Structure and Organization of the Mental Lexicon: A Network Approach Derived from Syntactic Dependency Relations and Word Associations

Simon De Deyne, Steven Verheyen, and Gert Storms

Abstract. Semantic networks are often used to represent the meaning of a word in the mental lexicon. To construct a large-scale network for this lexicon, text corpora provide a convenient and rich resource. In this chapter the network properties of a text-based approach are evaluated and compared with a more direct way of assessing the mental content of the lexicon through word associations. This comparison indicates that both approaches highlight different properties specific to linguistic and mental representations. Both types of network are qualitatively different in terms of their global network structure and the content of the network communities. Moreover, behavioral data from relatedness judgments show that language networks do not capture these judgments as well as mental networks.

1 Introduction

In cognitive science semantic networks, in which words are connected with each other through a set of links, have been introduced over 50 years ago in the work of Collins and Quillian (1969) and Collins and Loftus (1975) and have remained an influential theoretical model of the mental lexicon ever since. Until very recently, this model has been employed mainly as an elusive metaphor and idealized theoretical construct, since sizable implementations of such a network were missing. This has changed through a combination of factors such as the availability of large corpora, increased computational resources, and accelerated advances in network theory.

In this chapter two approaches towards constructing a large-scale network model of the mental lexicon are compared that make use of novel corpora. A first one is based on word associations and a second one is based on linguistic representations derived from a syntactically annotated text corpus.

Simon De Deyne
University of Adelaide, Adelaide, Australia
e-mail: simon.dedeyne@adelaide.edu.au

Steven Verheyen · Gert Storms
Faculty of Psychology and Educational Sciences, University of Leuven,
Tiensestraat 102, Box 3711 B-3000 Leuven, Belgium
e-mail: {steven.verheyen,gert.storms}@ppw.kuleuven.be

A. Mehler et al. (eds.), *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks*,
Understanding Complex Systems, DOI: 10.1007/978-3-662-47238-5_3

While previous work has looked into both types of corpora (De Deyne and Storms 2008; Kenett et al. 2011; Motter et al. 2002; Steyvers and Tenenbaum 2005; Solé et al. 2010), the interpretation of the findings is complicated by the lack of control for factors such as the number of tokens or network size. Another reason why such a comparison has been lacking is the limited size of around 5,000 nodes of the frequently used word association networks based on the University of Florida dataset (Nelson et al. 2004). In this chapter, a new word association corpus based on over 12,000 cues and over 3 million responses will be described, which for the first time enables a comparison with a similar-sized network derived from text resources. Apart from comparability, the choice of these two types of corpora also allows for a comparison between a representation based on purely linguistic materials from text and a representation that accesses more mental properties present in the lexicon by looking at word associations. In other words, by matching quantitative properties regarding the size of the network, the comparison allows for the identification of qualitative differences between the two networks.

This chapter will compare the networks' structure at a global and intermediate level by capitalizing on the innovative contributions of network science as a unifying formal framework to examine the structure at different levels simultaneously.

1.1 Macro-, Meso-, and Microscopic Properties of the Mental Lexicon

The fundamental strength of the network account lies in the way it addresses the structure of the lexicon at the macroscopic, mesoscopic, and microscopic level simultaneously. The ability to do so is an important feat of network science, since studies of complex systems indicate that different functional patterns emerge depending on the level of analysis and complexity of the network.

The macroscopic or network level reflects the combined role of all the connections between the nodes of the network. In naturally occurring networks, this pattern of connections is often very distinct from comparable random networks, for instance in the case of small-world networks. Over the past years, studies have revealed a small-world structure in both linguistic and word association networks (Steyvers and Tenenbaum 2005; De Deyne and Storms 2008). In these small-world networks, regardless of the starting node, any other node can be reached in less than four steps on average. Moreover, in contrast to comparable random networks, the networks also contain a small number of highly connected nodes or hubs. Similarly, the interconnectivity among neighboring nodes indicated by the clustering coefficient, tends to be much larger in these networks than in comparable random networks.

The way a network is organized at the macroscopic level provides insight in its robustness against damage and efficiency of information dissipation (Bullmore and Sporns 2012). It also captures various dynamic properties such as the gradual growth (Steyvers and Tenenbaum 2005), abrupt emergence of new cognitive functions during development, as well as the degradation of these functions with aging or neurodegenerative illness (Baronchelli et al. 2013).

The mesoscopic or group level involves the properties of a considerable subset of nodes in the network. The structure at the mesoscopic level in the mental lexicon is informative of the meaning of words. This is achieved by computing the distance between a set of words through a set of direct and indirect paths connecting them. These distances allow us to identify closely knit regions in the network. In network science, this method is called community detection. It has been successfully applied in cognitive science to uncover the community structure of phonological networks (Vitevitch 2008), and to identify different word senses in small word association networks (Lancichinetti et al. 2011). Identifying the communities in the mental lexicon might reflect similarity in meaning on a variety of grounds. For instance, this could be a taxonomic structure with groupings for different types of animals like birds, mammals, or fish (Rosch 1973). Communities could also be thematic, where different members of a community occur in a specific script, like a restaurant community consisting of members such as *eating*, *bill*, *waiter*, and *dessert* (Schank and Abelson 1977). Perhaps the communities group together words in a manner reflecting the neuro-anatomic properties of the brain leading to a distinction between living kinds and artefacts (Warrington and Shallice 1984), abstract and concrete words (Crutch and Warrington 2005) or categories grounded by emotional responses (Niedenthal et al. 1999). These are just a few examples, and it is quite likely that the investigating of a large network of words might point towards a structure different from these.

Focusing on just a pair of nodes rather than a larger subset, the mesoscopic level is also informative about how related or close two nodes are and what types of paths exist between them. Since the early propositional network model by Collins and Quillian (1969), the closeness between a pair of nodes has been shown to predict the time to verify sentences like *a bird can fly* (Collins and Quillian 1969). To accommodate for a larger set of behavioral data, the theory was extended to include the notion of spreading activation (Collins and Loftus 1975), in which both direct and indirect paths contribute to the closeness of pairs of words. In network theory, spreading activation is often thought of as a stochastic random walk, resulting in a measure of relatedness that reflects both the number and the length of paths connecting two nodes in the network. Such a random walk model allows us to infer additional information beyond the direct connection between two nodes, which has been shown to improve predictions of human similarity judgments (Capitán et al. 2012; Van Dongen 2000), and the extraction of categorical relations between words (Borge-Holthoefer and Arenas 2010).

A quintessential example of the role of these connections is the study of word priming. In priming tasks, the processing of a word is enhanced when it is preceeded by a related word. In the case of associative priming this involves the presentation of a prime such as *dog* which facilitates processing of the word *bone*. In network terms, such facilitation might be explained by the presence of an associative link between these words. Closely related is mediated priming, whereby one word primes another because they are connected through a mediated link, as in the example of *stripes – tiger – lion*. This type of priming is of particular theoretical importance, as it allows testing the assumption of activation spreading throughout the network (Hutchison 2003) similar to the original proposals by Collins and Loftus (1975). A final type of priming that is often considered distinct from the two previous ones, is semantic priming. Here, an ensemble of shared features or links rather than a single connection determines whether or not priming occurs. From the provided examples it will be clear that a network account provides an elegant way to understand many of the documented priming effects. In this area as well, such an account has been mostly influential at a theoretical level, rather than has made use of a fully implemented model of the mental lexicon.

The microscopic or node level of analysis of the network focuses on how a single node is connected with the rest of the network. Examples are node centrality measures, such as the number of in- or outgoing links. These type of centrality measures have been studied quite extensively in psycholinguistics and explain why certain words are processed more efficiently than others (Nelson and McEvoy 2000; Chumbley 1986; de Groot 1989; Hutchison 2003). In this case network-derived measures provide a structural explanation for many lexical properties of words which have been demonstrated to facilitate word processing.

Structural explanations have been given for the effects of variables such as age-of-acquisition (Steyvers and Tenenbaum 2005) and word frequency (Monaco et al. 2007). An interesting example is the finding that highly imageable words such as *chicken* will be processed faster and more accurately across a range of tasks, including naming and lexical decision, compared to more abstract words such as *intuition*. Such a finding can be explained by looking at the set-size (i.e., summed in- and out-degree) of a word. Researchers believe concrete words have smaller associate sets than abstract ones (Galbraith and Underwood 1973; Schwanenflugel et al. 1992) while others believe that concrete words have more semantic properties than abstract words (de Groot 1989; Plaut and Shallice 1993). A network approach has the potential to tease these two explanations apart.

1.2 Acquiring a Mental Lexicon through Language

The rationale of the current approach, in which the mental lexicon is implemented as a network derived from language, is that this lexicon should reflect a repository of shared subjective meaning, allowing language users to communicate efficiently. It is shared under the assumption that with increasing proficiency a speaker acquires a lexicon that mimics the linguistic properties of his or her environment. It is efficient, assuming that it is organized in a non-trivial fashion to meet information retrieval demands. Represented as a network or graph, the mental lexicon consists of nodes corresponding to lexicalized concepts, and links between these nodes indicate lexico-semantic relationships between these nodes.

We believe the mental lexicon acquires meaning through the continuous exposure to words in context, following similar ideas by Wittgenstein (2001) and Firth (1968), where word meaning is equated to its use in language. This is also the idea that underlies many large-scale models which track the co-occurrence of words at the document level (e.g. Landauer and Dutnais 1997) or at the sentence level (e.g., Lund and Burgess 1996). However, as many studies have shown, humans do not merely encode the surface level-properties of a single sentence or a larger discourse unit. Instead, it is assumed that a mental model is constructed that conveys the crucial information of the utterance beyond the verbatim format and involves comprehension of the syntactic nature of its constituents (Dennis 2005; Kintsch and Mangalath 2011) and the integration of its meaning with prior knowledge (Kintsch 1998; Prior and Bentin 2003).

Indeed, in addition to learning about which words co-occur in language, knowledge about different parts-of-speech and syntactic constructions are likely to be used by humans to capture additional information about the meaning of an utterance (Dennis 2005; Kintsch and Mangalath 2011). In many languages word meaning and part-of-speech characteristics are highly correlated, which allows one to infer what the actions (verbs), entities (nouns) and properties of these entities (adjectives) are. Similarly, syntactic relationships between a subject and an object might reveal something about agency. Furthermore, various studies have shown that linguistic models that incorporate this information provide a better account of human relatedness judgments compared to *n*-gram models that do not (Heylen et al. 2008; Padó and Lapata 2007). Altogether, this suggests that a language network derived from a syntactically annotated text corpus will lead to a representations that capture some key properties of the mental lexicon.

One limitation of this linguistic approach is the fact that language is not merely representational, as it is used to convey a message between a speaker and a listener. Utterances comprise pragmatic factors as well. Compared to a text-based network, this is one of the main reasons to assume that a word association model is likely to encode mental representations differently, as they are considered to be free from pragmatics or the intent to communicate some organized discourse, and believed to be simply the expression of thought (Szalay and Deese 1978). Moreover, these associations do not necessarily reflect propositional information derived from the linguistic environment, but might reflect imagery, knowledge, beliefs, attitudes, and affect as well (De Deyne and Storms 2008; De Deyne et al. 2013a; Szalay and Deese 1978; Rensbergen et al. 2014; De Deyne and Storms 2008; Simmons et al. 2008). In other words, word associations tap directly into the semantic information of the mental lexicon.

1.3 Chapter Outline

The remainder of this chapter starts with an explanation of how a language network is derived from a syntactically annotated text corpus, and how a mental network is derived from a large corpus of word associations. The language network chapter refers to a syntactic language network, where nodes are words and where two nodes are connected through a syntactic dependency relationship such as the adjective *red* modifying the noun *car*. The mental network refers to a network where nodes are also words but the relationship between them is determined by how strongly a specific word is evoked by a cue word in a word association task. Compared to the language networks, these responses are not constrained by syntax but reflect mental constraints of what prominently comes to mind. Both networks aim to capture the mental lexicon in an unsupervised way. This contrasts with the original handcrafted Collins and Loftus network (Collins and Loftus 1975) or WordNet (Fellbaum 1998), where the representations are derived manually by expert linguists. It also differs from connectionist approaches (Rogers and McClelland 2004), where the set of nodes and types of relations is decided in advance and connection weights are estimated using supervised learning.

The focus will be on the macroscopic and mesoscopic levels of the networks, as these have only been recently introduced in the context of studying structure in the mental lexicon (Baronchelli et al. 2013). First, the macroscopic stucture of the networks will be addressed. It will provide a characterization of their global organization and explore the nature of network hubs.

Next, community detection will be used to explore which types of clusters of meaning are present in language and mental networks at a mesoscopic level. An inspection of these communities can reveal what the underlying structural principles are and how various parts of the network relate to each other. For instance, one possibility is that the hubs identified in the previous analysis are indicative of the important domains of knowledge represented in the network. Another possibility is that certain nodes in the network play a special role by connecting different clusters in the graph, for instance in the case of polysemous words. In both cases, communities of limited size might allow us to interpret hubs much easier in comparison with hubs identified at the macro-level. Community members can also provide us with some information about the nature of the organization of the network. According to the dominant view in psychology, concepts are organized in a hierarchical taxonomy of natural categories (Rosch 1973) on the basis of shared perceptual properties, whereas other views attribute a larger role to a structure based on thematic relations of the lexicon (Lin and Murphy 2001).

To test whether the communities correspond with a taxonomic organization, the classification performance for basic-level categories such as birds or fish, obtained from human behavioral data, will be used. This allows us to evaluate whether language and mental networks make similar distinctions and provides the opportunity to discuss alternative interpretations if such structure wouldn't be evident.

The final part of the mesoscopic analysis complements the classification study but uses a more direct way of assessing the underlying mesoscopic properties of the network. This is accomplished by using network-derived similarity measures to predict human relatedness judgments. Considering various levels of abstraction and different types of semantic relations (e.g., relations at the basic and domain level, and thematic relations) allows us to generalize the results beyond concrete basic level nouns, which have dominated the field of cognitive science for a long time (Medin et al. 2000). However, because the large-scale networks in this chapter are extremely sparse such an evaluation poses a specific challenge as a simple overlap measure for relatedness that only takes into account shared neighbors between words might not suffice. To address this issue a spreading activation mechanism similar to the one originally conceived by Collins and Loftus (1975) will be proposed. One way of implementing this is by using Markov random walks over the network, as these also take into account indirect paths that exist between a pair of nodes. Just like dimension reduction in high dimensional semantic spaces like Latent Semantic Analysis (Landauer 2007), the spreading activation mechanism introduces a mechanism to infer indirect links. This allows us to deal with the sparsity associated with linguistic representations and is assumed to lead to more reliable estimates of relatedness. This sections ends with a brief discussion of the role of this spreading activation for both language and mental networks in predicting different types of semantic relations.

2 Constructing the Networks

In the following section, the derivation of several networks based on word association and text corpora are given. Both types of networks are implemented as a unipartite localist network, where nodes correspond to words, and are connected through weighted directed edges with other nodes. To make the networks comparable, the set of words will be restricted to those that occur in both the text corpus and word association data.

2.1 Mental Networks

The mental network was derived from a large scale word association study conducted between 2003 and 2010 at the University of Leuven.[1] This study is described in detail in De Deyne et al. (2013a). In short, it involved a total of 71,380 native Dutch speakers. The association procedure differed from most large-scale studies (e.g. Kiss 1968; Nelson et al. 2004) because it used a continued response format, where each participant generated three different responses for each cue instead of one. This allows one to get a better approximation of weak links in the network (De Deyne et al. 2013a). This way, a total of 300 responses were obtained from 100 participants per cue, corresponding to 100 primary, 100 secondary, and 100 tertiary associations. In order to be able to compare the results with previous work based on a single response procedure, an additional network will be derived which only includes the primary responses.

[1] The word association project is ongoing. In 2014, the project contained at least 300 responses per cue for 16,000 Dutch cues and 8,000 English cues. The studies can be accessed from `http://www.smallworldofwords.com`.

Data Preprocessing and Network Construction

The word association data consisted of 3.77 million responses for a total of 12,581 different cues. About 0.20 million different response types were represented in the data. From these data, two weighted directed networks were derived. The first network G_{asso1}, is based on the primary responses, comparable to the common single response datasets (Nelson et al. 2004). The second network, G_{asso123}, includes the secondary and tertiary responses as well. Reducing the network from a bipartite representation to a unipartite representation involves the removal of responses that were not members of the set of cues. The removal of these responses did not affect the coverage in terms of token too much, as about 87% and 83% of the response tokens were retained in G_{asso1} and G_{asso123}. To allow a comparison with the language networks which will be explained in the next section, a total of 11,252 cues (94% of the original) were retained. With a total of 0.85 million response tokens G_{asso1} and 2.41 million tokens in G_{asso123} it still represents a sizeable portion of nodes present in the original networks.

2.2 Language Networks

An advanced syntactic dependency parser was used to build a network from a small number of predefined syntactic relations (Heylen et al. 2008; Padó and Lapata 2007). This approach offers a number of advantages in comparison to simple *n*-gram models derived from raw text because it allows us to infer the part of speech of the words and the syntactic relation between the constituents of a sentence. Because many sentences exhibit a complex nested structure, a second advantage of this analysis is that it captures interesting relations between words even if they are not adjacent within an *n*-gram window.

Corpus

The corpus described in this chapter consists of a variety of language resources spanning three different registers (De Deyne et al. 2014): (1) text derived from Dutch articles in newspapers and magazines from the Twente Nieuws Corpus (Ordelman 2002) and the Leuven Newspaper Corpus (Heylen et al. 2008), (2) informal language retrieved from Internet web pages collected between 2005 and 2007 and the Dutch Wikipedia retrieved in 2008 (De Deyne 2008), and (3) spoken text from Dutch movie subtitles (Keuleers et al. 2010) and the Corpus of Spoken Dutch (Oostdijk 2000).

Each sentence in the corpus was parsed using Alpino, an advanced Dutch dependency parser (Bouma et al. 2000). Similar to Pereira et al. (1993) and Padó and Lapata (2007), two words were connected by a small number of predefined dependency paths. To reduce sparsity, part-of-speech tagged lemma forms provided by Alpino were used instead of word forms. In other words, plurals and inflections were all reduced to a more basic form. Next, all lemmas were counted and only adjectives, adverbs, nouns, and verbs occurring at least 60 times were retained. Applying this cutoff removed very infrequent words and aided in keeping the computations manageable. The resulting corpus vocabulary consisted of 157 million tokens and 103,842 different lemmas; 82.7% were nouns, 12.6% adjectives, 4.5% verbs, and 0.2% adverbs.

Table 1 Overview of the syntactic dependency paths *p* and examples

Abbreviation	**Full Path (p)**	**Example**
ObjHd	$V \xleftarrow{\text{object of head}} \underline{\mathrm{N}}$	We <u>need</u> some more *coffee*.
HdMod	$N \xrightarrow{\text{modification}} \underline{\mathrm{A}}$	This is, excuse me, damn <u>good</u> *coffee*.
HdModObj	$N \xrightarrow{\text{modification}} \mathrm{NP} \xrightarrow{\text{object of}} \underline{\mathrm{N}}$	Lucy takes a loud *sip* of <u>coffee</u>
SuObj	$N \xrightarrow{\text{subject of object}} \underline{\mathrm{N}}$	*Coffee* contains lots of <u>caffeine</u>.
SuHd	$N \xrightarrow{\text{subject of head}} \underline{\mathrm{V}}$	This *coffee* <u>tastes</u> delicious!
Cnj	$N \xleftrightarrow{\text{conjunction}} N$	Norma arrives with Cooper's *pie* and *coffee*.
SuPredc	$N \xrightarrow{\text{subject of predicative phrase}} \underline{\mathrm{N}}$	*Coffee* is a <u>drink</u>.
HdPredc	$V \xrightarrow{\text{predicative complement}} \underline{\mathrm{A}}$	This coffee *tastes* <u>delicious</u>!

Data Preprocessing and Network Construction

The syntactic relations coded as dependency paths, together with examples and the number of pairs for each of the eight paths are shown in Table 1. With the exception of the HdModObj pattern of length 2, all paths p had a length of 1. For each pattern a reverse path was created by transposing the path-dependent graph. For example, for pattern HdMod, the weight of a path for the adjective *good* and the noun *coffee* is derived from the transposed dependency matrix $G_{\mathrm{HdMod}'}$. An example of the obtained dependencies based on the sum of the original and transposed paths for the word *coffee* is shown in Table 2. It illustrates how the most frequent relations uncovered by the syntactic dependencies are interpretable as corresponding to distinctions in terms of function, attributes, and related entities.

To allow a comparison with the mental networks, the network G_{lex} consisted only of words that also occurred in the mental lexicon G_{asso123} which resulted in a set of 11,252 cues. The total number of tokens in G_{lex} was 83.87 million, while G_{asso1} and G_{asso123} contained only a fraction of this amount of tokens (0.85 million and 2.41 million respectively).

To further improve comparability, a new network G_{lex123} was derived to closely match the properties of G_{asso123}. This was accomplished by making two additional assumptions. First, apart from vocabulary size, the number of tokens in both networks was matched. This was achieved by sampling responses in a way that matched

Table 2 English translations of the 5 most frequent syntax dependencies derived for *coffee* in the G_{lex} network

ObjHd	HdMod	HdModObj	SuObj	SuHd	Cnj	SuPred	HdPred
drink	free	hand	visitor	serve	tea	coffee	ready
will	strong	man	person	offer	pastry	drink	cold
poor	fresh	taste	man	grow	tobacco	tea	free
get	fair	sugar	someone	drink	soda	water	delicious
sell	black	chance	company	cool	cookie	product	good

the out-strength (i.e. the total number of recorded association responses) of each cue in the mental network. In addition, because participants in the continued word association task were not able to provide the same associate twice, a sampling without replacement scheme was used.

3 Exploring the Structure of Language and Mental Networks

3.1 Macroscopic Structure

Previous studies have shown that a small-world structure is present in both language-derived networks and word association networks (De Deyne and Storms 2008; Solé et al. 2010; Steyvers and Tenenbaum 2005). In line with this work, such a structure should be present in all four networks derived in the previous section. By controlling the number of observations, the macroscopic network statistics of the language and mental networks can be directly compared. Moreover, since two different sampling regimes were applied, the effect of denser networks can be evaluated. Of particular interest is the clustering coefficient of the networks, as this measure provides an indication of the amount of structure present in the networks.

3.1.1 Network Statistics

For each of the four networks, the network statistics were calculated from the largest strongly connected component. The results are presented in Table 3. The largest difference between the two types of networks is based on their density D. In particular, the language network G_{lex} was over thirty times denser than G_{asso1}. The matched G_{lex123} had a higher density than the G_{asso123} network, which indicates that language-based representations are more heterogeneous in terms of connected nodes even when the total number of responses is matched to those of G_{asso123}. Presumably this reflects the fact that by definition most relations in the language network are syntagmatic (i.e., fulfilling a different syntactic role, e.g., *captain–ship*), while in word associations paradigmatic responses (i.e., fulfilling a similar syntactic role, e.g., *captain–boss*) are more common (Cramer 1968).

Density also differed between the single and continued word association networks. As indicated by Table 3, the single response network G_{asso1} had a density of 0.22%. Including the secondary and tertiary responses increased the density considerably, to 0.64% for G_{asso123}. This confirms that the continued procedure draws on a more heterogeneous response set through the inclusion of weaker links that might go undetected in single response procedures (De Deyne et al. 2013b). Despite this increase, the density remains very small in comparison to G_{lex} and G_{lex123}. Related to the observed differences in density, Table 3 also shows how the continued response procedure increases the in-degree (k^{in}) and out-degree (k^{out}) substantially, from 24.3 for G_{asso1} to 71.5 for G_{asso123}. These values are again considerably smaller than the corresponding ones for the language networks G_{lex} and G_{lex123}, reflecting the same heterogeneous distribution of edges in the networks.

Table 3 Descriptive network statistics for each of the four graphs

	$G_{\mathbf{asso1}}$		$G_{\mathbf{asso123}}$		$G_{\mathbf{lex}}$		$G_{\mathbf{lex123}}$	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
D	0.0022		0.0064		0.0611		0.0091	
L	3.77	0.824	2.85	0.57	1.98	0.33	2.68	0.61
$max(L)$	10		7		5		7	
k^{in}	24.3	51.9	71.5	140.9	687.0	870.5	102.9	226.2
k^{out}	24.3	8.3	71.5	16.41	687.0	870.1	102.9	41.4
CC	0.0046	0.0036	0.0015	0.0006	0.0005	0.0010	0.0009	0.0027
CC_{rand}	0.0004	0.0000	0.0001	0.0000	0.0000	0.0000	0.0001	0.0000

All networks were characterized by small average paths L (ranging from 1.98 to 3.77 steps) and network diameters $max(L)$ ranging between 5 and 10. In comparison to a matched random network (see CC_{rand}), the clustering coefficient CC for weighted directed networks (see Fagiolo 2007) was considerably higher for the real networks indicating an extensive degree of organization. Moreover, combined with the average short paths lengths, such structure indicates a small-world organization and replicates earlier results for the language and mental networks (Steyvers and Tenenbaum 2005; Solé et al. 2010).[2]

3.1.2 Network Hubs

A second way to characterize the macroscopic structure of the network is by looking at the most central nodes or hubs in the network. For each of the four networks, the

[2] Note that the absolute values are lower than that of previous reports. This is a side-effect of using a weighted form of the clustering coefficient as defined by Fagiolo (2007).

ten most central nodes in terms of in-strength and PageRank with α set to .80 (Page et al. 1998) are listed in Table 4 and illustrated in Fig. 1. Using these measures to identify network hubs allows us to evaluate the qualitative nature of the most central words in the networks.

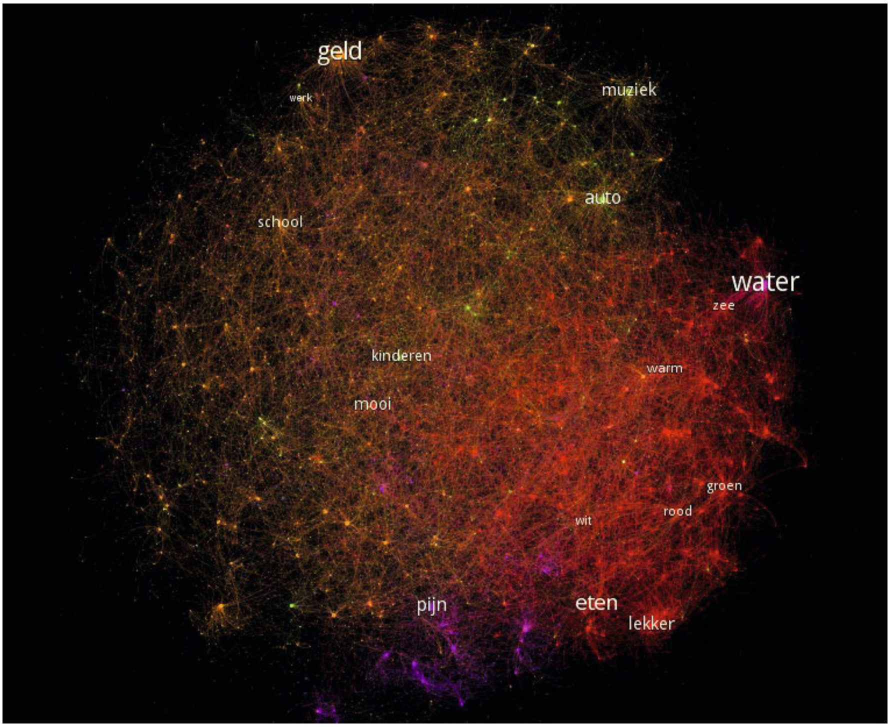

Fig. 1 Large-scale visualization of hubs and communities found in the $G_{asso123}$ network

If word associations are primarily based on associative learning from the linguistic environment, this should lead to hubs that closely match those in the language network. The hubs in the mental networks such as *water* (Dutch: *water*), *food* (*eten*), *money* (*geld*), *car* (*auto*), and *pain* (*pijn*) seem to reflect something about the basic human needs. The hubs in the language networks show some overlap with the mental networks' hubs, but tend to include more abstract words such as *year* (*jaar*), *new* (*nieuw*), *good* (*goed*), *human* (*mens*), *own* (*eigen*), *previous* (*vorig*), and *other* (*ander*).

Furthermore, despite the large differences in density, the hub nodes were quite similar in the mental networks and almost identical in the lexical graphs. The in-strength and PageRank measure of centrality capture slightly different patterns for the hub nodes, but were highly correlated overall, between .88 and .95. More than the type of centrality measure itself, the largest variability was due to the type of graph. In this case, only a moderate correlation existed between the centrality in mental and language networks (between .45 and .46 for in-strength and between .32 and .34 for PageRank) indicating a substantial difference in the identity of central nodes.

A final observation is that the hubs obtained here differ from those identified in previous reports. Where syntactic network hubs have been found to correspond to functional words, and semantic network hubs to polysemous words (Solé et al. 2010), the current results do not include functional hubs. This mainly reflects the fact that closed-form class words were excluded from the analysis as it would obscure any comparison between both types of graphs. In addition, hubs in both the language and mental networks cannot be considered polysemous in a classical sense, which likely reflects the fact that semantic networks reported in previous work (Solé et al. 2010), were based on linguistic expert knowledge derived from WordNet (Fellbaum 1998).

Table 4 Ten most central network hubs derived from in-strength and PageRank ($\alpha = .80$) centrality measures

In-strength				PageRank			
G_{asso1}	G_{asso123}	G_{lex}	G_{lex123}	G_{asso1}	G_{asso123}	G_{lex}	G_{lex123}
money	water	big	big	water	sun	big	big
water	money	human	human	warm	water	year	good
food	food	man	man	sun	warm	new	new
car	car	new	new	money	food	good	other
music	pain	good	good	green	money	other	year
pain	music	child	child	food	sea	human	human
child	pretty	other	other	car	pretty	man	man
school	school	woman	woman	fun	pain	previous	child
pretty	warm	year	year	sea	green	child	own
sea	sea	small	small	pretty	fun	own	woman

3.2 Mesoscopic Structure

The following analyses will compare clusters identified through community detection methods for language and mental networks. In particular, it will investigate the size and type of communities that can be derived from these graphs. Next, at the most detailed level of the community hierarchy, human data for basic-level categories will be used to explore to what degree these communities provide evidence for a hierarchical taxonomic structure of the kind proposed by Rosch and colleagues (Rosch 1973; Mervis and Rosch 1981) or supports alternative views based on thematic relations (Gentner and Kurtz 2005; Lin and Murphy 2001; Wisniewski and Bassok 1999). The last evaluation continues along these lines and uses human relatedness judgments to evaluate which relationships are best represented in the mental and language networks.

3.2.1 Community Clustering

To identify which clusters are represented at the mesoscopic level, the *Order Statistics Local Optimization Method* (OSLOM) community finding algorithm was applied (Lancichinetti et al. 2011). Using this method, communities (also called modules or clusters) can be identified by evaluating the likelihood that a found community can arise in a comparable random network (Lancichinetti et al. 2011). The proposal has a number of advantages in comparison to the many alternatives such as the Louvain method (Blondel et al. 2008). In particular, it operates on large, directed weighted graphs and allows for overlapping and hierarchical communities. Another advantage of OSLOM is that nodes that are not significantly associated with a community are not assigned. For each network, communities at different hierarchical levels were extracted.[3]

Hierarchical Organization and Interpretation of Communities

One of the interesting features of the OSLOM community procedure is that it identifies a hierarchical organisation by grouping smaller communities in larger ones by evaluating statistical evidence of such a structure to occur in random comparable networks. This allows us to investigate different levels of abstraction along the same lines of the hierarchical network as originally proposed by Collins and Quillian (1969) and taxonomy-based theories derived from the work of Rosch (1973). For G_{asso1} the hierarical structure had a depth of 4, while in G_{asso123} the depth was 5. The hierarchy was flatter for both language networks, with a depth of 3 in G_{lex} and a depth of 4 in G_{lex123}.

Starting at the highest level of the hierarchy, only a handful of communities were identified: 4 in G_{asso1}, 2 in and G_{asso123}. In the matched language networks the top level distinguished 2 communities in G_{lex} and 4 in G_{lex123}. In general, the large number of nodes in each community at the top level makes it difficult to interpret the meaning of these communities.

As the best community solution was found for G_{asso123} at the most detailed level (see Table 5 below), this network will be used to illustrate the structure of the communities at the higher levels of the hierarchy. To summarize the distinctions at the highest level, the most central words in each community were obtained by calculating the community specific in-strength. For each of the five levels of the hierarchy, the five most central items were computed and represented in Fig. 2.

[3] In contrast to the previous macroscopic analyses and similar to all subsequent analyses at the mesoscopic level, the weights in the networks were transformed using *positive pointwise mutual information* (PMI) weighting because of its good performance in the context of word co-occurrence models (Bullinaria and Levy 2007).

Table 5 Overview of community structure in the four networks at the lowest hierarchical level

	$G_{\mathbf{asso1}}$	$G_{\mathbf{asso123}}$	$G_{\mathbf{lex}}$	$G_{\mathbf{lex123}}$
# Communities	483	506	157	70
Average size	24	25	77	147
Standard dev. size	14	12	54	152
# Homeless nodes	1182	380	512	1721
# Overlapping nodes	3040	3624	2463	1509
Maximum overlap	8	5	5	15
Mean(p)	0.085	0.051	0.096	0.150

For illustration purposes, Dutch words that were synonymous in English (e.g., the Dutch words *fruit* and *vrucht*) were listed once in each community to convey a maximum of information.

At depth one, Fig. 2 shows the two distinct communities, with one of them containing highly central words with a negative connotation. To see whether this level distinguishes positive and negative words, a post-hoc test was set up using valence judgments for a large set of words from Moors et al. (2012). Ratings for a total of 3,642 non-overlapping words belonging to the two communities in the network were obtained. The difference in terms of valence was significant in an independent t-test ($t(3640) = 7.367$, $CI = [0.190, 0.327]$). This finding is in line with previous research that shows that valence is the most important dimension in semantic space (De Deyne et al. 2013; Samsonovic and Ascoli 2010) and proposals of emotion-based category structure (Niedenthal et al. 1999). However, a combination of factors might explain the observed high-level community structure and therefore strong conclusions might be preliminary.

From level 2 to 4, the interpretation of the communities becomes increasingly less abstract. For instance, level 2 shows that the "negative" community in level 1 also includes abstract words or words related to human culture (*knowledge*, *school*, *money*, *school*, *religion*, *time*,...) which is now differentiated from a pure negative community including community hubs like *negative*, *sadness* or *crossed*. The subdivisions of the "positive" community involve the central nodes *nature*, *music*, *sports*, and *food* which might be interpreted as covering sensorial information and natural kinds. At this level the communities point towards a distinction of concrete vs abstract words (Crutch and Warrington 2005) or natural kinds vs artifacts (Warrington and Shallice 1984) as structural principles of the lexicon. Clearly, such an interpretation is also suggestive, given the large size of the mental network communities and even larger size of the language network communities. More work is needed to confirm this result.

In order to be able to compare the different networks, the lowest level of the community structure provides us with the best chance of directly comparing results. An overview of the obtained community structure is shown in Table 5. In general, the average size of the communities was strongly related to the number of communities,

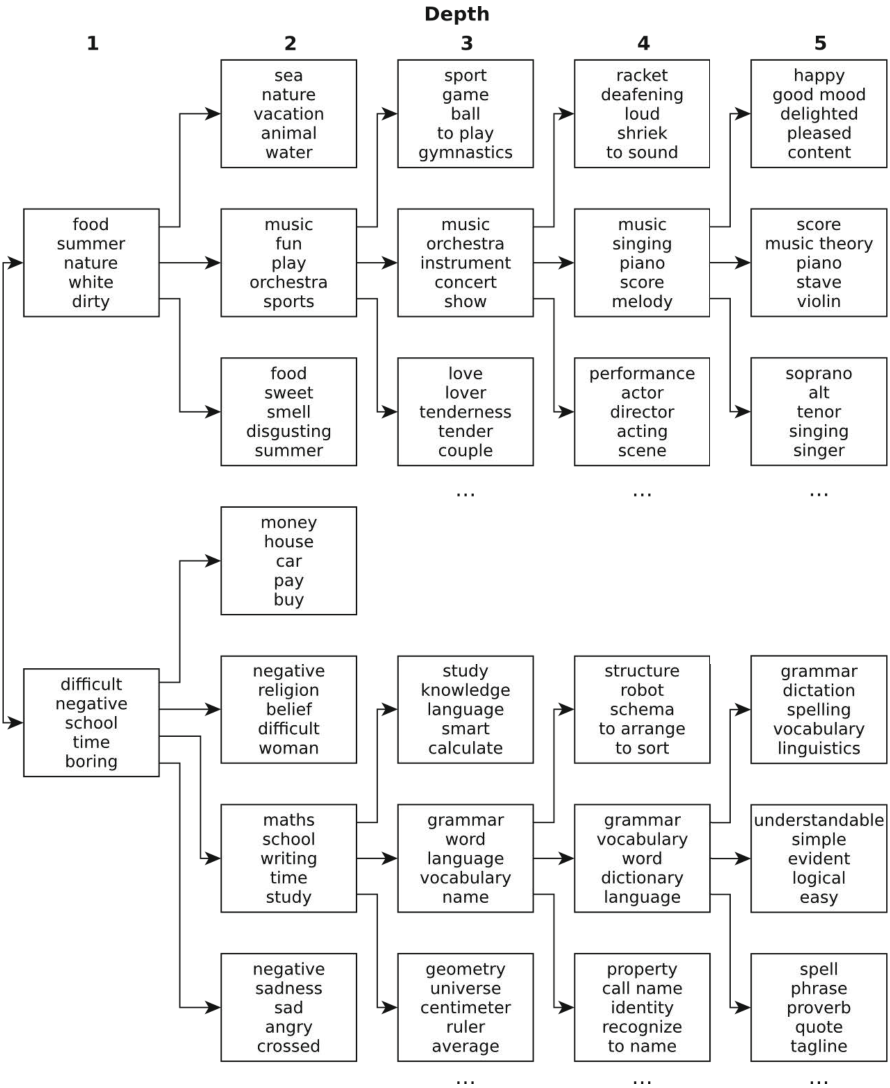

Fig. 2 Hierarchical tree visualization of communities in the $G_{asso123}$ network. Each community is indicated by five central members. At each depth beyond depth 2 a single example is shown of three descendant communities.

and the standard deviation for the community sizes in Table 5 was quite large. This is not surprising given that earlier studies show that in most networks the communities are not necessarily equal in size (Fortunato 2010).

Comparing the different networks, the most striking result is that both the number of communities, and the average significance p of the communities differ between the language and mental networks. The total number of communities was much smaller in the language networks than in the mental networks. The large difference between the two language networks (157 in G_{lex} vs 70 communities in G_{lex123}) can be explained by the difference in density between both graphs (see Table 3). The number of communities was quite similar in G_{asso1} and G_{asso123}, but the mean p-values of the identified communities indicate higher significance of identifying communities in the latter network when compared to a matched random network. The effect of increased density was also apparent for the language graphs, where in comparison with a random structure, the communities found in G_{lex} were more reliable, as the mean p was nearly half that of the sparser G_{lex123} network.

Similarly, there was a large difference in terms of the number of homeless nodes, with over three times more homeless nodes in the sparser networks (G_{asso1} and G_{lex123}). This could indicate that for these networks the density was simply too low to reliably assign nodes with either low in-strength and/or highly heterogeneous neighbors to a specific module. For example, in G_{asso123} the in-degree for homeless nodes was on average 17, compared to 71 for the entire graph and the clustering coefficient was 0.0013 compared to 0.0015 (see Table 3).

At all hierarchical levels, nodes could be assigned to multiple communities and a large number of overlapping nodes were also present at the lowest level. As can be seen from Table 5, networks with many and highly significant communities also assigned more nodes to multiple communities which could indicate the ability to distinguish different senses for a specific word at this level. Moreover, in various cases words belonging to more than a single community corresponded to homonyms or words with related senses. For example, in G_{asso123}, the Dutch word *bank* which means bank or couch in English, belonged to both a community indicating finance and a community for furniture and sitting. Similarly, the word *language* was attributed to four different communities related to nationality, speech, language education, and communication. Again, the mental networks provided the clearest example of this, while the communities in the language-based networks were too coarse to uncover some of the polysemy or homonymy present in the mental networks.

3.2.2 Taxonomic Structure Evaluation

As mentioned in previous sections, there are many different ways in which the mental lexicon can be structured at the mesoscopic level and the previous exploratory approach indicates that various factors might contribute to the organization of the mental lexicon. However, one of the most influential ideas in psychological theories about knowledge representation is that of a hierarchical taxonomy, in which concepts are grouped in progressively larger categories (Collins and Quillian 1969; Rosch 1973; Murphy 2002). An example of such a hierarchy would be *living-thing <animal <bird <sparrow <house sparrow*. In this hierarchy, one particular level, the basic level, is of special significance as categories at this level capture the psychological structure of concepts that is maximal informative in communication. In this example the basic level category is that of birds, because, this level of description provides the best compromise between maximizing within-category similarity (birds tend to be quite similar to each other as they share many features) and

minimizing between-category similarity (birds tend to be dissimilar to fish) (Medin and Rips 2005).

Despite the large number of studies who have looked at hierarchical taxonomic structures for concepts and explanations of basic-level effects, most of them have limited themselves to concrete nouns (Medin et al. 2000). Moreover, as suggested by the community structure in the mental graphs and literature on a thematically or emotionally organized lexicon (Szalay and Deese 1978; Niedenthal et al. 1999; Samsonovic and Ascoli 2010), the omnipresence of hierarchical taxonomies might be partly due to a selection bias. The goal of this section is to evaluate whether the communities identified at the most detailed level support the idea of a hierarchical taxonomy with a special status for basic-level categories.

Data from an exemplar generation task were used to members of basic level categories. In this task, 100 participants generated as many exemplars they could think of for a list of six artifact categories and seven natural kinds categories (Ruts et al. 2004). The names of the categories and the number of exemplars obtained through this procedure are presented in the first two columns of Table 6.

If the communities in each network group together different types of birds, vehicles, fruits, and so on, this would indicate a taxonomic organization of semantic

Table 6 F-values and corresponding community sizes for 13 basic level categories consisting of human-generated category members

	Category size					**F-values**			
Category	Human	G_{asso1}	$G_{asso123}$	G_{lex}	G_{lex123}	G_{asso1}	$G_{asso123}$	G_{lex}	G_{lex123}
Fruit	40	93	50	142	106	0.54	0.47	0.20	0.52
Vegetables	35	42	58	132	105	0.47	0.50	0.31	0.46
Birds	53	58	63	63	55	0.61	0.53	0.64	0.63
Insects	39	53	34	83	109	0.67	0.46	0.49	0.43
Fish	37	46	48	44	53	0.55	0.57	0.47	0.53
Mammals	61	32	21	217	212	0.30	0.20	0.38	0.34
Reptiles	23	18	22	83	109	0.59	0.62	0.19	0.18
Mean	41	49	42	109	107	0.53	0.48	0.38	0.44
Clothing	46	77	70	98	536	0.36	0.35	0.28	0.15
Kitchen Utensils	71	33	18	63	58	0.29	0.20	0.30	0.25
Musical Instrum.	46	62	24	104	69	0.56	0.37	0.59	0.71
Tools	73	51	56	51	151	0.26	0.25	0.31	0.25
Vehicles	46	25	28	135	195	0.23	0.16	0.28	0.20
Weapons	46	33	25	51	151	0.30	0.37	0.27	0.17
Mean	55	47	37	84	193	0.33	0.28	0.34	0.29

memory. Table 6 shows the size of the best matching communities and the Jaccard index or F-measure for clustering performance based on precision and recall for each basic level category (Ball et al. 2011). A good solution would be found for a clustering with high precision and recall through a high number of true positives and a low number of true and false negatives. Starting with the category size, Table 6 shows that on average the best matching communities were of comparable size in G_{asso1} and slightly smaller (and thus more specific) in G_{asso123}. The sizes of the language network communities were larger than the number of generated exemplars by humans. This indicates that in these networks the communities are too general, which will affect their F-values.

For each of the four graphs, the F-values are generally not very high, which indicates that the communities obtained from the language and mental networks do not provide convincing evidence for a general and strict taxonomic organization. Notable exceptions for natural kinds categories were birds (all networks except G_{asso123}), insects (G_{asso1}) and reptiles (G_{asso123}). For artifacts, the only indication of a possible taxonomic structure was musical instruments for G_{lex123}.

Table 7 Top 5 false positives ordered by module in-strength for words belonging to the communities derived from G_{asso123}

Category	**1**	**2**	**3**	**4**	**5**
Fruit	fruit	juicy	pit	pick	summer
Vegetables	vegetable	healthy	puree	sausage	hotchpotch
Birds	bird	beak	nest	whistle	egg
Insects	insect	vermin	beast	crawl	animal
Fish	fish	fishing	rod	slippery	water
Mammals	rodent	gnaw	tail	pen	marten
Reptiles	reptile	scales	animal	tail	amphibian
Clothing	clothing	fashion	blouse	collar	zipper
Kitchen Utensils	cooking	kitchen	stove	cooker hood	burning
Musical Instruments	wind instrument	to blow	fanfare	orchestra	harmony
Tools	tool	carpenter	carpentry	wood	drill
Vehicles	speed	drive	vehicle	motor	circuit
Weapons	sharp	stab	blade	point	stake

On average, natural kinds resulted in higher F-values compared to artifacts.This result supports previous findings, showing that the inter-category structure of artifacts does not have a generally accepted delineation compared to the natural kind categories (Ceulemans and Storms 2010). A contributing factor for the higher F-values for natural kind categories in the mental networks, is that many people are less familiar with certain members of these categories, and predominantly generate taxonomic associates in response to these words. For example, in the case of *swallow* the dominant response was *bird*. This would also explain the better performance of G_{asso1} in this evaluation, as this network only contains the first responses, which frequently correspond to the category-label.

If the communities do not primarily consist of category coordinates, but also contain other words, one might question what factors other than taxonomic ones contribute to the structure found at the most detailed hierarchical level. To address this issue, the five most central false positives for each of the 13 categories were derived by looking at the community specific in-strength as was done for Fig. 2. The results in Table 7 are quite illuminating. First of all, for 8 out of 13 categories the most central item was the category label, which is in line with what can be considered a basic-level category in the literature (Ruts et al. 2004; Rosch 1973). However, this table also shows categories where the representation was too specific, for instance in the case of *rodent* or *wind instrument*, which is also confirmed by the category sizes in Table 6. One could argue that the inclusion of these category-labels might indicate that the F-values are actually underestimates of potential taxonomic structure. Furthermore, the human generated exemplars are not necessarily exhaustive (despite the fact that 100 participants generated exemplars for each category) or correct. For example, *marten* was wrongly identified as a false positive, which suggests this word might have been too infrequent to be captured by 100 participants. However, it is unlikely that this explanation suffices, as the other false positives clearly indicate that related properties, actions, and other thematic information are central. For example, in the case of fruit, other central community members were *juicy*, *pick*, and *summer*. Other examples at the most detailed level in Fig. 2 (e.g., *score*, *music theory*, *piano*, *stave*, *violin*) support this as well. Altogether, the absence of a basic-level taxonomy even for biological categories and the widespread thematic structure across nearly all communities for both the language and mental networks strongly suggest that multiple factors contribute to structure in the mental lexicon, and thematic relations are a major one of them.

3.3 Semantic Relatedness Evaluation

So far, the community detection approach provided some valuable insights about how the mental lexicon might be structured. However, the lack of well-defined small communities in the language networks did not allow us to fully evaluate and compare the language-based and word association-based network. A common direct way to compare these networks and see what kind of relationships they capture uses human relatedness judgments for pairs of words (e.g., Borge-Holthoefer and Arenas 2010; Capitán et al. 2012; Hughes and Ramage 2007). By manipulating the taxonomic and semantic relations between words, it is possible to precisely quantify to what extent each network captures various aspects of the mesoscopic structure. Three studies that were set up for this purpose are described below. In all three studies participants provided relatedness judgments for pairs of words using a 20-point Likert scale. The nature of the pairs differed between studies. They either captured relations at the basic level, at a more general domain level, or thematic relations that do not follow a classical taxonomy.

In a first study, similarity judgments for exemplars from concrete and abstract basic level categories, derived from De Deyne et al. (2008) and Verheyen et al. (2011) respectively, were used. The data consists of similarity judgments for all pairwise combinations of exemplars from 5 animal categories (*birds*, *fish*, *insects*, *mammals*, and *reptiles*), 6 artifact categories (*clothing*, *kitchen utensils*, *musical instruments*, *tools*, *vehicles*, and *weapons*) and 6 abstract categories (*art forms*, *crimes*, *diseases*, *emotions*, *media*, *sciences*, and *virtues*). Because the comparisons were performed at a basic-category level, they required an evaluation of nuanced and detailed properties (for instance, when comparing *hamster* and *mouse* or *kindness* and *helpfulness*).

In contrast to the information encoded at the basic category level, it is possible that the networks cover semantics at a wider range and capture a more course structure. According to this scenario, the networks would only capture a small amount of the variability of the relatedness structure within basic-level categories, but are well suited to distinguish between categories, at the domain level. This would mean that for instance natural kinds and artifacts can be distinguished at a high level in the hierarchy, perhaps at level 2 or 3 in Fig. 2. This might be especially true for the language networks. Since they tend to have broader clusters, they might adequately capture domain distinctions.

To test whether the networks differ in terms of how they capture domain differences apart, a second dataset was included. In this dataset, items from the 5 basic-level animal or 6 basic-level artifact categories introduced previously were paired, leading to pairs such as *butterfly* and *eagle* or *accordion* and *fridge*. If the networks are primarily sensitive to domain-level differences, this would lead to better predictions compared to basic-level categories. Since it is not feasible to present to participants all the pairwise combinations of the combined set of artifact or animal items, only five items from each of the artifact and animal categories were selected. Both items that were central to the category (e.g., *swallow* is a typical bird and thus a central member) and items that were not (e.g., *bat* is an atypical member of the mammals set, and is closely related to birds) were included.

As suggested by the findings on the network communities, it is quite likely that the lexicon reflects a thematic rather than taxonomic organization. If this is the case, this would suggest a high degree of agreement for human judgments of thematic pairs, compared to the basic-level pairs and domain-level pairs. In contrast to the previous pairs, thematic pairs can be closely related without necessarily belonging to a common category or domain. To test these hypotheses thematically related pairs, such as *boat* and *captain* and *rabbit* and *carrot* were used. The set of pairs included among others the items from the study by Miller and Charles (1991), a widely used benchmark test in computational linguistics.

For each of the three studies the number of pairs are listed in the first row of Table 8. An average of 17 participants provided relatedness judgments for a pair of words. The average judgments proved very reliable with Spearman-Brown split-half correlations ranging between .85 and .99. For details and stimuli see De Deyne et al. (2014).

Besides addressing how the networks predict judgments for distinct types of semantic relations, an important issue that remains is the role of sparsity in each graph. While all networks are extremely sparse, the network statistics in Table 3 indicate large differences in terms of network connectivity. The small out-degrees in the mental networks and the matched lexical network G_{lex123} hint at potential limitations when overlap measures of similarity are used based on common neighbors. To investigate if indirect paths between nodes can contribute to model derived estimates of relatedness by reducing sparsity, a random-walk based measure for relatedness will be proposed.

3.3.1 Network Relatedness Measures

A widely used measure of similarity is the cosine measure. This distributional overlap measure captures the extent to which two nodes in the network share the same immediate neighbors. Two nodes that share no neighbors have a similarity of 0, and nodes that are linked to the exact same set of neighbors have similarity 1. Formally, it is defined as follows. Let **A** denote a weighted adjacency matrix, whose element a_{ij} contains a count of the number of times word j is given as an associate of word i in a word association task or the times it occurs in a syntactic dependency relationship. Each row in **A** is therefore a vector containing the associate / syntactic dependency frequencies for word i. The cosine measure of similarity is obtained by first normalizing each row so that all of these vectors are of length 1. This gives us a new matrix **G**, where $g_{ij} = a_{ij}/(\sum_j a_{ij}{}^2)^{1/2}$, and the matrix of all pairwise similarities is now:

$$\mathbf{S} = \mathbf{G}\mathbf{G}^T \tag{1}$$

The cosine measure defined in the previous section depends solely on the *local* structure of the graph: the similarity between two words is assessed by looking only at the words to which they are immediately linked. A different approach to similarity aims to take into account the overall structure of the entire network graph, and thus reflects a broader view of the relationship between two nodes. In this approach two nodes are similar if they share many direct or indirect paths. These paths are explored by a random walker, which stochastically follows local links in the network until the proportion of time it visits each node in the limit converges to a stationary distribution (Hughes and Ramage 2007).

Formally, this random walk corresponds to the regular equivalence measure by Leicht et al. (2006) and is specified by beginning with the weighted adjacency matrix **A**. This time, however, we normalize the rows so that each one expresses a probability distribution over words. That is, we use the matrix **P** where $p_{ij} = a_{ij}/\sum_j a_{ij}$, and then calculate

$$\mathbf{G}' = (\mathbf{I} - \alpha\mathbf{P}^{-1}) \tag{2}$$

where $\mathbf{I}$ is a diagonal identity matrix and the α parameter governs the decay in spread of activation by determining the relative contribution of short and longer paths. A path of length r is assigned a weight of α^r, so when $\alpha < 1$, longer paths get less weight than shorter ones.[4]

The resulting network G' can be thought of as a network of weighted paths. The similarity of two nodes in this network corresponds to the similarity of their stationary distributions. The value of α was fixed at 0.80 (similar to the α for PageRank used in previous sections). This represents a reasonable trade-off between some degree of decay and a non-trivial contribution of longer paths. As in the local relatedness measure above, a cosine measure can then be used to derive a pairwise similarity matrix $\mathbf{S}$ using these distributions. In contrast to the local relatedness measure, such random-walk based measure involves the entire network and is therefore sensitive to the global or macroscopic structure of the network.

3.3.2 Results

For each of the four graphs, relatedness measures were derived as defined above. The measures of relatedness were correlated with the human judgments after standardizing the measures for each category.

Table 8 Results of the similarity analyses for the four datasets (concrete, abstract, domain and thematic) and four graphs

		Basic Level		Domain Level	Thematic
		Concrete	Abstract		
	N	4437	694	1470	126
Local Overlap (Cosine)	G_{asso1}	.528	.505	.711	.567
	G_{asso123}	.593	.617	.786	.769
	G_{lex}	.338	.423	.623	.463
	G_{lex123}	.337	.319	.620	.481
Spreading Activation	G_{asso1}	.564	.616	.793	.802
	G_{asso123}	.590	.660	.824	.827
	G_{lex}	.370	.433	.718	.523
	G_{lex123}	.361	.375	.653	.504

The results for the local overlap measure presented in the top part of Table 8 show moderate to strong correlations between human judgments of relatedness and network-derived measures. One of the most striking patterns in Table 8 is the systematic difference between the amount of variability accounted for by the four graphs. Regardless of the dataset, the denser G_{asso123} network shows substantial

[4] This approach is very similar to the PageRank measure ($\mathbf{X} = (\mathbf{I} - \alpha\mathbf{P}^{-1})\mathbf{1}$).

better agreement than all other graphs. Moreover, even the sparse mental network G_{asso1} outperforms the lexical networks in all cases. Since the G_{lex} network is almost 20 times denser than G_{lex123} (see Table 3), one would expect a better result for this denser graph, however, a significantly different correlation was only found for the basic-level abstract words $z = 2.25$, $p < .05$. A closer look at the different semantic relations indicates that the networks primarily capture the domain judgments, followed by thematic, and basic-level judgments. In line with the community clustering results, this confirms that the networks organize meaning in a thematic way but also include some taxonomic structure.

Next, the role of spreading activation in predicting human relatedness judgments was investigated. The results for the random walk-based spreading activation measure show a consistent improvement for all networks and datasets. The only exception were the results for the concrete words in G_{asso123}, where a slightly lower correlation was found. In this case, the setting of $\alpha = 0.80$ might have resulted in a detrimental contribution of longer paths. When α was systematically varied, the correlation improved to .601 for $\alpha = 0.6$, indicating that the optimal value of this parameter may depend on the type of relationships under consideration. However, in general, the correlation changes were very moderate across various parameter settings. Similar to the overlap measure, the large difference in density between G_{lex} and G_{lex123} did not systematically affect the performance in these language graphs, as only for the domain dataset the correlation values were significantly different $z = 3.33$, $p < .05$.

In conclusion, the use of human relatedness judgments to compare how different taxonomic and thematic relations are represented in the language and mental networks, resulted in findings similar to those from the community clustering of these networks described earlier. Language and mental networks capture primarily the domain level relations between words followed by the thematic relations. The mental networks also capture the basic-level conceptual structure, but the strength of this correlation was moderate. Regardless of the dataset, the mental networks provided a clearly better prediction of human judgments. Using longer indirect paths derived through a stochastic random walk led to systematic improvements in both types of networks, but did not alter the basic findings regarding the relationships captured by these networks.

4 Discussion

In this chapter, the main goal was to compare the macroscopic and mesoscopic properties of language and mental graphs, derived from text corpora and word associations, respectively. One of the key results was that representations systematically differ between both graphs. These differences in itself provide us with important pointers about what processes operate on the linguistic input humans are exposed to.

At a global, macroscopic level, the network-based approach unveiled a highly structured representation that is characterized by short average path-lengths and a significant degree of clustering in both language and mental graphs. This indicates that both graphs have a small-world structure. While there is some overlap between what constitutes a hub in the respective graphs, systematic differences between node centrality emerged. In mental graphs, a larger role for nodes that are presumably of psychological importance exists, while in the language networks hub nodes appear to be more abstract. The latter might reflect the frequency of words typical for language derived from newspaper and other written sources. Moreover, what constitutes a central hub in the mental network seems to be a universal property shared among multiple languages. For instance, for a similar ongoing word association project in English, the ten largest hubs in terms of in-strength in a network with 7,000 nodes corresponded to *money*, *food*, *water*, *love*, *work*, *car*, *music*, *time*, *happy*, and *green*.

Furthermore, the structure of the network argues against the view of the mental lexicon as exclusively and strictly taxonomically organised, where words are grouped in coherent semantic domains and categories. First of all, a substantial number of words were part of multiple communities, which argues against mutually exclusive categories. Second, while the representations can be described in a hierarchical clustered decomposition of the graph, most clusters or communities are characterized by thematic coherence rather than reflecting the type of structure that underlies thesauri, natural taxonomies, or WordNET.

The thematic structure was wide-spread, showing up in nearly all investigated communities at various depths of the hierarchy. The finding that many words from domains like animals, which traditionally are considered taxonomic, are thematically clustered at the lowest level of the hierarchy, corroborates the idea that the networks are organized along primarily thematic rather than categoric lines. In addition, evaluating the obtained structure in the language and mental networks through human relatedness judgments also confirmed the thematic nature of the networks as indicated by the large proportion of variance that was explained for thematic compared to basic-level judgments. This converges with recent evidence that highlights the role of thematic representations even in domains such as animals (Wisniewski and Bassok 1999; Lin and Murphy 2001; Gentner and Kurtz 2006) and the fact that a taxonomic organization of knowledge might be both heavily culturally defined (Lopez et al. 1997), a consequence of formal education (Sharp et al. 1979) or reflect different levels of expertise (Medin et al. 1997).

A number of explanations can account for why thematic structure was so central in both language and mental networks. One possible explanation is the wide coverage of all kinds of words in the network in terms of their abstractness, emotional connotation, and part of speech (verbs, adjectives, and nouns). By not restricting the type of words in the network, the risk of a selection bias towards concrete nouns (Medin et al. 2000) is reduced and the likelihood of identifying thematic relationships increases. In addition, it is quite likely that this reflects an inherent property of language, where most words are taxonomically related to only a small number of other words, but might occur in a variety of thematic settings. This is in line with previous findings showing that Zipf's law reflects the tendency to avoid excessive synonymy in semantic networks (Manin 2008). Clearly, many of these claims remain speculative, but given their potential implications for understanding the mental lexicon, it is hoped they will motivate future work.

One of the key features in many psychological network proposals is the idea of spreading activation. The current study showed that such a mechanism is of importance as it makes use of the network as a whole. The results show that by including not only direct paths that exist between two nodes (neighbors) but also indirect paths, leads to an improved ability to predict human judgments of relatedness. While this measure led to improvements in all networks, the current results also showed that the gain of indirect paths in predicting relatedness was modulated by the sparsity in the original graph, which is well exemplified by comparing the gains for G_{asso1} to those of G_{asso123}.

Similar to the spreading activation account at the mesoscopic level, access at the microscopic level might be governed by more than just the in-strength of a specific node. Measures such as eigen-centrality and PageRank make it conceptually clear that central nodes are those nodes which are easily reached among many possible paths in the graph. These measures are examples of recursive centrality measures, in which centrality is not only influenced by the neighbors of a node, but also takes into account the centrality of the neighbors themselves. This might result in similar benefits found for the spreading activation mechanism operating on sparse graphs. Support for this idea comes from recent studies showing that PageRank accounts for more variance than simple measures of in-strength (Griffiths et al. 2007) and detailed theoretical accounts that explain word frequency advantages in word recognition through higher level structural properties of the network (Monaco et al. 2007). Again, this illustrates the benefits of a network approach which simultaneously describes a macro-, meso- and microscopic level.

4.1 Relationship between Language and Word Associations

A number of studies have tried to predict word associations from text corpora (e.g. Griffiths and Steyvers 2003). While this prediction is often used as a yardstick to compare different text-based models, one of the striking patterns is the overall poor prediction. For instance, in a study by Griffiths and Steyvers (2003), the median rank for predicting the first word association in the University of Florida norms (Nelson et al. 2004) using a text-based topic model was 32. Prediction of the Dutch word association norms (which are considerably larger than the University of Florida norms) on the basis of G_{asso123} resulted in a median rank of 129, and the correct prediction of the first associate in only 5.4% of the cases. Similar results were found when the overlap was calculated in terms of relatedness. Here both the association graphs were strongly correlated (.99), and so were the lexical graphs (.88). Crucially, between both types of graphs, the agreement was quite small: .14 for G_{asso123} and G_{lex} and .11 for G_{asso123} and G_{lex123}. Similar comparisons of microscopic measures of centrality showed only moderate correlations between language and word association graphs. The choice of network - language or mental - might thus lead to different conclusions about findings that show how semantic rich nodes (i.e., those with a high degree, or high clustering) are processed more efficiently in naming or word recognition (Buchanan et al. 2001; Pexman et al. 2003; Pexman et al. 2008).

The limited agreement between the networks and the systematic differences in how they account for specific types of words (especially concrete ones) provide further support to the idea that the association task does not rely on the same properties as common language production, but should rather be seen as tapping into the semantic information of the mental lexicon (Mollin 2009; McRae et al. 2011). This view resonates with the original ideas of Collins and Loftus (1975), in which the network depends both on semantic similarity and lexical co-occurrence in language, and other works that highlight the role of imagery and affect in the production of word associations (Szalay and Deese 1978). As mentioned in the introduction, the role of pragmatics in natural language explains why mentally central properties (e.g., the fact that bananas are yellow or apples are round) are very strong responses in word association data but much less prominently expressed in conventional written and spoken language. To some extent, this might also be the reason why the language networks did not fully capture the human judgments for concrete words (see Table 8). Of course, one could also argue that the language networks in this study are simply too limited due to the vocabulary size restrictions. It seems unlikely that this explanation can account for the entire set of findings. First of all, the results for G_{lex} and G_{lex123} showed that a sampled network based on only a fraction of the tokens produced comparable results for a number of domains. Second, a comparable study involving a language network consisting of a vocabulary of over 100,000 lemmas and the same human relatedness judgments, produced highly similar results (De Deyne et al. 2014). Naturally, this is not to say that additional data and pragmatics are irrelevant. In understanding a story, for instance, where representations that go beyond the word level are required, pragmatics are likely to play a more central role.

4.2 Final Words

In this chapter, a view of the mental lexicon, as a weighted directed graph, with words for nodes, has been advanced as a useful way to explore the structure and processing of word meaning. This account is limited in various ways and by no means complete. For instance, further studies are needed to investigate whether qualitatively different links could lead to a better model of the lexicon through differential weighting of different types of relations in the language network, either syntactic (conjunctions, modification of nouns, etc.) or semantic (hyperonymy, meronymy, etc.).

Similar to the language network, the connections in the word association network are presumably governed by a set of latent relations. In this area as well, the use of a multi-network representation with different weights for various types of relations is likely to explain additional properties of the data. On the basis of these relations, new studies might reveal distinct types of comparison processes as suggested by previous work on thematic and taxonomic comparisons (Wisniewski and Bassok 1999). In particular, a first type of process could be based on the integration of a word in a thematic context (e.g., *doctor* and *hospital*) while a second type might involve the alignment of shared properties between similar entities (e.g., *cat* and *tiger*). Presumably these processes might reflect a highly probable path in the former situation, while some kind of summation over a large number of different paths could be involved in the second process. Knowing something about the properties of nodes on a path (e.g., whether they refer to similar physical entities, a function, or thematic property) requires the derivation of a multi-network as mentioned earlier, and could inform us how such a differential comparison process takes place.

Of course, there are many other areas in which a network approach is likely to contribute in future studies of the lexicon, for instance by studying the development of the lexicon through dynamic networks (Beckage et al. 2010), the networks of individuals (Morais et al. 2013) or by comparing the networks of healthy individuals with clinical populations (Kenett et al. 2013). Presumably, better assumptions about how representations are extracted from the statistical regularities in the language environment will play an important role in these endeavors. In this respect, the application of a syntax-based dependency model represents a first, but certainly not the last step to build a more appropriate mental model of the lexicon. The close relation with empirical indices of mental organization such as human relatedness judgments, but potentially also online measures such as priming (Chumbley and Balota 1984) and word centrality (De Deyne et al. 2013a), suggests that a mental network derived from word associations represents a valuable alternative to model cognitive functions at various levels of abstraction offered through a network science framework.

Acknowledgements. My gratitude goes to Kris Heylen, Dirk Speelman, and Dirk Geeraerts for making the LeNC corpus available, and to Yves Peirsman for collaboration during the early stages of this project. This work was supported by Research Grant G.0436.13 from the Research Foundation - Flanders (FWO) to the first author and by the interdisciplinary research project IDO/07/002 awarded to Dirk Speelman, Dirk Geeraerts, and Gert Storms. Steven Verheyen is a postdoctoral fellow at the Research Foundation - Flanders. Comments may be sent to `simon.dedeyne@adelaide.edu.au`

References

Ball, B., Karrer, B., Newman, M.E.J.: Efficient and principled method for detecting communities in networks. Physical Review E 84(3), 036103 (2011)

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., Christiansen, M.H.: Networks in cognitive science. Trends in Cognitive Sciences 17(7), 348–360 (2013)

Beckage, N.M., Smith, L.B., Hills, T.: Semantic network connectivity is related to vocabulary growth rate in children. In: The Annual Meeting of The Cognitive Science Society (CogSci), pp. 2769–2774 (2010)

Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008(10), P10008+ (2008)

Borge-Holthoefer, J., Arenas, A.: Categorizing words through semantic memory navigation. The European Physical Journal B-Condensed Matter and Complex Systems 74(2), 265–270 (2010)

Bouma, G., van Noord, G., Malouf, R.: Alpino: Wide Coverage Computational Analysis of Dutch. In: Eleventh Meeting of Computational Linguistics in the Netherlands, CLIN, Tilburg, pp. 45–59 (2000)

Buchanan, L., Westbury, C., Burgess, C.: Characterizing the neighbourhood: Semantic neighbourhood effects in lexical decision and naming. Psychonomics Bulletin and Review 8, 531–544 (2001)

Bullinaria, J.A., Levy, J.P.: Extracting semantic representations from word co-occurrence statistics: A computational study. Behavior Research Methods 39, 510–526 (2007)

Bullmore, E., Sporns, O.: The economy of brain network organization. Nature Reviews Neuroscience 13(5), 336–349 (2012)

Capitán, J.A., Borge-Holthoefer, J., Gómez, S., Martinez-Romo, J., Araujo, L., Cuesta, J.A., Arenas, A.: Local-based semantic navigation on a networked representation of information. PloS One 7(8), e43694 (2012)

Ceulemans, E., Storms, G.: Detecting intra-and inter-categorical structure in semantic concepts using HICLAS. Acta Psychologica 133(3), 296–304 (2010)

Chumbley, J.I.: The roles of typicality, instance dominance, and category dominance in verifying category membership. Journal of Experimental Psychology: Learning, Memory, and Cognition 12, 257–267 (1986)

Chumbley, J.I., Balota, D.A.: A word's meaning affects the decision in lexical decision. Memory & Cognition 12, 590–606 (1984)

Collins, A.M., Loftus, E.F.: A spreading-activation theory of semantic processing. Psychological Review 82, 407–428 (1975)

Collins, A.M., Quillian, M.R.: Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior 9, 240–247 (1969)

Cramer, P.: Word Association. Academic Press, New York (1968)

Crutch, S.J., Warrington, E.K.: Abstract and concrete concepts have structurally different representational frameworks. Brain 128, 615–627 (2005)

De Deyne, S., Verheyen, S., Storms, G.: The role of corpus-size and syntax in deriving lexico-semantic representations for a wide range of concepts. Quarterly Journal of Experimental Psychology 26, 1–22 (2015), doi:10.1080/17470218.2014.994098

De Deyne, S., Voorspoels, W., Verheyen, S., Navarro, D.J., Storms, G.: Accounting for graded structure in adjective categories with valence-based opposition relationships. Language, Cognition and Cognitive Processes 29, 568–583 (2013)

De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M., Voorspoels, W., Storms, G.: Exemplar by Feature Applicability Matrices and Other Dutch Normative Data for Semantic Concepts. Behavior Research Methods 40, 1030–1048 (2008)

De Deyne, S.: Proximity in Semantic Vector Space. Unpublished Doctoral Dissertation (2008)

De Deyne, S., Navarro, D.J., Storms, G.: Better explanations of lexical and semantic cognition using networks derived from continued rather than single word associations. Behavior Research Methods 45, 480–498 (2013a)

De Deyne, S., Navarro, D.J., Storms, G.: Associative strength and semantic activation in the mental lexicon: evidence from continued word associations. In: Knauff, M., Pauen, M., Sebanz, N., Wachsmuth, I. (eds.) Proceedings of the 33rd Annual Conference of the Cognitive Science Society, pp. 2142–2147. Cognitive Science Society, Austin (2013b)

De Deyne, S., Storms, G.: Word Associations: Network and Semantic properties. Behavior Research Methods 40, 213–231 (2008)

de Groot, A.M.B.: Representational Aspects of Word Imageability and Word Frequency as Assessed Through Word Association. Journal of Experimental Psychology: Learning, Memory, and Cognition 15, 824–845 (1989)

Dennis, S.: A Memory-Based Theory of Verbal Cognition. Cognitive Science 29(2), 145–193 (2005)

Fagiolo, G.: Clustering in Complex Directed Networks. Physical Review E 76, 026107 (2007)

Fellbaum, C.: WordNet: An electronic lexical Database. MIT Press, Cambridge (1998)

Firth, J.R.: Selected papers of JR Firth, 1952-59. Indiana University Press (1968)

Fortunato, S.: Community detection in graphs. Physics Reports 486(3), 75–174 (2010)

Galbraith, R.C., Underwood, B.J.: Perceived frequency of concrete and abstract words. Memory & Cognition 1, 56–60 (1973)

Gentner, D., Kurtz, K.: Relational categories. In: Ahn, W.K., Goldstone, R.L., Love, B.C., Markman, A.B., Wolff, P.W. (eds.) Categorization Inside and Outside the Lab, pp. 151–175. American Psychology Association, Washington, DC (2005)

Gentner, D., Kurtz, K.: Relations, objects, and the composition of analogies. Cognitive Science 30, 609–642 (2006)

Griffiths, T.L., Steyvers, M.: Prediction and Semantic Association. In: Advances in Neural Information Processing Systems, vol. 15, pp. 11–18. MIT Press, Cambridge (2003)

Griffiths, T.L., Steyvers, M., Firl, A.: Google and the Mind. Psychological Science 18, 1069–1076 (2007)

Heylen, K., Peirsman, Y., Geeraerts, D.: Automatic synonymy extraction. In: A Comparison of Syntactic Context Models. LOT Computational Linguistics in the Netherlands 2007, pp. 101–116 (2008)

Hughes, T., Ramage, D.: Lexical Semantic Relatedness with Random Graph Walks. In: Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), pp. 581–589. Association for Computational Linguistics, Prague (2007)

Hutchison, K.A.: Is semantic priming due to association strength or feature overlap? Psychonomic Bulletin and Review 10, 785–813 (2003)

Kenett, Y.N., Kenett, D.Y., Ben-Jacob, E., Faust, M.: Global and local features of semantic networks: Evidence from the Hebrew mental lexicon. PloS One 6(8), e23912 (2011)

Kenett, Y.N., Wechsler-Kashi, D., Kenett, D.Y., Schwartz, R.G., Ben-Jacob, E., Faust, M.: Semantic organization in children with cochlear implants: computational analysis of verbal fluency. Frontiers in Psychology 4 (2013)

Keuleers, E., Brysbaert, M., New, B.: SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. Behavior Research Methods 42(3), 643–650 (2010)

Kintsch, W.: Comprehension: A paradigm for cognition. Cambridge University Press (1998)

Kintsch, W., Mangalath, P.: The construction of meaning. Topics in Cognitive Science 3(2), 346–370 (2011)

Kiss, G.R.: Words, Associations, and Networks. Journal of Verbal Learning and Verbal Behavior 7, 707–713 (1968)

Lancichinetti, A., Radicchi, F., Ramasco, J.J., Fortunato, S.: Finding statistically significant communities in networks. PloS One 6(4), e18961 (2011)

Landauer, T.K., Dumais, S.T.: A solution to Plato's Problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. Psychological Review 104, 211–240 (1997)

Landauer, T.K.: LSA as a theory of meaning. In: Landauer, T.K., McNamara, D.S., Dennis, S., Kintsch, W. (eds.) Handbook of Latent Semantic Analysis, pp. 3–35. Lawrence Erlbaum Associates, Mahwah (2007)
Leicht, E.A., Holme, P., Newman, M.E.J.: Vertex similarity in networks. Psychical Review E 73, 026120 (2006)
Lin, E.L., Murphy, G.L.: Thematic Relations in Adults' Concepts. Journal of Experimental Psychology: General 1, 3–28 (2001)
Lopez, A., Atran, S., Coley, J.D., Medin, D.L., Smith, E.E.: The tree of life: Universal and cultural features of folkbiological taxonomies and inductions. Cognitive Psychology 32(3), 251–295 (1997)
Lund, K., Burgess, C.: Producing high-dimensional semantic spaces from lexical co-occurrence. Behavior Research Methods, Instruments, and Computers 28, 203–208 (1996)
Manin, D.Y.: Zipf's law and avoidance of excessive synonymy. Cognitive Science 32(7), 1075–1098 (2008)
McRae, K., Khalkhali, S., Hare, M.: Semantic and associative relations: examining a tenuous dichotomy. In: Reyna, V.F., Chapman, S., Dougherty, M., Confrey, J. (eds.) The Adolescent Brain: Learning, Reasoning, and Decision Making, pp. 39–66. American Psychological Association, Washington, DC (2011), doi:10.1037/13493-002
Medin, D.L., Lynch, E.B., Coley, J.D., Atran, S.: Categorization and reasoning among tree experts: Do all roads lead to Rome? Cognitive Psychology 32(1), 49–96 (1997)
Medin, D.L., Lynch, E.B., Solomon, K.O.: Are there kinds of concepts? Annual Review of Psychology 51(1), 121–147 (2000)
Medin, D.L., Rips, L.J.: Concepts and categories: memory, meaning, and metaphysics. In: Holyoak, K., Morrison, R. (eds.) The Cambridge Handbook of Thinking and Reasoning, pp. 37–72. Cambridge University Press, Cambridge (2005)
Mervis, C.B., Rosch, E.: Categorization of natural objects. Annual Review of Psychology 32, 89–115 (1981)
Miller, G.A., Charles, W.G.: Contextual Correlates of Semantic Similarity. Language and Cognitive Processes 6, 1–28 (1991)
Mollin, S.: Combining corpus linguistics and psychological data on word co-occurrence: Corpus collocates versus word associations. Corpus Linguistics and Linguistic Theory 5, 175–200 (2009)
Monaco, J.D., Abbott, L.F., Kahana, M.J.: Lexico-semantic structure and the word-frequency effect in recognition memory. Learning & Memory 14, 204–213 (2007)
Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.-L., De Schryver, M., De Winne, J., Brysbaert, M.: Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. Behavior Research Methods, 1–9 (2012)
Morais, A.S., Olsson, H., Schooler, L.J.: Mapping the structure of semantic memory. Cognitive Science 37(1), 125–145 (2013)
Motter, A.E., de Moura, A.P.S., Lai, Y.-C., Dasgupta, P.: Topology of the conceptual network of language. Physical Review E 6, 065102 (2002)
Murphy, G.L.: The big book of concepts. MIT Press, Cambridge (2002)
Nelson, D.L., McEvoy, C.L.: What is this thing called frequency? Memory & Cognition 28, 509–522 (2000)
Nelson, D.L., McEvoy, C.L., Schreiber, T.A.: The University of South Florida free association, rhyme, and word fragment norms. Behavior Research Methods, Instruments, and Computers 36, 402–407 (2004)
Niedenthal, P.M., Halberstadt, J.B., Innes-Ker, A.H.: Emotional response categorization. Psychological Review 106(2), 337 (1999)

Oostdijk, N.: The Spoken Dutch Corpus: Overview and first evaluation. In: Piperidis, S., Stainhaouer, G. (eds.) Proceedings of Second International Conference on Language Resources and Evaluation, vol. 2, pp. 887–894. ELRA, Paris (2000)

Ordelman, R.J.F.: Twente nieuws corpus (TWNC). Tech. rep. Parlevink Language Technology Group, University of Twente (2002)

Padó, S., Lapata, M.: Dependency-Based Construction of Semantic Space Models. Computational Linguistics 33(2), 161–199 (2007)

Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: Bringing order to the web. Tech. rep. Computer Science Department, Stanford University (1998)

Pereira, F., Tishby, N., Lee, L.: Distributional clustering of English words. In: Proceedings of the 31st Annual Meeting on Association for Computational Linguistics, pp. 183–190. Association for Computational Linguistics, Columbus (1993)

Pexman, P.M., Hargreaves, I.S., Siakaluk, P.D., Bodner, G.E., Pope, J.: There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. Psychonomic Bulletin & Review 15(1), 161–167 (2008)

Pexman, P.M., Holyk, G.G., Monfils, M.-H.: Number-of-features effects in semantic processing. Memory & Cognition 31, 842–855 (2003)

Plaut, D.C., Shallice, T.: Deep dyslexia: A case study of connectionist neuropsychology. Cognitive Neuropsychology 10, 377–500 (1993)

Prior, A., Bentin, S.: Incidental formation of episodic associations: the importance of sentential contex. Memory & Cognition 31, 306–316 (2003)

Van Rensbergen, B., De Deyne, S., Storms, G.: Cue-association correspondence on valence, dominance, and arousal. Manuscript submitted for publication (2014)

Rogers, T.T., McClelland, J.L.: Semantic cognition: A parallel distributed processing approach. MIT Press, Cambridge (2004)

Rosch, E.: Natural Categories. Cognitive Psychology 4, 328–350 (1973)

Ruts, W., De Deyne, S., Ameel, E., Vanpaemel, W., Verbeemen, T., Storms, G.: Dutch norm data for 13 semantic categories and 338 exemplars. Behaviour Research Methods, Instruments, and Computers 36, 506–515 (2004)

Samsonovic, A.V., Ascoli, G.A.: Principal semantic components of language and the measurement of meaning. PloS One 5(6), e10921 (2010)

Schank, R.C., Abelson, R.P.: Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. Lawrence Erlbaum, Hillsdale (1977)

Schwanenflugel, P.J., Akin, C., Luh, W.M.: Context availability and the recall of abstract and concrete words. Memory & Cognition 20, 96–104 (1992)

Sharp, D., Cole, M., Lave, C., Ginsburg, H.P., Brown, A.L., French, L.A.: Education and cognitive development: The evidence from experimental research. In: Monographs of the society for Research in Child Development, pp. 1–112 (1979)

Simmons, W.K., Hamann, S.B., Harenski, C.N., Hu, X.P., Barsalou, L.W.: fMRI evidence for word association and situated simulation in conceptual processing. Journal of Physiology - Paris 102, 106–119 (2008)

Solé, R.V., Corominas-Murtra, B., Valverde, S., Steels, L.: Language networks: Their structure, function, and evolution. Complexity 15(6), 20–26 (2010)

Steyvers, M., Tenenbaum, J.B.: The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. Cognitive Science 29, 41–78 (2005)

Szalay, L.B., Deese, J.: Subjective meaning and culture: An assessment through word associations. Lawrence Erlbaum, Hillsdale (1978)

Van Dongen, S.: Graph clustering by flow simulation. PhD thesis. University of Utrecht (2000)

Verheyen, S., Stukken, L., De Deyne, S., Dry, M.J., Storms, G.: The generalized polymorphous concept account of graded structure in abstract categories. Memory & Cognition 39, 1117–1132 (2011)
Vitevitch, M.S.: What can graph theory tell us about word learning and lexical retrieval? Journal of Speech, Language, and Hearing Research 51(2), 408–422 (2008)
Warrington, E.K., Shallice, T.: Category specific semantic impairments. Brain 107, 829–854 (1984)
Wisniewski, E.J., Bassok, M.: What Makes a Man Similar to a Tie? Cognitive Psychology 39, 208–238 (1999)
Wittgenstein, L.: Philosophical Investigations: 50th Anniversary Commemorative Edition. Blackwell Publishing, Incorporated, Oxford (2001)