Exploring the Effects of Language Skills on Multilingual Web Search

Abstract. Multilingual access is an important area of research, especially given the growth in multilingual users of online resources. A large body of research exists for Cross-Language Information Retrieval (CLIR); however, little of this work has considered the language skills of the end user, a critical factor in providing effective multilingual search functionality. In this paper we describe an experiment carried out to further understand the effects of language skills on multilingual search. Using the Google Translate service we show that users have varied language skills which are non-trivial to assess and impacts their multilingual searching experience and search effectiveness.

1 Introduction

As globalisation and the Internet have facilitated the exchange and accessibility of ideas and information in a variety of languages, the field of Cross-Language Information Retrieval (CLIR) has emerged as an area of focus in the IR community. Many experiments have been conducted under the Cross-Language Evaluation Forum (CLEF[1]), mostly focusing on evaluating the retrieval of news articles from an unknown language collection based on a query submitted in a user's native language.

However, in reality, individuals' needs are not always so simplistic or limited only to this type of situation. There are other scenarios in which cross-language information needs may vary based upon different factors. For example, users may wish to access multilingual material that is not plain text (e.g. web pages or images). Furthermore, individuals can have a range of both passive and active linguistic abilities based on their mother tongue and other languages they may have studied for any length of time.

The present experiment was designed to expand upon previous CLIR research by focusing on the role language skills play in a multilingual web searching context whilst also considering the importance of other factors inherent to the interactive search process (such as user satisfaction.) Participants were asked to find a variety of web pages in three different languages: their native language, one that could be passively understood, and one which was completely foreign. The popular, well-known Google web search engine and associated Google Translate service for search results were chosen as representative systems for testing. Search behaviours, functionalities used, and overall performance were compared in each of the three language conditions. As expected, many of these varied depending on the target language and the type of query submitted. However, the findings provide useful input regarding how

[1] http://www.clef-campaign.org/

best to design cross-language support in information retrieval systems that will meet the real-world needs of a range of end users.

2 Background

Any study examining cross-language search must consider its users' language skills. Unknown and native languages are the two endpoints of a spectrum of language knowledge; foreign language ability can vary greatly within these two extremes. However, it can manifest itself in different ways: Ringbom [1] points out the distinction between (passive) comprehension and (active) production ability. Laufer & Goldstein [2] suggest that this passive/active distinction is too simplistic, and propose a 4-tiered continuum of knowledge strength that includes recall and recognition. According to Gibson & Hufeisen [3], prior knowledge of a language has been shown to assist understanding of an unfamiliar but related one (e.g. German and Swedish.)

As argued by Gonzalo [4], there are two different situations relating to a user's language skills that carry different design implications for cross-language systems. If a user is monolingual, full translation assistance is needed in a CLIR context (e.g. back translation of query terms and document translation). If the user has some passive language skills, then special assistance in describing translations and document content is less likely to be used or desired. Language ability, therefore, is an important variable to consider when designing a system that will cater for a range of users with different needs.

Further past work has focused on user behaviour and these studies have typically addressed either texts or images. In the first category, Zazo Rodriguez et al. [5] examined the effect of users' language abilities on the types of functionalities they used for a question-answering exercise. Compared to individuals with "good" foreign language skills, users with poor skills were found to be more likely to enter queries in their native language and then have it automatically translated to the document language. These "poor" users were also more likely to use and appreciate a functionality which translated the document summaries into their native language.

Petrelli et al. [6] also acknowledged that users are not always monolingual and looked beyond this typical view by investigating how polyglots interacted with a cross-language text retrieval system. However, completely bilingual users with excellent language skills were studied, and thus little insight was given into how the system could have served users with moderate or passive language abilities.

Artiles et al. [7] studied which CLIR functionalities were employed when searching for images with a system that offered three query translation options: no translation, automatic translation, and assisted translation (where the machine translated result could be viewed and edited). Translation was typically selected in cases where the search was precision-oriented and geared towards finding something specific. Overall, the assisted translation mode was the most popular, although the possibility of changing translation was largely unexploited (perhaps partly due to the tasks assigned).

Research by Kralisch & Berendt [8] found that the linguistically-determined cognitive effort involved in processing information in a foreign language can be mediated or lessened in cases where domain knowledge is high. Similarly, Gaspari [9] asserted

that some users may understand specialized terms relevant to their field of interest, even if their general foreign language ability is limited..

Other studies have looked at how users interact with cross language functionality, even if language skills are not explicitly considered as a variable. For example, Dorr et al. [10] noted that giving users control over the translations produced by the machine led to a more satisfying overall experience. However, it was also found that "translation selection provides an extra interaction anchor that users found helpful; however, its effect is not as great as query reformulation." What this study did not examine, however, was the role that knowledge (or lack thereof) of the target language played and how this could have affected users' behaviour in this case.

To examine the best way of displaying machine translations to a monolingual user, He et al. [11] tested two different approaches: pure back translations and more contextual translations (showing the keyword in the context of a sample sentence). Overall, the potential utility of each approach was deemed to depend on factors such as the characteristics of the topic, the collection, and the translation resources. Even if query translation is offered, it may not necessarily be used if it is not perceived as providing some benefit. For example, research conducted as part of the European TRANSLIB project revealed that "people made little use of the title translation capabilities in TRANSLIB because they tended to use the system only to find documents in languages that they could read." (cited in [12]).

Many of the aforementioned studies focused on individuals searching for text-only articles. Web pages are different from texts because they often contain images or other cues to help provide additional (non-verbal) information about the content. Little is known about how people may conduct cross-lingual search using mainstream Web-based systems, especially in a variety of languages; hence, these areas will be the focus of the present investigation.

3 Methodology

12 participants (3 female and 9 male) were involved in the study. They were predominantly computer science postgraduate students or researchers with a mean age of 30 years. To investigate the influence of cross-linguistic similarity, individuals with Romance language skills were specifically recruited. Because these languages share a common origin, it was highly likely that each participant had some latent, interchangeable passive knowledge.

Five of the participants were native or near-native speakers of Spanish, four of Portuguese, two of French and one of Italian. Before beginning, participants completed a questionnaire relating to search engine use and reading/writing ability levels in all languages they could speak. Based on this, a person's active, passive, and unknown languages were determined. An active language (L1) was counted as their native language or a language spoken at near-native fluency. A passive language (L2) was defined as a Romance language similar to the individual's L1, but for which their reading/writing self-rated abilities were "beginner" or below. The unknown language (L3) was a language the participant could not understand, selected at random from the possibilities of German, Japanese, and Russian.

The Google Translate "search results" translation service[2] was used for these experiments. It was chosen over other similar systems because Google's search engine draws upon a large index, and its widespread use means it is familiar to most individuals. This system provides a wide range of functionalities, including automatic query translation, snippet translation, web page translation, and possibility of viewing and editing the query's translation. As such, it provides the set of "ideal" cross-language search functionalities advocated by Zhang & Lin [13].

Participants were assigned tasks based on the following scenario: "You are a high school teacher and you have three new students entering your class. They speak no English so you would like to find some web pages in their own languages for them. For each topic, please find and bookmark 3 pages which you think are relevant. You will have a maximum of 5 minutes for each topic, after which you will be asked to move on to the next topic. To find this information, you may either use the Google search engine (including any localised versions, e.g. google.es) or you may use the Google Translate site." None of the participants had used Google Translate before; therefore the basic functionalities and features of the site were demonstrated to them beforehand.

There were 12 topics in total (4 for each language). All were related to cultural heritage, ranging from proper names and titles to more general subjects and fairly specific terms. Half of the topics were considered "hard" for translation (that is, they were incorrectly translated by Google Translate), and half considered "easy" for translation. "Hard" topics were not always identical across languages because the automatic translation system did not make the same types of mistakes in all languages. Nonetheless, types of errors leading to "hard" (incorrectly translated) queries had various characteristics corresponding to three main categories of "performance issues" in CLIR (cf. [14]): lack of coverage (out of vocabulary terms - e.g. Etruscan,) translation ambiguity (Hamlet being translated as "small village" instead of the title of a play) and incorrect translation of phrases ("still life" translated word-for-word).

The language orders and the task-language combinations were assigned based on a Latin square arrangement, with 2 hard and 2 easy topics for each language. After each set of 4 questions (one language set), the participants filled out a brief questionnaire to assess their difficulty of the task and confidence with finding relevant sources for each topic. At the end of the experiment, participants filled out a language test for their "passive" language to assess the correspondence with their self-reported levels. They were also asked to assess the overall process.

4 Results

4.1 Languages Used for Web Search

Except for one individual, none of the participants were native speakers of English. However, they reported using English to search on the Internet between 48 and 95% of the time (mean 75.5%). This may be because all participants were currently studying or working in the UK and therefore may have needed to, grown accustomed to,

[2] http://www.google.com/translate_s?hl=en

felt more confident, or had more success using English to search on a regular basis. Responses indicated that users predominantly search in English or their native language, using other foreign languages relatively infrequently.

Foreign language abilities in reading and writing were self-reported on a scale from 1 (beginner) to 5 (advanced.) Across all responses in all languages, the mean value of reading skills (3.59) was slightly higher than that of writing skills (2.96), suggesting that people judged themselves to be better at reading than at writing (this difference was not statistically significant).

4.2 Sites and Functionalities Used

There are two ways of evaluating interactive retrieval systems, according to Oard [12]: qualitatively and quantitatively. The former approach aims to observe users' behaviour (e.g. do they use a feature in the way envisioned?) and to make generalisations based upon this. The latter approach asks and measures specific questions (e.g. do users perform some task more quickly when using a new feature?) Both methods were used in this study. One type of qualitative observation made focused on which of the functionalities were used in the various conditions.

Table 2. Number of topics searched with each site, by language

	Google only	**Google Translate only**	**Google and Google Translate**
L1	35 (72.9%)	4 (8.3%)	9 (18.8%)
L2	9 (18.8%)	15 (31.3%)	22 (45.8%)
L3	2 (4.1%)	26 (54.2%)	20 (41.7%)

In general, as language unfamiliarity increased, the use of Google Translate increased. Many searches were conducted with a combination of Google and Google Translate (Table 2). Often, participants switched from one to the other after a few unsatisfactory query modifications, thinking that the second system would yield different results (in reality there was no difference; Google Translate results were the same as those obtained from using Google).

As shown in Table 3, reliance upon query translation functionalities increased with language unfamiliarity: users were more likely to look at the translated versions of pages for L3, and the original versions for L2. Query editing occurred only 3 times out of all 144 topics, and these were exclusively in the L2 condition. In this case, based on the tools available, users were much more likely to reformulate or edit the query in the source language than to deal with the machine translation, a behaviour also noted by Dorr et al. [10].

Table 3. Frequency of use of Google Translate functionalities for each topic, by language

	Query translation	**Translated query editing**	**Original links viewed**	**Translated links viewed**	**Both links viewed**
L1	13 (27.1%)	0	4 (8.3%)	1 (2.1%)	1 (2.1%)
L2	37 (77.1%)	3 (6.3%)	26 (54.2%)	2 (4.2%)	4 (4.2%)

L3	46 (96.0%)	0	14 (29.2%)	19 (39.6%)	9 (18.8%)

4.3 Performance Measures

The following quantitative measures were used to assess user performance on the tasks:

1. Relevant Items: the number of pages bookmarked (0-3) and
2. Time: the length of time taken to do so.
3. Modifications: the number of times the query was modified (something else was typed into the search box) per task
4. Links viewed: the number of page links selected (in original language and in target language)
5. Percent Chosen: the number of links bookmarked as a proportion of total links clicked on.
6. Success: a number calculated in order to give a relative indication of how easy the task was, determined by dividing the number of bookmarks by time (a higher number means the person was more "successful" at completing the task).
7. Difficulty: a rating of task difficultly supplied by the user (this referred to all four searches for a given language) (1=very difficult, 7=very easy).
8. Confidence: a rating of user confidence that sources found were relevant (1=not at all confident, 7=very confident).

Perhaps not surprisingly, more people successfully completed the task of book marking three pages in the L1 condition (67%) as opposed to the L2 and L3 conditions (33% and 19%, respectively.) Within each language, more bookmarks were made for the easy topics than the hard topics. However, nearly 30% of participants found three bookmarks they felt were relevant using easy queries in the L3 condition.

4.3.1 Language Effects

A one-way ANOVA was carried out to determine the effect of search language on the quantitative measures mentioned above (Table 4). The tasks in L1 were self rated as significantly easier than those of L2, which were in turn rated as significantly easier than those of L3. The significant differences between the language groups with respect to mean values for relevant items found, time, percent chosen, success, and confidence were between L1-L2 and L1-L3. The differences in time and success seem to be in accordance with findings by Kralisch & Berendt [8] that non-L1 information processing requires more cognitive effort than L1 information processing.

Table 4. Effects of search language on various measures.

Measure	Language	Mean	Significance
Relevant Items*	L1	2.458ab	a =.001
	L2	1.646a	b=.000
	L3	1.437b	
Time*	L1	3.985ab	a=.002
	L2	4.584a	b=.000

	L3	4.766b	
Modifications	L1	1.655	
	L2	2.000	
	L3	2.313	
Links Viewed	L1	4.479	
	L2	4.333	
	L3	3.812	
Percent Chosen*	L1	0.588ab	a=.006
	L2	0.395a	b=.010
	L3	0.408b	
Success*	L1	0.7193ab	a=.000
	L2	0.4077a	b=.000
	L3	0.3433b	
Difficulty*	L1	5.667ab	a=.000
	L2	4.000ac	b=.000
	L3	3.167bc	c=.003
Confidence*	L1	5.958ab	a=.000
	L2	4.213a	b=.000
	L3	4.106b	

*differences significant at p>.01

4.3.2 Effects of Topic Difficulty

An independent samples t-test was performed to compare mean results between easy and hard topics (see Table 5.) Significant differences were found between these two groups with respect to the number of pages bookmarked, number of query modifications made, success, and confidence. The significantly reduced number of modifications made for easy queries corresponds with an assertion by Och et al. [15] that better quality machine translations result in reduced post-editing effort.

Table 5. Effects of topic difficulty on various measures.

Measure	**Topic type**	**N**	**Mean**	**Sig. (2-tailed)**
Relevant Items	E	76	2.105*	.003
	H	68	1.559*	
Time	E	76	4.310	.061
	H	68	4.596	
Modifications	E	76	1.526*	.000
	H	68	2.500*	
Links Viewed	E	76	4.461	.065
	H	68	3.927	
Percent Chosen	E	76	0.506	.078
	H	68	0.414	
Success	E	76	.5678*	.016
	H	68	.4033*	
Difficulty	E	76	4.408	.303
	H	68	4.132	
Confidence	E	76	5.461*	.000
	H	68	3.970*	

*differences significant at p>.05

The effect of topic difficulty on confidence was also significant on the results for L1, L2, and L3 when analysed separately using independent samples t-tests (see Ta-

ble 6). Within each language, users were significantly more confident with the results they found for the easy queries as opposed to the hard queries. This easy-hard distinction also emerged, surprisingly, in the L1 condition (in which occurrences of query translation were much lower). Since Google Translate exploits the web as a parallel corpus, perhaps what helps to make a query easily translatable or not is influenced by the number of pages available on that topic. If the hard topics were less well-represented even in English, then the likelihood or speed of finding relevant results could be reduced compared to more popular, "easy" topics. There was no significant interaction between language and difficulty.

Table 6. Mean confidence rating for easy vs. hard topics, by language.

Language	Topic type	N	Mean	Sig. (2-tailed)
L1	E	28	6.393*	.007
	H	20	5.350*	
L2	E	24	5.083*	.002
	H	23	3.304*	
L3	E	24	4.750*	.024
	H	23	3.435*	

4.3.3 Quantitative Measures from Final Questionnaire

The average mean ratings of the usefulness of the three translation aids offered by Google Translate ranged from 3.90 for query editing to 5.08 for query translation to 5.50 for translated snippets (with 7 being most useful.) The rated usefulness of the various features corresponded with their frequency of use (as shown in Table 3). That is, since both query and snippet translations were actually used more often than query editing, it is not surprising that they were also rated as more useful.

Table 7. Frequency of reported most useful functionality for each language in experiment.

	Query translation	Translated snippets	Query editing
L1	1	4	5
L2	5	5	1
L3	5	5	1

The usefulness ratings of the various functionalities (query translation, snippet translation, and query editing) varied based on the language being considered. Obviously, query translation was not viewed as particularly necessary when searching in L1. It is unlikely that the other two features were useful either, but they were probably chosen just by default. For non-native languages L2 and L3, majority opinion on the most useful feature was split equally between query translation and translated snippets (Table 7). The mean usefulness ratings of proposed additional functionalities (dictionary support and greater control over the query) were 6.25 and 5.41, respectively. However, dictionary support with back translations or pictures was viewed as more helpful than just showing the alternative translations in the target language with no further explanation or assistance.

4.3.5 Comments and Observations

Responses to the post-experiment question: "Would you use Google Translate again? Why or why not?" revealed three common attitudes:

1. Negative - not necessarily a useful tool (4) "Interesting tool but not sure I need to use it", "Just when no other means to get the information are available."
2. Positive with reservations – translation quality could be improved (5) "I'm not quite sure about how accurate the translations were." "It's not very reliable when doing translations"
3. Positive - helpful in some situations (2) "Useful to translate words into different languages" "The searching environment is very useful"

What emerges here is a feeling that many people could not normally envision a reason to use a system like Google Translate. However, none of them were aware of the system's existence prior to using it in the study. They may have not realized its potential use in some situations (i.e. when planning foreign travel, to broaden the scope of a search, etc). The experiment by its very nature created a somewhat artificial, restricted situation in which users were only allowed to use two specific sites. It is unclear to what extent they would voluntarily use (or need to use) Google Translate in their everyday search behaviour. Further research could examine this question in a more open-ended and naturalistic context.

Other observations made of search behaviour indicate that people using a machine translation system expect it to operate in the same way as an ordinary search engine (with regards to query syntax and formulation). In the case of Google Translate, this was not so. Adding quotation marks to mark phrases and refining queries with supplementary terms, while conventions for Google search, did not have a positive effect with the machine translation system. Some users employed creative strategies to find information when the Google Translate results were unsatisfactory (e.g., using the localised pages in Wikipedia.)

This suggests that the automatic translation, while beneficial, still produces some errors (and this was recognized by the users). The means of dealing with these errors was not sufficiently developed in Google Translate for the "hard" queries, leading to a lower level of search precision. It should be noted that shortly after the present experiment was conducted, a dictionary service was added to the Google Translate pages to allow the lookup of words or phrases in a limited set of language pairs. This no doubt can help the user to identify the correct translation for their query. However, the dictionary service is located in a separate tab and thus is still not as user-friendly as it could be if it were integrated into the main "translated search" interface itself or integrated into the search service to automatically display alternative translations in a seamless fashion.

5 Discussion and Design Implications

The use of available functionalities often depended on the foreign language condition:

Passive Language (L2)	Unknown Language (L3)
• Some (limited) use of translated query editing • Frequently able to view snippets/ web pages in original language (L2); document translation less necessary	• Translated query editing not useful • More likely to view automatically translated versions of snippets/web pages • Search perceived as more difficult than L2

It therefore appears that a query editing feature is not very useful in these situations. Users with passive reading skills still struggle to produce language themselves. Even if they could identify a translation as erroneous, they would not necessarily know how to correct it. This implies that suggestions of alternatives are needed when a query is ambiguous or incorrect. Different functionalities are needed for different levels of language knowledge; therefore, it is important to provide a range of tools to cater to these differing conditions. Alternatively, frequent users could create a set of tools and preferences to personalize their experience based on language profiles.

Observation of the types of translation errors that made some queries "hard," combined with participants' feedback, highlights the importance of dictionary support when using machine translation. The concept of depicting dictionary terms pictorially (language-independently) is a novel idea that warrants further investigation.

In general, users employed the same strategies for L2 and L3 search as they did for L1 search. If the initial results set did not seem adequate, they tended to employ the same strategies and mental models in all three conditions; namely, by adding quotation marks around words they wished to translate together, and by modifying queries by adding supplementary terms. However, these either had no effect or had a negative effect on the searches conducted with the Google translation system, sometimes unbeknownst to the users. For individuals searching in a completely unknown language, providing query translation and document translation are essential. Query translation is still important when an individual has passive knowledge of the target language, although the need for document translation may be reduced. Overall, it was encouraging that many measures were not significantly different between L2 and L3; given the appropriate tools for assistance, people can still find basic relevant information in a foreign language, be it partially or completely unknown.

6 Conclusions

This study expanded upon previous work in cross-language information retrieval by examining the effect of language skills on web search behaviour using Google Translate as a sample web-based query and results translation system. Whereas the majority of CLIR-based research has focused only on how people retrieve material in unknown languages, the present study indicates that many individuals also have passive language skills. They behaved closer to native language ability when using a passive language as opposed to one which was unknown, although these differences were not

statistically significant. Overall, as might be expected, the perceived and actual difficulty of the task increased as language unfamiliarity increased. However, the accuracy of the query translation also seemed to have an effect across all the language conditions, so that it was harder to find relevant information (in any language) for queries that were incorrectly translated by the system. The problem of incorrect translations was further compounded when queries were modified by adding extra terms.

One limitation of the study may have been the five minute time limit placed upon each task. While this was put into place to keep the experiment down to a reasonable length, some users felt it was "artificial" and it may have caused them to feel pressure and led them to bookmark some less appropriate sites just to feel that they were able to complete the task in time. Whilst Google Translate was clearly able to provide enough support to help participants locate at least some relevant material in both passive and unknown languages, there are ways in which it (or any similar cross-language searching system) could be improved. Aside from creating translation systems that produce fewer mistakes, it would be beneficial to offer: (1) phrase recognition and translation (either automatically detected or manually indicated) and (2) integrated dictionary support to identify alternative translations for ambiguous terms, and some means of displaying these in a way that is understandable to the user.

As web pages were being searched for, the associated pictures and visual cues also helped the participants to make relevance judgments. Future work could focus on the cross-language functionalities that would assist users searching for other types of media (e.g. images or video), as these may differ from those used in a purely text-based situation. Overall, the present experiment provided insight into the behaviours and strategies of individuals searching for material in a variety of languages. Findings can help to influence ways in which future systems can be designed to provide personalized and tailored searching support based on an individual's varying abilities and cross-language needs.

Acknowledgments

Work partially supported by European Community under the Information Society Technologies (IST) programme of the 6th FP for RTD - project MultiMATCH contract IST-033104. The authors are solely responsible for the content of this paper. It does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of data appearing therein.

References

[1] Ringbom, H.: Cross-linguistic similarity in foreign language learning. Multilingual Matters, Clevedon (2007)

[2] Laufer, B., & Goldstein, Z.: Testing vocabulary knowledge: Size, strength and computer adaptiveness. Language Learning, 54(3) (2004) 399-436.

[3] Gibson, M., & Hufeisen, B.: Investigating the role of prior foreign language knowledge: Translating from an unknown into a known foreign language. In:

Cenoz, J., Hufeisen, B., & Jessner, U. (ed.): The Multilingual Lexicon. Kluwer Academic Publishers, Dordrecht (2003)

[4] Gonzalo, J.: Scenarios for interactive cross-language information retrieval systems. Proceedings of SIGIR 2002 Workshop on Cross-Language IR (2002).

[5] Zazo Rodriguez, A.F., Figuerola, C., Alonso Berrocal, J.L., & Fernandez Marcial, V.: iCLEF 2005 at REINA-USAL: Use of free on-line machine translation programs for interactive cross-language question answering. (2005).

[6] Petrelli, D., Levin, S., Beaulieu, M., & Sanderson, M.: Which user interaction for cross-language information retrieval? Design issues and reflections. JASIST 57 (2006) 709-722.

[7] Artiles, J., Gonzalo, J., Lopez-Ostenero, F., & Peinado, V.: Are users willing to search cross-language? An experiment with the Flickr image sharing repository. (2006).

[8] Kralisch, A., Berendt, B.: Language-sensitive search behaviour and the role of domain knowledge. New Review of Hypermedia and Multimedia 11 (2005) 221-246.

[9] Gaspari, F.: Online machine translation services and real users' needs: An empirical usability evaluation. LNCS 3265 (2004) 74-85.

[10] Dorr, B., He, D., Luo, J., & Oard, D.: iCLEF at Maryland: Translation selection and document selection. In C. Peters (Ed.), Working Notes for the CLEF 2003 Workshop (2003).

[11] He, D., Wang, J., Oard, D., & Nossal, M.: Comparing User-Assisted and Automatic Query Translation. LNCS 2785 (2003) 400-415.

[12] Oard, D.: Evaluating interactive cross-language information retrieval: Document selection. Lecture Notes in Computer Science 2069 (2001) 57-71.

[13] Zhang, J., & Lin, S.: Multiple language supports in search engines. Online Information Review, 31(4) (2007) 516-532.

[14] Hull, D., & Grefenstette, G.: Experiments in multilingual information retrieval. Proceedings of 19th ACM International Conference on Research and Development in Information Retrieval (SIGIR '96) (1996) 49-57.

[15] Och, F., Zens, R, & Ney, H.: Efficient search for interactive statistical machine translation. Proceedings of 38th Annual Meeting of Association for Computational Linguistics (ACL) (2003) 440-447.