients' data. In their study, feature distinctiveness and visual complexity are among the chief factors that predicted the major selective semantic impairments trends found in patients (Cree, McRae, 2003, p.163). Feature-based models, such as the attractor network trained on semantic-features of objects, have been shown to predict human similarity judgments for words with high accuracy, as evidenced by the results from semantic priming tasks (Cree & McRae, McNorgan, 1999).

Analogous to this account is Vigliocco, Vinson, Lewis, & Garrett's (2004) Future and Unitary Semantic Space semantic (FUSS) similarity measures. Trained in the low-dimensional structure of the elementary features of 456 common object and action words, the FUSS similarity measures outperformed distributional word co-occurrences model Latent Semantic Analysis in human performances in word similarity judgments and semantic categorization (Vigliocco, Vinson, Lewis, & Garrett 2004; Vinson, Vigliocco, Cappa, & Siri, 2003). In light of these findings, Vigliocco, Vinson, Lewis, & Garrett (2004) argued that modality-specific featural representation is partly encoded in word meanings for concrete entities.

Congruent with their proposal, cross-linguistic studies show that Spanish, English and Chinese native speakers produced similar patterns of similarity judgments for concrete objects (e.g., vase, bottles) despite the cross-linguistic variability for these concepts, and the mismatches commonly found between linguistic categories and perceptually continuous domains such as color (Malt et al., 1999; Vigliocco, Vinson, Lewis, & Garrett, 2004). Consistent with the notion that word meaning for concrete entities is partially grounded in modality-specific featural representation that is largely unconstrained by the variabilities in languages, the present study shows that the correspondence between the semantic representation of concrete object concepts as conceptualized by WordNet's hierarchical and non-hierarchical relations, and the neural representation of thematic images of the same concepts remains robust even after extrapolating across different language profiles.

Weak correspondence between distributed semantic representation and neural representation of concrete object concepts

The results in the RSA show the structure in the time-varying neural representation of object concept corresponds to the structure in the distributed semantic structure produced by Word2Vec. Intriguingly, despite GloVe's robust performance in modelling human performances in psycholinguistic tasks (Pereira et al, 2018; Pennington, Socher & Manning, 2014) and in neural activity patterns associated with cognitive contents (Carlson et al., 2013; Sassenhagen & Fiebach, 2019), its' correlations with the neural RDMs never rose above the chance level. While common to both models is the principle that similar words appear in similar contexts, the nature of semantic representation produced by these models might be subjected to differences in the types of training text corpora, vector length and the size of the context window. As GloVe leverages the global statistics of word co-occurrence in the document and also considers the relations between multiple word pairs, it has been suggested that GloVe captures the more intricate, nuanced relations between concepts than local context prediction-based models, such as Word2Vec (Pennington, Socher & Manning, 2014).

Notably, the current results contradict previous fMRI study by Carlson, Simmons, Kriegeskorte & Slevc (2013) on the organization of object representation in the inferior temporal cortex and Sassenhage & Fiebach's (2019) ERP study on induced neural activity from English and German concrete nouns. In Carlson et al. (2013), it was found that the emergent semantic structure captured by distributional co-occurrence models, Latent Semantic Analysis (LSA) (Landauer & Dumais,1997) and Correlated Occurrences Analogue to Lexical Semantics (COAL) (Rohde, Gonnerman & Plaut, 2005), corresponded to the geometric structure in object representation in inferior temporal cortex more so than WordNet-based similarity measures PATH, and the gloss vector measure LESK. Similarly, Sassenhagen & Fiebach (2019) also found brain activity induced by English and German nouns encoded distributed semantic representation of word meanings produced by GloVe and Word2Vec more so than hierarchical lexical models WordNet and GermanNet, which demonstrated significant but weaker correlations with the neural activity in comparison to GloVe and Word2Vec.

Reconciling the disparities between the current results and these two streams of results is complicated by several factors. First, the characteristics of the stimuli set used in these two studies varied. For example, in Carlson et al. (2013), a stimulus set of 92 cropped colour images depicting 67 object concepts (natural and artificial objects, human faces and bodies, and animals) was used and, as discussed previously, cropped images and human faces tend to drive activity in V1 and upper visual systems strongly to the extent that might affect the analysis outcome whereas the stimulus set used in Sassenhagen & Fiebach (2019) consisted of word labels of 960 English and 150 German nouns, which introduces orthographical and lexical complexities into the stimuli set. Second, the neuroimaging data used by these studies were either entirely or partially adopted from previous studies. For example, the fMRI recordings used in Carlson et al. (2013) were adopted from a previous study by Kriegeskorte et al. (2008), recorded for a relatively small sample (n = 4) while they performed a colour-discrimination task. Likewise, the Event-Related Potential (ERP) data for the English sample used in Sassenhagen & Fiebach (2019) was adopted from a previous study by Dufau, Grainger, Midgley & Holcomb (2015) on word reading, recorded from 75 participants while they performed a lexical decision go/no go task. The main issue with adopted data, according to Herbart et al. (2019), is that the data is always influenced by previous studies' stimulus selection criteria that serve to test a specific set of hypotheses. Finally, the quality and the nature of the neural representation obtained also differs according to the modalities of brain-activity measurement used. In Carlson et al. (2013), fMRI was used to measure brain activity in IT and the primary visual cortex, whereas Sassenhagen & Fiebach (2019) used EEG to capture whole-brain activity responses. These fundamentally different neuroimaging techniques and the brain regions being sampled mean that the neural representation captured could reasonably be said to differ from those used here, such that the spatial/ temporal trade-off in fMRI data means that one single fMRI image is likely to double the time as the entire time down-sampled in the current decoding and

RSA analyses (Beres, 2017; Biasiucci, Franceschiello & Murray, 2019). Together, the issues of stimulus characteristics, pre-adopted data, and relating activity patterns captured by different modalities of brain-activity measurement (fRMI and EEG) each poses a challenge for reconciling the disparities in results from different studies. These challenges, however, are not unique to the present study but other research areas that utilize different brain-activity measurements (fMRI, M/EEG) and computational modelling and as well.

Limitations in the study: the categorical structure of the stimuli set and the assumption driven nature of machine learning

The present study has two key limitations. Namely, the strong categorical structure in the stimuli set and the inherent challenges for applying machine learning in research areas that currently, lack a fundamental understanding of the underlying neural mechanisms that drive the decoding classification. First, a top-down validation was carried out using on all concepts using WordNet word sense disambiguation to remove highly similar object concepts and synonyms from the THINGS database (Herbart et al., 2019). Hence, while these object concepts demonstrate a high level of conceptual distinctiveness, the word disambiguation process itself may inadvertently introduce a strong categorical structure in the stimuli set that contribute to the WordNet’s hierarchical models' performance in the RSA analysis.

The second limitation comes from the challenge common to machine learning approaches in cognitive neuroscience. According to Carlson et al. (2018), precisely due to the simple patterns produced by machine learning, it is easy to over-interpret the data and see phenomena that are not there. Furthermore, Carlson et al. (2018) also point out that decodable information may not be the same as the underlying cognitive mechanism used for constructing a representation. Hence, for research areas where the neural source that drives the decoding performance has yet to be identified, such as memory and language, one needs to be prudent about differentiating information that is merely decodable and the information the brain uses to represent the information content in question.

Conclusion

The core aim in the present study was to investigate the extent to which the structure in the time-varying neural representation of concrete object concepts resembles the structure in semantic representation produced by linguistic models that is primarily acquired from streams of imprecise and noisy data in natural language. The current results show the structure in these neural representations primarily correspond to the structure in the semantic representation produced by WordNet models, based on the concepts' explicit dictionary definitions and their hierarchical and non-hierarchical relations in WordNet taxonomy. By comparison, the correspondence with the neural data is relatively weaker for the distributed semantic representation produced by distributional models trained on word co-occurrences statistics in text corpora. The current results do not necessarily mean that distributed semantics do not play a part in representing conceptual knowledge, as the results in RSA analysis show, all linguistic models display similar correlation paths with the neural data within the time window where robust decoding performances for object concept categories were observed. This temporal synchrony between the linguistic models coupled with the potential influence from non-hierarchical relations in WordNet suggests the rapid transition from perception to representation compatible with language and conceptual thoughts, is underpinned by concept category distinctive features.

References

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. Psychological review, 116(3), 463.

Baroni, M., Murphy, B., Barbu, E., & Poesio, M. (2010). Strudel: A corpus□based semantic model based on properties and types. Cognitive science, 34(2), 222-254.

Barsalou, L. W. (2008). Cognitive and neural contributions to understanding the conceptual system. Current Directions in Psychological Science, 17(2), 91-95.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. Trends in cognitive sciences, 7(2), 84-91.

Beres, A. M. (2017). Time is of the essence: A review of electroencephalography (EEG) and event-related brain potentials (ERPs) in language research. Applied psychophysiology and biofeedback, 42(4), 247-255.

Biasiucci, A., Franceschiello, B., & Murray, M. M. (2019). Electroencephalography. Current Biology, 29(3), R80-R85.

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.

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433-436.

Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. Behavior research methods, 39(3), 510-526.

Capitani, E., Laiacona, M., Mahon, B., & Caramazza, A. (2003). What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. Cognitive Neuropsychology, 20(3-6), 213-261.

Caramazza, A., Hillis, A. E., Rapp, B. C., & Romani, C. (1990). The multiple semantics hypothesis: Multiple confusions?. Cognitive neuropsychology, 7(3), 161-189.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. Journal of cognitive neuroscience, 10(1), 1-34.

Caramazza, A., & Mahon, B. Z. (2003). The organization of conceptual knowledge: the evidence from category-specific semantic deficits. Trends in cognitive sciences, 7(8), 354-361.

Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal resolution decoding of object position and category. Journal of vision, 11(10), 9-9.

Carlson, T. A., Schrater, P., & He, S. (2003). Patterns of activity in the categorical representations of objects. Journal of cognitive neuroscience, 15(5), 704-717

Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: the first 1000 ms. Journal of vision, 13(10), 1-1.

Carlson, T. A., Simmons, R. A., Kriegeskorte, N., & Slevc, L. R. (2014). The emergence of semantic meaning in the ventral temporal pathway. Journal of cognitive neuroscience, 26(1), 120-131.

Carlson, T., Goddard, E., Kaplan, D. M., Klein, C., & Ritchie, J. B. (2018). Ghosts in machine learning for cognitive neuroscience: Moving from data to theory. NeuroImage, 180, 88-100.

Carlson, T. A., Grootswagers, T., & Robinson, A. K. (2019). An introduction to time-resolved decoding analysis for M/EEG. arXiv preprint arXiv:1905.04820.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nature neuroscience, 2(10), 913.

Chen, L., & Rogers, T. T. (2014). Revisiting domain□general accounts of category specificity in mind and brain. Wiley Interdisciplinary Reviews: Cognitive Science, 5(3), 327-344.

Cichy, R. M., Pantazis, D., & Oliva, A. (2016). Similarity-based fusion of MEG and fMRI reveals spatio-temporal dynamics in human cortex during visual object recognition. Cerebral Cortex, 26(8), 3563-3579.

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. Cognitive Science, 23(3), 371-414.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). Journal of experimental psychology: general, 132(2), 163.

Clarke, A., Taylor, K. I., & Tyler, L. K. (2010). The Evolution of Meaning: Spatio-temporal Dynamics of Visual Object Recognition.

Clarke, A., Taylor, K. I., Devereux, B., Randall, B., & Tyler, L. K. (2012). From perception to conception: how meaningful objects are processed over time. Cerebral Cortex, 23(1), 187-197.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. Journal of verbal learning and verbal behavior, 8(2), 240-247.

Damasio, A. R. (1989). The brain binds entities and events by multiregional activation from convergence zones. Neural computation, 1(1), 123-132.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.

DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. Trends in cognitive sciences, 11(8), 333-341.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition?. Neuron, 73(3), 415-434.

Dufau, S., Grainger, J., Midgley, K. J., & Holcomb, P. J. (2015). A thousand words are worth a picture: Snapshots of printed-word processing in an event-related potential megastudy. Psychological science, 26(12), 1887-1897.

Durda, K., Buchanan, L., & Caron, R. (2009). Grounding co-occurrence: Identifying features in a lexical co-occurrence model of semantic memory. Behavior Research Methods, 41(4), 1210-1223.

Duarte, L. R., Marquié, L., Marquié, J. C., Terrier, P., & Ousset, P. J. (2009). Analyzing feature distinctiveness in the processing of living and non-living concepts in Alzheimer's disease. Brain and cognition, 71(2), 108-117.

Farah, M. J., Hammond, K. M., Mehta, Z., & Ratcliff, G. (1989). Category-specificity and modality-specificity in semantic memory. Neuropsychologia, 27(2), 193-200.

Farah, M. J., & McClelland, J. L. (2013). A computational model of semantic memory impairment: Modality specificity and emergent category specificity (Journal of Experimental Psychology: General, 120 (4), 339–357). Exploring Cognition: Damaged Brains and Neural Networks, 79-110.

Fischer-Baum, S., Jang, A., & Kajander, D. (2017). The cognitive neuroplasticity of reading recovery following chronic stroke: a representational similarity analysis approach. Neural plasticity, 2017.

Forde, E., & Humphreys, G. (Eds.). (2005). Category specificity in brain and mind.

Funnell, E., & Sheridan, J. (1992). Categories of knowledge? Unfamiliar aspects of living and nonliving things. Cognitive Neuropsychology, 9(2), 135-153.

Gainotti, G. (1996). Cognitive and anatomical locus of lesion in a patient with a category-specific semantic impairment for living beings. Cognitive Neuropsychology, 13(3), 357-390.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. Journal of memory and language, 43(3), 379-401.

Goldberg, Y., & Levy, O. (2014). word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. arXiv preprint arXiv:1402.3722.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. Psychological review, 114(2), 211.

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. Journal of cognitive neuroscience, 29(4), 677-697.

Grootswagers, T., Robinson, A. K., & Carlson, T. A. (2019). The representational dynamics of visual objects in rapid serial visual processing streams. NeuroImage, 188, 668-679.

Hebart, M. N., Dickter, A. H., Kidder, A., Kwok, W. Y., Corriveau, A., Van Wicklin, C., & Baker, C. I. (2019). THINGS: A database of 1,854 object concepts and more than 26,000 naturalistic object images. bioRxiv, 545954.

Howard, M. W., Shankar, K. H., & Jagadisan, U. K. (2011). Constructing semantic representations from a gradually changing representation of temporal context. Topics in Cognitive Science, 3(1), 48-73.

Humphreys, G. W., Price, C. J., & Riddoch, M. J. (1999). From objects to names: A cognitive neuroscience approach. Psychological research, 62(2-3), 118-130.

Humphreys, G. W., & Forde, E. M. (2001). Hierarchies, similarity, and interactivity in object recognition:"Category-specific" neuropsychological deficits. Behavioral and Brain Sciences, 24(3), 453-476.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. Journal of Experimental Psychology: Human Learning and Memory, 6(1), 1.

Ishibashi, R., Pobric, G., Saito, S., & Lambon Ralph, M. A. (2016). The neural network for tool-related cognition: an activation likelihood estimation meta-analysis of 70 neuroimaging contrasts. Cognitive Neuropsychology, 33(3-4), 241-256.

Keysers, C., Xiao, D. K., Földiák, P., & Perrett, D. I. (2001). The speed of sight. Journal of cognitive neuroscience, 13(1), 90-101.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. Vision research, 46(11), 1762-1776.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3?.

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., ... & Bandettini, P. A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron, 60(6), 1126-1141.

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: integrating cognition, computation, and the brain. Trends in cognitive sciences, 17(8), 401-412.

Kubilius, J., Schrimpf, M., Nayebi, A., Bear, D., Yamins, D. L., & DiCarlo, J. J. (2018). CORnet: modeling the neural mechanisms of core object recognition. *BioRxiv*, 408385.

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*, *104*(2), 211.

Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. Neuron, 62(2), 281-290.

Locke, J. (1841). *An essay concerning human understanding*.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*, *8*, 213.

Louwerse, M. M. (2007). Symbolic or embodied representations: A case for symbol interdependency. Handbook of latent semantic analysis, 107-120.

Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of object knowledge in the brain?. Trends in cognitive sciences, 15(3), 97-103.

Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. Annual review of psychology, 60, 27-51.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, *40*(2), 230-262.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. Journal of neuroscience methods, 164(1), 177-190.

Marti, S., & Dehaene, S. (2017). Discrete and continuous mechanisms of temporal selection in rapid visual streams. Nature communications, 8(1), 1955.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. Science, 270(5233), 102-105.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. Current opinion in neurobiology, 11(2), 194-201.

Martin, A. (2007). The representation of object concepts in the brain. Annu. Rev. Psychol., 58, 25-45.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. Nature, 379(6566), 649.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. Nature reviews neuroscience, 4(4), 310.

McRae, K., De Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. Journal of Experimental Psychology: General, 126(2), 99.

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.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. Advances in neural information processing systems, 3111-3119.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.

Mikolov, T., Yih, W. T., & Zweig, G. (2013, June). Linguistic regularities in continuous space word representations. In Proceedings of the 2013 Conference of the North

American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 746-751).

Miller, G. A., & Charles, W. G. (1991). Contextual correlates of semantic similarity. Language and cognitive processes, 6(1), 1-28.

Miller, G. A. (1995). WordNet: a lexical database for English. Communications of the ACM, 38(11), 39-41.

Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K. M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. science, 320(5880), 1191-1195.

Moore, C. J., & Price, C. J. (1999). A functional neuroimaging study of the variables that generate category-specific object processing differences. Brain, 122(5), 943-962.

Murphy, B., Talukdar, P., & Mitchell, T. (2012, June). Selecting corpus-semantic models for neurolinguistic decoding. In Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation(pp. 114-123). Association for Computational Linguistics.

Noppeney, U., Patterson, K., Tyler, L. K., Moss, H., Stamatakis, E. A., Bright, P., ... & Price, C. J. (2007). Temporal lobe lesions and semantic impairment: a comparison of herpes simplex virus encephalitis and semantic dementia. Brain, 130(4), 1138-1147.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical neurophysiology*, *112*(4), 713-719.

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. Frontiers in neuroinformatics, 10, 27.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. Nature Reviews Neuroscience, 8(12), 976.

Paz-Caballero, D., Cuetos, F., & Dobarro, A. (2006). Electrophysiological evidence for a natural/artifactual dissociation. Brain Research, 1067(1), 189-200.

Pedersen, T., Patwardhan, S., & Michelizzi, J. (2004, May). WordNet:: Similarity: measuring the relatedness of concepts. In Demonstration papers at HLT-NAACL 2004 (pp. 38-41). Association for Computational Linguistics.

Pelli, D. G., & Vision, S. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*, 437-442.

Pennington, J., Socher, R., & Manning, C. (2014, October). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).

Pereira, F., Detre, G., & Botvinick, M. (2011). Generating text from functional brain images. Frontiers in human neuroscience, 5, 72.

Pereira, F., Gershman, S., Ritter, S., & Botvinick, M. (2016). A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data. Cognitive neuropsychology, 33(3-4), 175-190.

Pilehvar, M. T., & Collier, N. (2016). De-conflated semantic representations. arXiv preprint arXiv:1608.01961.

Potter, M. C. (1976). Short-term conceptual memory for pictures. Journal of experimental psychology: human learning and memory, 2(5), 509.

Potter, M. C., Wyble, B., Hagmann, C. E., & McCourt, E. S. (2014). Detecting meaning in RSVP at 13 ms per picture. Attention, Perception, & Psychophysics, 76(2), 270-279.

Proverbio, A. M., Del Zotto, M., & Zani, A. (2007). The emergence of semantic categorization in early visual processing: ERP indices of animal vs. artifact recognition. BMC neuroscience, 8(1), 24.

Pulvermüller, F. (2001). Brain reflections of words and their meaning. Trends in cognitive sciences, 5(12), 517-524.

Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. Journal of cognitive neuroscience, 17(6), 884-892.

Randall, B., Moss, H. E., Rodd, J. M., Greer, M., & Tyler, L. K. (2004). Distinctiveness and correlation in conceptual structure: Behavioral and computational studies. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30(2), 393.

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. arXiv preprint cmp-lg/9511007.

Resnik, P. (1999). Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. Journal of artificial intelligence research, 11, 95-130.

Richardson, R., Smeaton, A., & Murphy, J. (1994). Using WordNet as a knowledge base for measuring semantic similarity between words.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. Nature neuroscience, 2(11), 1019.

Riordan, B., & Jones, M. N. (2011). Redundancy in perceptual and linguistic experience: Comparing feature□based and distributional models of semantic representation. Topics in Cognitive Science, 3(2), 303-345.

Rogers, T. T., & Wolmetz, M. (2016). Conceptual knowledge representation: A cross-section of current research. Cognitive neuropsychology, 33(3-4), 121-129.

Rohde, D. L., Gonnerman, L. M., & Plaut, D. C. (2006). An improved model of semantic similarity based on lexical co-occurrence. *Communications of the ACM*, *8*(627-633), 116.

Rosch, E. (1975). Cognitive representations of semantic categories. Journal of experimental psychology: General, 104(3), 192.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive psychology, 8(3), 382-439.

Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization of natural images. Nature neuroscience, 5(7), 629.

Sacchett, C., & Humphreys, G. W. (1992). Calling a squirrel a squirrel but a canoe a wigwam: A category-specific deficit for artefactual objects and body parts. Cognitive Neuropsychology, 9(1), 73-86.

Sahlgren, M. (2008). The distributional hypothesis. Italian Journal of Disability Studies, 20, 33-53.

Sassenhagen, J., & Fiebach, C. J. (2019). Traces of Meaning Itself: Encoding distributional word vectors in brain activity. bioRxiv, 603837.

Schendan, H. E., & Maher, S. M. (2009). Object knowledge during entry-level categorization is activated and modified by implicit memory after 200 ms. Neuroimage, 44(4), 1423-1438.

Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. Cognitive psychology, 1(1), 1-17.

Sim, E. J., & Kiefer, M. (2005). Category-related brain activity to natural categories is associated with the retrieval of visual features: Evidence from repetition effects during visual and functional judgments. Cognitive Brain Research, 24(2), 260-273.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. Psychological review, 81(3), 214.

Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. Neuroimage, 65, 69-82.

Stewart, F., Parkin, A. J., & Hunkin, N. M. (1992). Naming impairments following recovery from herpes simplex encephalitis: Category-specific?. The Quarterly Journal of Experimental Psychology, 44(2), 261-284.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. Science, 268(5217), 1632-1634.

Teichmann, L., Grootswagers, T., Carlson, T., & Rich, A. N. (2018). Decoding digits and dice with magnetoencephalography: evidence for a shared representation of magnitude. Journal of cognitive neuroscience, 30(7), 999-1010.

Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., & Levy, J. P. (2000). Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. Brain and language, 75(2), 195-231.

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. Trends in cognitive sciences, 5(6), 244-252.

Tranel, D., Damasio, H., & Damasio, A. R. (1997). A neural basis for the retrieval of conceptual knowledge. Neuropsychologia, 35(10), 1319-1327.

Tranel, D., Logan, C. G., Frank, R. J., & Damasio, A. R. (1997). Explaining category-related effects in the retrieval of conceptual and lexical knowledge for concrete entities: Operationalization and analysis of factors. Neuropsychologia, 35(10), 1329-1339.

Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. Journal of artificial intelligence research, 37, 141-188.

Vanrullen, R., & Thorpe, S. J. (2001). The time course of visual processing: from early perception to decision-making. Journal of cognitive neuroscience, 13(4), 454-461.

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.

Vinson, D. P., Vigliocco, G., Cappa, S., & Siri, S. (2003). The breakdown of semantic knowledge: Insights from a statistical model of meaning representation. Brain and Language, 86(3), 347-365.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. Brain, 107(3), 829-853.

Warrington, E. K., & McCarthy, R. A. (1987). Categories of knowledge: Further fractionations and an attempted integration. Brain, 110(5), 1273-1296.

Yang, D., & Powers, D. M. (2005, January). Measuring semantic similarity in the taxonomy of WordNet. In Proceedings of the Twenty-eighth Australasian conference on Computer Science-Volume 38 (pp. 315-322). Australian Computer Society, Inc..