Exploring Sentence Variations with Bilingual Corpora

Jin, Z.; Barrière, Caroline

National Research Council Canada
Conseil national de recherches Canada

Institute for Information Technology
Institut de technologie de l'information

NRC · CNRC

*Exploring Sentence Variations with Bilingual Corpora **

Jin, Z., and Barrière, C.
July 2005

* published at the Corpus Linguistics 2005 Conference. Birmingham, United Kingdom. July 14-17, 2005. NRC 48511.

Canada

Exploring sentence variations with bilingual corpora

Zhenglin Jin and Caroline Barrière
Interactive Language Technology Group,
Institute for Information Technology,
National Research Council Canada
caroline.barriere@nrc-cnrc.gc.ca

Abstract

We propose a system for retrieving similar sentences from a corpus which treats sentences as pure strings. The advantage of such an approach compared to more linguistically motivated approaches is that the system can quickly retrieve similar sentences from a large size corpus (over one million sentences), work well with ill-structured sentences, and work across different human languages. The system has been tested using English, French and Chinese corpora and the results have been manually evaluated. The application suggested in this paper is to use our similar sentence search engine within a language-learning context to help language learners improve their writing skills and better understand grammar rules of their second language by studying different sentence variants from realistic examples. We further suggest using the system with bilingual parallel corpora to help translation students enhance their translation skills by accessing professional translations.

1. Introduction

Learning from examples, referred to as Data-Driven Learning (Johns, 1994), has been promoted in recent years as a valuable way of learning for intermediate and advanced students. It is made possible by large corpora now being available to language learners. The emphasis is that students can now learn from authentic language as opposed to examples made-up by teachers. Monolingual corpora are built based on real written or oral communication by native speakers. Aligned bilingual corpora are constructed with examples of professional translations. Both corpora are invaluable sources for language or translation learners to understand and learn from real world data (McEnery and Wilson, 2004).

Opposition to this idea emphasizes the danger for students to get lost among too many examples, and not knowing where to go and what to do. Access to a useful but huge corpus can be overwhelming. Valuable examples might not be so obvious to find if hidden among collections of numerous examples (often millions).

As an example of a methodology for guiding students during their search of examples, concordancers, much used in terminology for finding word collocations, have found their way into language learning, as the favourite form of data-driven learning aid. Many L2 researchers and teachers have looked into concordancers (Aston, 2001).

We propose a different, novel way of searching in corpora, at the sentence level rather than the word or expression level as used in corcondancers. We suggest starting with an input sentence and looking through a corpus for finding similar sentences. If the input sentence is taken from a text for a reading task, finding similar sentences will help for its comprehension. If the input sentence is created by a learner in a writing exercise, finding similar sentences will help for structuring it correctly or finding slight variants of meaning. Since we suggest a pure string approach to establishing sentence similarity, the learner's input sentence could be ill-structured, the system would still be able to find similar sentences in a corpus. This pure string approach pays less emphasis on linguistic features of a sentence and therefore has advantages of being quite fast (important factor when looking at large corpora) and of being language independent.

We developed a system which provides a user interface to find similar sentences to an input sentence. When used on a monolingual corpus, the system shows sentences similar to a source language. When used on a bilingual aligned corpus, it shows pairs of similar sentences in both source language and translated target language.

The focus of this paper is first to present, in section 2, a view on the concept of similarity as well as different algorithms normally used in the information retrieval domain which were adapted, implemented and tested to perform sentence similarity ranking. In section 3, a small human evaluation is performed to establish the value of the algorithms. From these observations, we reduce the set of algorithms to be further used in our application system which we present in section 4. Section 5 suggests possible use of the system in language learning contexts. Section 6 gives conclusions and points to future work.

2. Exploring Sentence Similarity

Research in cognitive science has noted the importance of comparative settings in the learning process and indicated the importance of finding similarity and noting differences among items (Tversky, 1977). Human categorization is based on the idea of grouping similar objects into categories and then understanding differences between objects in each category. Inspired by this idea, we aim at finding similar sentences in a corpus. Only among a group of similar sentences, can the small differences be noted and understood by the learner.

In a contrast model, Tversky (1977) states that an object can be represented by a list of features, and the similarity between two objects a and b can be generally defined as

$$S(a,b) = \theta f(A \cap B) - \alpha f(A-B) - \beta f(B-A) \qquad (1)$$

Where $A \cap B$ are common features of both a and b; A-B are features that belong to a but not to b; B-A are features that belong to b but not to a. θ, α, and β are the weights of how similarity is measured by a combination of common and non-common features. We consider the simplest case of $\theta=1$ and $\alpha=0$, $\beta=0$, where the similarity of two objects is measured by their common features, that is, $S(a,b) = f(A \cap B)$ (Tversky 1977)

Considering an object as a sentence and a feature as a word in the sentence, similarity between sentences can be defined as

$$\text{Similarity}(a,b) = f(A \cap B) \qquad (2)$$

where each sentence is represented as a series of words. $A \cap B$ are words which are common to both a and b.

To find $A \cap B$, we define the word level equality measure as follows. Each pair of words taken from a pair of sentences is defined to be equal if they are constructed from exactly the same sequence of strings.

If there is a sentence A, and a list of sentences $B_{set} = \{B_1, B_2, B_3, \ldots B_n\}$, a similarity ranking function $f(A \cap B_i)$ can assign a similarity value between A and each sentence in B_{set}. The larger the value given by f, the more similar $A \cap B_i$ is. Thus sentences in B_{set} can be sorted based on their similarity value. The most similar sentence in B_{set} goes to the top of the list. There are many ways to define the function $f(A \cap B_i)$ and some of them will be described in 2.2.

2.1 Sentences as strings

Sentence similarity in the literature is usually of interest in the context of Example-Based Machine Translation and Machine-Aided Human Translation application such as translation memories. Compared with other available resources, existing translations contain more solutions to variant translation problems (Isabelle et al., 1993). Extracting similar sentences from aligned corpora can help reuse existing translations.

Somers (1999) reviewed several sentence distance or similarity measures that were linguistically motivated. Different linguistic components of a sentence (e.g. characters, words, or structures) can be used as comparison units. So far, character-based matching (Sato, 1992), word-based matching (Nagao, 1984), structure-based matching (Matsumoto, 1993), and syntax-matching (Sumita and Tsutsumi, 1988) have been used.

These approaches consider sentences as linguistic entities and algorithms are tied to a specific language. What will happen if we treat a sentence as a pure string? First, since Unicode[1] allows the software manipulation of many languages as pure strings, we can use the same set of algorithms to retrieve sentences written in different languages. Second, each sentence can be treated as a small piece of document and information retrieval similarity ranking algorithms can be adapted to calculate sentence similarities. Third, in a language-learning context, sentence correctness is difficult to predict and pure string comparison is a more tolerant approach than a syntax-based approach.

[1] Information about this standard is given at: http://www.unicode.org/

2.2 Looking at the algorithms

Some well-known similarity-ranking functions in the information retrieval process are Dice coefficient, Vector space model (cosine), and Lin's information theory similarity measure (referred to as Lin hereafter). Besides these three algorithms, we also look at BLEU which is a metric used for Machine Translation systems evaluation. Since BLEU works by comparing pure strings, we decided to test it here for sentence similarity. All four algorithms are tested in order to find a good similarity function for the system. We briefly review the algorithms hereafter, but refer the reader to appropriate references for more details about the equations.

The Dice Coefficient is a word-based similarity measure. The similarity value is related to a ratio of the number of common words for both sentences and the number of total words of the two sentences.

When comparing two sentences Q and S, if N_{common} is the count of common words, N_Q is the total count of words of sentence Q, and N_S is the total count of words of sentence S, the Dice coefficient can be expressed as follows (Hersh, 2003).

$$Dice(Q,S) = \frac{(2 * N_{common})}{(N_Q + N_S)} \qquad (3)$$

In the Vector space model, documents (S) and queries (Q) are decomposed into smaller word units. All words are used as elements in the vectors that will represent Q and S. Both vectors contain weights assigned to each word corresponding to the number of occurrence of that word within them (Jurafsky, 2000).

The formula is given in Equation (4) (Salton et al., 1983) for t words.

$$COSINE(Q,S) = \frac{\sum_{k=1}^{t}(w_{qk} \cdot w_{sk})}{\sqrt{\sum_{k=1}^{t}(w_{qk})^2 \cdot \sum_{k=1}^{t}(w_{sk})^2}} \qquad (4)$$

An information-theoretic definition of similarity has been proposed in recent years and the similarity measure is applicable when there exists a probabilistic model. Based on certain assumptions, the similarity between A and B is measured by the ratio of the amount of information needed to state the commonality of A and B and the amount of information needed to fully describe what A and B are (Lin, 1998)

For objects, which can be described by a set of independent features *w*, Lin derives the following instantiation of this principle (Aslam, 2003).

$$IT - Sim(Q,S) = \frac{2\sum_{w \in Q \cap S} \log \pi(w)}{\sum_{w \in Q} \log \pi(w) + \sum_{w \in S} \log \pi(w)} \qquad (5)$$

where $\pi(w)$ is the probability of feature *w*. For sentence similarity, we assign all words in Q and S as possible features.

BLEU (Papineni et al., 2002) is a method for automatic evaluation of machine translation. We use this algorithm to rank similar sentences by comparing the input sentence Q and only one reference sentence S. The implementation is based on the following formula:

$$\text{Log BLEU} = \min(1 - r/c, 0) + \sum_{n=1}^{N} w_n \log p_n \qquad (6)$$

Where c is the length of Q and r is the length of S. The second term is calculating the geometric average of the modified n-gram precision p_n. If p_n is zero, a constant value ε is added to make p_n a non zero value.

3. Evaluation of the output

By manually evaluating the outputs and analysing the ranking agreement between human ratings and the above functions, we may suggest the best $f(A \cap B_i)$ (in reference to section 2) which can accurately and quickly rank a list of similar sentences for language and translation learning purpose.

3.1 Evaluation process

Four similarity functions, Dice coefficient, Cosine, Lin and BLEU (equations (3)-(6)) are tested with the Canadian Hansard (English-French), Xinhua corpus (Chinese), and corpora for NIST MT evaluation (English-Chinese[2]. For each function the system outputs the top four most similar sentences found by that function to the input query sentence. The evaluation focuses on monolingual output in order to find the most suitable similarity functions.

To evaluate the accuracy of the system outputs, we adapt a grade scheme proposed by Sato (1990), as shown in Table 1. "The sentence" in the table refers to the output sentence.

[2] All corpora can be found at: http://www.nist.gov/speech/tests/mt/

Grade	Explanation	Category
4	The sentence exactly matches the input	Same
3	The sentence provides enough information about the whole input	Very Similar
2	The sentence provides information about some part of the input	Partly similar
1	The sentence provides no information about the input	Completely Different

Table 1. Accuracy grades

Each sentence given as output by the system is manually graded using the four categories given in Table 1. If an output sentence belongs to the first three categories, it is regarded as a useful or relevant sentence to the L2 or translation learner. If the sentence falls in the last category, it is regarded as useless from the point of view of L2 learning or translation assistance. Appendix 1 shows a sample of the evaluation scheme given to the evaluators.

Ten bilingual evaluators were involved in the evaluation. Each of Chinese-English pair evaluators received more than ten years of education in China and minimum five years education in Canada. Each of French –English pair evaluators are all received bilingual (French and English) education since their childhood in Canada. Seven of the evaluators have university degree and three of them are third year university students. One of the evaluators who evaluated English-French pair has human translation experience. Table 2 gives the detailed information regarding the human evaluation. The word "package" in the third column refers to the number of input sentences for which 16 output sentences had to be evaluated (4 algorithms * 4 highest ranked sentences for each).

Language	Number of Reports	Number of packages	Corpus from which sentences are taken
French	5	10	English–French Hansard
English	7	20	English–French Hansard
Chinese	4	10	Xinhua

Table 2. Human evaluation information

3.2 Human evaluation result

For each similarity function, the average score received among sixteen reports are compared and shown in Table 3. We find that the simple Cosine algorithm has the best performance with average score of 2.75. For BLEU it seems that it gives high rank to some sentences not very relevant and so it receives the lowest average score of 2.64.

Algorithm	Average similarity score received (Highest score is 4)
Dice	2.73
Cosine	2.75
Lin	2.73
BLEU	2.64

Table 3. Average similarity score for different algorithms across languages

Although on the basis of the different average grades in Table 3 we may say that the different algorithms perform differently, the question is how significant these differences should be. The Student`s *t*-test is a useful tool to check the difference between two sets of experimental results with a quantitative measure. The *t*-test results for Cosine & Lin, Lin & Bleu, Cosine & bleu are show in Table 4. For each algorithm there are 16 experiments (human evaluation), so the degree of freedom is (16+16-2)=30. According to the *t*-value and the degree of freedom, the probability of assuming the null hypothesis, or the confidence level, can be obtained. The *t*-test shows that there is no significant difference between the Cosine and Lin's algorithms, or in other words, their difference can be neglected. However, the Bleu is certainly different from the other two algorithms with over 97% confidence levels. Thus Cosine, Dice coefficient, Lin can be selected for future study.

	Mean value	*t*-value	Degree of freedom	Confidence level of difference between two algorithms
Cosine & Lin	2.75 & 2.73	0.616	30	46%
Lin & Bleu	2.73 & 2.64	2.18	30	97%
Cosine & Bleu	2.75 & 2.64	2.60	30	99%

Table 4. Student`s *t*-test for significance of difference among three algorithms

Table 5 gives the similarity score of the Cosine for different languages. Chinese receives higher score than English and French receive. This is probably because there is morphological analysis module, adapted in the current system. Such module is important in processing English**,** even more so in processing French, but not required for processing Chinese. For example, "is" and "was" are treated as different words without considering morphological module, and this may certainly affect the similarity retrieval.

Algorithm	Average similarity score received (Highest score is 4)
English	2.76
French	2.73
Chinese	2.80

Table 5 Average similarity score of Cosine by different languages

As a different type of evaluation, we look into the agreement between the ranking of each algorithm and human ranking of output. The system outputs the top four similar sentences based on the automatic ranking by each of the similarity function. Each function`s output is in decreasing value. If a human ranking is also in descending order then it is considered as agreement with the automatic ranking. Table 6 shows the percentage of agreement between each algorithm and human rating in terms of similarity ranking. The Dice Coefficient is in 100% agreement with human rating but the Lin's ranking only agrees 67% of the human rating.

Algorithm	Percentage of agree with human rating
Dice	100%
Cosine	93%
Lin	67%
Bleu	80%

Table 6 Percentage of each algorithm agreeing with human rating

As a compromise of the average similarity score received by each similarity function and the agreement in terms of ranking made by each similarity function and human rating, the Cosine is preferentially used as the algorithm of the current system implementation. However, such evaluation is at quite a small scale and given the closeness between Lin`s, Dice and Cosine, as shown by the t-test, we have built the system, as presented in the next section, with the flexibility to access any similarity function defined.

4. Sentence Similarity Module

We introduce an independent sentence similarity module, which can be integrated within a language learning system. In a comprehension situation, such a system would provide texts or examples understandable to the students. In a production situation, a student would be writing on a particular topic and expecting to see some examples written by native speakers. In translation learning situation, a student can find translation of similar sentence by requesting bilingual output. A highlighting of a sentence in either situation could prompt a search in an external bilingual corpus and monolingual or bilingual similar sentences will be returned.

4.1 Corpus requirement

The English-French corpus in use is the Hansard corpus starting from 1999 and ending 2003. It contains 1,458,500 sentence pairs. The Chinese-English corpora are the ones used for the NIST MT evaluations for the years 2002-2004. It contains about 5,000 sentence pairs. The Chinese Xinhua corpus which contains 849,720 sentences is used to test the Chinese monolingual output[3]. All corpora were pre-processed so that they contain one sentence per line and all the sentences are properly tokenized.

The performance of the system is known to be corpus dependent. Thus to make this system useful in language learning domain, it may require corpora specifically for learning purpose.

4.2 Design of sentence similarity module

Figure 1 gives the architecture of the module. The user specifies the source language and provides the input sentence by typing or highlighting through graphical user interface. The module then collects relevant sentences by searching through the indexed corpus.

[3] The details regarding these corpora can be found at the web page of NIST evaluation (http://www.nist.gov/speech/tests/mt/).

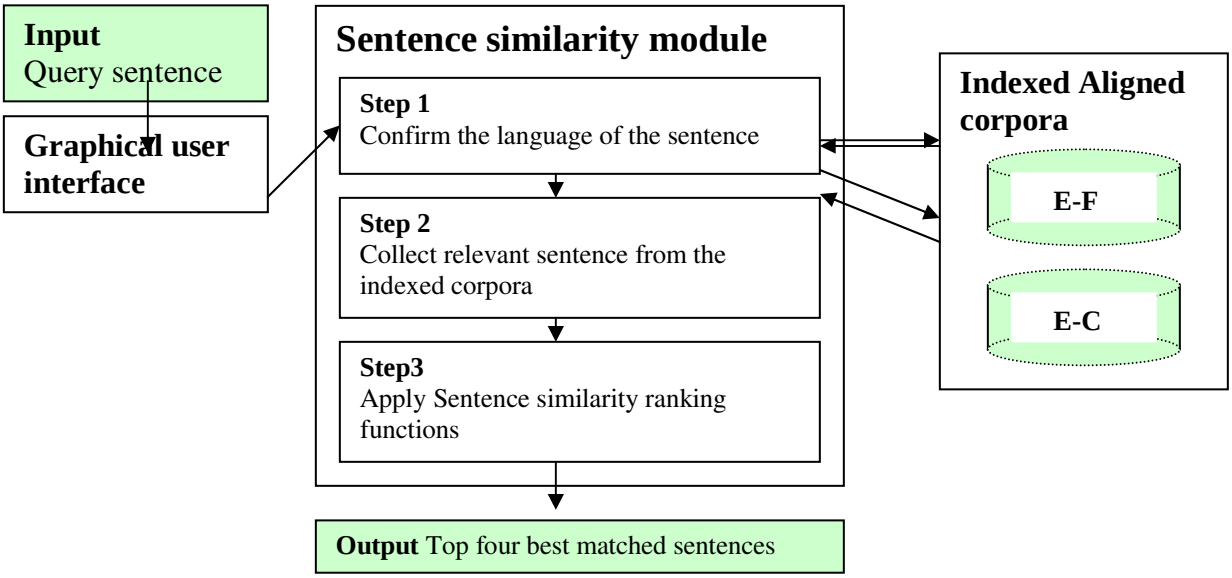

Figure1. System architecture

Figure 2 gives a simple interface of the sentence similarity module. For the input sentence, a sentence can be typed in, or it can be highlighted from a text chosen by the user and displayed on the left side of the screen. The right side outputs the similar sentences in decreasing order of similarity.

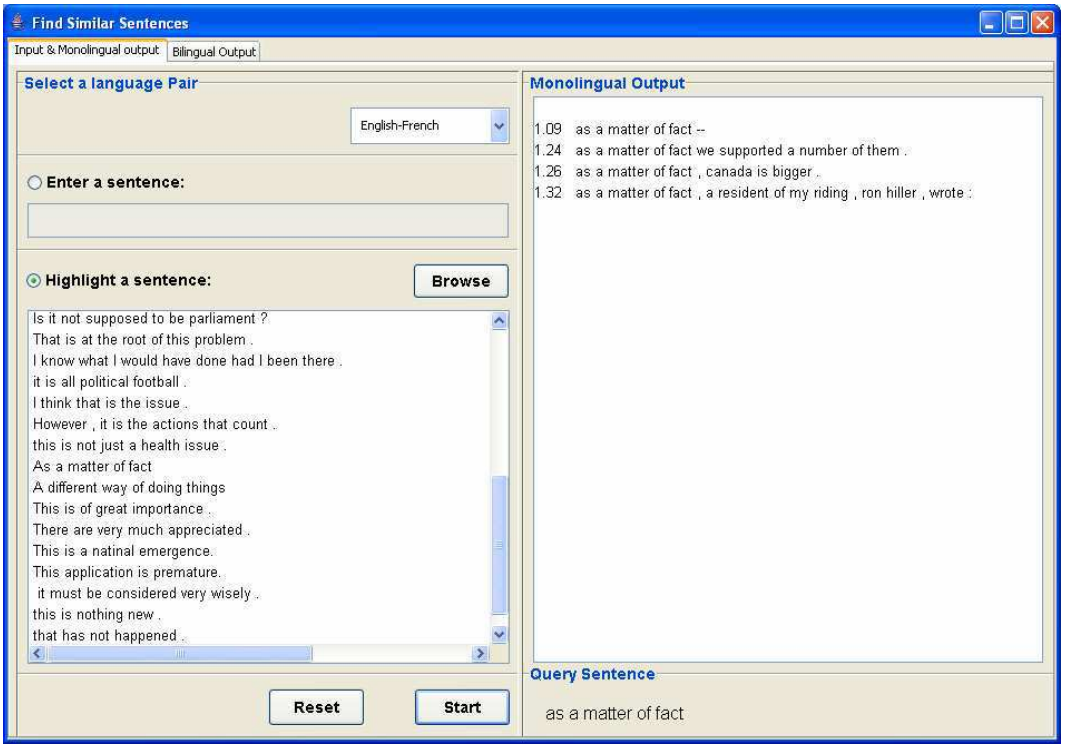

Figure 2. The graphical user interface of sentence similarity module

4.3 Corpus Indexing

Given its large size, normally with over a million sentences, the corpus is broken down into smaller files and is indexed by using Apache Lucene[4] to speed up the search. Lucene can index all the files under a given directory. Thus every corpus is broken into files with 20 sentences. Since a sentence can be seen as a list of words, Lucene will return all the files which contain those words. Only those sentences returned by Lucene' search will be used to perform similarity ranking.

Nevertheless, even with such indexing approach, access time for short sentences (a few seconds) is not acceptable for an application, and future work will investigate better indexing techniques.

5. Example usage of the system

We present three possible use of the system in language learning and translation learning environment. The Cosine similarity function[5] is used to rank sentence retrieved from the corpus and the top four best matches are given as output.

5.1 Production setting for students of English as a second language

Given the input "I believe that the world changed", the system returns the following output as shown in Table 7.

1	I believe that the world has changed.
2	the world has changed, as I have said.
3	since the events of September 11, the world has changed, I believe.
4	the world has certainly changed.

Table 7 English output to help teach writing in English as a second language

From a known pattern used by the learner "I believe that X " now, the access to examples in the corpus show that "X, as I have said" or "X, I believe" are possible variations. It also shows a possible adverb "certainly" which can be introduced before "changed".

5.2 Learning grammar rules by similar variants

A small corpus with hundreds of sentences is created using examples collected from English grammar books. For the input sentence "He washed the car. He polished it", the system returns output sentences as shown in Table 8.

[4] Information about Lucene can be found at: http://lucene.apache.org/java/docs/

[5] The modularity of the system, as shown in Figure 1, allows any similarity function to be used, if a particular function is designed for a particular type of application.

1	he washed the car. he polished it.
2	he washed the car and polished it.
3	he washed the car and then polished it.
4	he not only washed the car, but polished it too.

Table 8 English output for help of learning English grammar

By studying the examples given by Table 8, the learner can learn how to apply grammatical rules. The examples returned by the program sound more coherent and they will teach the user other grammatically correct but different ways of expressing the similar meaning.

5.3 Learning translation using English-French bilingual output

The use of specialized monolingual native-language corpus has shown to improve subject-field understanding of students in translation and improves the quality of the translation output when the task is to translate articles in a specialized field (Bowker, 1998). Given the success of monolingual corpora in translation learning, we should investigate the possibility that bilingual corpora would aid students more.

For instance, if a student needs to translate an English sentence: “This is not just a health issue". The system will return bilingual output as shown in Table 9. The student might find interesting that the word “just” has two French equivalent, “simplement” and “uniquement”, and try to see further if those two adverbs are really synonyms in this case or if there are variations in their meaning. This brings us back to our original idea inspired by Tversky that differences can be found only among similar items.

English	French
This is not just a health issue.	Ce n'est pas uniquement une question de santé.
Mental illness is not just a health issue.	La maladie mentale n'est pas simplement une question de santé.
This is not a partisan issue.	Ce n'est pas une question partisane.
This is not just a workplace safety and health violation anymore.	Il ne s'agit plus simplement d' une infraction au règlement sur la santé et la sécurité au travail .
Is it a health issue?	S'agit il d'une question de santé ?

Table 9 English-French bilingual output for learning translation

6. Conclusions and Future work

Our paper focused on presenting the idea of using sentence similarity as a navigation tool for corpora exploration. We first show that there are existing algorithms in the literature that we can use, and second get a sense of which ones provide better results as judged by human evaluators.

The small intrinsic evaluation of the sentence similarity algorithms, as performed by humans, is based on is the definition proposed by Sato (1990). This experiment has shown that algorithms such as Dice Coefficient, Cosine, Lin, for which there were no significant difference in our evaluation, could give access to sentences in the corpus, which humans thought were similar to an input sentence.

This intrinsic evaluation of the technology (sentence-similarity) should now be complemented by an application-based evaluation by teachers and learners. Our insights are that sentence similarity corpus navigation will have many useful usages within language learning and we have given some examples of what we envisage. L2 language learners can know variants of a sentence and better understand grammar rules, and they can learn to write proper expressions by taking reference writings made by native speakers. Students in translation field can learn from practical examples produced by professional translators. However, as for now no evaluation of usefulness within a language learning environment has been done, and how well this system can be used for a language learning purpose needs to be evaluated by experts in the field.

These preliminary results show that the pure string approach can produce acceptable output in terms of finding similar sentences from large corpus. We believe adding simple linguistically oriented processing (such as morphological analysis) on top of it will improve its accuracy and make it much useful system for the language learning purpose. However, such processing would render our algorithm language-dependent, something to consider.

To guide our future work, we would like to quote one of our evaluator's comments about the system: "Traditionally, people use dictionaries or grammar books as tools in learning and translating process. However, even a good dictionary, which collects detailed expressions of a word, cannot tell you how to write a sentence; and a grammar book only tell you the rules how to put words together. You never know from these tools how to write a sentence in a good style and with a living expression. This system can help you to realize these functions. Even at its preliminary stage, it offers different style of writing in a second language after you give what you want to write in your first language. It is worth to continue improving this system, to make it cover multiple languages, to have processing more efficient, and to make outputs friendlier to a general user. "

References

Aslam, J. Frost, M. (2003) An Information theoretic Measure for Document Similarity, In *Proceedings of the 26the Annual International ACM SINGIR Conference on Research and Development in Information Retrieval* (ACM Press) 449-450.

Aston, G. (2001) *Learning with Corpora* (Athelstan).

Bowker, L. (1998) Using specialised monolingual native-language corpora as a translation resource: A pilot study. *Meta*, 43(4), 631-651. Available on-line from http://www.erudit.org/revue/meta/1998/v43/n4/002134ar.pdf (accessed Dec. 9th 2004)

Hersh, W. (2003) *Information Rerieval: A Health & Biomedical Perspective* (Second Edition, Springer-Verlag), chap. 8.

Isabelle, P. Dymetman, M. and Foster, G. (1993) Translation Analysis and Translation Automation. *Proceedings of TMI'93*, 201-21.

Johns, T. (1994) From printout to handout: Grammar and vocabulary teaching in the context of Data-driven Learning. In T. Odlin (ed.) *Perspectives on Pedagogical Grammar* (New York: Cambridge University Press).

Jurafsky, D. and Martin, J. (2000) *Speech and Language Processing* (Prentice Hall).

Lin, D. (1998) An Information-Theoretic Definition of Similarity, In J. Shavlik (ed.) *Proc. 15 the International Conf. on Machine Learning*, San Francisco, CA, 296-304. (Morgan Kauffman)

Matsumoto, Y. Ishimoto, H. and Utsuro, T. (1993) Structural Matching of Parallel Texts, *31st Annual Meeting of the Association for Computational Linguistics*, Columbus, Ohio, 23-30.

McEnery, T. and Wilson, A. (2004) Corpus Linguistics (Edinburgh: Edinburgh University Press).

Nagao M. (1984) A framework of a mechanical translation between Japanese and English by analogy principle, in A. Elithorn and R. Banerji (eds) Artificial and Human intelligence (North Holland), 173-180.

Papineni, K. Roukos, S. Ward, T. Zhu, W. (2002) BLEU: a Method for Automatic Evaluation of Machine Translation, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL),* Philadelphia, 311-318

Salton, G. McGill, M. (1983) *Introduction to Modern Information Retrieval,* McGraw-Hill Book Company.

Sato, S. Nagao, M. (1990) Toward Memory-based Translation, *Proceedings of Coling*, 247-252.

Somers, H. (1999) *Review Article: Example-based Machine Translation*, Machine Translation, vol. 14, 113-157.

Sumita, E. Tsutsumi, Y. (1988) A Translation Aid System Using Flexible Text Retrieval Based on Syntax-Matching, *TRL Research Report.*

Tversky, A. (1977) Features of similarity, *Psychological Review,* vol. 84, no. 4, 327-352.

Appendix 1 Output examples in English and French.

English output example --- Input sentence: We support some aspects of the bill .

---

2 DiceCoefficient:
we support some aspects of the bill .
the alliance supports some of the aspects of the bill .
i will support some of the components of the bill .
we in the ndp support many aspects of the bill unequivocally .

3 Cosine:
we support some aspects of the bill .
we in the ndp support many aspects of the bill unequivocally .
the alliance supports some of the aspects of the bill .
i will support some of the components of the bill .

4 Lin:
we support some aspects of the bill .
we in the ndp support many aspects of the bill unequivocally .
the canadian alliance supports some aspects of the bill .
the alliance supports some of the aspects of the bill .

5 Bleu:
we support some aspects of the bill .
the canadian alliance supports some aspects of the bill .
these are positive aspects of the bill .
these are the positive aspects of the bill .

French output --- Input sentence: Je suis d' accord avec cela .

---

2 DiceCoefficient:
je suis d' accord avec cela .
je suis entièrement d' accord avec cela .
je ne suis pas d' accord avec cela .
je suis d' accord avec eux .

3 Cosine:
je suis d' accord avec cela .
je suis entièrement d' accord avec cela .
je ne suis pas d' accord avec cela .
je suis d' accord avec elle .

4 Lin:
je suis d' accord avec cela .
je suis entièrement d' accord avec cela .
je ne suis pas d' accord avec cela .
je suis d' accord avec lui .

5 Bleu:
je suis d' accord avec cela .
je suis d' accord avec vous .
je suis d' accord avec eux .
je suis d' accord avec lui .