TEXTUAL CONSTRAINTS IN L2 LEXICAL DISAMBIGUATION

VILSON J. LEFFA

Catholic University of Pelotas, CNPq, Brazil

ABSTRACT

Reference: LEFFA, Vilson Jose. Textual Constraints In L2 Lexical Disambiguation. *System*, England, v. 26, n. 2, p. 183-194, 1998.

ABSTRAC:

Part of understanding a foreign language text involves the ability to solve lexical ambiguities that are not found in the first language. Traditionally, it has been claimed that the resolution of lexical ambiguity is done through schema activation. The hypothesis investigated here is that collocation may be a more dependable source than the reader's previous knowledge. Twenty ambiguous words were selected, disambiguated through rules based on collocation, and then tested with a concordancer, using an English language corpus of 20,000,000 words of expository text. The results showed that more than 94% of the ambiguities were solved by using syntactic and semantic restrictions between the ambiguous word and a related disambiguating word that co-occurred in the same sentence. The interpretation offered for these results is that collocation replaces with many advantages the use of encyclopedic knowledge to solve lexical ambiguities.

TEXTUAL CONSTRAINTS IN L2 LEXICAL DISAMBIGUATION

INTRODUCTION

One of the basic assumptions in reading comprehension research is that the understanding of a text involves the cooperation of different knowledge sources. These knowledge sources reside in the reader's mind and have to be activated so that the text can be understood. Among these sources, one of the most important is the ability to recognize the meaning of the different lexical items that make up the text. This is a complex task in itself and one aspect that adds up to this complexity is the fact that words have many different meanings, out of which, usually, only one can be chosen.

The following example, taken at random from *The Columbia Dictionary of Quotations*, can be used to demonstrate the complexity involved in assigning meaning to each of the lexical items.

Consul. In American politics, a person who having failed to secure an office from the people is given one by the Administration on condition that he leave the country. (Ambrose Bierce (1842–1914), *The Devil's Dictionary*)

Looking up each word in *The American Heritage Dictionary*, and counting their meanings produces the following results (number of meanings for each word shown in brackets):

Consul(3). In(22) American(6) politics(6), a(25) person(9) who(3) having(15) failed(12) to(21) secure(16) an(3) office(8) from(5) the(8) people(9) is(10) given (21) one(13) by(16) the(8) Administration(7) on(25) condition(12) that(15) he (2) leave(13) the(8) country(8).

The average meaning for each word is 11.34. If the meaning of each word depended on the meaning of every other word, we would have, in the small passage above, the astronomical sum of 11^{29} possibilities. Obviously, language does not work like that. Words do not multiply their meanings when in use, but have them restricted, typically to only one. The simple presence of other words, either to the right or to the left, has an inhibiting effect. No matter how many meanings a word has in the dictionary, it can carry only one when in the company of other words in a text. The pressure from the text is so great that the word may even loose its individuality, sometimes to the point of being divested

of all its original dictionary meanings and imposed a new one.

Among the many resources available to the reader to arrive at the one meaning supported by a given text is collocation. How far can collocation go in helping the reader to disambiguate the words is the point addressed in this study.

THE ILLUSION OF SIMPLICITY

A problem faced by readers when dealing with a text in a foreign language is that the different meanings a word presents in the new language are not symmetrical to the ones presented in the first language. This happens with practically every word. The apparently unambiguous *he*, in the passage above, when translated into Portuguese, has a different rendering in each of the following sentences (translation in brackets):

The cat is a he (macho).

He who seeks equity must do equity (aquele).

It was he who pushed the controversial points (ele).

What kind of knowledge sources does the L2 reader use to deal with this problem? The traditional approach is to emphasize the role of world knowledge, represented in our minds through frames, scripts, or schemas (Minsky, 1975; Schank and Abelson, 1977; Rumelhart, 1981; Anderson, 1983; Arbib, Conklin, and Hill, 1987). When an ambiguous word is found, the reader solves the ambiguity by activating the adequate schema. In the passage above, for example, the ambiguous word *office* would be interpreted as a public position, not room or building, because the election or political schema is invoked.

The problem is that schemas, to the extent that they represent larger chunks of knowledge, are not refined enough to solve many of the ambiguities that are encountered when two languages are involved. The verb *fail* in the sentences below can refer to the same all-including negotiations schema, but in each sentence it has a different meaning.

The delegates failed

The delegates failed to reach an agreement.

The delegates cleverly failed to show the statistics.

The delegates failed their own people.

The negotiations schema, although adequate, does not seem to be specific enough to discriminate between all the meanings. Subschemas would have to be created, probably subdivided into still smaller subschemas. Although such a fragmentation is possible from a theoretical point of view, in practice it results in unmanageable complex systems. Studies in artificial intelligence have demonstrated that the use of these subschemas leads to combinatory explosions when dealing with unrestricted texts.

The use of background knowledge is complex not because it deals with specialized knowledge — which, to the extent that it refers to specific domains, is usually manageable — but because it deals with everyday knowledge: common sense notions of time, space, causality, events, etc. Attempts to provide a computer with this kind of knowledge have resulted in decade-long projects involving teams of researchers and assistants (e.g., Guha and Lenat, 1994). As stated by Minsky:

Common sense is not a simple thing. Instead, it is an immense society of hard-earned practical ideas — of multitude of life-earned rules and exceptions, dispositions and tendencies, balances and checks (Minsky, 1985, p. 22).

The immense number of schemas needed to deal with background knowledge is only part of the problem, however. Two other problems still have to be considered: the dynamic nature of schemas and the lack of symmetry between the schemas and the language used to instantiate them.

Schemas are not only almost infinite in number but also extremely variable, changing constantly in space and time (Kintsch and Greene, 1978; Johnson, 1981), impairing communication between individuals from different cultures and times (such as the difficulty we many have when reading a book that was written many years ago). To the point that schemas represent the individual's theory of the world, changing constantly within each individual to adapt to his or her mutable viewpoints, communication between individuals from the same

culture is also affected (such as the frequent misunderstandings in conversation).

But what seems to make schemas recalcitrant to any computational treatment, beyond the problems generated by their abundance and dynamic nature, is the fact that they are independent of linguistic realization. A given schema can be instantiated over a wide range of different linguistic choices, both in terms of syntax and lexis (Hatch and Brown, 1995, p. 158).

Profusion, dynamic nature, and asymmetrical relation with language do not seem to make schemas a good choice to solve ambiguity problems, for they place a too high demand on memory capacity and processing time. Even if only distinctive features are stored, the data necessary to discriminate one schema from all the others surpass the capacity of known artificial systems when dealing with unrestricted texts. The same is true of the processing cost for instatiating one schema; there are so many variables with so many possibilities that combinatory explosions are inevitable.

A WORD IS KNOWN BY THE COMPANY IT KEEPS

One interesting hypothesis is to move away from world knowledge, assumed by schema theory, and get closer to the text, exploiting in more detail, not only how sentences relate to each other, but also how the words themselves behave, their preferences and hostilities towards the company of other words.

One way to approach this problem is by assigning roles to the lexical items that make up the sentences in a text. In the sentence “The cat is a he”, for example, the meaning of “he” is determined by the role this word plays in the sentence. Fillmore’s (1968) case grammar, Jackendoff’s (1985) conceptual structure hypothesis, and Chomsky’s (1981) theta roles, as developed by Levin (1985), are some of the substitutes for the schema theory, making Word Sense Disambiguation (WSD) more manageable for automatic language treatment and a cognitively more economic alternative for the human reader. Some studies in both automatic language processing (Dagan and Itai, 1994; Brill and Resnik, 1994; Justeson and Katz, 1995; Eizirik et al., 1993; Hindle and Rooth, 1993) — investigating problems related to machine translation, information retrieval, and man-machine interfaces — and human language processing (Britt, 1994; Dopkins et al., 1992; Trueswell et al., 1994; MacDonald, 1993) — related to reading comprehension — have advanced along this more textual line when dealing with lexical disambiguation.

The closest we can get to the text, however, is through the use of collocation. The lexical items in the text are not described in terms of the world they may

represent, but simply in terms of the words they tend to co-occur with, along with Firth's idea that the key to a word's meaning is the company it keeps. Some investigators have already suggested that collocates can be efficient word sense disambiguators (Sinclair 1991; Sinclair 1994; Clear, 1994; Smadja, 1993; Grefenstette, 1994). Clear, for example, argued that the ambiguous word *bow*, will tend collocate with *tie, tied*, etc. in the sense "type of knot"; with *arrow, string*, in the sense of "weapon"; with *stern, wave, starboard*, in the sense of "the front part of the ship"; and so on.

Some studies in both automatic and human processing of language have used collocations to investigate WSD. These studies have generally added some extra information beyond simple co-occurrence, including selection restrictions ("*he* when used for pronominal reference cannot be preceded by a determiner") , frequency data ("*pen* in the sense of a writing instrument is more frequent than *pen* meaning *enclosure*"), domain tags (*arm* as weapon and *arm* as part of the body receive different tags or markers when the word is entered in the dictionary), and even syntactic patterns (a verb may have a different meaning depending on whether or not it is followed by a direct object). Since all this extra information is built up on co-occurrence, it can the argued that the studies are essentially collocational. The use of tags, for example, is basically an economical procedure in which one tag can replace many words (e.g. the *animate* tag for all animals and humans beings).

The issue addressed here then is to what extent the presence or absence of certain word types to the right or to the left of the ambiguous term can determine its meanings. Replacing world knowledge with textual constraints to resolve lexical ambiguities would lead to less processing, thus producing faster and more economical results.

Studies in this area are scarce and usually limited to parts of speech that emphasize the role of collocation such as prepositions (Britt, 1994; Brill and Resnik, 1994) or adjectives (Justeson and Katz, 1995). Prepositions tend to have their meanings determined either by the verb that precedes them (e.g. "I depend on you") or by the noun that follows them (e.g. "I left on Sunday"), and adjectives may depend on the noun they modify before they have their meanings determined ("old" would mean "aged" before human beings and "not new" before objects.

Studies on nouns, which can occupy the head position in a phrase, are even scarcer and typically restricted to very few words. The long investigation conducted by Ng and Lee (1996), for example, uses only one word ("interest").

While studies on prepositions and adjectives can be argued to overrate the importance of collocation on meaning assignment, studies on the use of nouns, when restricted to one word, can be argued to be unreliable. Nouns may be more or less dependant on the immediate context for disambiguation, the word "interest" probably being the dependant type. An examination of its occurrences on the Collins-COBUILD corpus (COBUILD, 1995) shows that its most frequent collocates (*rates, rate, cut, high* and *public*) are typically placed either immediately after it or immediately before it. It is reasonable to assume that, while "interest" tends to have nearby collocates (thus making it easier to be disambiguated through collocation), other nouns might present a different pattern, based on distant collocates (and probably diminishing the significance of collocation on word sense disambiguation).

Another problem found in the literature on WSD concerns the methodology. Most studies on lexical disambiguation are conducted by first counting the different meanings of ambiguous words in standard monolingual dictionaries, (e. g. Ng and Lee, 1996), and then examining occurrences of the words in texts. The use of dictionaries poses some serious problems. First, there is no agreement on the number of meanings: for the word *interest*, for example — for which Ng and Lee (1996) distinguished 6 meanings — *The American Heritage Dictionary* lists 9, *Collins-COBUILD* 10, *Wordnet* 11, and The *Webster's Third* 12. The second problem is that it is very difficult to discriminate between different meanings when the word is used in texts. As Ng and Lee concede: "It involves making subtle human judgement, such that there are many genuine cases where two humans will not agree on the best sense assignments" (Ng and Lee, 1996, p. 45).

In a bilingual situation it is possible to use a more objective methodology, based on the criterion that a new meaning is involved only when a different translation has to be used. Although this procedure may look too coarse to capture many of the subtleties of the language, since many of the ambiguities are carried over from one language to another, it offers the advantage of being easily implemented, less dependent of subjective judgements. In practice, it may end up by capturing differences which are not detected in monolingual situations, as can be demonstrated with the sentences below:

The England team failed to win a place in the finals.

She failed in her attempt to swim to France.

The word "fail" in these sentences are presented as examples of the same

meaning ("not to succeed") in Collins-COBUILD. When the two sentences are translated into Portuguese, however, two different words have to be used.

The present investigation was designed to assess the role of textual constraints, as opposed to world knowledge, in resolving lexical ambiguities. Assuming that disambiguation becomes a more serious problem when we move from one language to another, this study will include two languages: English and Portuguese. Considering that different words have different collocational patterns, even when belonging to the same part of speech, different words will be included.

METHODOLOGY

The methodology used to collect the data involved three steps: (1) selecting a list of ambiguous words; (2) compiling examples of use for each word from a corpus; (3) processing the examples in a computer system, using syntactic and semantic constraints.

The first step was making up a list of words in English that would produce different translations in Portuguese. The first criterion for selecting these words was that they belonged to the same part of speech: ambiguities between “answer” as a noun and “answer” as a verb, for example, were discarded. Another criterion was that the ambiguous words should not depend too much on the immediate context to have the ambiguity solved such as prepositions, which are usually resolved at the syntactic level (“I depend on you”). The reason for trying to select words that depended on a larger context to be disambiguated was the assumption that it would result in a more reliable test of the hypothesis. The part of speech that seemed to depend less on the immediate context, all other things being equal, was the noun. A previous list of 52 words was constructed, collected from experience in teaching English to Portuguese-speaking students. From this list, 32 words were discarded because of high variance in the distribution of meanings (*pen*, for example, always occurred in the sense of a writing instrument in the corpus). The final shortlist included the following target nouns: *arm, bank, bar, bill, board, chip, coat, coach, corner, driver, gum, letter, nail, page, plane, record, room, table, time, wall.*

The second step was compiling examples of use for each word. The source for these examples was a corpus of 20,000,000 words of expository text. Occurrences of each word were recorded using the *Oxford Microconcord* (Scott, 1992). Since many of the ambiguous words belonged to different parts of speech, any part other than the noun was discarded. The occurrences were finally reduced to 200 examples for each word, using a random selection

procedure, which resulted in a total of 4,000 examples. Each example was 140 characters long, producing segments of text with about 20 words each.

The third step was testing the disambiguating rules for the target nouns in the examples. A computer system that is being developed for machine translation was used for this purpose. This system has disambiguating rules, based on textual constraints, for the treatment of ambiguity. Figure 1 shows an example of such a rule to disambiguate between the different meanings of “left”, including pairs such as:

When he *left* the house he was ready.

When he *left* the house was ready.

As can be seen, the rules do not assume world knowledge of the type typically described in schema theory.

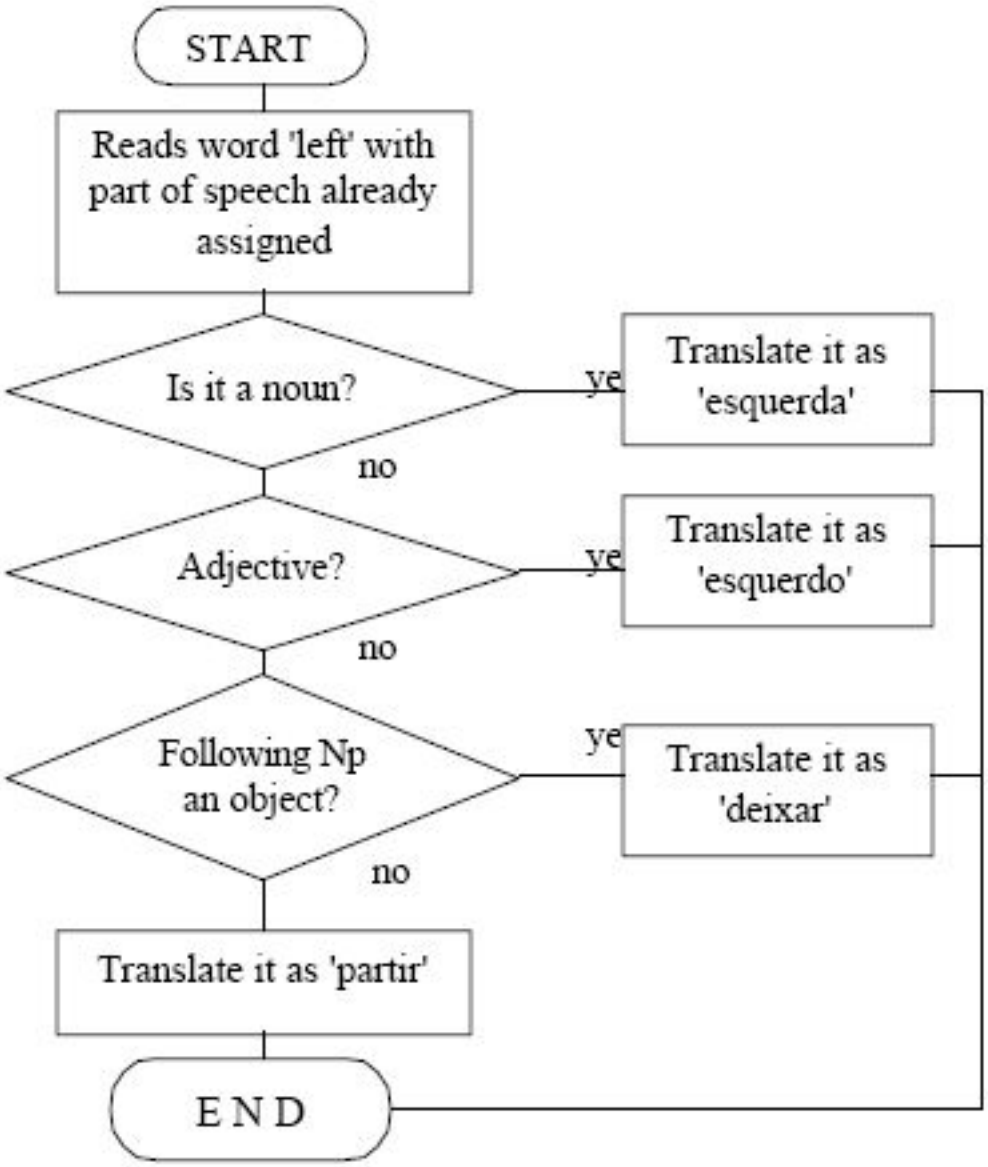

Figure 1: Example of a disambiguating rule.

The 4,000 segments of text with the target words were fed into the system and processed. The output was a tentative Portuguese translation of these English segments, at the lexical level. The morphological attributes and syntactical rules of the Portuguese language were not included here.

ANALYSIS

A first look at the output (Figure 2) shows that, in terms of translation, a lot of garbage was produced. The original English text, segmented arbitrarily in chunks of 140 characters, produced incomplete sentences and even incomplete words. As the segments were all put together, the program treated them as parts of the same text, connecting parts that should not be connected. Besides the 20 target words examined here, there were also many other ambiguities in the

textual segments for which the program was not yet prepared to cope with. Although all these difficulties, when put together, place an unfair demand on the system itself, it is also believed that in terms of this investigation they contribute to a more robust testing of the hypothesis.

...kness, spasticity, and atrophy, usually starting in the hands and *arms* and then spreading to other parts of the body. Difficulty with speak...	...kness, spasticidade, e atrofia, geralmente começar no mão e *braço* e então espalhar para outro parte do corpo. Dificuldade em falar...
...hat aid took the form of the government's handing over munitions, *arms*, and clothing to the playwright Caron de BEAUMARCHAIS and his fake "H...	...chapéu auxílio tomar o forma do governo estar entregar munição, *arma*, e vestuário para o dramaturgo Caron de BEAUMARCHAIS e seu fake "h...

Figure 2 - Examples of disambiguated target words (in italics)

Table 1 shows the results for the 4,000 segments with the 20 ambiguous words. Resolution rate varied from 83% to 98%, providing an average of 94.5% of correct disambiguations, with a low standard deviation, which means that results tended to be similar. More variation was noticed in terms of distance between the ambiguous word and the term that disambiguates it, which we will refer to as *collocate*. Examples of collocates are the words *hand* and *munitions*, shown in Figure 1 and used by the system to disambiguate *arm*.

A given word may belong to a phrase, which was separated into corpus phrases, based solely on frequency of occurrences in the corpus, and dictionary phrases, if entered in typical learning dictionaries. Examples of dictionary phrases with *arm* include *arm and leg, arm of the law, arm's length*, etc. Examples of corpus phrases are *take up arms, bear arms, left arm, small arms.*

Table 1 – General results

Target Words	Resolution %	Distance in words	Corpus phrases	Dictionary phrases

arm	98	1.5	4	8
bank	98	2.0	8	13
coach	98	2.0	7	2
gum	98	2.1	10	7
nail	98	2.2	5	9
bill	97	2.1	8	14
coat	97	3.0	10	6
page	97	2.0	3	5
plane	97	2.1	11	6
chip	96	3.5	11	2
record	96	2.9	9	18
time	96	2.0	7	39
table	95	3.7	7	17
board	93	2.5	12	15
driver	92	4.0	6	3
letter	92	4.1	2	16
room	92	6.0	7	10
wall	90	8.0	5	9
bar	87	6.0	9	11
corner	83	11.0	11	6
mean	94.5	3.4		

SD	4.16	1.89		

All dictionary phrases and most of corpus phrases, when felt to be unambiguous, were treated as discrete units of meaning and entered in the lexicon as such. Thus the string *take up arms* is an entry in the system lexicon, but *have in* is not, because sometimes it can make up a unit of meaning (e.g., *They have in guests*) and sometimes it cannot (e.g., *This is the house they have in Leeds)* When some of these phrases are included in the lexicon, they are not treated as collocation and consequently not included in the resolution column (Table 1). Considering that this procedure contributes to lower the scores, since many correct resolutions are not taken into account, it is believed that it should increase the reliability of the results.

In terms of the distance between the ambiguous word and the collocate, there is a high correlation (r = .92), as expected. An obvious explanation for this, is that, as distance increases, intervening factors can affect the results. All other things being equal, it is much easier to disambiguate *arms* in a segment such as *military arms,* where the collocate is juxtaposed to the target word, than in segments such as *the military gathered in the plaza raised their arms,* where the collocate is more distant.

The average of 94.5 of correct disambiguations reflects the application of rules as they had been previously introduced into the system. Some of these rules can be improved and eventually produce better results — as can be seen in the following sentence:

The music is set in duple meter (2 beats to the bar) and is based on about 50 standard calypso melodies.

In the sentence above, *bar* was incorrectly interpreted by the system as a room where drinks are served and music can be played, instead of the measure used in music to divide a staff. The problem was that the collocate *music* was not specific enough to discriminate between these two meanings of bar. There are, however, other more restricting collocates in the sentence — which could have been fed into the system to solve the problem such as *duple meter* and even *beats,* although an ambiguous word itself.

CONCLUSION

Words, before they are used in a text, are just a set of possibilities, pointing imprecisely to a bank of concepts we have stored in dictionaries or in our minds. In terms of dictionaries, the number of meanings assigned to every word, as they

are used in a current text, is around 11 meanings per word. In terms of what we have stored in our minds the number is probably much higher, including hundreds or maybe thousands of recessive meanings, meanings that are hidden behind the dominant one and that come to life when certain conditions are met in a text. Once, however, a recessive meaning becomes dominant all the others become recessive, discounting for the rare cases when double meaning is intentionally used. The main finding in this investigation is that this drastic reduction to one meaning is due to syntactic and semantic restrictions imposed by the neighboring words. It is claimed that the use of these restrictions is advantageous to the process of reading, making it more efficient.

Of course we can always build examples in which a given word may be disambiguated only by the activation of a given schema. The data analyzed here, however, suggest that in cases where both schemas and textual restrictions can be used, the application of syntactic and semantic restrictions is more precise and economical.

It can also be argued that once a schema is activated it guides disambiguation, sometimes to the point of predicting what is coming next. If someone is reading a text about hand care and meets a sentence that starts with the words “when you cut your ...”, the meaning of *nail*, as part of the finger, is probably activated even before the word is read.

The problem here, it seems, is to decide which comes first, schema activation or data from the text. It is true that in some cases schemas may be previously induced, such as in a classroom situation where the teacher prepares the student for the reading of a text. Most often, however, it seems that schemas are activated as data are processed from the text.

Another argument that could be used to favor schemas against textual constraints is that schemas are more powerful to help the reader guess the meaning of unknown meanings for ambiguous words. The reader, for example, may be familiar with the meaning of pen as a writing instrument and may have problems when he or she meets the word used in the sense of an enclosure for animals.

Again, it seems that the reader does not need the broader context provided by a schema to guess the new meaning of the word. A collocate, such as *sheep* or *chicken*, if found in the vicinity of the ambiguous word, would probably be enough.

The use of syntactic and semantic constraints, as compared to world knowledge,

occur automatically, below the level of consciousness. The sequential and more time-consuming strategies are replaced by faster, automatic processing, where activities are performed in parallel. The result, whenever we move a subprocess to these lower, more automatic levels, is a general gain in reading efficiency.

The finding that syntactic and semantic constraints can be used to disambiguate word senses should also be of consequence to reading instruction and language development, in general. It can be argued that what the students have to learn is not a representation of the world, but rather the language that has to be used to represent that world. Schema-oriented approaches, by providing only one context at a time, may offer language learners less than the minimum exposure they need to acquire the language. One way to solve this problem is to complement the use of full passages with different activities based on text fragments, as provided by concordance printouts. The learner would have the advantage of approaching a specific problem from multiple contexts, without sacrificing the textual clues necessary for comprehension, which, as suggested by this study, can be found in the neighboring words.

REFERENCES

Anderson, J. (1983) *The architecture of cognition*. Harvard University Press, Cambridge, Mass.

Arbib, M. A., Conklin, E. J., Hill, J. C. (1987) *From schema theory to language*. Oxford University Press, New York.

Brill, E. and Resnik, P. (1994) A rule-based approach to prepositional phrase attachment disambiguation. In *Proceedings of COLING-94*, pp. 1198-1204.

Britt, M. A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. *Journal of Memory and Language*, **33**, 251-283.

Chomsky, N. (1981) *Lectures on government and binding*. Dordrecht: Foris.

Clear, J. H. (1994). I Can't See the Sense in a Large Corpus. In F. Kiefer, G. Kiss, J. Pajzs (eds.) . *Research Institute for Linguistics*, Budapest: Hungarian Academy of Sciences, pp. 33-48.

COBUILD, *English Collocations on CD-ROM*. (1995). Harper Collins, London.

Dagan, I.; Itai, A. (1994) Word sense disambiguation using a second language monolingual corpus. *Computational linguistics*, **20**(4), 563-596.

Dopkins, S.; Morris, R. K.; Rayner, K. (1992). Lexical ambiguity and eye fixations in reading: A test of competing models of lexical ambiguity resolution. *Journal of memory and language*, **31**, 461-476.

Eizirik,L. M. R.; Barbosa, V. C.; Mendes, S. B. T. (1993) A Bayesian-network approach to lexical disambiguation. *Cognitive Science*, **17**, 257-283.

Fillmore, C. J. (1968). The case for case. In: E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory*. Holt, Rinehart, & Winston, New York.

Grefenstette, G. (1994). *Explorations in automatic thesaurus discovery*. Boston: Kluwer Academic Publishers.

Guha R. V., and Lenat, B. L. (1994). Enabling agents to work together. *Communications of the ACM*, **37**(7), 127-142.

Hatch E. and Brown, C. (1995). *Vocabulary, semantics, and language education*. Cambridge University Press, Cambridge.

Hindle D. and Rooth, M. (1993) Structural ambiguity and lexical relations.

Computational Linguistics, **19**(1), 103-120.

Johnson, P. (1981) Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly*, **15**(1), 169-181.

Jackendoff, R. (1985) *Semantics and cognition*. Cambridge: Cambridge University Press.

Justeson, J. S. and Katz, S. M. (1995). Principled disambiguation: Discriminating adjective senses with modified nouns. *Computational Linguistics*, **21**(1), 1-27.

Kintsch, W. and Greene, E. (1978) The role of culture-specific schemata in the comprehension and recall of stories. *Discourse Processes*, **1**, 1-13.

Levin, B. (1985) *Lexical semantics in review*. Cambridge, Mass.: Center for cognitive sciences, MIT.

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, **32**, 692-715.

Minsky, M. (1975) A framework for representing knowledge. In: Winston, P. (Ed.) *The psychology of computer vision*. McGraw-Hill, New York. pp. 211-277.

Minksy, M. (1985). *The society of mind*. Simon & Schuster, New York.

Ng, H. T. and Lee, H. Integrating multiple knowledge sources to disambiguate word sense: An exemplar-based approach. In *Proceedings of ACL-96*, pp. 40-47.

Raskin, V. (1987) Linguistic heuristics of humor: a script-based semantic approach. *International Journal of the sociology of Language*, **65**, 11-25.

Rumelhart, D. E. (1981) Schemata: The building blocks of cognition. In: Guthrie, J. T. (Ed.) *Comprehension and teaching: Research reviews*. International Reading Association, New Haven. pp. 3-25.

Schank, R. C. and Abelson, R. P. (1977) *Scripts, plans, goals, and understanding*. Lawrence Erlbaum, Hillsdale, N. J.

Scott, M. and Johns, T. (1992) *Microconcord.* Oxford University Press, Oxford (Oxford English Software).

Sinclair, J. (1991) *Corpus, concordance, collocation* (Describing English Language Series). Oxford University Press, Oxford.

Sinclair, J. (1994) Trust the text. In Coulthard, M. (ed.) *Advances in written text analysis*, pp. 13-25. Routledge, London.

Smadja, F. (1993) Retrieving Collocations from Text: Xtract. *Computational Linguistics*, **19**,143-177,.

Trueswell, J. C.; Tanenhaus, M. K.; Garnsey, S. M. (1994) Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of memory and language*, **33**, 285-318.