MIRACLE's hybrid approach to bilingual and monolingual Information Retrieval

José M. Goñi-Menoyo[1], José C. González[1,3], José L. Martínez-Fernández[2,3],
Julio Villena-Román[2,3], Ana M. García-Serrano[1], Paloma Martínez-Fernández[2],
César de Pablo-Sánchez[2], Javier Alonso-Sánchez[3]
[1] Universidad Politécnica de Madrid
[2] Universidad Carlos III de Madrid
[3] DAEDALUS - Data, Decisions and Language, S.A.
jmg@mat.upm.es, jgonzalez@dit.upm.es, agarcia@dia.fi.upm.es
joseluis.martinez@uc3m.es, julio.villena@uc3m.es, cesar.depablo@ uc3m.es, paloma.martinez@uc3m.es
jalonso@daedalus.es

Abstract

The main goal of the bilingual and monolingual participation of the MIRACLE team at CLEF 2004 was testing the effect of combination approaches to information retrieval. The starting point is a set of basic components: stemming, transformation, filtering, generation of n-grams, weighting and relevance feedback. Some of these basic components are used in different combinations and order of application for document indexing and for query processing. Besides this, a second order combination is done, mainly by averaging or by selective combination of the documents retrieved by different approaches for a particular query.

1. Introduction

The MIRACLE team is constituted by three university research groups located in Madrid (UPM, UC3M and UAM) along with a company, DAEDALUS, started up in year 1998 as a spin-off from two of them. DAEDALUS is a leading company in linguistic technologies in Spain[1], and acts as the coordinator of the MIRACLE team. This is the second participation in CLEF, after year 2003 [6] [7]. Besides bi and monolingual tasks, the team has participated in the ImageCLEF and Q&A tracks.

The main purpose of the bi and monolingual participation was testing the effect of combination approaches to information retrieval. The starting point is a set of basic components: stemming, transformation (transliteration, elimination of diacritics and conversion to lowercase), filtering (elimination of stop and frequent words), generation of n-grams, weighting (giving more importance to titles) and relevance feedback. Some of these basic components are used in different combinations and order of application for document indexing and for query processing. Besides this, a second order combination is done, mainly by averaging or by selective combination of the documents retrieved by different approaches for a particular query. When evidence is found about better precision of one system at one extreme of the recall level (i.e. 1,0), complemented by better precision of another system at the other recall end (i.e. 0,0), then both of them are combined to benefit from their complementary results.

On the other hand, our group has been developing during the last year an indexing system based on the *trie* data structure [5]. Tries [4] are successfully used by the MIRACLE team for an efficient storing and retrieve of huge lexical resources, combined with a continuation-based approach to morphological treatment. However, the adaptation of these structures to manage efficiently document indexing and retrieval for commercial applications has been a hard task. The currently available prototype shows a strong improvement of performance (both indexing and retrieval times are considerably reduced). However, this system was not fully operative for this CLEF campaign. So, the Xapian [3] indexing system, robust, efficient, and well suited for our purposes, was used as in the last campaign.

---

[1] DAEDALUS clients include leading companies in different sectors: media (EL PAÍS), publishing (Grupo SM), telecommunication (Grupo Telefónica), digital rights management (SGAE), photography (StockPhotos) and the reference institution for the Spanish language, Instituto Cervantes. Its portfolio of solutions includes STILUS® (professional spell, grammar and style checking of texts in Spanish), K-Site® (information retrieval, fuzzy search and knowledge management), LUCAS (universal locator of audiovisual contents, an Internet spider), etc.

For this year, we have submitted runs for the following tracks:

a) Monolingual Russian.
b) Monolingual French.
c) Bilingual Dutch to French.
d) Bilingual German to French.

2. Description of the Tools in MIRACLE's Tool Box

The Xapian system has been the basic indexing and retrieval tool for bilingual and monolingual experiments for the MIRACLE group. Before being indexed, document collections have been pre-processed using different combinations of scripts, each one oriented to a particular experiment. For each one of these, topic queries have been also processed by the same combination of scripts (although in some cases some variants have been used, as will be described later).

The baseline approach to processing documents and topic queries is composed of the following sequence of steps:

1. **Extraction:** Ad-hoc scripts are run on the files that contain particular documents or topic queries collections, to extract the textual data enclosed in XML marks. We used all the marks permitted for automatic runs (depending on the particular collection, all of the existing TEXT, TITLE, LEAD1, TX, LD, TI, or ST for document collections, and the contents of the TITLE and DESC marks for topic queries – NARR marks contents were systematically ignored). The contents inside these marks were concatenated to feed the following steps. However, in some experiments only the titles were extracted (including in the run identifier the strings *titnormal, titnostem* or *titngrams*), and in some *normal* experiments in monolingual Russian, the terms appearing inside the TITLE marks were given more importance by repeating them several times more (these experiments include in the identifiers the strings *normaltit1, normaltit2* or *normaltit3,* when the titles terms are included one, two or three times more).
2. **Parsing:** A simple parsing process is made for eliminating punctuation signs and detecting basic indexing chunks (usually words, by some basic entities can be detected, as compounds, proper nouns, and so on). We think that the quality of this step is of paramount importance in all the document processing. A high-quality entity recognition (proper nouns or acronyms for people, companies, countries, locations, and so on) could improve the precision and recall figures of the overall retrieval, as well as it could a proper recognition and normalization of dates, times, numbers, etc.
3. **Lowercase words:** All document words are normalized by converting all uppercase letters to lowercase.
4. **Stopwords filter:** All the words known as stopwords are eliminated from the document. Stopwords in the target languages were initially obtained from [1], but were completed using other several sources and using own knowledge and resources.
5. **Stemming:** The process known as stemming is applied to each one of the words of the document. The stemmer used is the one referenced in [2].
6. **Remove accents:** All document words are normalized by eliminating accents in stemmed words. Note that this process can be done before stemming, but resulting *lexemes* are different. In spite of that, some experiments have been made doing this step before stemming.
7. **Final use:**
    a. **Indexing:** When all the documents processed by the former steps are to be indexed, they are fed to a Xapian ad-hoc front-end to build the Xapian document database.
    b. **Retrieval:** When all the documents processed by the former steps are topic queries, they are fed to a Xapian ad-hoc front-end for searching the previously built Xapian index. In our experiments for this year we have only used OR combinations of the search terms.

In the case of the Russian language the basic processing steps described above is slightly changed, due to the different encodings of the Russian files and the resources used for Russian: while document collection and topics files were encoding in UTF8, as well as stopwords resources, the stemming resources worked in KOI8, so some recoding steps were added in appropriate processing points. In addition to that, some other tools did not work properly with the UTF8 encoding, so some workarounds had to be added: (a) The parsing process was simplified even more, using a *sed* script to achieve basic punctuation processing, and (b) a transliteration of the files to the

ASCII charset was needed in order to get the XAPIAN indexing system to work. The transliteration script used was the one available in reference [1].

In addition to the baseline experiments (identified with the suffix *normal* in the run identifiers), other experiments have been also defined as variations of these: If the stemming step is not made, we identify the resulting experiments with the suffix *nostem*, where the actual word forms appearing in the documents are used for indexing and retrieval. We also tried a variant of the *nostem* experiments, where we obtained particular *n-grams* from each of the actual word forms in the documents. We denoted these with identifiers of the form *ngramsXY*, where X is the length of the n-grams and Y the number of characters that overlap between two consecutive n-grams. (For example, in an experiment denoted by an identifier with the suffix *ngrams54*, from *president* we would obtain the n-grams: "*_pres*", "*presi*", "*resid*", "*eside*", "*siden*", "*ident*", and "*dent_*". The symbol "_" is introduced to denote word boundaries. Note that four characters overlap between two consecutive n-grams).

In the case of the topic queries, an additional variation is introduced: the FW (*Frequent Words*) filter is applied by filtering out of the queries the 20 most frequent words, or stems, that appeared in the corpora, as well as some typical query terms. These variants were identified by using the *FW* string in the run identifier.

The Xapian engine allowed us to use relevance feedback, so we used such technique in several experiments. When the terms of the first documents retrieved in the first retrieval step are fed back to a second retrieval step, we used the strings *R1*, *R2*, *R3*, *R4* or *R5*, in the run identifier depending on the actual number of documents used. Note that using relevance feedback does not affect to an indexing process, and can be used in any of the variants used for processing the documents collections or the topics queries.

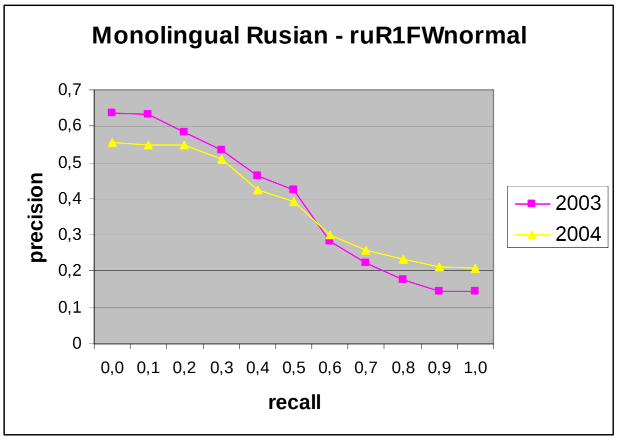
For translation purposes, the SYSTRAN system was used. Our tests done on the collections and topics of CLEF 2003, showed that SYSTRAN outperformed other on-line translators on the selected pairs of languages (Dutch to French and German to French) when used to find documents in the French collections from queries in Dutch or German. As other pairs as Finnish and Swedish to French where not available on-line in SYSTRAN, other translators where tested, with very poor results.

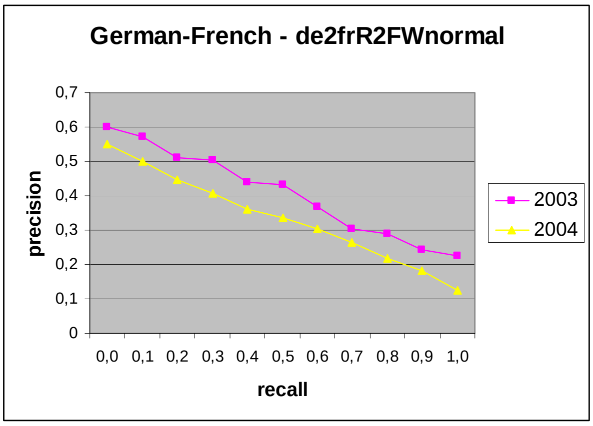
3. Description of the Baseline Experiments

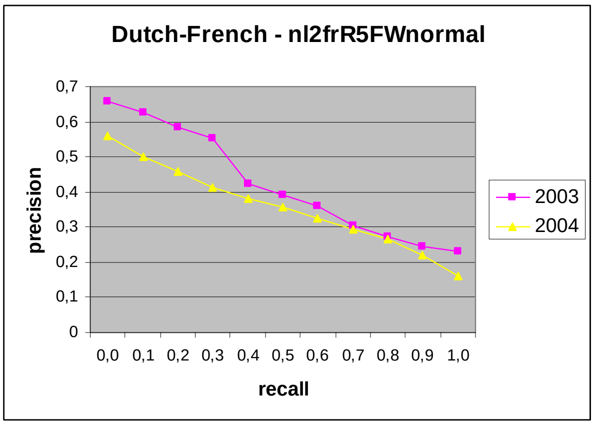
Not all the possible combinations of the variants described in the previous section were tried in the experiments, due to evident limitations of resources and time. The experiments were tried in a rather intuitive, non-systematic way, trying to test a wider and richer set of trials. To compare these approaches, we used these techniques following the instructions given for CLEF 2003 (corpora and topic queries) and using the appropriate *qrels* available at the beginning of this campaign. The appendix includes all the data that compares the results that we obtained in the experiments. The tables show the precisions at recall points 0 and 1, the average precision, the percentage of the latter with respect to the best average precision experiment (the first one in each table) for each of the experiments. The best value is marked with the symbol "*". The *comb* column in each table indicates if the experiment is a combined experiment, what will be described in the next section, and the *sel* column shows what experiments were selected for CLEF 2004, usually the ones with a better result in precision (regarding CLEF 2003 experiments).

The appendix also includes the results for the same experiments for the CLEF 2004 campaign (the relevant tables mark the experiments submitted in the *sub* column), once the *qrels* for this campaign have been available.

In the following figures, the results obtained by the best baseline experiments submitted to CLEF 2004 are compared with the results obtained by exactly the same system when applied to the 2003 tasks. The comparison shows qualitative differences between the 2003 and 2004 topics. No figure is presented for French, as all the submitted runs were in this case obtained through combination (see the next section).







4. Description of Combined Experiments

In this campaign, we have made some tests that try to combine the results from the basic experiments in different ways. We thought that to some extent, the documents that have a good score in almost all experiments are more suitable to be relevant that other documents that have good score in one experiment but a bad one in others. We have tried two strategies for combining experiments:

Average: The relevance figures obtained in the Xapian probabilistic retrieval in all the experiments to be combined for a particular document in a given query are added. This approach combines the relevance figures of the experiments without giving more importance to a particular experiment.

Asymmetric DWX combination: In this particular type of combination, two experiments are combined in the following way: The relevance of first D documents for each query of the first experiment is preserved for the resulting combined relevance, whereas the relevance for the remaining documents in both experiments are combined using weights W and X. We have only ran experiments labeled "101" and "201", that is, the ones that get the one (or two) documents more relevant from the first basic experiment and all the remaining documents retrieved from the second basic experiment, then re-sorting all these results using the original relevance.

Average combinations get better figures in average precision or in precisions at 0 or 1 points of recall, than the original basic experiments. The reason could be that *good* relevant documents that appear with a high score in the experiments combined are strengthened. The asymmetric "101" or "201" that we have used do not get improvements as the average combinations do.

The combined experiments have been the following:

- Monolingual French:

Experiment	Combination	Basic experiments				
Frcomb1s101	Asym101	frFWnormal	frR1FWnormal			
Frcomb1s201	Asym201	frFWnormal	frR1FWnormal			
Frcomb2s101	Asym101	frFWnormal	frR2FWnormal			
Frcomb2s201	Asym201	frFWnormal	frR2FWnormal			
Frav3	Average	frR2FWnormal	frFWnormal	frnormal		
Frav5	Average	frR2FWnormal	frFWnormal	frnormal	frR1FWnormal	frnormalinv
Frav7	Average	frR2FWnormal frFWnostem	frFWnormal frngrams54	frnormal	frR1FWnormal	frnormalinv
Frav8	Average	frR2FWnormal frFWnostem	frFWnormal frngrams54	frnormal frnostem	frR1FWnormal	frnormalinv
Frav9	Average	frR2FWnormal frFWnostem	frFWnormal frngrams54	frnormal frnostem	frR1FWnormal Frtitnormal	frnormalinv

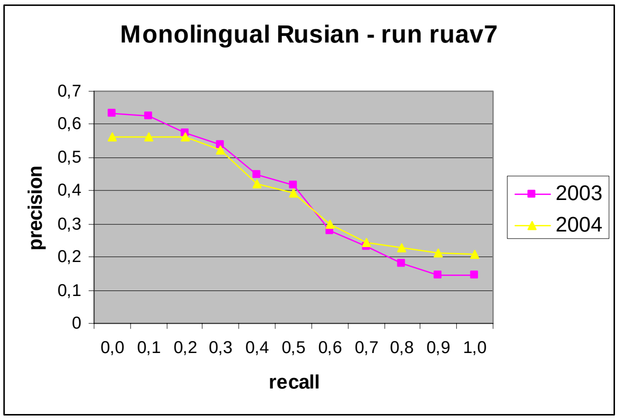
- Monolingual Russian:

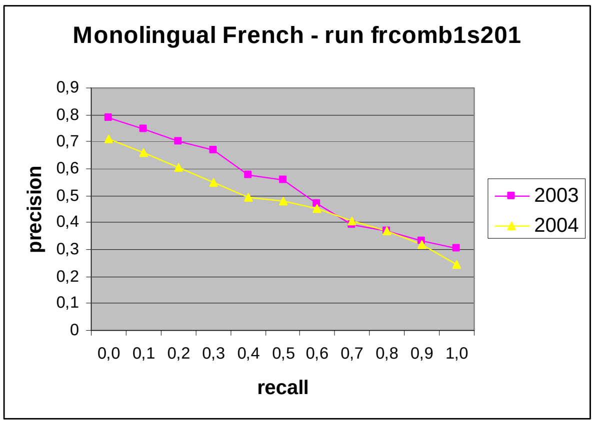
Experiment	Combination	Basic experiments				
Rucomb1s101	Asym101	runormaltit3	ruR1FWnormal			
Rucomb1s201	Asym201	runormaltit3	ruR1FWnormal			
Ruav5	Average	runormaltit3	ruR1FWnormal	ruFWnormal	Runormal	rungrams54
Ruav7	Average	runormaltit3 runormaltit1	ruR1FWnormal ruR2FWnormal	ruFWnormal	Runormal	rungrams54
Ruav8	Average	runormaltit3 runormaltit1	ruR1FWnormal ruR2FWnormal	ruFWnormal runostem	Runormal	rungrams54

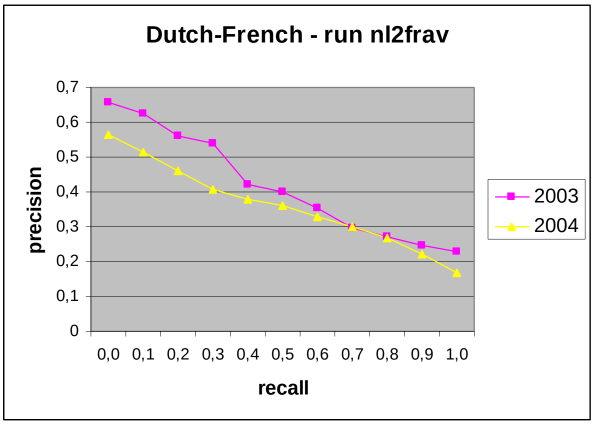
- Bilingual Dutch to French and German to French:

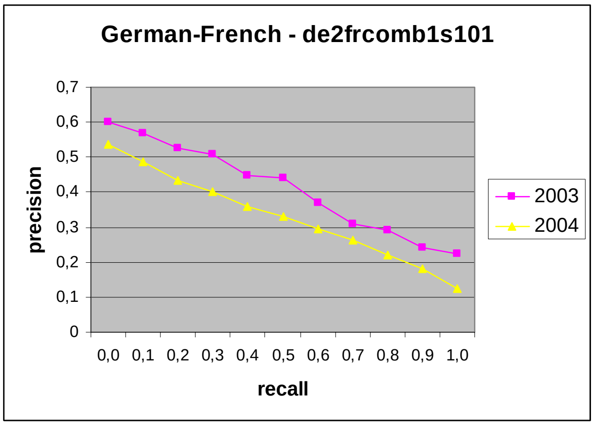
Experiment	Combination	Basic experiments		
nl2frcomb1s101	Asym101	nl2frFWnormal	nl2frR4FWnormal	
nl2frav	Average	nl2frR1FWnormal nl2frR4FWnormal	nl2frR2FWnormal nl2frR5FWnormal	nl2frR3FWnormal nl2frFWnormal
de2frcomb1s101	Asym101	de2frFWnormal	fe2frR3FWnormal	
de2frcomb2s201	Asym201	de2frR3FWnormal	De2frFWnormal	
de2frav	Average	de2frR1FWnormal de2frR4FWnormal	de2frR2FWnormal de2frR5FWnormal	de2frR3FWnormal de2frFWnormal

In the following figures, the results obtained by the best combined experiments submitted to CLEF 2004 are compared with the results obtained by exactly the same systems when applied to the 2003 tasks. The comparison shows again the qualitative differences between the 2003 and 2004 topics.









5. Conclusions

The combination approach seems to improve slightly the precision results for IR retrieval tasks, although an in-depth analysis of the reasons for that is still needed. The differences shown between the 2003 and 2004 experiments seem to be highly idiosyncratic, dependent to a great extent on the different topics selected each year. This is particularly true in the case of Russian, due to the low number of documents relevant for the topics set. Regarding the basic experiments, the general conclusions were known in advance: retrieval performance can be improved by using stemming, filtering of frequent words, appropriate weighting and relevance feedback with a few documents. On the other side, n-grams performed worse than expected.

Future work of the MIRACLE team in these tasks will be directed to several action lines: (a) Getting better performance in the indexing and retrieval phases, to be able to make experiments in a more efficient way (indexing times for huge documents collection is now excessive for a flexible scheduling of experiments). This will be achieved using our own trie-based libraries for the indexing and retrieval phases. (b) Improving the first parsing step: we think that this is one of the most critical processing steps that can improve the overall results of the IR process. A good entity recognition and normalization is still missing in our processing scheme for these tasks.

6. Acknowledgements

This work has been partially supported by the projects OmniPaper (European Union, 5th Framework Programme for Research and Technological Development, IST-2001-32174) and RIMMEL (*Multilingual and Multimedia Information Retrieval and its Evaluation*", Spanish Ministry of Science and Technology, years 2004-2007).

The participation of the MIRACLE team in year 2003 was partially funded by the Regional Government of Madrid through the research project "*MIRACLE: Multilingual Information Retrieval System and its Evaluation under the CLEF European Initiative*" (07T/0055/2003) and through its Entrepreneurship Innovation Programme (Madrid Innova, project PIE/594/2003).

References

[1] University of Neuchatel page of resources for CLEF (Stopwords, transliteration, stemmers …). On line http://www.unine.ch/info/clef/. [Visited 11/08/2004]

[2] Snowball stemmers and resources. On line http://www.snowball.tartarus.org. [Visited 11/08/2004]

[3] Xapian: an Open Source Probabilistic Information Retrieval library. On line http://www.xapian.org. [Visited 11/08/2004]

[4] Jun-Ichi Aoe, Katsushi Morimoto, Takashi Sato: An Efficient Implementation of Trie Structures. Software Practice and Experience 22(9): 695-721, 1992.

[5] Goñi-Menoyo, José Miguel; González-Cristóbal, José Carlos and Fombella-Mourelle, Jorge. An optimised trie index for natural language processing lexicons. MIRACLE Technical Report. Universidad Politécnica de Madrid, 2004.

[6] Martínez, J.L.; Villena-Román, J.; Fombella, J.; García-Serrano, A.; Ruiz, A.; Martínez, P.; Goñi, J.M. and González, J.C. (Carol Peters, Ed.): Evaluation of MIRACLE approach results for CLEF 2003. Working Notes for the CLEF 2003 Workshop, 21-22 August, Trondheim, Norway.

[7] Villena-Román, J.; Martínez, J.L.; Fombella, J.; García-Serrano, A.; Ruiz, A.; Martínez, P.; Goñi, J.M. and González, J.C. (Carol Peters, Ed.); MIRACLE results for ImageCLEF 2003. Working Notes for the CLEF 2003 Workshop, 21-22 August, Trondheim, Norway.

Appendix: Tables and figures

Table 1: CLEF 2004 results for Monolingual Russian

at0	at1	avgp	%	run id	comb	sub
0.5707	0.2184*	0.3754*	0.00%	ruav5	X	
0.5742*	0.2143	0.3697	-1.52%	runormaltit3		
0.5638	0.2100	0.3695	-1.57%	ruav7	X	X
0.5706	0.2108	0.3685	-1.84%	runormaltit2		
0.5717	0.2080	0.3676	-2.08%	runormaltit1		
0.5553	0.2092	0.3672	-2.18%	ruR1FWnormal		X
0.5683	0.2014	0.3660	-2.50%	runormal		
0.5693	0.2094	0.3648	-2.82%	rucomb1s101	X	X
0.5574	0.2050	0.3641	-3.01%	ruav8	X	X
0.5597	0.2094	0.3608	-3.89%	rucomb1s201	X	
0.5558	0.1940	0.3584	-4.53%	ruFWnormal		
0.5225	0.1762	0.3309	-11.85%	ruR2FWnormal		
0.5102	0.1883	0.3195	-14.89%	rungrams54		
0.4906	0.1790	0.3125	-16.76%	rungrams43		
0.4885	0.1771	0.3012	-19.77%	ruFWnostem		
0.4731	0.1827	0.2907	-22.56%	rungrams76		
0.4757	0.1642	0.2884	-23.18%	runostem		
0.1715	0.0128	0.0764	-79.65%	rutitngrams43		
0.1538	0.0109	0.0723	-80.74%	rutitngrams54		
0.1166	0.0049	0.0433	-88.47%	rutitFWnormal		
0.1119	0.0003	0.0383	-89.80%	rutitnormal		
0.0876	0.0004	0.0245	-93.47%	rutitnostem		

Table 2: CLEF 2003 results for Monolingual Russian

at0	at1	avgp	%	run id	comb	sel
0.6384*	0.1459	0.3799*	-0.00%	ruav8	X	X
0.6379	0.1465	0.3750	-1.29%	ruR1FWnormal		X
0.6323	0.1471	0.3706	-2.45%	ruav7	X	X
0.6344	0.1463	0.3697	-2.68%	ruav5	X	
0.6234	0.1593*	0.3695	-2.74%	rucomb1s101	X	X
0.6276	0.1575	0.3695	-2.74%	rucomb1s201	X	
0.6230	0.1563	0.3695	-2.74%	runormaltit1		
0.6234	0.1593*	0.3695	-2.74%	runormaltit3		
0.6228	0.1585	0.3694	-2.76%	runormaltit2		
0.6254	0.1430	0.3653	-3.84%	ruFWnormal		
0.6194	0.1423	0.3645	-4.05%	runormal		
0.6044	0.1482	0.3605	-5.11%	ruR2FWnormal		
0.5789	0.1318	0.3418	-10.03%	rungrams54		
0.5579	0.1438	0.3323	-12.53%	rungrams43		
0.5609	0.1052	0.3046	-19.82%	ruFWnostem		
0.5609	0.1052	0.3046	-19.82%	runostem		
0.5172	0.1058	0.2753	-27.53%	rungrams76		
0.2922	0.0584	0.1382	-63.62%	ruFWtitnormal		
0.2910	0.0584	0.1381	-63.65%	rutitnormal		
0.2716	0.0661	0.1377	-63.75%	rutitngrams43		
0.2378	0.0462	0.1125	-70.39%	rutitngrams54		
0.2277	0.0476	0.1112	-70.73%	rutitnostem		

Table 3: CLEF 2004 results for Monolingual French

at0	at1	avgp	%	run id	comb	sub
0.7070	0.2444	0.4677*	0.00%	frav5	X	
0.7111	0.2459	0.4673	-0.09%	frcomb1s201	X	X
0.7107	0.2438	0.4670	-0.15%	frav3	X	
0.7100	0.2477*	0.4670	-0.15%	frcomb2s201	X	X
0.7032	0.2477*	0.4654	-0.49%	frR2FWnormal		
0.7242*	0.2349	0.4654	-0.49%	frFWnormal		
0.6986	0.2459	0.4653	-0.51%	frR1FWnormal		
0.6986	0.2459	0.4653	-0.51%	frcomb1s101	X	
0.6998	0.2477*	0.4639	-0.81%	frcomb2s101	X	
0.7170	0.2425	0.4635	-0.90%	frav9	X	
0.7186	0.2338	0.4628	-1.05%	frnormalinv		
0.7169	0.2378	0.4624	-1.13%	frav7	X	X
0.7172	0.2352	0.4596	-1.73%	frnormal		
0.7113	0.2371	0.4589	-1.88%	frav8	X	X
0.6634	0.2060	0.4206	-10.07%	frngrams54		
0.6797	0.2036	0.4187	-10.48%	frnostem		
0.6685	0.2014	0.4177	-10.69%	frFWnostem		
0.6393	0.0719	0.3263	-30.23%	frtitnormalinv		
0.6278	0.0719	0.3254	-30.43%	frtitnormal		

0.6066	0.0619	0.2999	-35.88%	frtitngrams54		
0.5932	0.0650	0.2985	-36.18%	frtitnostem		

Table 4: CLEF 2003 results for Monolingual French

at0	at1	avgp	%	run id	comb	sel
0.8053	0.3271*	0.5312*	0.00%	frav7	X	X
0.7993	0.2987	0.5288	-0.45%	frcomb2s201	X	X
0.8091*	0.3202	0.5287	-0.47%	frav8	X	X
0.7902	0.3049	0.5220	-1.73%	frcomb1s201	X	X
0.7707	0.2987	0.5207	-1.98%	frcomb2s101	X	
0.7707	0.2987	0.5207	-1.98%	frR2FWnormal		
0.7951	0.3029	0.5200	-2.11%	frav5	X	
0.7927	0.3017	0.5191	-2.28%	frav3	X	
0.7731	0.3049	0.5162	-2.82%	frcomb1s101	X	
0.7731	0.3049	0.5162	-2.82%	frR1FWnormal		
0.7954	0.2980	0.5124	-3.54%	frFWnormal		
0.7855	0.2987	0.5083	-4.31%	frnormal		
0.7717	0.2749	0.4913	-7.51%	frav9	X	
0.7281	0.2958	0.4875	-8.23%	frnormalinv		
0.7313	0.2778	0.4722	-11.11%	frngrams54		
0.6896	0.2753	0.4579	-13.80%	frFWnostem		
0.6806	0.2618	0.4452	-16.19%	frnostem		
0.6241	0.1725	0.3315	-37.59%	frtitnormal		
0.5850	0.1516	0.3117	-41.32%	frtitngrams54		
0.4939	0.1213	0.2288	-56.93%	frtitnostem		

Table 5: CLEF 2004 results for Bilingual Dutch to French

at0	at1	avgp	%	run id	comb	Sub
0.5591	0.1716	0.3519*	0.00%	nl2frR2FWnormal		
0.5628	0.1668	0.3505	-0.40%	nl2frav	X	X
0.5558	0.1637	0.3486	-0.94%	nl2frR3FWnormal		
0.5458	0.1739	0.3483	-1.02%	nl2frR1FWnormal		
0.5598	0.1593	0.3483	-1.02%	nl2frR5FWnormal		X
0.5583	0.1595	0.3472	-1.34%	nl2frR4FWnormal		X
0.5653*	0.1717	0.3469	-1.42%	nl2frnormal		
0.5430	0.1750*	0.3451	-1.93%	nl2frFWnormal		
0.5515	0.1595	0.3449	-1.99%	nl2frcomb1s101	X	X

Table 6: CLEF 2003 results for Bilingual Dutch to French

at0	at1	avgp	%	run id	comb	sel
0.6766*	0.2323*	0.4159*	-0.00%	nl2frR4FWnormal		X
0.6564	0.2296	0.4112	-1.13%	nl2frR5FWnormal		X
0.6528	0.2323*	0.4087	-1.73%	nl2frcomb1s101	X	X
0.6583	0.2285	0.4069	-2.16%	nl2frav	X	X
0.6518	0.2286	0.4043	-2.79%	nl2frR3FWnormal		
0.6684	0.2230	0.4016	-3.44%	nl2frFWnormal		
0.6423	0.2321	0.3997	-3.90%	nl2frR2FWnormal		
0.6533	0.2225	0.3986	-4.16%	nl2frR1FWnormal		
0.6478	0.2159	0.3862	-7.14%	nl2frnormal		

Table 7: CLEF 2004 results for Bilingual German to French

at0	at1	avgp	%	run id	comb	sub
0.5419	0.1195	0.3217*	0.00%	de2frR5FWnormal		
0.5289	0.1209	0.3208	-0.28%	de2frR4FWnormal		
0.5485	0.1263	0.3201	-0.50%	de2frR2FWnormal		X
0.5340	0.1241	0.3199	-0.56%	de2frR3FWnormal		X
0.5349	0.1244	0.3178	-1.21%	de2frav	X	
0.5439	0.1250	0.3174	-1.34%	de2frR1FWnormal		
0.5381	0.1241	0.3166	-1.59%	de2frcomb1s101	X	X
0.5447	0.1265*	0.3134	-2.58%	de2frFWnormal		
0.5265	0.1265*	0.3116	-3.14%	de2frcomb2s201	X	X
0.5505*	0.1221	0.3100	-3.64%	de2frnormal		

Table 8: CLEF 2003 results for Bilingual German to French

at0	at1	avgp	%	run id	comb	sel
0.6064	0.2255	0.3999*	-0.00%	de2frR3FWnormal		X
0.6007	0.2255	0.3975	-0.60%	de2frcomb1s101	X	X
0.6017	0.2246	0.3942	-1.43%	de2frR2FWnormal		X
0.5931	0.2244	0.3938	-1.53%	de2frav	X	
0.5912	0.2273	0.3931	-1.70%	de2frR5FWnormal		
0.5867	0.2178	0.3899	-2.50%	de2frR1FWnormal		
0.5795	0.2288*	0.3890	-2.73%	de2frR4FWnormal		
0.6082*	0.2093	0.3837	-4.05%	de2frcomb2s201	X	X
0.5962	0.2093	0.3816	-4.58%	de2frFWnormal		
0.5857	0.2030	0.3770	-5.73%	de2frnormal		