

Corpus-Based Computation of Reverse Associations

Reinhard Rapp

Aix-Marseille Université, Laboratoire d'Informatique Fondamentale
163 Avenue de Luminy, 13288 Marseille, France
reinhardrapp@gmx.de

Abstract

According to psychological learning theory an important principle governing language acquisition is co-occurrence. For example, when we perceive language, our brain seems to unconsciously analyze and store the co-occurrence patterns of the words. And during language production, these co-occurrence patterns are reproduced. The applicability of this principle is particularly obvious in the case of word associations. There is evidence that the associative responses people typically come up with upon presentation of a stimulus word are often words which frequently co-occur with it. It is thus possible to predict a response by looking at co-occurrence data. The work presented here is along these lines. However, it differs from most previous work in that it investigates the direction from the response to the stimulus rather than vice-versa, and that it also deals with the case when several responses are known. Our results indicate that it is possible to predict a stimulus word from its responses, and that it helps if several responses are given.

Keywords: word associations, reverse associations, associative learning, multiword association

1. Introduction

Free word associations are the words human subjects spontaneously come up with upon presentation of stimulus words. For example, on presentation of *black* subjects typically respond with *white*, and on presentation of *funny* they respond with *laugh*. Psychological learning theory has already more than a century ago hypothesized that associations are learned by memorizing the contiguities of perceived objects (James, 1890), and later work, such as Schvaneveldt et al. (1989), Wettler & Rapp (1989), and Church & Hanks (1990) showed that contiguity learning also applies if words are the objects of study.

With the advent of corpus linguistics people started to compare word co-occurrences as observed in large text corpora to the word associations as produced by human subjects in the word association experiment (Rapp & Wettler, 1991). The findings confirmed the theory: The simulated (i.e. corpus-based) associations were almost indistinguishable from the associations produced by humans (Wettler et al., 2005; Tamir, 2005). In particular, quantitative evaluations showed that they lie well within the natural range of variation as observed for human associations.

Given this success, people started to apply the same methodology in the case when several stimulus words are known. For example, given the stimulus words *King* and *daughter*, subjects would frequently come up with *princess*, or given *circus* and *laugh* they would come up with *clown*. This behavior could also be replicated to some extent, but the results were not as convincing as with single stimulus words. The respective work often comes under headwords such as *multiword associations* (e.g. Rapp, 2008) or (in psychology) the *remote association test*. A recent notable publication on the remote association test which gives pointers to other related work, is Smith et al. (2013) who apply this for problems that require consideration of multiple constraints, such as choosing a job based on salary, location, and work des-

cription. Another one is Griffiths et al. (2007) who assume that concept retrieval from memory can be facilitated by inferring the gist of a sentence, and using it to predict related concepts and disambiguate words. They implement this by using a topic model.

A problem when trying to simulate multiword associations is that the human responses upon multiword stimuli show a much higher variation than those on single stimulus words. This makes the overlap between simulated and human associations so low that quantitative evaluation measures would only be reliable if large amounts of human data (i.e. multiword-stimuli together with their responses) could be considered. However, it is difficult to collect such data from human subjects, especially as the subjects describe the task of multiword association to be considerably more demanding than the task of single word association, making it more likely that it is not conducted properly. Using recent developments such as *crowdsourcing* or *games with a purpose* it should nevertheless be possible to collect large quantities of multiword associations. However, to our knowledge this has not happened yet. As a consequence, the gold standards available for optimizing the algorithms were rather small (see e.g. Rapp, 2008) so that their limited reliability presumably lead to suboptimal results.

In the current paper we suggest to bypass this problem by focusing on a related but different and particularly well defined task where data is comparatively plentiful, namely the *reverse association* task. With this we mean the production of a stimulus word when given its responses. For example, given the stimulus word *cold* the top five most frequent responses as produced by test persons were *hot*, *ice*, *warm*, *water*, and *freeze*. We now simply reverse the task, i.e. we assume that the five responses are given, and try to compute the stimulus. Doing so has the advantage that we can use any of the previously collected association norms as a gold standard. For example, the Edinburgh Associative Thesaurus alone will give us 8400 test items, which is about two orders of magnitude larger than e.g.

the multiword association test sets used in Griffiths et al. (2007) and in Rapp (2008).

Let us mention that the proposed reversal could lead to terminological confusion for the following reason: The words which were the stimulus words in the human experiments become the associative responses in the reverse association task. And the words which were the associative responses become the stimulus words. To minimize this confusion, in the remainder of this paper when we talk about the reverse association task we will not use the terms *stimulus word* and *associative response*, but instead the terms *given word* and *target word*.

2. Approach

We use a vector space approach as described in our previous work (Rapp, 2008; Rapp, 2013). It involves computing a co-occurrence matrix of the words found in a text corpus and applying a standard association measure (in our case the log-likelihood ratio, see Dunning, 1993) to the co-occurrence counts. In the resulting association matrix, the strongest association to a given word can be retrieved by locating in its association vector the highest value. The word relating to this value is considered to be the strongest associative response to the given word.

In Rapp (2008) we extended this algorithm to multiword associations, i.e. to the case when several stimulus words are given and the aim is to compute their common associative response. In Rapp (2013) we applied this approach to the reverse association task. Let us briefly review the core findings of these studies.

Rapp (2013) assumes that reverse associations can be computed using exactly the same algorithm as suitable for multiword associations. The results will usually be even better as the task is somewhat easier. The reason for the relative simplicity is that in the reverse association task all given words point to the same target word, and often nicely disambiguate each other. For example, for the stimulus word *palm* the EAT lists associative responses such as *hand* and *finger*, but also *tree* and *oil*. The former two relate to the *body part* sense of *palm*, the latter to its *plant* sense. But when we reverse the task, the four given words nicely point to *palm* as probably there are not many other English words with senses relating to both plants and body parts.

In Rapp (2008) we had described a simple yet effective algorithm for computing multiword associations, which, as argued above, is also suitable for computing reverse associations. The underlying assumption was that a target word must have strong connections to all given words, and that strong connections to only some of them do not suffice. Such a behaviour can be put into practice using a multiplication.

However, we do not multiply the associative weights as our association measure of choice, namely the log-likelihood ratio, has an inappropriate (exponential) value characteristic. This value characteristic has the effect that a weak association to one of the stimuli can easily be overcompensated by a strong association to another stimulus, which is not desirable. Instead of multiplying the asso-

ciation strengths, we therefore multiply their ranks. This improves the results significantly.

Such considerations lead us to the following basic procedure: Given an association matrix of vocabulary V containing the association strengths (log-likelihood ratios) between all possible pairs of words, to determine the target word triggered by the given words a, b, c, \dots the following steps are conducted:

- 1) For each word in V look up the ranks of the words a, b, c, \dots in its association vector, and compute the product of these ranks.
- 2) Sort the words in V according to these products, with the sort order such that the lowest value obtains the top rank (i.e. conduct a reverse sort).

Note that this procedure, which we call the *product-of-ranks algorithm*, is computationally somewhat demanding as these computations are required for each word in a possibly very large vocabulary.¹ On the plus side, the procedure is in principle applicable to any number of stimulus words,² and when we have more of them there is only a slight increase in the computational load.

3. Resources

3.1 Corpora

As our corpora used for extracting word co-occurrence information we use the *British National Corpus* (Burnard & Aston, 1998), the *ukWaC* corpus (Ferraresi et al., 2008), and the *deWaC* corpus (Baroni et al., 2009). The former is a balanced sample of current day British English comprising about 100 million words. The latter two are large collections of texts downloaded from the web with each comprising in the order of two billion words. Hereby, as the names suggest, *ukWaC* contains British English texts and *deWaC* German texts.

From these corpora the function words were removed using lists of English and German stopwords. Then the corpora were lemmatized by applying the lookup procedure described in Rapp (1999) which utilizes large lists of inflected word forms paired with their lemmas. This procedure does not consider the context of a word. Therefore word forms with several possible lemmas (a case which is not very common in English and German) remained unchanged. Although this limitation is usually a drawback, in our setting it has the advantage that the same procedure is also applicable for isolated words without context. We could therefore apply it in the same way to the association norms described below.

3.2 Association norms

As our source of human reference data we use the *Edinburgh Associative Thesaurus* (EAT; Kiss et al. 1973) which is the largest classical collection of its kind. The EAT comprises about 100 associative responses as re-

¹ For this reason in the experimental part of this paper we always somewhat restrict the vocabulary to be considered.

² In section 4 we present results for up to 30 stimulus words.

requested from British students for each of altogether 8400 stimulus words. As some of these stimulus words are multiword units or function words, both of which we did not want to include here, we removed these from the associative thesaurus. Also, the above mentioned lemmatization led in a few cases to the duplication of existing stimulus words (e.g. if a plural form is reduced to a singular form and this singular form is already covered). Altogether, this led to a reduction of the number of items in the EAT from 8400 to 7918. This is the subset to be used in section 4.2.

A type of resource similar to the EAT are the Minnesota word association norms (Russel & Jenkins, 1954; Jenkins, 1970) which we use in another experiment (see section 4.3). Although they comprise only the 100 standard stimulus words suggested by Kent & Rosanoff (1910), they have the advantage that the same type of association experiment was also conducted for German, thereby presenting the German test persons translations of the English stimulus words (Russell & Meseck, 1959; Russell, 1970). This German data we also use in an experiment.

4. Results

4.1 Previous results

As a baseline for comparisons, let us recapitulate some previous results from Rapp (2013). These were obtained using a pre-processed version of the British National Corpus as described above. For counting the word co-occurrences a window size of plus and minus two words from a given word was considered. The vocabulary used for the rows and columns of the co-occurrence matrix were all 34,324 words which in the lemmatized BNC had a corpus frequency of 100 or higher.

To give a first impression, Table 1 shows some sample results (Rapp, 2013). For example, the EAT lists *apple* and *juice* as the top responses when given the stimulus word *fruit*, but our algorithm, when provided with *apple* and *juice*, computes that *orange* would be the best target word. This is not as expected, but also has some plausibility. The expected target word *fruit* shows up on the 8th position of the computed list of words.

For a quantitative evaluation, we only looked at a subset of the EAT. This comprises the 100 stimulus words as suggested by Kent & Rosanoff (1910), together with their responses as taken from the EAT. Analogous to the BNC, the EAT data was also lemmatized using the same procedure. For the Kent & Rosanoff subset, we counted in how many cases the expected target word is ranked first in the list of computed words. This leads to conservative numbers as only exact matches are considered as correct. For example, the last item in Table 1, where *whisky* instead of *whiskey* is on rank 1, would count as wrong.

Figure 1 shows the percentage of correctly computed target words depending on the number of given words (i.e. associative responses from the EAT) that are taken into account. As can be seen, the quality of the results improves up to seven given words where it reaches 54% accuracy, and from then on degrades. This means that, on

average, already the eighth response word is not helpful for determining the respective stimulus word.

TOP 2 RESPONSES FROM EAT: apple (1385), juice (1613)
STIMULUS WORD FROM EAT: fruit (3978)
COMPUTED TARGET WORDS: orange (2333), grape (273), lemon (1019), lime (612), pineapple (220), grated (423), apples (792), fruit (3978), grapefruit (113), carrot (359)
TOP 3 RESPONSES FROM EAT: water (33449), tub (332), clean (6599)
STIMULUS WORD FROM EAT: bath (415)
COMPUTED TARGET WORDS: rinsed (177), bath (2819), soak (315), rinse (288), wash (2449), refill (138), rainwater (160), polluted (393), towels (421), sanitation (156)
TOP 4 RESPONSES FROM EAT: grass (4295), blue (9986), red (13528), yellow (4432)
STIMULUS WORD FROM EAT: green (10606)
COMPUTED TARGET WORDS: green (10606), jersey (359), ochre (124), bright (5313), pale (3583), violet (396), purple (1262), greenish (136), stripe (191), veined (103)
TOP 5 RESPONSES FROM EAT: drink (7894), gin (507), bottle (4299), soda (356), Scotch (621)
STIMULUS WORD FROM EAT: whiskey (129)
COMPUTED TARGET WORDS: whiskey (1451), whiskey (129), tonic (511), vodka (303), brandy (848), Whisky (276), scotch (151), lemonade (229), poured (1793), gulp (196)

Table 1: Top ten computed target words for various numbers of given words. Numbers in brackets refer to the respective words' corpus frequencies in the BNC.

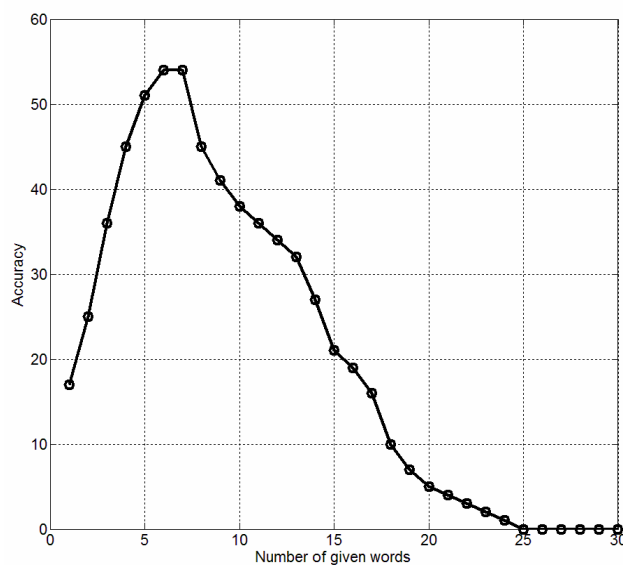


Fig. 1: Percentage of correctly predicted target words depending on the number of given words. The maximum accuracy of 54% is achieved for six and likewise for seven given words.

Let us mention a detail: For the case of single stimulus words (leftmost point in curve) we did not use the product-of-ranks algorithm. Instead simply the word showing the highest log-likelihood score in conjunction with the given word is considered as the associative response of the system. The reason is that with the product-of-ranks algorithm for a single given word it is relatively common that for several responses this word ends up on the same rank,³ in which case the system would make an arbitrary and thus suboptimal decision. See Rapp (2013) for details. The same consideration also applies to the results shown in Figures 2 to 4.

4.2 Results for a large EAT-derived dataset

Given that the previous results shown in Figure 1 reflect only a 100 word subset of the EAT, in the following we provide analogous results for almost the full EAT as described in section 3.2.

Like in section 4.1 the co-occurrences were counted using the lemmatized BNC with function words removed. The window size was again plus and minus 2, and the association measure of choice was the log-likelihood ratio. The only difference was the vocabulary used. Due to the large data set, for reasons of time efficiency we had to somewhat reduce the vocabulary.

Actually we used two vocabularies: One for the rows of the co-occurrence matrix and one for the columns. For the rows we used all 7918 stimulus words occurring in our large EAT subset. For the columns we used all 19,973 unique words appearing in the lemmatized EAT as responses to the 7918 stimulus words.

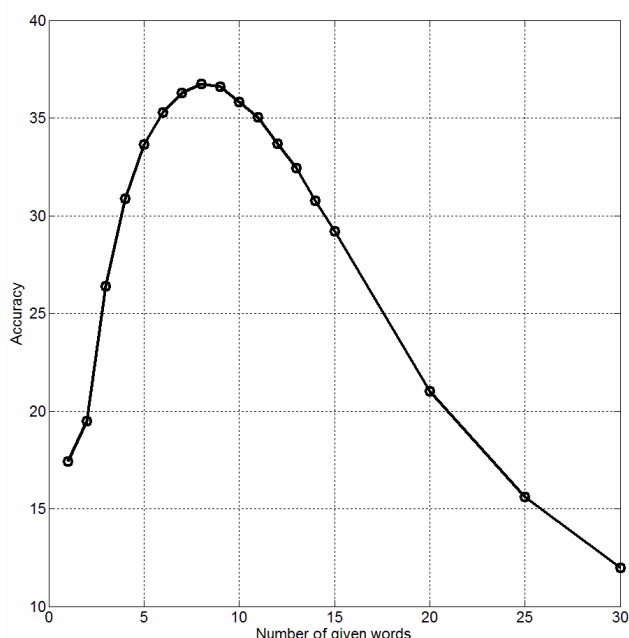


Fig. 2: Percentage of correctly predicted target words depending on the number of given words. The maximum accuracy of 36.7% is achieved for eight given words.

³ For example, the given word *white* might end up on rank 1 for two potential target words, namely *black* and *snow*.

The limited vocabularies mean that the task of the product-of-rank algorithm is slightly facilitated. When computing a target word, rather than choosing from (potentially) all words of the English vocabulary, it had to choose from only the 7918 word subset. The respective results are shown in Figure 2.

4.3 Results for other data

In order to see in how far the quality of the results is affected when using different data, we did a further experiment. This time, instead of the BNC we used the ukWaC corpus, and instead of the EAT we used the Minnesota association norms, as described in section 3. Except for the vocabulary, all computational parameters (lemmatization, window size, association measure) remained the same. As our vocabulary, to be applied for both rows and columns of our co-occurrence matrix, we used all 3884 words occurring in the lemmatized Minnesota association norms. Figure 3 shows the results.

Although the Minnesota norms provide only data for the 100 Kent & Rosanoff (1910) stimulus words, their total vocabulary (i.e. including stimuli and responses) is relatively large as they are based on an about ten times larger number of test persons per stimulus word than in the case of the EAT. Whereas for the EAT the responses from 100 test persons were collected for each stimulus word, for the Minnesota norms the number of test persons amounted to 1008 per stimulus word.

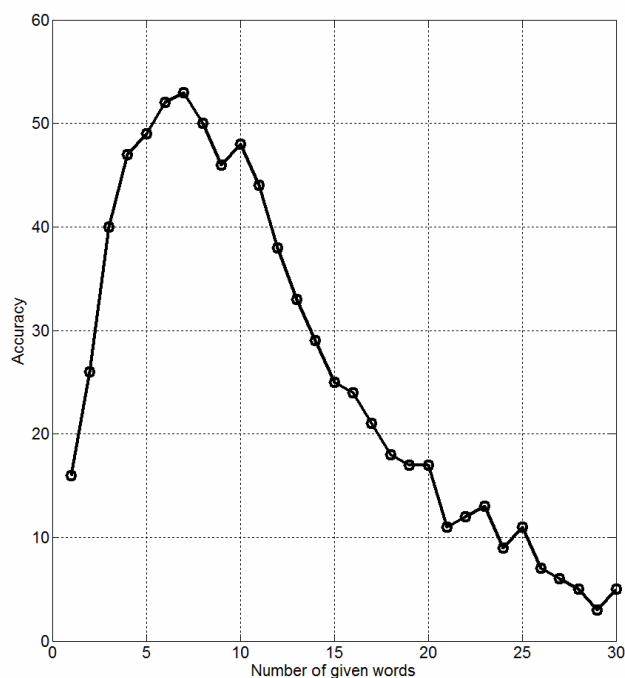


Fig. 3: Percentage of correctly predicted target words depending on the number of given words. The maximum accuracy of 53% is achieved for seven given words.

4.4 Results for another language

The theory of associative learning underlying this work should in principle be applicable to all languages. To investigate this, we conducted an experiment as similar as possible to the one described in the previous section but

for German. There we had already chosen our language resources in such a way that similar resources are available for German. In the case of the Minnesota norms, the German counterpart are the norms published by Russell & Meseck (1959) where the 100 stimulus words used are German translations of the stimulus words used in the Minnesota norms. A difference is that in the German experiment the associative responses were collected from 331 rather than 1008 students.

Concerning the corpus used, the German counterpart of the ukWaC corpus is the deWaC corpus which is of similar size (see section 3.1). Note that for the BNC (used in section 4.1) no German counterpart exists, which is why in section 4.3 we went for the ukWaC corpus. We also had a lemmatization procedure similar to the one for English available for German (Rapp, 1999) which we applied on the deWaC corpus as well as on the German association norms.

All computational parameters were chosen in analogy to section 4.3. That is, the window size (plus an minus two words) and the association measure (log-likelihood ratio) remained the same, and as the vocabulary we used all 4977 words occurring in the lemmatized German association norms for both the rows and the columns of our matrix.

The outcome of the respective experiment is shown in Figure 4.

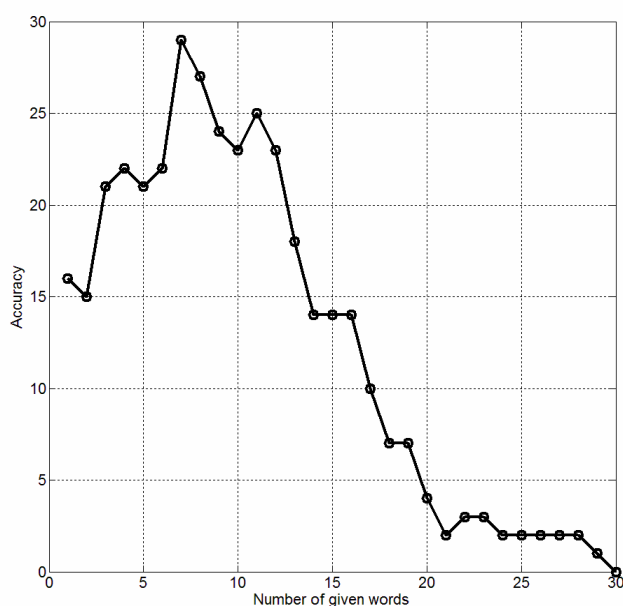


Fig. 4: Percentage of correctly predicted target words depending on the number of given words. The maximum accuracy of 29% is achieved for seven given words.

5. Discussion and conclusions

When predicting the target words in the reverse association task, an interesting question is in how far the quality of the results depends on the number of given words. To answer this question, we conducted evaluations for various numbers of given words. Beforehand, we had three expectations:

- The more subjects have given a particular associative response, the more salient it is and the more precisely it should trigger the desired target word.
- In contrast, responses given by only one or very few subjects might be of arbitrary nature and therefore not helpful for predicting the target word.
- Considering a larger number of salient associative responses should improve the results.

The shapes of all curves (Fig. 1 to 4) show an accuracy maximum for a range between six and eight given words, with accuracies significantly decreasing for lower and for higher numbers. This provides some evidence that the above expectations might be correct.

With accuracies of 54% and 53%, respectively, the best results were achieved in the experiments described in sections 4.1 and 4.3. Hereby it may seem surprising that the results of section 4.1 are marginally better despite the fact that the corpus used in section 4.3 (ukWaC) is about 20 times larger than the one used in 4.1 (BNC). However, it should be mentioned that the BNC is better balanced. Also, the Minnesota norms used in section 4.3 reflect American language use, which might be harder to predict on the basis of a British language corpus.

Another similarity between the outcome of the two experiments is that, although they are based on different corpora and different association norms, the shapes of the curves in Figures 1 and 3 show good agreement. In both cases the maximum accuracy is more than three times higher when compared to the case with only a single given word.

An explanation for the rather good performance is that in the reverse association task typically all clues (given words) point to the same target word. On the other hand, the task seems even non-trivial for humans, and sometimes there are several plausible options how the given words might disambiguate each other. For example, given *apple* and *juice* (see Table 1), the solution our system came up with, namely *orange*, seems quite as plausible as the expected solution *fruit*. However, in our evaluation *orange* is counted as wrong, and this is true for many others of the incorrect results.

With a maximum of 36.7%, the accuracies resulting from the large set of 7918 EAT items (see section 3.2) are generally somewhat lower than the ones reported in sections 4.1 and 4.3 for the 100 items from the Kent & Rosanoff (1910) list of words. This discrepancy can be explained by the fact that the Kent & Rosanoff (1910) words are mostly easy and frequent words which are likely to be well covered in any text corpus. In comparison, the EAT contains a higher proportion of rare and difficult words.

Concerning the application of the approach to another language, namely German, the overall shape of the curve turned out to be similar again. But it is somewhat erratic which might be explained by the observation that in German the problem of data sparsity is likely to be more severe than it is in English due to its extensive compounding and the higher number of inflectional variants. The maximum of the curve is also obtained for seven given words,

but with a value of 29% the accuracy is at a considerably lower level than the comparable English results from section 4.3 (53%). For explanation, it should be noted that, as described in Rapp (1996), the associations collected from English subjects are about twice as homogeneous than those from German subjects, so the discrepancy in accuracies might simply reflect this observation.

In conclusion, although (in comparison to related work) our algorithm does not require sophisticated processing involving e.g. *Latent Semantic Analysis* or *Topic Modeling*, its results seem rather good. For example, Griffiths et al. (2007) reported 11.54% correctly predicted target words. Likewise, our own previous work (Rapp, 2008), despite presenting a number of evaluations using various corpora and data sets, achieved accuracies which were all below 10%. The paper by Smith et al. (2013) which was mentioned in the introduction, does not give quantitative results at all.

Concerning applications, we see a number of possibilities: One is the tip-of-the-tongue problem, where a person cannot recall a particular word but can nevertheless think of some of its properties and associations. Another application is in information retrieval where the system can help to sensibly expand a given list of search words, and with the expanded list in turn used to conduct a better search. As a further application, the system could be used in multiword semantics to measure in how far the components of a multiword unit can predict each other. And finally, if in the context of natural language understanding in an utterance a word is missing or uncertain, we could try to predict this word by considering all other content words in the utterance (or within a somewhat wider context) as multiword input for our algorithm.

Acknowledgment

This research was supported by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme.

References

- Baroni, M.; Bernardini, S.; Ferraresi, A.; Zanchetta, E. (2009). The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation* 43(3): 209–226
- Burnard, L.; Aston, G. (1998): *The BNC Handbook: Exploring the British National Corpus with Sara*. Edinburgh: University Press.
- Church, K.W.; Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics* 16 (1), 22–29.
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19 (1), 61–74.
- Ferraresi, A.; Zanchetta, E.; Baroni, M.; Bernardini, S. (2008). Introducing and evaluating ukWaC, a very large web-derived corpus of English. In S. Evert, A. Kilgarriff and S. Sharoff (eds.) *Proceedings of the 4th Web as Corpus Workshop (WAC-4) – Can we beat Google?*, Marrakech.
- Griffiths, Thomas L.; Steyvers, Mark, Tenenbaum, Joshua B. (2007). Topics in semantic representation. *Psychological Review*, Vol. 114, No. 2, 211–244.
- James, W. (1890). *The Principles of Psychology*. New York: Holt. Reprinted New York: Dover Publications, 1950.
- Jenkins, J.J. (1970). The 1952 Minnesota word association norms. In: L. Postman; G. Keppel (eds.): *Norms of Word Association*. New York: Academic Press, 1–38.
- Kent, G.H.; Rosanoff, A.J. (1910). A study of association in insanity. *American Journal of Insanity*, 67, 37–96, 317–390.
- Kiss, G.R., Armstrong, C., Milroy, R., and Piper, J. (1973). An associative thesaurus of English and its computer analysis. In Aitken, A.J., Bailey, R.W. and Hamilton-Smith, N. (Eds.): *The Computer and Literary Studies*. Edinburgh: University Press, 153–65.
- Rapp, R. (1996). *Die Berechnung von Assoziationen*. Hildesheim: Olms.
- Rapp, R. (1999). Automatic identification of word translations from unrelated English and German corpora. In: *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics 1999*, College Park, Maryland. 519–526.
- Rapp, R. (2008). The computation of associative responses to multiword stimuli. *Proceedings of the Workshop on Cognitive Aspects of the Lexicon (COG-ALEX at Coling 2008, Manchester)*. 102–109.
- Rapp, R. (2013): From stimulus to associations and back. *Proceedings of the 10th Workshop on Natural Language Processing and Cognitive Science*, Marseille, France.
- Rapp, R., Wettler, M. (1991). Prediction of free word associations based on Hebbian learning. In: *Proceedings of the (IEEE and INNS) International Joint Conference on Neural Networks*, Singapore, Vol. 1, 25–29.
- Russell, W.A. (1970). The complete German language norms for responses to 100 words from the Kent-Rosanoff word association test. In: L. Postman, G. Keppel (eds.): *Norms of Word Association*. New York: Academic Press, 53–94.
- Russell, W.A.; Jenkins, J.J. (1954). *The Complete Minnesota Norms for Responses to 100 Words from the Kent-Rosanoff Word Association Test*. Technical Report No. 11, University of Minnesota, Minneapolis.
- Russell, W.A.; Meseck, O.R. (1959). Der Einfluß der Assoziation auf das Erinnern von Worten in der deutschen, französischen und englischen Sprache. *Zeitschrift für experimentelle und angewandte Psychologie*, 6, 191–211.
- Schvaneveldt, R. W., Durso, F. T., & Dearholt, D. W. (1989). Network structures in proximity data. In G. Bower (ed.): *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 24. New York: Academic Press, 249–284.

- Smith, Kevin A.; Huber, David E.; Vul, Edward (2013). Multiply-constrained semantic search in the Remote Associates Test. *Cognition* 128, 64–75.
- Tamir, R. (2005). A Random Walk through Human Associations. *Proceedings of the Fifth IEEE International Conference on Data Mining (ICDM'05)*, 442–449.
- Turney, P.T.; Pantel, P. (2010). From frequency to meaning: vector space models of semantics. *Journal of Artificial Intelligence Research* 37, 141–188.
- Wettler, M., 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.; Sedlmeier, P. (2005). Free word associations correspond to contiguities between words in texts. *Journal of Quantitative Linguistics* 12(2), 111–122.