

A Graph-Based Approach for Computing Free Word Associations

Gemma Bel Enguix, Reinhard Rapp, Michael Zock

Aix-Marseille Université, Laboratoire d'Informatique Fondamentale
UMR 7279, Case 901, 163 Avenue de Luminy, 13288 Marseille, France
gemma.belenguix@gmail.com, reinhardrapp@gmx.de, Michael.Zock@lif.univ-mrs.fr

Abstract

A graph-based algorithm is used to analyze the co-occurrences of words in the British National Corpus. It is shown that the statistical regularities detected can be exploited to predict human word associations. The corpus-derived associations are evaluated using a large test set comprising several thousand stimulus/response pairs as collected from humans. The finding is that there is a high agreement between the two types of data. The considerable size of the test set allows us to split the stimulus words into a number of classes relating to particular word properties. For example, we construct six saliency classes, and for the words in each of these classes we compare the simulation results with the human data. It turns out that for each class there is a close relationship between the performance of our system and human performance. This is also the case for classes based on two other properties of words, namely syntactic and semantic word ambiguity. We interpret these findings as evidence for the claim that human association acquisition must be based on the statistical analysis of perceived language, and that when producing associations the detected statistical regularities are replicated.

Keywords: free word associations, association norms, language acquisition, acquisition of associations

1. Introduction

Graph theory is important for studying association networks, and both have a long history. Their combination along with statistical information can help us to better understand our learning and using of language. Associations go back at least to Aristotle,¹ and as every mathematician knows, graph theory is rooted in Euler's analysis (Euler, 1735) of walking the seven bridges in Königsberg.² The results of his analysis laid the foundation of modern graph theory (Euler, 1735). Graphs³ are mathematical structures modelling pairwise relations between objects. Since objects can be anything (cities, people, words, webpages), graphs are very useful abstractions allowing us to model all kinds of things and phenomena, be they *static* (structures, i.e. organisation, topology, distance, ...) or *dynamic* (connectionism, word

access, evolution of social networks, etc.). This is probably also one of the reasons why they have become so popular. Nowadays they are used in many domains (social network, neurosciences, computational linguistics, etc.).

The very notion of associations has been very common in psycholinguistics in the middle of the last century (see Hörman, 1972, Wettler, 1980), and associationism has been a major paradigm in psycholinguistics until Chomsky's devastating critique on Skinner's book 'Verbal Behavior' and behaviorism (Chomsky, 1959) which was equated to associationism. Since then, there has been a revival in the form of connectionism (Rumelhart et al., 1986). Nowadays, associations are used for many tasks: support navigation to access words (Zock et al., 2010), brainstorming, memorisation (Meara, 2009), etc. Encouraged by this revival and their usefulness, psychologists have started again to build association lists, making them available to the community to be used freely. Actually, psychologists had built such lists already decades ago (Deese, 1965; Schvaneveldt et al. 1989), but in those days we did not use computers, neither did we know how to exploit the richness of corpora. Things have changed a lot. Association lists are now freely available on the web. For example, there is the *Edinburgh Associative Thesaurus*⁴ and the compilation of Nelson et al.⁵ These are for English, but there are also compilations available for other languages like Dutch,⁶ French,⁷ German^{8,9} and Japanese.¹⁰ As one can see, associations are considered nowadays as very

¹ Actually the very notion of association goes back at least to Aristotle (350BC), but it is also inherent in work done by philosophers (Locke, 1689; Hume, 1739), physiologists (James & Stuart Mills, 1904), psychologists (Galton, 1880; Freud, 1901; Jung and Riklin, 1906) and psycholinguists (Deese, 1965; Jenkins, 1970; Schvaneveldt, 1989). For surveys in psycholinguistics see (Aitchison, 2003; Hörmann, 1972), or more recent work (Spitzer, 1999). The notion of association is also implicit in work on semantic networks (Quillian, 1968), hypertext (Bush, 1945), the web (Nelson, 1967), connectionism (Dell et al., 1999) and, of course, in WordNet (Miller et al., 1993; Fellbaum, 1998).

² http://en.wikipedia.org/wiki/Seven_Bridges_of_Königsberg

³ For a good historical overview of graph theory and its relation to the brain, see (Sporn, 2011 and 2012; Bullmore & Sporns, 2009, Stam & Reijneveld, 2007). For its relation to the mental lexicon take a look at (Vitevitch, 2008). A very useful survey concerning the use of graph-based methods in NLP and IR can be found in (Mihalcea & Radev, 2011).

⁴ <http://www.eat.rl.ac.uk/>

⁵ <http://cyber.acomp.usf.edu/FreeAssociation>

⁶ <http://www.smallworldofwords.com>

⁷ <http://www.jeuxdemots.org/jdm-accueil.php>

⁸ <http://www.schultheimwalde.de/resource.html>

⁹ <http://www.coli.uni-saarland.de/projects/nag/>

¹⁰ <http://www.valdes.titech.ac.jp/~terry/jwad.html>

valuable resources.

Like associations, probability-based models (Markov processes) have been severely criticized by Chomsky, yet they also have seen an impressive revival. While not being perfect, these methods are nevertheless very powerful. They allow us not only to solve a number of difficult problems, for example, machine translation, the whole grail of NLP, but also to make certain predictions concerning human language behavior, a point we try to check in more detail in this paper.

2. Related work

The word association test as introduced by Galton (1879) has been used by psychoanalysts in the hope to reveal certain aspects of the humans' subconscious. In this kind of experiment, a person typically hears or reads a word, and is asked to come up with the first other word that comes to mind. Among others, Kent & Rosanoff (1910) used this method, introducing a standard list of test words to facilitate comparisons. They conducted the first large scale study of word associations (1000 test persons) and came to the conclusion that there was uniformity in the organization of associations and that people shared stable networks of connections among words (Istifci, 2010).

In this paper, we are less interested in psychoanalytical aspects than in associative learning, and in simulating the observed behavior. According to associationist learning theory (Schwartz & Reisberg, 1991) the associative strength between two events increases by a constant fraction of the maximally possible increment whenever they co-occur, and it decreases in the opposite case. Wettler et al. (2005) showed that this well established learning procedure can be replicated by looking at the co-occurrence-frequencies of words in large text collections.

This had been done with the rise of corpus linguistics even before this connection was known. Two of the pioneers, Church & Hanks (1990), introduced the idea of using *mutual information* for computing association strength. Wettler & Rapp (1989) compared a number of association measures for the purpose of finding search terms in information retrieval. It should be noted, though, that there has been quite some work prior to the above mentioned. Despite the fact that it was not corpus-based, some of it was rather influential.

For example, Collins & Loftus (1975) used associative semantic networks to extract the gist of a set of words. Rosenzweig (1961:358) made the claim that all speakers, regardless of their culture and language, share many verbal associations, even though their verbal forms are different (Ekpo-Ufot, 1978). Although this work is important, having a significant impact on the field, we will concentrate here only on corpus-based work.

In order to derive associations from a corpus, an association measure is needed. Given the number and diversity of possible measures, Evert & Krenn (2001) felt the need to introduce some criteria and methods for their qualitative evaluation. Pecina & Schlesinger (2006) compared 82 different association measures using the task of collocation extraction, while Hoang et al. (2009) classified them. Michelbacher et al. (2011) investigated the potential of asymmetric association measures, and Wash-tell & Markert (2009) tried to answer the question whe-

ther or not word associations should be computed via window-based co-occurrence counts. They came up with a windowless approach that measures the distances between words.

While there have been several studies which quantitatively compare human word associations with computed associations as derived from corpus statistics (e.g. Wettler et al., 2005; Tamir, 2005; Seidensticker, 2006), apart from our own related work (e.g. Bel Enguix et al., 2014) to our knowledge none of them uses an approach based on graph analysis techniques.¹¹ In this paper we try to close this gap by applying such an approach to the problem of word association.

3. Resources

For simulating the human associative behavior on the basis of corpus evidence, a large text corpus is required. This should as far as possible replicate the language environment of people. We decided to use the *British National Corpus* (BNC, Burnard & Aston, 1998) as it is a balanced corpus of modern English comprising about 100 million words.

The corpus was lemmatized, i.e. inflected forms (e.g. *wheels*) were replaced by their base forms (e.g. *wheel*). The purpose of this is to reduce data sparseness and to improve evaluation. Note that evaluation is based on exact string matching. Therefore, if our simulation would produce *wheels* in response to *car*, then normally we would consider this as an error as the primary associative response of the test persons is *wheel*. Lemmatization solves this problem.

As our gold standard for evaluation we used the associations as collected in the *Edinburgh Associative Thesaurus*, henceforth EAT (Kiss 1975; Kiss et al. 1973). These association norms were produced by presenting stimulus words to about 100 subjects each, and by collecting their responses. The subjects were 17 to 22 year old British students. Table 1 shows the associations which were produced by at least five participants in response to the stimulus words *bath* and *cold* together with the number of participants who produced them.

The EAT lists the associations upon altogether 8400 stimulus words. But, as in this study we are particularly interested in nouns, verbs, and adjectives, we removed all other words, and also some multiword units (e.g. '*a lot*') which had made it into the EAT. We also lemmatized the data. This way 5910 test items remained which is considerably more than the usually around 100 stimulus words used in many previous studies (e.g. Wettler et al., 2005).

STIMULUS	OBSERVED RESPONSE	NUMBER OF SUBJECTS
bath	water	20
	tub	8
	clean	5
	hot	5
cold	hot	34
	ice	10
	warm	7
	water	5

¹¹ See Mihalcea & Radev (2011).

Table 1: Extract from the EAT for the stimulus words *bath* and *cold*.

4. Approach

In contrast to previous approaches (Wettler et al. 2005; Church & Hanks, 1990) which used standard vector-based algorithms, we use a graph-based approach to compute word associations. As for this study we were interested in nouns, verbs, and adjectives only, in a pre-processing step we removed all other words from the British National Corpus. Subsequently, a graph was built where the nouns, verbs, and adjectives in the corpus are the nodes, and where the edges between these nodes are zero at the beginning, and are incremented by one whenever the two respective words co-occur in the corpus as direct neighbors. That is, after processing the corpus the weight of each edge represents the number of times the respective words (nodes) co-occur.

As illustrated in Fig. 1, the associations to a given stimulus word are calculated by searching the nodes in the direct neighborhood of this word, and by ranking the neighborhood nodes according to the weights. Given a graph $G=V,E$; with $V=\{i,j,\dots,n\}$ as its set of vertices and E its set of edges linking pairs of nodes over V , we express by $N(i)$ the neighborhood of a node $i \in V$, being $N(i)$ every $j \in V \mid e_{i,j} \in E$.

From its construction, it could be expected that the system will retrieve only syntagmatically related words. But surprisingly it turns out that also many paradigmatic associations are computed.

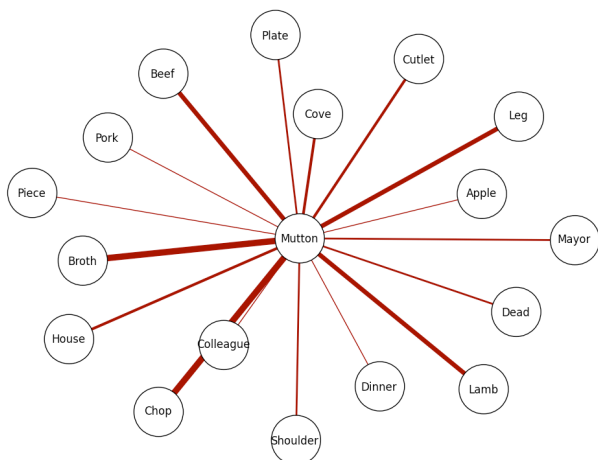


Fig. 1: Results when computing the associations for the word *Mutton* using the BNC. The widths of the connections were computed by normalizing the weights.

5. Results

To give an idea concerning the outcome, Table 1 shows some sample results. Despite the fact that, apart from '*Black* → *White*' the computed primary responses are different from the ones given by humans in the EAT, they seem perfectly plausible. Hence the question: Does this lie within the bandwidth of variation exhibited by human associative behavior?

We measured the quality of our results by counting (for all 5910 items) how many subjects in the EAT answered with the response our system came up with. This number is 6.2 on average. In comparison, the

number of other subjects who also come up with the answer of an average test person is 5.8. If the numbers were identical, our system would be perfectly within the range of variation of the human associative responses, i.e. our system's answers could hardly be distinguished from the human answers. This is actually the case. Our system's answers show even slightly more overlap with the test persons than the test persons among each other.

In what follows we describe further experiments based on word classes derived from three different criteria: saliency, syntactic ambiguity, and semantic ambiguity.

Stimulus Word	Human Primary Response	Computed Primary Response
Afraid	Fear	Person
Anger	Hate	Frustration
Baby	Boy	Mother
Bath	Water	Shower
Beautiful	Ugly	Woman
Bed	Sleep	Hospital
Bible	Book	God
Bitter	Sweet	Taste
Black	White	White
Blossom	Flower	White

Table 1: Comparison between human and computed associations for the first 10 alphabetically sorted words of the Kent/Rosanoff (1910) list.

5.1 Saliency classes

In this experiment we classified our 5910 EAT stimulus words into six saliency classes (SC). Hereby we define *saliency* via the proportion of subjects uttering the primary associative response (PAS; i.e. the most frequent response).

- SC 1: less than 10% producing the PAR (10.7%)
- SC 2: 10 to 20% producing the PAR (36.0%)
- SC 3: 20 to 30% producing the PAR (24.3%)
- SC 4: 30 to 40% producing the PAR (13.3%)
- SC 5: 40 to 50% producing the PAR (8.0%)
- SC 6: more than 50% producing the PAR (7.6%)

The percentages at the end of each line denote the proportion of words belonging to the respective saliency class. All classes are reasonably well covered. Here are some sample words for each class:

- SC 1: black, aunt, woman
- SC 2: aid, cell, gasoline
- SC 3: driver, monarchy, tornado
- SC 4: chief, jungle, kiss
- SC 5: horse, mountain, semaphore
- SC 6: leader, professor, yellow

As it seems impossible to tell them apart, it appears that our intuitions do not easily allow us to make predictions concerning the saliency classifications of words.

Figure 2 (blue curve) shows how well our system performs for each of these saliency classes. For the words in each class it was counted how many human subjects came up with the same associative response as

computed by the system. There is a strong effect that the system's performance is best for very salient words, and decreases for less salient words. Note that this is a desired result if we want to simulate human associative behavior: For words with homogeneous human responses our system is likely to produce the same responses. And for words where the human responses are heterogeneous, our system is likely to produce different responses.

The pink curve in Figure 2 shows for each saliency class how many other test persons respond with the associative answer of an average test person. As can be seen this line is almost identical to the line for the system's performance. This means that with regard to saliency the system's behavior is very similar to the behavior of an average test person.

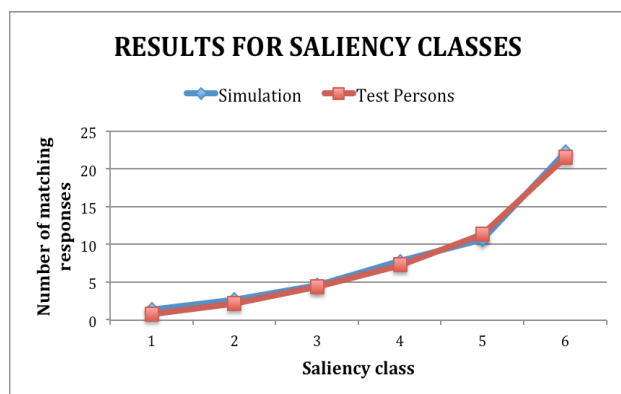


Fig. 2: Quality of our system's (blue curve) and an average test person's (pink curve) performance (measured as the number of matching responses found in the EAT) depending on the saliency class of the given word.

5.2 Syntactic ambiguity classes

In this experiment we took WordNet as a reference. As our network has only nouns, verbs and adjectives, we extracted from WordNet lists of these parts of speech, and classified the words in three syntactic ambiguity classes (SAC):

SAC1: No ambiguity at all. These words appear in WordNet only as a noun, as a verb, or as an adjective.

SAC2: The words which, according to WordNet, belong to any two out of the three parts of the speech.

SAC3: The words which, according to WordNet, are threefold ambiguous, i.e. can be a noun, an adjective, and a verb.

We could not use all of our 5910 EAT test items because some stimulus words were missing in WordNet, so we discarded the respective items. Another problem was that in the BNC and in WordNet words are lemmatized, but not so in the EAT. For example, the EAT contains not only *acid* as a stimulus word, but also *acids*. This is why we also lemmatized the EAT stimulus words. Hereby we eliminated any resulting double occurrences of stimuli, i.e. items where lemmatization led to a stimulus word which was already present. Moreover, compound forms like *a lot* were discarded. Finally, we took the

intersection (common words) between BNC, EAT, and WordNet. This resulted in a test set of 5522 words, distributed as follows:

Number of words in SAC1: 3129 – 56.6%
 Number of words in SAC2: 2151 – 38.9%
 Number of words in SAC3: 242 – 4.38%

The numbers on the right side of each line are the absolute and relative numbers of items belonging to the respective class. Here some sample words from each class:

SAC1: algebra, hardness, require, weak
 SAC2: complex, awake, refund
 SAC3: beat, gross, perfect, stretch

After performing the experiment, the results obtained (Figure 3) show that this parameter does not seem to be very important for word association. The figure demonstrates that, even though the results are slightly better for the words belonging to only one class, the impact of syntactic ambiguity is only small. Nevertheless, even this small effect is correctly replicated by our system.

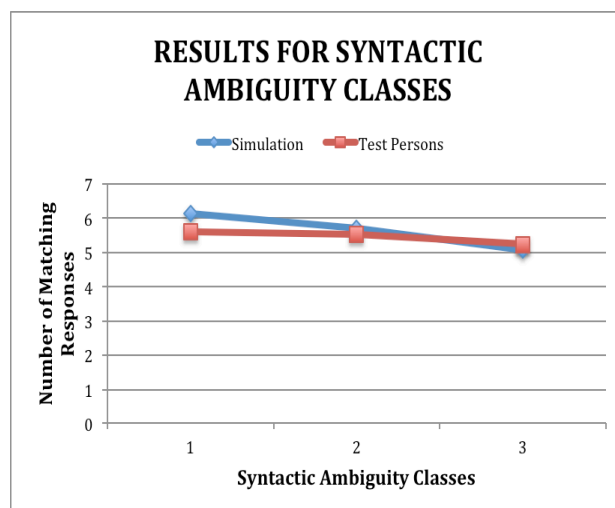


Fig. 3: Quality of our system's (blue curve) and an average test person's (red curve) performance (measured as the number of matching responses found in the EAT) depending on the syntactic ambiguity classes.

5.3 Semantic ambiguity classes

Our objective when performing this experiment was to explore the effect which semantic word ambiguity has on the generation of word associations. It is motivated by the observation that many words have more than one meaning. As described in the previous section, from our list of 5910 test items we discarded those which were not covered in WordNet. For each of the remaining 5522 stimulus words we retrieved its number of senses from WordNet, thereby not distinguishing between senses as nouns, verbs, or adjectives (i.e. we accumulated the respective numbers). This led to the following semantic ambiguity classes (SA):

SA1: Words with only 1 sense in WN. 996 – 18.03%

SA2: Words with 2 to 5 senses in WN. 2910 – 52.69%
 SA3: Words with 6 to 10 senses in WN. 1068 – 19.34%
 SA4: Words with 11 to 20 senses in WN. 447 – 8.09%
 SA5: Words with 21 to 40 senses in WN. 84 – 1.52%
 SA6: Words with more than 40 senses in WN. 17 – 0.30%

These are some sample words for each class:

SA1: accordion, beware, ledge, stealthy, wrist
 SA2: agony, childhood, extent, tract, yawn
 SA3: accent, diamond, mix, slump, yellow
 SA4: bull, march, present, stone, wrong
 SA5: check, dip, follow, open, post
 SA6: call, draw, light, set, take

The results obtained using this classification can be seen in Figure 4. Apparently, semantic ambiguity has a strong effect on word associations. The accuracy significantly decreases when the number of meanings increases, e.g. there is only little agreement between test persons when dealing with words of more than twenty senses. This general tendency can be expected as the more ambiguous a word is, the more responses it can trigger. In the simulation this tendency is clearly confirmed. Moreover, although we do not have a good explanation for the little bump in the curve for ambiguity class 3, it is surprising that the simulation even confirms this bump.¹²

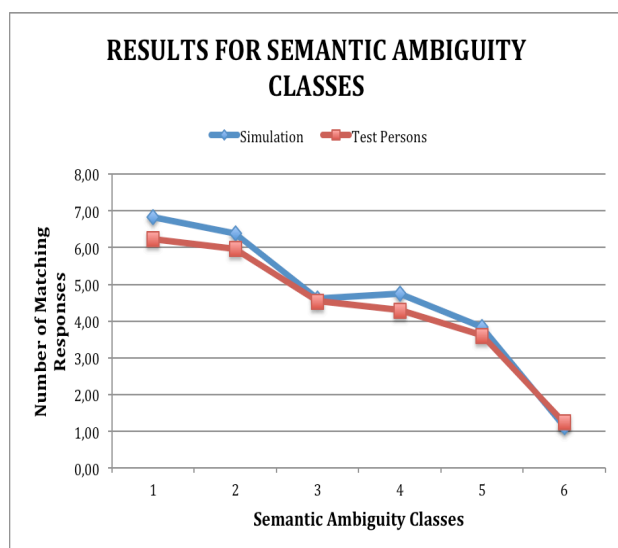


Fig. 4: Quality of our system's (blue curve) and an average test person's (red curve) performance (measured as the number of matching responses found in the EAT) depending on the semantic ambiguity classes.

6. Conclusions

We have presented a graph-based algorithm for the computation of word associations. The results were evaluated with a large test set comprising all nouns, verbs, and adjectives occurring in the Edinburgh Associative The-

saurus.

Contrary to what could be expected our system predicts not only syntagmatic but also paradigmatic relations. For instance, *bitter-taste* and *man-woman* are stimulus/ response pairs which were correctly computed. This shows that in texts not only word pairs in syntagmatic relation co-occur, but also word pairs having a paradigmatic relation. The results indicate that statistical co-occurrence-based methods are suitable for tasks that traditionally were supposed to require more sophisticated symbolic approaches.

Thanks to the considerable size of our test set, as a novel feature of our evaluation we were able to distinguish six saliency classes of words and found that for stimulus words where human subjects show a high degree of uniformity our system produces considerably better results than for stimulus words where the human responses show a lot of variance. This behavior almost exactly matches the one of an average test subject from the EAT. Similarly encouraging results were obtained for classes based on syntactic and semantic word ambiguity.

To conclude, not only is it possible to predict thousands of associations correctly. It could also be shown that the predictions for salient words are much better than for non-salient words, and that for unambiguous words they are better than for ambiguous words. All this to a degree which matches the behavior of human subjects. In conclusion, our results provide evidence that human associative behavior as observed in the association experiment is governed by the observed co-occurrences of words in perceived language.

Acknowledgments

This research was supported by the Marie Curie Intra European Fellowships DynNetLAc and AutoWordNet within the 7th European Community Framework Programme

References

- Aitchison, J. (2003). *Words in the Mind: an Introduction to the Mental Lexicon* (3d edition). Blackwell, Oxford.
- Aristotle, S. (350 avant JC). *De memoria et reminscientia*. In: Parva Naturalia, Vrin.
- Bel Enguix, G., Rapp, R. and Zock, M. (2014). How well can a corpus-derived co-occurrence network simulate human associative behavior? *Proceedings of the EACL 2014 Workshop on Cognitive Aspects of Computational Language Learning (CogACLl)*, Gothenburg, Sweden.
- Bullmore, E. und O. Sporns., O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* 10 (3):186–198 (2009)
- Burnard, L. and Aston, G. (1998). *The BNC Handbook: Exploring the British National Corpus*. Edinburgh: Edinburgh University Press.
- Bush, V. (1945). As we may think. *The Atlantic Monthly*, 176:101–108.
- Chomsky, N. (1959). A review of B. F. Skinner's Verbal Behavior. *Language*, 31(1), 26–58.
- Church, K.W. and Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Compu-*

¹² Let us nevertheless speculate: This bump could indicate that the words in ambiguity class 3 should in reality have more senses, i.e. it could be an indication for a systematic problem of WordNet concerning the granularity of the senses.

- tational Linguistics* 16 (1), 22–29.
- Collins, A. M. and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review* 8.
- Collins, A. and M.R. Quillian (1969). *Retrieval time from semantic memory*. *Journal of verbal learning and verbal behavior* 8 (2): 240–248.
- Cramer, P. (1968). *Word association*. Academic Press, New York.
- Deese, J. 1965. *The structure of associations in language and thought*. Johns Hopkins Press.
- Dell, G., Chang, F. and Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, 23: 517–542.
- Ekpo-Ufot, A. (1978). Word associations: a comparative study among college students in Nigeria and the United States. *Journal of Cross-Cultural Psychology*, Vol. 9, No. 4, December 1978, 455-468.
- Evert, S. and Krenn, B. (2001) Methods for qualitative evaluation of lexical association measures. In *Proceedings of the 39th Annual Meeting of the Association of Computational Linguistics*, Toulouse, France, 188-915.
- Fellbaum, C. (ed.) (1998). *WordNet: An Electronic Lexical Database and some of its Applications*. MIT Press.
- Freud, S. (1901). *Psychopathology of everyday life*. Payot, Paris, 1997 edition.
- Galton, F. (1879). Psychometric experiments. *Brain* (2), 149–162.
- Hoang, H., Kim, S. N. and Kan, M.Y. (2009). A re-examination of lexical association measures. *Proceedings of the Workshop on Multiword Expressions, ACL-IJCNLP 2009*, Suntec, Singapore, 31–39.
- Hörmann, H. (1972). *Introduction à la psycholinguistique*. Larousse, Paris, France.
- Hume, D. (1739): *A Treatise of Human Nature*, volumes; 1-2. London, John Noon. <http://www.rbjones.com/rbjpub/philos/classics/hume/index.htm>
- Istifci, I. (2010). Playing with words: a study on word association responses. *The Journal of International Social Research*, Volume 3 / Winter 2010.
- Jenkins, J. (1970). The 1952 Minnesota word association norms. In L. Postman and G. Kepper, editors, *Norms of Word Association*. Academic Press, New York, 1–38.
- Jung, C. and F. Riklin. (1906). Experimentelle Untersuchungen über Assoziationen Gesunder. In: C. G. Jung (ed.): *Diagnostische Assoziationsstudien*. Barth, Leipzig, Germany, 7–145.
- Kent, G.H. and Rosanoff, A.J. (1910). A study of association in insanity. *American Journal of Insanity*, 67, 37–96, 317–390.
- Kiss, G. R. (1975). An associative thesaurus of English: Structural analysis of a large relevance network. In: A. Kennedy, A. Wilkes (eds.): *Studies in Long Term Memory*. London: Wiley, 103–121.
- Kiss, G.R., Armstrong, C., Milroy, R., and Piper, J. (1973). An associative thesaurus of English and its computer analysis. In: A. Aitken, R. Beiley, N. Hamilton-Smith (eds.): *The Computer and Literary Studies*. Edinburgh University Press.
- Locke, J. (1689). *An Essay Concerning Human Understanding*. London, T. Basset. www.rbjones.com/rbjpub/philos/classics/locke/cib2c33.htm
- Meara, P. (2009). *Connected Words: Word Association and Second Language Vocabulary Acquisition*. John Benjamins, Amsterdam, Holland.
- Michelbacher, L., Evert, S. and Schütze, H. (2011). Asymmetry in corpus-derived and human associations. *Corpus Linguistics and Linguistic Theory*, Vo. 7, No. 2, 245–276.
- Mihalcea, R. and Radev, D. (2011). *Graph-Based Natural Language Processing and Information Retrieval*. Cambridge University Press, Cambridge, UK.
- Miller, G. (1990). Wordnet: An on-line lexical database. *International Journal of Lexicography*, 3(4).
- Miller, G., Beckwith, R., Fellbaum, C., Gross, D. and K. Miller. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography*, 3(4), pages 235–244.
- Mills, J.S.(1904): *Nature, The Utility of Religion and Theism*. London, Rationalist Press. www.marxists.org/reference/archive/mill-john-stuart/1874/nature.htm.
- Nelson, D., McEvoy, C. and T.A. Schreiber. (1998). The university of South Florida word association, rhyme, and word fragment norms. <http://www.usf.edu/FreeAssociation>
- Quillian, R. (1968). Semantic memory. In M. Minsky, (ed.): *Semantic Information Processing*. MIT Press, Cambridge, MA, 216–270.
- Pecina, P. and Schlesinger, P. (2006). Combining association measures for collocation extraction. *Proceedings of the 21th International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (COLING/ACL 2006)*, Sydney, Australia, 651–658.
- Rosenzweig, M. R. (1961). Comparisons among word-association responses in English, French, German, and Italian. *The American Journal of Psychology*, Vol. 74, No. 3, 347–360.
- Rumelhart, D.E., McClelland, J. and the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*, Cambridge, MA: MIT Press.
- Schvaneveldt, R. editor. (1989). *Pathfinder Associative Networks: studies in knowledge organization*. Ablex, Norwood, New Jersey, US.
- Schwartz, B. and Reisberg, D. (1991). *Learning and Memory*. New York: Norton.
- Seidensticker, P. (2006). *Simulation von Wortassoziationen mit Hilfe von mathematischen Lernmodellen in der Psychologie*. Dissertation an der Universität

Paderborn.

- Spitzer, M. (1999). *The mind within the net: models of learning, thinking and acting*. MIT Press, Cambridge, MA.
- Sporns, O. (2012) *Discovering the Human Connectome*. MIT Press, Cambridge.
- Sporns, O. (2011) *Networks of the Brain*. MIT Press, Cambridge.
- Stam, C. and Reijneveld, J. (2007). Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomedical Physics*, 1:3
- Tamir, R. (2005). A Random Walk through Human Associations. *Proceedings of ICDM 2005*: 442–449.
- Vitevitch, M. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language, and Hearing Research*, 51:408–422.
- Washtell, J. and Markert, K. (2009). A comparison of windowless and window-based computational association measures as predictors of syntagmatic human associations. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing (EMNLP '09)*, Volume 2, 628–637.
- Wettler, M. (1980). *Sprache, Gedächtnis, Verstehen*. Berlin, de Gruyter.
- Wettler, M. and Rapp, R. (1989). A connectionist system to simulate lexical decisions in information retrieval. In: R. Pfeifer, Z. Schreter, F. Fogelman, L. Steels (eds.): *Connectionism in Perspective*. Amsterdam: Elsevier, 463–469.
- Wettler, M., Rapp, R. and Sedlmeier, P. (2005). Free word associations correspond to contiguities between words in texts. *Journal of Quantitative Linguistics* 12(2), 111–122.
- Zock, M., Ferret, O. and D. Schwab, (2010). Deliberate word access: an intuition, a roadmap and some preliminary empirical results. *International Journal of Speech Technology*, 13(4), 107-117. Heidelberg: Springer.