Comparing different approaches to treat Translation Ambiguity in CLIR: Structured Queries vs. Target Co-occurrence Based Selection

Xabier Saralegi and Maddalen López de Lacalle
R & D
Elhuyar Foundation
Usurbil, Spain
e-mail: {xabiers,maddalen}@elhuyar.com

***Abstract*— Two main problems in Cross-language Information Retrieval are translation selection and the treatment of out-of-vocabulary terms. In this paper, we will be focusing on the problem concerning the translation selection. Structured queries and target co-occurrence-based methods seem to be the most appropriate approaches when parallel corpora are not available. However, there is no comparative study. In this paper we compare the results obtained using each of the aforementioned methods, we specify the weaknesses of each method, and finally we propose a hybrid method to combine both. In terms of mean average precision, results for Basque-English cross-lingual retrieval show that structured queries are the best approach both with long queries and short queries.**

Crosslingual information retrieval; structured query translation; co-occurrence statistics

I. INTRODUCTION

The importance of Cross-language Information Retrieval (CLIR) nowadays is patent in multiple contexts. In fact, communication is more global, and the access to multilingual information is more and more widespread within this globalized society. However, unless some lingua franca is established in specific geographic areas and discourse communities, it is still necessary to facilitate access in the native speaker's language.

In our case, we are developing a CLIR system to allow Basque speakers to access texts in other languages. Since Basque has relatively few speakers (about 1,000,000) CLIR is an attractive technology for providing Basque speakers access to those global contexts. Even though lately most extended CLIR approaches are based on parallel corpora, Basque is a less resourced language, and that is why we have to turn our gaze to parallel corpora free approaches. The work presented in this paper compares the performance of two methods for the translation selection problem which do not require the use of parallel corpora. In addition, we have also designed and evaluated a hybrid algorithm that combines both methods in a simple way.

The CLIR topic and its problematic are introduced in the next section. Section 3 addresses the specific problem of the translation selection. The experimental setup is described in subsection A. The two approaches proposed for dealing with the translation ambiguity are presented in subsections B and C. Following (D. subsection), we propose a simple combination of both methods. Then, in Section 4 we evaluate and compare the different methods for the Basque-English pair, in terms of MAP (Mean Average Precision) and using CLEF (Cross Language Evaluation Forum) collections and topics. Finally, we present some conclusions and future works in Section 5.

II. THE TRANSLATION METHODS FOR CLIR

CLIR does not differ too much from Information Retrieval (IR) and only the language barrier requires specific techniques, which are mainly focused on the translation process. The different approaches differ essentially with respect to which available information is translated (queries, documents or both), and in the method used to carry out the translation.

There are three strategies for tackling a cross-language scenario for IR proposes: a) translating the query into the language of the target collection, b) translating the collection into the language of the source query, and c) translating both into an interlingua. The majority of the authors have focused on translating queries mainly due to the lower requirements of memory and processing resources [1]. However, richer context information is useful for dealing with disambiguation problems, and it has been proved that the quality of the translation and retrieval performance improve when the collection is translated, [2]. Translating both queries and documents into an interlingua provides even better results [3], [4].

As for the translation methods, they can be classified into three main groups: Machine Translation (MT)-based, parallel corpus-based, and bilingual Machine Readable Dictionary (MRD)-based. In general, authors point out that using MT systems is not adequate for several reasons: the quality of precision is often poor and the system requires syntactically well formed sentences, while in IR systems the queries are often sequences of words [1].

The corpus-based approach implies the use of parallel (and also comparable) corpora to train statistical translation models [5]. The main problem is the need for large corpora. The available parallel corpora are usually scarce, especially for minority languages and restricted domains. The advantage of this approach is that the translation ambiguity can be solved by translating the queries by statistical translation models. Comparable corpora, which are easier to obtain, can be used in order to improve the term coverage [6].

Lastly, MRD-based translation guarantees enough recall but does not solve the translation ambiguity problem. Thus, two main problems arise when using dictionaries to translate: ambiguities in the translation, and also the presence of some out-of-vocabulary terms. Many papers have been published about these two issues when queries are translated [7], [8], [9].

Among the displayed alternatives, the MRD-based approach has been explored, because of the lack of sufficient parallel corpora for Basque, and because we assume that this situation will be similar for other minority languages. Specifically, we have concentrated on testing two methods to deal with translation ambiguity: structured queries and co-occurrence-based methods. Although the influence level of the errors derived from using dictionaries depends on the quality of the resources used and the tasks done, Qu et al. point out that the wrong translation selection is the most frequent error in an MT-Based translation process [10]. So, we assume that this error distribution will be similar in MRD-based systems.

We have translated only the queries in our experiments. The reasons for this decision are, on the one hand, that the methods we want to analyze have been tested in such an experimental setup. On the other hand, the results of this research will be used for the development of a commercial web searcher, and so the processing and memory consumption are also important factors.

III. Selecting the Correct Translation from a Dictionary

In order to deal with the translation selection problem affecting queries derived from bilingual dictionaries (MRD), there are several methods proposed in the literature. An extended approach to tackle the problem of ambiguity is by using structured queries, also called Pirkola's method [11]. All the translation candidates are treated as a unique token in the calculation of relevances estimating term frequency (*TF*) and document frequency (*DF*) statistics separately. Thus, the disambiguation takes place implicitly during the retrieval instead of during the query formulation. A more advanced variant of this algorithm, known as probabilistic structured queries [12], allows to weight the different translation candidates offering better performance.

Other approaches to tackle ambiguity in query translation are based on exploiting statistically monolingual corpora in the target language. Specifically, these methods try to select the most probable translation of the query, choosing the set of translation candidates that most often co-occur in the target collection. The algorithms differ in the way the global association is calculated and in the translation unit used (e.g., word, noun phrases...):

In [7] a co-occurrence method and a technique using parallel corpora are compared, leading to the conclusion that the co-occurrence method is significantly better at disambiguating than the parallel corpus-based technique. In [13], the basic co-occurrence is extended by adding a decaying factor that takes into account the distance between the terms when calculating their Mutual Information. Hence, if the distance between the terms increases, the decaying factor does too. In the basic co-occurrence model, when calculating the coherence for a translation candidate, not only are the selected translations taken into account, but also those which are not selected. Yi Liu et al., propose a statistical model called "maximum coherence model" that estimates all the translations of all query terms simultaneously and these translations maximize the overall coherence of the query [14]. In this case, the coherence of a translation candidate is independent from the selection of other query terms translations. This new model is compared with a co-occurrence model similar to the one proposed in [8], which takes into account all the translations of the rest of words in the query. The model that they propose performs substantially better, but it is computationally very expensive. Jang et al. propose a co-occurrence method that only takes into account the consecutive terms when calculating the mutual information [15]. Monz and Dorr introduces an iterative co-occurrence method which combines term association measures with an iterative machine learning approach based on expectation maximization [9].

This work compares two alternatives proposed in the literature which do not require parallel corpora. The unique resources used are a bilingual MRD and a corpus in the target language for the co-occurrence-based method, which makes them suitable for less resourced languages like Basque. We have chosen a specific method for each approach: Pirkola's method, and a co-occurrence-based method. Among all the co-occurrence-based algorithms we have chosen the Monz and Dorr's algorithm assuming that being iterative yields better estimations, although we do not have any references that confirm this. In addition, we have designed an algorithm that combines both approaches. In this last case, we have used Darwish and Oard's probabilistic structured queries as a framework and Monz and Dorr's algorithm to estimate the weights of the translation candidates.

A. *Experimental Setup*

The collection used in our experiments is composed by LA Times 94 and Glasgow Herald 95 (CLEF 2001). In the development phase only the LA Times 94 collection was used. We translated from English to Basque two sets of topics: one for development (41-90) and the other one for test purposes (250-350). MAP values are calculated automatically with respect to existing human relevance judgments for queries and documents of the collections. The translation of the topics was carried out by professional translators and correctors of the Elhuyar foundation. The process was done in two steps: firstly, a translator translated the English topics into Basque, and then a corrector corrected the translations in order to minimize the possible bias -and the possible lack of naturalness- caused by the translation process.

We used the Indri [16] as ranking model and the Porter Stemmer both for collections and translated topics. Before applying the proposed translation methods we removed words like *"dokumentuak...(documents)"* and selected the content words manually. Specifically, nouns, adjectives, verbs and adverbs. Postpositions like "*artean (between),* *buruz (about)...*" were also removed. We used a Basque-English MRD which includes 34167 entries. For the treatment of OOV (Out-Of-Vocabulary) words we looked for their cognates in the target collection. Transliteration rules (see Figure 1) were applied and then LCSR (Longest Common Sequence Ratio) was computed. Those which reached a threshold (0.8) were taken as translation candidates in the translation phase.

ph- → f-, phase=fase
-tion → -zio, action=akzio

Figure 1. Example of transliteration rule

B. Dealing with Ambiguous Translations using Structured Queries

The basic idea is to group together the translation candidates of a source word, thus making a set and treating them as if they were a single word in the target collection [11]. Hence, when estimating the term frequency (*TF*) and document frequency (*DF*) statistics for query terms, the occurrences of all the words in the set are counted as occurrences of the same word. We assume that s_i is a query term, D_k is a document term, d is a document and $T(s_i)$ is the set of translation candidate terms of s_i given by the MRD.

$$TF_j(s_i) = \sum_{\{k|D_k \in T(s_i)\}} TF_j(D_k)$$

$$DF(S_i) = \left| \bigcup_{\{k|D_k \in T(S_i)\}} \{d \mid D_k \in d\} \right|$$

where $TF_j(s_i)$ is the term frequency of s_i in document j, and $DF(s_i)$ is the number of document that contain s_i.

If the translation candidates are correct or semantically related, the effect is an expansion of the query. The problem arises especially when wrong translations that are common words occur, because *DF* of the #syn set can take high scores and the correct translation loses weight in the retrieval process. *TF* statistics can also be altered when wrong translations appear in the retrieval documents. But the probability that many wrong translations occur in retrieved documents is low. That is what we call retrieval time translation selection.

In order to test this method the following experiment was carried out in the development phase. First, we have calculated the MAP for different numbers of translation candidates from the MRD (Figure 2), because a high coverage of translations and the precision level of the MRD affects the performance of this method [17]. Moreover, the translation equivalents of source words are usually ordered by frequency use in a MRD. Therefore, we can exploit that order to prune the least probable translations in the interests of query translation precision.

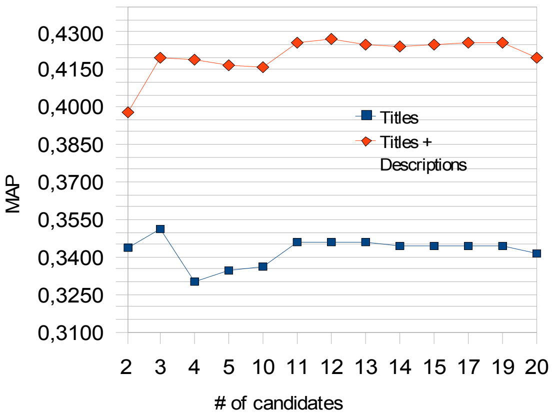

Figure 2. MAP values for different numbers of translation-candidates (41-90 topics)

In the graph (Figure 2), we can see how the number of translation candidates from the MRD accepted for each source word affects the MAP. MAP curves are similar for both titles and titles+descriptions queries. They have local maximum in near points but the global maximum is reached by taking more candidates with the title+description set. The maximum MAP is achieved by taking the first three candidates for short queries, and the twelve first candidates for the long queries. This seems logical because there are more context words that can improve the retrieval-time disambiguation.

C. Target Co-occurrence-based Selection

Define abbreviations and acronyms the first time they are used in the text, e As explained above, structured queries do not really do translation selection, and translations and statistics (*TF* and *DF*) can be wrong in some cases and decrease the retrieval performance. An alternative to executing the translation selection without using parallel corpora is to guide the selection by using statistics of the co-occurrence of the translation candidates in the target collection. The basic idea is to choose the ones that co-occur more frequently, assuming that the correct translation equivalents of query terms are more likely to appear together in target document collection than incorrect translation equivalents. The main problem of this idea is to compute that global correlation in an efficient way, because the maximization problem is *NP-hard*.

The algorithm we have used for the translation selection is the one introduced by Monz and Dorr [9]. Basically, it selects the translation candidates combination which maximizes the global coherence of the translated query by means of an *EM* (Expectation Maximization) type algorithm.

Initially, all the translation candidates are equally likely. Assuming that t is a translation candidate for a query term s_i given by the MRD, then:

Initialization step:

$$w_T^0(t \mid s_i) = \frac{1}{|tr(s_i)|}$$

In the iteration step, each translation candidate is iteratively updated using the weights of the rest of the candidates and the weight of the link connecting them.

Iteration step:

$$w_T^n(t \mid s_i) = w_T^{n-1}(t \mid s_i) + \sum_{t' \in inlink(t)} w_L(t,t') \cdot w_T(t' \mid s_i)$$

where $inlink(t)$ is the set of translation candidates that are linked to t.

After re-computing each term weight they are normalized.

Normalization step:

$$w_L^n(t \mid s_i) = \frac{w_L^n(t \mid s_i)}{\sum_{m=1}^{|tr(s_i)|} w_L^n(t_{i,m} \mid s_i)}$$

The iteration stops when the variations of the term weights become smaller than a predefined threshold.

There are different association measures to compute the association strength between two terms ($w_L(t,t')$). We experimented with Mutual Information and Log Likelihood Ratio, and obtained the best results with the second one. That is the measure we use in the evaluation.

The question is whether by choosing the best translation of each query term we obtain a better MAP than grouping all the translation candidates by means of structured queries. As mentioned before, although in the structured queries some weights and translations can be wrong, an expansion that can benefit the MAP is also produced. For example, for the Basque query "*gene gaitz*", when we select the best English translation "*gene disease*" and run it, we obtain an AP of 0.5046. However, when all the translation candidates given by the MRD are put in sets with the #syn operator, *gene #syn(harm disease flaw ailment hurt malady defect difficult)*, even if we incorporate incorrect translations, we get a greater AP value, 0.5548. So, in this example it is clear that the noise expanded translation gives a higher AP score than the best translation. Nevertheless, for the Basque query *"gose greba"* we construct a translated query like *#syn(hunger yearning desire famine urge ravenous craving famished hungry) #syn(#1(work stoppage) strike walkout)* obtaining an AP of 0.0741. Whereas if we choose the best translation manually, we get the query "*hunger strike"* and obtain an AP of 0.6743. Looking at this example, it seems that our co-occurrence method could provide a margin for improving the MAP compared with structured queries when query terms have many incorrect translation candidates. In order to estimate whether this case is general, a lexicographer manually disambiguated some Basque queries (built from 41-90 CLEF queries) translated into English by an MRD. We preprocessed the queries by keeping only the lemmas of content words and then translated them using the MRD. The work by the lexicographer was to select the best translation candidate for each source term of the queries (Example on Table 1.).

TABLE I. SELECTING THE BEST TRANSLATION OF THE STRUCTURED QUERY

English query	Tainted-Blood Trial
Basque query	kutsatuko odolaren epaia
Basque query (content words)	kutsatu odol epai
Structured translation into English	#syn(pollute impregnate infect) #syn(blood kinship) #syn(sentence crest judgment ridge notch scratch mark cut incision)
Best manual translation	#syn(pollute impregnate infect) #syn(blood kinship) #syn(sentence crest judgment ridge notch scratch mark cut incision)
Best manual translations	#syn(pollute infect) blood sentence #syn(pollute ~~impregnate~~ infect) #syn(blood ~~kinship~~) #syn(sentence ~~crest judgment ridge notch scratch mark cut incision~~)

Then, we calculated the MAP by processing Basque queries (Table 2.) (titles and titles+description separately) for the different translation methods including the manual-based one. The MAP results show the MAP obtained by manual disambiguation does not reach that obtained using structured queries. So it seems that there is no margin for improvement for the co-occurrences-based method. However, the co-occurrences-based method outperforms structured queries when we are dealing with short queries. It even outperforms the theoretical threshold marked by the manual disambiguation. It could be due to a more statistical selection of short queries, more adequate for relevances in that collection.

TABLE II. MAP RESULTS FOR 41-90 TOPICS (DEVELOPMENT SET)

Translation method	**MAP**	
	Titles	***Titles+ description***
English monolingual	0.4639	0.4912
Structured query (3 and 13 candidates)	0.3510	0.4274
Structured query (all candidates)	0.3352	0.4200
Best manual translation	0.3218	0.4127
Co-occurrence-based	0.3564	0.3908
Best manual translations	0.3471	0.4308
Probabilistic structured query	0.3568	0.4268
Probabilistic structured query+threshold (0.8)	0.3594	0.4249

D. Combining Structured Queries and Co-occurrence-based Algorithm

We think that we could take advantage of both techniques. Structured queries contribute to the translation less restrictiveness and query expansion in the retrieval phase, and the co-occurrence-based method contributes translation selection and weighting capability. To do this, we propose that probabilistic structures queries (Darwish and Oard, 2003) [12] be used, and the weights be estimated according to Monz and Dorr's algorithm. Thus, assuming $w_L(D_k \mid s_i)$ as the weight for the translation candidate D_k of a term s_i of a source query s we estimate *TF* and *DF* in this way:

$$TF_j(s_i)=\sum_{\{k|D_k\in T(s_i)\}} TF_j(D_k)w_L(D_k \mid s_i)$$

$$DF(s_i)=\sum_{\{k|D_k\in T(s_i)\}} TF_j(D_k)w_L(D_k \mid s_i)$$

As we did in subsection C, in order to estimate the possible improvement margin of this method, a lexicographer manually removed the wrong translations of the development queries, while maintaining only the correct ones (See Table 1.). We maintained all the possible candidates since this method is capable of selecting more than one candidate. Thus, for the Basque query *"gene gaitz"* (*"gene disease"* on English*)* we obtained a query (*gene #syn(disease ailment malady)*) achieving an AP of 0.5946. A higher score than the one achieved taking all candidates. However, contrary to what we expected, the MAP for 41-90 topics is not much higher than that achieved without doing any kind of selection (although pruning some translations of the MRD can be considered to be a general disambiguation method) for long queries and for short queries it is even worse (Table 2). Therefore, better quality in the translations does not seem to imply a big improvement in MAP. A further analysis will be conducted in the next section.

IV. Evaluation and discussion

The runs were done by taking the titles as queries (short queries), and also by taking the titles and descriptions as queries (long queries) and carrying out Basque to English translation:

1) Monolingual: Titles and titles+descriptions of CLEF 250-350 English topics.
2) First translation: First translation from dictionary
3) Structured query: Group translation candidates from the dictionary in a #syn set using Pirkola's method.
4) Structured query (Optimized dictionary): first translation candidates of the dictionary grouped in a #syn set (three for titles and twelve for the titles+descriptions maximize MAP on development experiments) using Pirkola's method.
5) Co-occurrence-based translation: Best translation selected by Monz and Dorr's co-occurrence-based algorithm.
6) Probabilistic structured query: all translation candidates of the dictionary grouped in a #wsyn set using Darwish and Oard's method, and weighted according Monz and Dorr's co-occurrence-based algorithm.
7) Probabilistic structured query +threshold: Best translations selected according to a threshold and weighted by Monz and Dorr's co-occurrence-based algorithm and grouped by #wsyn set using Darwish and Oard's method.

The results are presented in Table 3.

TABLE III. MAP Values for 250-350 Topics (Test Set)

Run	**MAP**		**% of Monolingual**		**Improvement Over First %**	
	Short	***Long***	***Short***	***Long***	***Short***	***Long***
English monolingual	0.3176	0.3778				
First	0.2118	0.2500	67	66		
Structured query	0.2342	0.2959	74	78	9.56*	15.51*
Structured query (optimized dictionary)	0.2359	0.2960	74	78	10.22*	15.54*
Co-occurrences-based	0.2338	0.2725	74	72	9.41*	8.26*
Probabilistic structured queries+threshold	0.2404	0.2920	76	77	11.9*	14.38*
Probabilistic structured queries	0.2371	0.2941	75	78	10.67	14.99*

The achieved MAP is higher with long queries than with short queries in both cases, monolingual and cross-lingual. In the cross-lingual retrieval the translation methods proposed also offer greater improvement with long queries. This is logical because more context words help in the translation selection. Unlike the results in the development experiments, the methods do not show a different performance depending on the length of the queries. We have examined the queries translated by Monz and Dorr's method and the quality is quite adequate except for a few cases due to false associations. For example, the Basque query *"kutsatu odol epai"* is translated as *"infect blood cut"* by Monz and Dorr's method instead of *"infect blood sentence*". We can assume that it happens due to the stronger relation between *epai* source word's translation candidate and *infect* and *blood* and *cut -epai* source word's translation candidate- than between *infect* and *blood* and *sentence* -another translation candidate for *epai*-. It seems to be because of the the limited representativity of the target collection where some words rarely co-occur. So this could be mitigated by using a bigger corpus. For short queries, too, the hybrid method shows the best results, but statistically does not outperform Pirkolas's method significantly. Pirkolas's method achieves the best results when dealing with long queries. The optimized MRD improves the MAP but not significantly. All improvements that are statistically significant according to the Paired

Randomization Test with α=0.05 are marked with an asterisk in Table 3.

It seems that selecting and weighting translation candidates by means of Monz and Dorr's method in order to include them in structured queries do not imply a significant improvement in MAP terms with respect to Pirkola's method. As in the earlier case, the queries translated by the hybrid method are adequate except for a few cases of false associations. In any case, as we have seen in subsection D, improving the quality of the translation does not always improve the MAP.

TABLE IV. SELECTING THE BEST CANDIDATES FROM THE STRUCTURED QUERY (TOPICS 46 AND 81)

Translation phase	**query**	**AP**
English query (46)	**Embargo on Iraq**	
Basque query (46)	Irakeko bahitura	
Basque (content words)	Irak bahitura	
Structured translation	Iraq #syn(seizure mortgage kidnapping confiscation)	0.2989
Best translations	Iraq #syn(seizure)	0.1302
English query (81)	**The reserve in the Antarctic in which hunting for whales is forbidden**	
Basque query (81)	Baleak ehizatzea debekatuta dagoen Antarktikako erreserba	
Basque (content words)	balea erreserba antarktika ehiza debekatu	
Structured translation	whale #syn(reservation reserve) Antarctica #syn(game hunting prey) prohibit	1.000
Best translations	whale #syn(~~reservation~~ reserve) Antarctica #syn(~~game~~ hunting ~~prey~~) prohibit	0.3333

In our opinion, apart from the query expansion effect and retrieval time selection, another positive effect produced with structured queries is that the weight of some non-relevant terms are smoothed. It is a collateral effect that happens because non-relevant words tend to be common words which inflate the DF statistic. We have examined the differences between AP values corresponding to 41-90 queries (when titles and descriptions are taken) translated by taking all translations of the MRD and by pruning the wrong ones manually. In theory, all the AP values corresponding to each query will be better with the pruned ones. However, there are 6 queries where AP is significantly higher when all translation candidates are taken, despite many of them being wrong (Figure 3).

If we analyze these queries more deeply, we can detect two factors that explain this effect:

1. Wrong translations can turn out to be relevant terms: In the example (46) of Table 4. among all the translation candidates of the Basque source word *bahitura* only *kidnapping* appears in the relevant documents of the collection for that query.
2. Wrong translations can reduce non relevant or noise producer source term weight: in the example (81) of Table 4. Any of the translations of *erreserba* and *ehiza* appear in the relevant documents. Thus, taking all candidates decreases the weight of these irrelevant sets, leading to a better AP score.

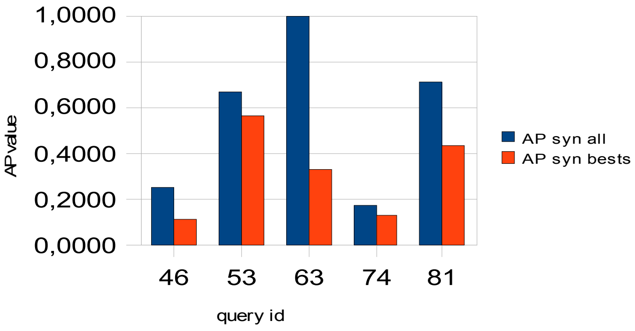

Figure 3. AP values for queries with significantly improved AP when taking all translations candidates

V. CONCLUSIONS

We've seen that query translation guided by MRD is useful for the Basque-English pair. Structured queries seem to be a useful method to deal with translation ambiguity. In fact, this method outperforms significantly both first translation method and selection method based on target collection co-occurrence in terms of MAP. Although the co-occurrences-based method significantly outperforms first translation selection, the translation probabilities used in probabilistic structured queries do not improve the MAP achieved when using simple structured queries. Otherwise, the MAP is close to the MAP of monolingual retrieval (74% and 78% for short and long queries, respectively) applying only the synonymy expansion provided by the dictionary.

REFERENCES

[1] D.A. Hull. and G. Grefenstette, "Querying across languages: a dictionary-based approach to multilingual information retrieval". Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval,p. 49-57. 1996.

[2] D.W Oard, "A comparative study of query and document translation for cross-language information retrieval". In Proceedings of the Third Conference of the Association for Machine Translation in the Americas (AMTA) Philadelphia,PA.P.208-214. 1998.

[3] J. S. McCarley, "Should we translate the documents or the queries in cross-language information retrieval?". Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics, pages 208-214, 1999.

[4] A.Chen and F.C. Gey, "Combining query translation and document translation in cross-language retrieval". 4th Workshop of the Cross-Language Evaluation Forum,p. 108-121.2003.

[5] D. Hiemstra, "Using language models for information retrieval". University of twente.2000.

[6] T. Talvensaari, “Comparable corpora in cross-language information Retrieval”. Thesis .2008

[7] L.Ballesteros and W.Bruce Croft, “Resolving ambiguity for cross-language retrieval”. Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval, p.64–71. 2008

[8] J Gao, JY Nie, E Xun, J Zhang, M Zhou, C Huang, “Improving query translation for cross-language information retrieval using statistcal models”. In Proceedings of the 24th annual international ACM SIGIR conference on Research an development in information retrieval, p. 96-104. 2001

[9] C. Monz and B.J. Dorr, “Iterative translation disambiguation for cross-language information retrieval”. Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval. Pages 520-527, 2005.

[10] Y. Qu, Alla N. Eilerman, Hongming Jin, David A. Evans, “The effects of pseudo-relevance feedback on MT-based”. RIAO 2000. 2000.

[11] A. Pirkola, “The effects of query structure and dictionary setups in dictionary-based cross-language information retrieval”. Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval. P. 55-63.1998

[12] K. Darwish and D. W.Oard, “Probabilistic structured query methods”. Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval.P.338–344. 2003.

[13] J.Gao, J.Y. Nie, H. He, W. Chen and M. Zhou, “Resolving query ambiguity using a decaying co-occurrence model and syntactic dependence relations”. In Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, p.183–190,. 2002

[14] Y Liu, R Jin, JY Chai, “A maximum coherence model for dictionary-based cross-language information retrieval” .Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval, p. 536-543.2005

[15] M. G Jang,., S. H Myaeng,. and S.Y Park, “Using mutual information to resolve query translation ambiguities and query term weighting”. In proceedings of 37th Annual Meeting of the Association for Computational Linguistics, p. 223-229. 1999

[16] H. Turtle Strohman, D. Metzler and W.B. Croft, “Indri: a language model-based search engine for complex queries”. In Proceedings of the International Conference on Intelligence Analysis. 2005

[17] S. Leah, J.A.Larkey, M.E. Connell, A. Bolivar, and C. Wade, UMass at TREC 2002: Cross Language and Novelty Tracks. Ellen M. Voorhees and Lori P. Buckland (Eds.) The Eleventh Text Retrieval Conference, TREC 2002, NIST Special Publication 500-251, pp 721-732. 2003.