A Computational Model of Children's Semantic Memory

Guy Denhière (denhiere@up.univ-mrs.fr)
L.P.C & C.N.R.S. Université de Provence
Case 66, 3 place Victor Hugo
13331 Marseille Cedex, France

Benoît Lemaire (Benoit.Lemaire@upmf-grenoble.fr)
L.S.E., University of Grenoble 2, BP 47
38040 Grenoble Cedex 9, France

Abstract

A computational model of children's semantic memory is built from the Latent Semantic Analysis (LSA) of a multisource child corpus. Three tests of the model are described, simulating a vocabulary test, an association test and a recall task. For each one, results from experiments with children are presented and compared to the model data. Adequacy is correct, which means that this simulation of children's semantic memory can be used to simulate a variety of children's cognitive processes.

Introduction

Models of human language processing are usually based on a layer of basic semantic representations on top of which cognitive processes are described. For instance, the construction-integration model (Kintsch, 1998) describes processes that operate on a network of propositions. These basic representations can just be descriptions of what the human memory looks like, in order for the upper models to be explicitly stated, but they can also be operationalized so that the model can be tested on a computer. In the first case, these representations are usually designed by hand, but this method prevents large-scale simulations.

This was the case with Kintsch's construction-integration model until 1998. Before that, researchers had to code propositions by hand and guess relevant values to code the strength of links between nodes. Then Kintsch (1998) used the Latent Semantic Analysis (LSA) model (Deerwester et al., 1990; Landauer et al., 1998) which provides a way to automatically build these basic representations. This was a major step since the construction-integration processes could then be tested on a large variety of inputs, while being less dependent on idiosyncratic codings. Such a mechanism for automatically constructing basic semantic representations should be carefully designed and tested in order to simulate as good as possible human semantic memory.

LSA is nowadays considered as a good candidate for modeling an adult semantic memory based on a large corpora of representative texts: Bellissens et al. (2002), Kintsch (2000) and Lemaire & Bianco (2003) used it for modeling metaphor comprehension; Pariollaud et al. (2002) used it for modeling the comprehension of idiomatic expressions; Howard & Kahana (2002) relied on it to model free recall and episodic memory retrieval; Laham (1997) did the same for modeling categorization processes; Landauer & Dumais (1997) designed a model of vocabulary acquisition based on LSA; Lemaire & Dessus (2001), Rehder et al. (1998) and Wolfe et al. (1998) used it for modeling knowledge assessment; Quesada et al. (2001) modeled complex problem solving by means of LSA basic representations; Wolfe & Goldman (2003) worked on a model of reasoning about historical accounts based on LSA. However, to our knowledge, no computational basic representations were made that mimic full children's semantic memory.

This paper aims at presenting such a model. First, we present LSA. We then describe our corpus, which is supposed to mimic the kind of texts children are exposed to. Finally, we present three experiments which aim at validating the model.

Latent Semantic Analysis

Basic semantic representations

There are many ways of constructing basic semantic representations that can be processed by a computer. The first one is to build them by hand. Powerful formalisms like description logic (Borgida, 1996) or semantic networks (Sowa, 1991) have been designed to accurately represent concepts, properties and relations. However, in spite of huge efforts (Lenat, 1995), no full set of symbolic representations has been made that can be considered a reasonable model of human semantic memory. Hand-coding semantic information is tedious and, as we mention later, symbolic representations might not be the best formalism for that.

Another strategy is to rely on corpora to get the semantic information. Artificial intelligence researchers have designed sophisticated syntactic processing tools for automatically describing the knowledge using the kind of symbolic formalisms mentioned earlier. They usually refer to them as ontologies or knowledge bases (Vossen, 2003). However, in spite of great strides, this approach still cannot be the means to form the basic semantic representations that cognitive researchers need. First, it cannot be fully automatized, except for specific domains, thus preventing complete descriptions of the language. Second and quite paradoxically, since the descriptions are quite elaborated, it is very hard to design reasoning processes on top of them. For instance, a simple process like estimating the degree of semantic association is very hard to operationalize on complex structures like semantic networks.

Instead of relying on symbolic representations, a third approach consists in (1) analyzing the co-occurrence of words in large corpora in order to draw semantic similarities and (2) relying on very simple structures, namely high-dimensional vectors, to represent meanings. In this approach, the unit is the word. The meaning of a word is not defined per se, but rather determined by its relationships with all others. For instance, instead of defining the meaning of bicycle in an absolute manner (by its properties, function, role, etc.), it is defined by its degree of association to other words (i.e., very close to bike, close to pedals, ride, wheel, but far from duck, eat, etc.). This semantic information can be established from raw texts, provided that enough input is available. This is exactly what human people do: it seems that most of the words we know, we learn by reading (Landauer & Dumais, 1997). The reason is that most words appear almost only in written form and that direct instruction seems to play a limited role. Therefore, we would learn the meaning of words mainly from raw texts, by mentally constructing their meaning through repeated exposure to appropriate contexts.

Relying on direct co-occurrence

One way to mimic this powerful mechanism would be to rely on direct co-occurrences within a given context unit. A usual unit is the paragraph which is both computationally easy to identify and of reasonable size. We would say that:

R1: words are similar if they occur in the same paragraphs.

Therefore, we would count the number of occurrences of each word in each paragraph. Suppose we use a 5,000-paragraph corpus. Each word would be represented by 5,000 values, that is by a 5,000 dimension vector. For instance:

avalanche: (0,1,0,0,0,0,1,0,2,0,0,0,0,0,0,1,1,0,1,0,1,0,0,0,0,0,0…)
snow: (0,2,0,0,0,0,0,0,1,1,0,0,0,0,0,0,2,1,1,0,1,0,0,0,0,0,0…)

This means that the word avalanche appears once in the 2^{nd} paragraph, once in 7^{th}, twice in the 9^{th}, etc. One could see that, given the previous rule, both words are quite similar: they co-occur quite often. A simple cosine between the two vectors can measure the degree of similarity. However, this rule does not work well (Perfetti, 1998; Landauer, 2002): two words should be considered similar even if they do not co-occur. French & Labiouse (2002) think that this rule might still work for synonyms because writers tend not to repeat words, but use synonyms instead. However, defining semantic similarity only from direct co-occurrence is probably a serious restriction.

Relying on higher-order co-occurrence

Therefore, another rule would be:

R1: words are similar if they occur in similar paragraphs.*

This is a much better rule. Consider the following two paragraphs:

Bicycling is a very pleasant sport. It helps keeping a good health.

For your fitness, you can practice bike. It is very nice and good to your body.

Bicycling and *bike* appear in similar paragraphs. If this is repeated over a large corpus, it would be reasonable to consider them similar, even if they never co-occur within the same paragraph. Now we need to define paragraph similarity. We could say that two paragraphs would be similar if they share words, but that would be restrictive: as illustrated in the previous example, two paragraphs should be considered similar although they do not have words in common (functional words are usually not taken into account). Therefore, the rule is:

R2: paragraphs are similar if they contain similar words.

Rules 1* and 2 constitute a circularity, but this can be solved by a specific mathematical procedure called singular value decomposition, which is applied to the occurrence matrix. This is exactly what LSA does. To state it in other words, LSA is not only based on direct co-occurrence, but rather on higher-order co-occurrence. Kontostahis & Pottenger (2002) have shown that these higher-order co-occurrences do appear in large corpora.

LSA consists in reducing the huge dimensionality of direct word co-occurrences to its best N dimensions. All words are then represented as N-dimensional vectors. Empirical tests have shown that performance is maximal for N around 300 for the whole general English language (Landauer et al., 1998; Bellegarda, 2000) but this value can be smaller for specific domains (Dumais, 2003). We will not describe the mathematical procedure which is presented in details elsewhere (Deerwester, 1990; Landauer et al., 1998). The fact that word meanings are represented as vectors leads to two consequences. First, it is straight-forward to compute the semantic similarity between words, which is usually the cosine between the corresponding vectors, although others similarity measures can be used. Examples of semantic similarities between words from a 12.6 million word corpus are (Landauer, 2002):

cosine(doctor, physician) = .61
cosine(red, orange) = .64

Second, sentences or texts can be assigned a vector, by a simple weighted linear combination of their word vectors. This is a powerful feature of a semantic representation to be able to go easily from words to texts. An example of semantic similarity between sentences is:

cosine(the cat was lost in the forest, my little feline disappeared in the trees) = .66

Modeling children's semantic memory

Semantic space

As we mentioned before, our goal was to rely on LSA to define a reasonable approximation of children's semantic memory. This is a necessary step for simulating a variety of children cognitive processes.

LSA itself obviously cannot form such a model: it needs to be applied to a corpus. We gathered French texts that approximately correspond to what a child is exposed to:

stories and tales for children (~1,6 million words), children's productions (~800,000 words), reading textbooks (~400,000 words) and children's encyclopedia (~400,000 words). This corpus is composed of 57,878 paragraphs for a total of 3.2 million word occurrences. All punctuation signs were ruled out, capital letters were transformed to lower cases, dashes were ruled out except when forming a composed word (like *tire-bouchon*). This corpus was analyzed by means of LSA and the occurrence matrix reduced to 400 dimensions, which appears to be an optimal value as we will see later. The resulting semantic space contains 40,588 different words. This step took 15 minutes on a 2.4 Ghz computer with 2 Gb RAM.

Tests

In order to test whether this semantic space can be an acceptable approximation of the semantic memory of children, we tested three features: its *extent*, its *organization* and its *use*. For each one, we relied on a specific task and compared the data from the simulation of the task to data obtained from children on the exact same task.

The *extent* feature has to do with the size of lexical knowledge. Does our semantic space *knows* the kind of words that a child knows? We used a vocabulary task for that: given a word, the goal is to find the correct definition from four of them. By comparing the model data with children's data at various ages, our goal is to approximately identify the kind of children we are mimicking.

The *organization* feature concerns the way words are associated to others in memory. Do we correctly mimic the semantic neighborhood of words? The task we used for testing that feature is an association task :given a word, the goal is to provide the most associated one. We will compare children's association norms to association measures in the semantic space.

The *use* feature has to do with the way semantic memory is used. Is our semantic space adequate enough so that it can account for a process that uses it? We used a recall task for studying the text comprehension process which obviously largely relies on semantic representations.

These three experiments cover different tasks and different grain sizes of language entities, from words to texts: the first one consists of word comparisons, the second one compares a word and a sentence and the third one compares texts. We expect a good match between human data and model data. In addition, we hypothesize that results will be higher with our children corpus than with adult corpora.

Experiment 1

The first experiment, which aims at validating the model, involves a vocabulary task. The design of the material as well as the experiments with children were realized by Denhière et al. (in preparation). Material consists of 120 questions, each one composed of a word and four definitions: the correct one, a close definition, a far definition and an unrelated definition. For instance, given the word *nourriture* (*food*), translations of the four definitions are:

- *what is used to feed the body (correct);*
- *what can be eaten (close)*;
- *matter which is being spoiled (far)*;
- *letter exchange (unrelated).*

Participants were asked to select what they thought was the correct definition. This task was performed by four groups of children: 2nd grade, 3rd grade, 4th grade and 5th grade. These data were compared with the cosines between the given word and each of the four definitions. For instance, the four cosines on the previous examples were: .38 (correct), .24 (close), .16 (far) and .04 (unrelated). 116 questions were used because the semantic space did not contain four rare words.

The first measure we used was the percentage of correct answers. Figure 1 displays the results. The percentage of correct answers is .53 for the model, which is exactly the same value as the 2nd grade children. Except for unrelated answers, the model data globally follow the same pattern as the children's data.

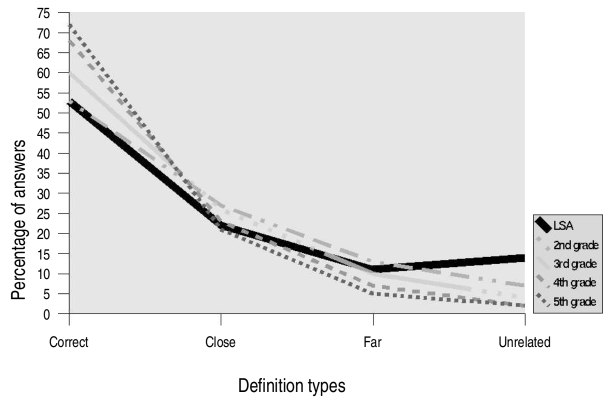

Figure 1: Percentage of answers for different types of definitions

In order to compare our semantic spaces with adult semantic spaces, we defined a measure which integrates the four values. We used a d measure, which is a normalized difference between the cosines for correct and close definitions together and the cosines for far and unrelated definitions together. The higher this measure, the better the result. Given a word W, four definitions (correct, close, far and unrelated) and a global standard deviation S, the formula is the following:

$$d=\frac{\frac{\cos(W,correct)+\cos(W,close)}{2}-\frac{\cos(W,far)+\cos(W,unrelated)}{2}}{S}$$

We also compared these results with several adult corpora, in order to test whether our semantic space was specific to children. We used five corpora: a literature corpus, composed of novels from the XIXth and XXth centuries and four corpora from the French daily newspaper *Le Monde*, of the years 1993, 1995, 1997 and 1999. Table 1 shows the results.

Table 1: Comparison between children's semantic space and adult semantic spaces

Semantic space	Size (in million words)	Percentage of correct answers	***d***
Children	3.2	.53	.69
Literature	14.1	.38	.52
Le Monde 1993	19.3	.44	.23
Le Monde 1995	20.6	.37	.21
Le Monde 1997	24.7	.40	.28
Le Monde 1999	24.2	.34	.25

In accordance with the previous experiment, the children's semantic space has the better results, although its size is much smaller. Student tests have shown that the children semantic space is significantly different from others ($p < .05$) except for the percentage of correct answers when compared to the *Le Monde* 1993 corpus ($p < .1$).

Experiment 2

This second experiment is based on verbal association norms published by de La Haye (2003). Two-hundred inducing words (144 nouns, 28 verbs and 28 adjectives) were proposed to 9 to 11-year-old children. For each word, participants had to provide the first word that came to their mind. This resulted in a list of words, ranked by frequency. For instance, given the word *cartable* (*satchel*), results are the following for 9-year-old children:

- école (*school*): 51%
- sac (*bag*): 12%
- affaires (*stuff*): 6%
...
- classe (*class*): 1%
- sacoche (*satchel*): 1%
- vieux (*old*): 1%

This means that 51% of the children answered the word *école* (*school*) when given the word *cartable* (*satchel*). The two words are therefore strongly associated for 9-year-old children. These association values were compared with the LSA cosine between word vectors: we selected the three best-ranked words as well as the three worst-ranked (like in the previous example). We then measured the cosines between the inducing word and the best ranked, the 2nd best-ranked, the 3rd best ranked, and the mean cosine between the inducing word and the three worst-ranked. Results are presented in Table 2.

Table 2: Mean cosine between inducing word and various associated words for 9-years-old children

Words	Mean cosine with inducing word
Best-ranked words	.26
2nd best-ranked words	.23
3rd best ranked-words	.19
3 worst-ranked words	.11

Student tests show that all differences are significant ($p < .03$). This means that our semantic space is not only able to distinguish between the strong and weak associates, but can also discriminate the first-ranked from the second-ranked and the latter from the third-ranked.

Measure of correlation with human data is also significant (r(1184 =.39, p<.001). Actually, two factors might have lowered this result. First, although we tried to mimic what a child has been exposed to, we could not control all word frequencies within the corpus. Therefore, some words might have occurred with a low frequency in the corpus, leading to an inaccurate semantic representation. When the previous comparison was performed on the 20% most frequent words, the correlation was much higher (r(234 =.57, p<.001).

The second factor is the participant agreement: when most children provide the same answer to an inducing word, there is a high agreement, which means that both words are very strongly associated. However, there are cases when there is almost no agreement: for instance the three first answers to the word *bruit* (*noise*) are *crier* (*to shout*) (9%), *entendre* (*to hear*) (7%) and *silence* (*silence*) (6%). It is not surprising that the model corresponds better to the children data in case of a high agreement, since this denotes a strong association that should be reflected in the corpus. In order to select answers whose agreement was higher, we measured their entropy. The formula is the following:

$$entropy(item) = \sum_{answer} freq(answer) . \log\left(\frac{1}{freq(answer)}\right)$$

A low entropy corresponds to a high agreement and vice versa. When we selected the 20% items with the lowest entropy, the correlation also raises (r(234)=.48, p<.001).

All these results show that the association degree between words defined by the cosine measure within the semantic space seems to correspond quite well to children's judgement of association.

We also compared these results with the previous adult semantic spaces. Results are presented in Table 3.

Table 3: Correlations between participant child data and different kinds of semantic spaces

Semantic space	Size (in million words)	Correlation with child data
Children	3.2	.39
Literature	14.1	.34
Le Monde 1993	19.3	.31
Le Monde 1995	20.6	.26
Le Monde 1997	24.7	.26
Le Monde 1999	24.2	.24

In spite of much larger sizes, all adult semantic spaces correlate worse than the children's semantic space with the data of the participants in the study. Statistical tests show that all differences between the child model and the other semantic spaces are significant (p<.03).

Experiment 3

The third experiment is based on recall or summary tasks. Children were asked to read a text and write out as much as they could recall, immediately after reading or after a fixed delay. We used 7 texts. We tested the ability of the semantic representations to estimate the amount of knowledge recalled. This amount is classically estimated by means of a propositional analysis: first, the text as well as the participant production are coded as propositions. Then, the number of text propositions that occur in the production is calculated. This measure is a good estimate of the knowledge recalled. Using our semantic memory model, this is accounted for by the cosine between the vector representing the text and the vector representing the participant production.

Table 4 displays all correlations between these two measures. They range from .45 to .92, which means that the LSA cosine applied to our children's semantic space is a good estimate of the knowledge recalled.

Table 4: Correlations between LSA cosines and number of propositions recalled for different texts.

Story	Task	Number of participants	Correlations
Poule	Immediate recall	52	.45
Dragon	Delayed recall	44	.55
Dragon	Summary	56	.71
Araignée	Immediate recall	41	.65
Clown	Immediate recall	56	.67
Clown	Summary	24	.92
Ourson	Immediate recall	44	.62
Taureau	Delayed recall	23	.69
Géant	Summary	105	.58

In an experiment with adults, Foltz et al. (1996) have shown that LSA measures can be used to predict comprehension. Besides validating our model of semantic memory, this experiment shows that an appropriate semantic space can be used to assess text comprehension in a much faster way than propositional analysis, which is a very tedious task.

Conclusion

A model of the development of children's semantic memory

Our model is not only a computational model of children's semantic memory, but of its *development*. Other computational models of human memory have been developed but some of them are based on inputs that do not correspond to what humans are exposed to. They are good models of the memory itself, but not of the way it is mentally constructed. In order to be cognitively plausible, models of the construction of semantic memory need to be approximately based on the kind of input to humans.

LSA is such model. Its performance is similar to those of human people. It needs an input of a few million words, which is comparable to what humans are exposed to (Landauer & Dumais, 1997). On the contrary, PMI-IR (Turney, 2001) is a good model of semantic similarities, even better than LSA in modeling human judgement of synonymy, but it is based on an input of thousands of millions of words, since it relies on all the texts published on the web. This is of course cognitively unplausible. HAL (Burgess, 1998) is another model of human memory. It is quite similar to LSA except that it does not rely on a dimension reduction step. It is currently based on a corpus of 300 million words, which is closer to the human inputs than PMI-IR, although this could be considered quite overestimated.

Further investigations

Our semantic space provides a means for researchers studying children's cognitive processes to design and simulate computational models on top of these basic representations. In particular, computational models of text comprehension could be tested using the basic semantic similarities that the space provides. It would also be possible to investigate the development of semantic memory by looking at the evolution of various semantic similarities according to the size of the corpus in detail. In particular, Landauer & Dumais (1997) claim that we learn the meaning of a word through the exposition to texts that do not contain it. Our semantic space gives the opportunity to test this assertion by checking the kind of paragraphs that cause an increase of similarity through incremental exposure to the corpus.

Improvements

Our semantic space could be improved in many ways. Its composition (50% stories, 25% production, 12.5% reading textbooks, 12.5% encyclopedia) is very rough and work has to be done to better know the amount and nature of texts that children are exposed to. Several studies led us to think that lemmatization could significantly improve the results, especially for the French language that has so many forms for some verbs. We did perform the previous experiments on a lemmatized version of the corpus (using the Brill tagger on the French files developed by ATILF, and the Flemm lemmatizer written by Fiametta Namer). Results were worse than with the non-lemmatized version. In order to know more about this surprising result, we distinguished between verbs and nouns. We found that the overall decrease is mainly due to a decrease for the nouns. One reason could be that the singular and plural forms of a noun are not arguments of the same predicates. For instance, the word *vague* (*wave*) is generally used in its plural form in the context of the *sea*, but more frequently in the singular form in its metaphorical meaning (*a wave of success*). Therefore, if both forms are grouped into the same one, this affects the co-occurrence relations and modifies the semantic representations.

Another way of improvement would have to deal with syntax. LSA does not take any syntactic information into account: all paragraphs are just bags of words. A slight improvement would consist in considering a more precise unit of context than a whole paragraph. A sliding context window (like in the HAL model for instance) would take into account the local context of each word. This might improve the semantic representations, while being cognitively more plausible. We are working in that direction.

For the moment, our model is an estimation. We cannot precisely identify to which age it corresponds. Our goal is to stratify it so that we would have a model for each age. Developmental models would then be able to be simulated.

Acknowledgements

This work was done while the second author was in sabbatical at the university of Aix-Marseille. We would like to thank D. Chesnet, E. Lambert and M.-A. Schelstraete, for providing us with parts of the corpus, F. de la Haye for the association data as well as M. Bourguet and H. Thomas for the design of the vocabulary test. We also thank P. Dessus and E. de Vries for their comments on a previous version.

References

Bellegarda, J. (2000) Exploiting Latent Semantic Information in statistical language modeling. *Proceedings of IEEE*, 88 (8), 1279-1296.

Bellissens, C., Thiesbonenkamp, J. & Denhière, G. (2002). Property attribution in metaphor comprehension: simulations of topic and vehicle contribution within the LSA-CI framework. *Twelfth Annual Meeting of the Society for Text and Discourse*, June 2002.

Borgida, A. (1996). On the relative expressive power of description logics and predicate calculus, *Artificial Intelligence, 82*, 353-367.

Burgess, C. (1998). From simple associations to the building blocks of language: modeling meaning in memory with the HAL model. *Behavior Research Methods, Instruments, & Computers*, 30, 188-198.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. & Harshman, R. (1990). Indexing by Latent Semantic Indexing. *Journal of the American Society for Information Science* , 41-6, 391-407.

de la Haye, F. (2003). Normes d'associations verbales chez des enfants de 9, 10 et 11 ans et des adultes. *L'Année Psychologique*, 103, 109-130.

Dumais, S. D. (2003). Data-driven approaches to information access. *Cognitive Science 27* (3), 491-524.

Foltz, P. W. (1996). Latent Semantic Analysis for text-based research. *Behavior Research Methods, Instruments and Computers*, 28-2, 197-202.

French, R.M. & Labiouse, C (2002). Four problems with extracting human semantics from large text corpora. *Proceedings of the 24th Annual Conference of the Cognitive Science Society*, NJ:LEA.

Howard, M.W. & Kahanna, M.J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language* 46, 85-98.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review,* 95, 163-182.

Kintsch, W. (2000). Metaphor comprehension: a computational theory. *Psychonomic Bulletin & Review,* 7-2, 257-266.

Kontostathis, A. & Pottenger, W.M. (2002). Detecting patterns in the LSI term-term matrix. *Workshop on the Foundation of Data Mining and Discovery, IEEE International Conference on Data Mining*.

Laham, D. (1997). Latent Semantic Analysis approaches to categorization. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the 19th annual meeting of the Cognitive Science Society* (p. 979). Mawhwah, NJ: Erlbaum.

Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem : the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104, 211-240.

Landauer, T. K., Foltz, P. W. & Laham, D. (1998). An introduction to Latent Semantic Analysis. *Discourse Processes* , 25, 259-284.

Landauer (2002). On the computational basis of learning and cognition: Arguments from LSA. In N. Ross (Ed.), *The psychology of Learning and Motivation, 41*, 43-84.

Lemaire, B & Dessus, P. (2001). A system to assess the semantic content of student essay. *Journal of Educational Computing Research,* 24-3, 305-320.

Lemaire, B. & Bianco, M. (2003). Contextual effects on metaphor comprehension: experiment and simulation. *Proc. of the 5th International Conference on Cognitive Modeling (ICCM'2003)*, Bamberg, Germany.

Lenat, D.B. (1995). CYC: A large-scale investment in knowledge infrastructure. *Comm. of the ACM*, 11, 32-38.

Pariollaud, F., Denhière, D. & Verstiggel, J.C. (2002) Comprehension of idiomatic expressions: effect of meaning salience. *Proc. of the 9th Int. Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU), Annecy, July 1-5.*

Perfetti, C. A. (1998). The limits of co-occurrence: tools and theories in language research. *Discourse Processes, 25,* 363-377.

Quesada, J., Kintsch, W., & Gomez E. (2002) A theory of complex problem solving using Latent Semantic Analysis. In W. D. Gray & C. D. Schunn (Eds.) *Proceedings of the 24th Annual Conference of the Cognitive Science Society.* Lawrence Erlbaum Associates, Mahwah, NJ.

Rehder, B., Schreiner, M. E., Wolfe, M. B. W., Laham, D., Landauer, T. K. & Kintsch, W. (1998). Using Latent Semantic Analysis to assess knowledge: some technical considerations. *Discourse Processes*, 25, 337-354.

Sowa, J.F. (1991). *Principles of Semantic Networks: Exploration in the Representation of Knowledge*, Morgan Kaufmann.

Turney, P. (2001). Mining the Web for synonyms: PMI-IR versus LSA on TOEFL. In De Raedt, L. and Flach, P., (Eds). *Proceedings of the 12th European Conference on Machine Learning (ECML-2001)* , 491-502, Freiburg.

Vossen, P. (2003). Ontologies. In R. Mitkov (Ed.) *The Oxford Handboof of Computational Linguistics.* Oxford, 464-482.

Wolfe, M. B. W., Schreiner, M. E., Rehder, B., Laham, D., Foltz, P. W., Kintsch & W., Landauer, T. K. (1998). Learning from text: matching readers and texts by Latent Semantic Analysis. *Discourse Processes*, 25, 309-336.

Wolfe, M.B.W. & Goldman, S.R. (2003) Use of Latent Semantic Analysis for predicting psychological phenomena: two issues and proposed solutions. *Behavior Research Methods Instruments & Computer*, 35(1), 22-31.