

An Evaluation of How Search Engines Respond to Greek Language Queries

Efthimis N. Efthimiadis
 Information School
 University of Washington
 Seattle, WA, USA
 +1 (206) 616-6077
 efthimis@u.washington.edu

Nicos Malevris, Apostolos Kousaridas,
 Alexandra Lepeniotou, and Nikos Loutas
 Department of Informatics
 Athens University of Economics and Business
 Athens, Greece
 +30 (210) 820-3126
 {ngm, akousar, alex, nloutas}@aueb.gr

Abstract

Over 20 billion Web pages from around the world have been indexed by search engines [10]. This study investigates how search engines respond to non-English queries and more specifically to Greek language queries. To address this we conducted an evaluation using Greek queries in ten search engines: five “global” (A9, AltaVista, Google, MSN Search, and Yahoo!) and five Greek (Anazitisi, Ano-Kato, Phantis, Trinity, and Visto). A set of navigational queries for known Greek organizations was created. The organizations correspond to ten categories: government departments, universities, colleges, travel agencies, museums, media (TV, radio, newspapers), transportation, and banks. Searches were performed using the Greek and its corresponding English, Latin, or transliterated name of each organization. The ideal retrieval would be to get the website of that organization ranked first in the result set. The results of this evaluation are presented in this paper, together with a report on how the engines respond to Greek and Anglicized queries, and on the best performing global and Greek search engines.

1. Introduction

The web continues to expand and the dominant search engines, Google and Yahoo! claim to have indexed more than 20 billion pages [10]. Recent statistics on Internet usage by language show that 31.2% is English and 68.8% is non-English [7]. As the non-English web usage increases there are an increasing number of non-English queries that need to be handled by the search engines.

The goals of this research are (a) to evaluate how well search engines respond to Greek language queries; and, (b) to assess whether the Greek or global search engines are more effective in satisfying the user requests.

2. Related Work

Bar-Ilan and Gutman [2] explored how search engines respond to queries in four non-English languages, Russian, French, Hungarian and Hebrew. For each of the languages

they searched in three global search engines, AltaVista, FAST and Google, and in three local for that language engines. The local engines were the Russian Yandex, Rambler, Aport; the French Voila, AOL France, La Toile de Quebec, the Hungarian Origo-vizsla, Startlap, Heureka, and the Hebrew Morfix and Walla. For each of the four languages the authors developed queries that emphasized specific linguistic characteristics of that language. The first ten results of each search were evaluated not for relevance, but for whether the exact word form or a morphological variant of the query was retrieved. They found that the search engines ignored the special language characteristics and do not handle diacritics well.

Moukdad [11] studied how three global search engines, AltaVista, AllTheWeb, and Google, handle Arabic queries compared to three Arabic engines, Al bahhar, Ayna, and Morfix. He employed the same methodology used by Bar-Ilan and Gutman [2]. A set of eight Arabic search terms was selected and run in the six search engines. He found that the global search engines had shortcomings in handling Arabic.

Lazarinis [9] evaluated the performance of eleven search engines, seven global (AlltheWeb, AltaVista, AOL, ASK, Google, MSN, Yahoo) and four Greek (Anazitisi, In.gr, Pathfinder, Robby), with the use of six Greek language queries. He employed thirty one users who were divided into six groups and each searched one query. Each group member retrieved twenty results. The retrieved results by all group members were evaluated for relevance collectively by the members of each group. Lazarinis reports that the precision of all engines is very similar. Based on the six queries the study further investigated how engines handle upper and lower case input, diacritics, stemming, and stop words. The study noted that there were variations in the handling of Greek.

Moukdad and Cui [12] investigated how Chinese language queries are handled by Google and AlltheWeb, as well as Sohu and Baidu, the Chinese search engines. They created ten queries by selecting terms from a Chinese-English dictionary. The terms emphasized certain linguistic characteristics of Chinese. The queries were searched in the Simplified Chinese script which is in use in mainland China. The results were evaluated based on the

number of retrieved documents, word segmentation, and correct display of Chinese characters. They found that the global search engines did not use any linguistic processing and thus were not able to process the Chinese queries satisfactorily, which lead to the introduction of unexpected results.

3. The Greek Language

The Greek language uses a different script to that of Latin-based languages. The Greek alphabet set has twenty four upper case letters, twenty five lower case letters and a number of diacritics or accent marks depending on the form used (see Figure 1).

Figure 1: The Greek Language Script

Upper Case Letters (24):	Α Β Γ Δ Ε Ζ Η Θ Ι Κ Λ Μ Ν Ξ Ο Π Ρ Σ Τ Υ Φ Χ Ψ Ω
Lower Case Letters (25):	α β γ δ ε ζ η θ ι κ λ μ ν ξ ο π ρ σ [ς] τ υ φ χ ψ ω
	e.g.: ΣΩΣΤΟΣ, σωστός (correct)
Accent (`): Α ά Ε έ Η ή Ι ί Ο ό Υ ύ Ω ώ	
	e.g.: νερό (water)
Diaeresis (` `): μικροϋπολογιστής (microcomputer)	
Combined accent & diaeresis (` `): καΐκι (sailboat)	
Character combinations - Diphthongs:	
Vowels:	αι, αυ, ει, ευ, ηυ, οι, ου, υι
Consonants:	γγ, γκ, γξ, γλ, μπ, ντ

The most commonly known forms of the Greek language are ancient or classical Greek, Katharevousa, and Demotic Greek (Dhimotiki). Depending on the system of accents used Greek is either polytonic or monotonic. The polytonic orthography system for Greek uses three accents, two breathings, iota subscripts and diaeresis. The polytonic system was used since the ancient times and was simplified into the monotonic system in 1982. The monotonic Greek language system uses one accent and the diaeresis, in order to signify that two adjacent vowels are pronounced separately and not as a diphthong.

Transliteration of Greek to Latin letters is common but adds to the complexity of processing Greek because of the different transliteration standards. Furthermore, individuals often ignore the standards and apply their own phonetic interpretation. The widespread use of computers and the Internet coupled with the slow progress in adopting non-Latin-based scripts has given rise to Greeklish, which is a form of transliteration used to exchange email messages and post to discussion fora.

Alevizos *et al.* [1] discuss the challenges faced by search systems in handling Greek. Kalamboukis [8] introduces the inflectional aspects of Greek and presents a stemming approach.

4. Methodology

4.1. Selecting the Search Engines

For the study we selected ten search engines based on their popularity and market share. These were divided into two groups, five global or international in scope, and five Greek search engines. The global search engines are: A9, AltaVista, Google, MSN (Live) Search, and Yahoo!. The Greek engines are: Anazitisis, Ano-Kato, Phantis, Trinity, and Visto. Appendix 9 lists the engines and their corresponding URLs.

4.2. User Needs and Task Definition

There has been a threefold increase in the numbers of Greeks using the Internet between 2000 and 2006, jumping from 9.1% to 33.5% respectively [6]. Similarly, the Greek web has proliferated with an increasing presence of governmental and commercial entities. In 2004, most of the Greek web pages (63.5%) were in the Greek language [4]. Though most Greeks learn a second language to some degree of proficiency, it is reasonable to assume that they would search in Greek to find information in the Greek web. Following the Broder [3] classification of web queries we selected the “navigational” class as the basis of a user task definition. We assume that a user will search to find the specific site of an organization. To that respect our methodology relates to that of Hawking et al. [4].

4.3. Queries and Subject Categories

We identified ten popular broad categories in which we selected organizations to search for. The categories are: government departments, universities, colleges, travel agencies, museums, media (TV, radio, newspapers), transportation, and banks. Using professional and business directories we selected two hundred and seventeen (217) organizations that had a web presence. For each organization we established the formal name in Greek, its non-Greek equivalent if available (usually in English or other Latin-based language) and the URL(s) of the web site.

Table 1 lists the subject categories and the corresponding numbers of Greek organizations. There were a total of 217 organizations, of which 92 had a corresponding English or other non-Greek equivalent name, thus, resulting in 309 queries.

Searches were submitted automatically to the engines in August 2006, and we used as queries the Greek and English names of each organization. Examples of the queries are given in Table 2.

Table 1: Subject categories searched and number of queries.

Subject Categories		Organizations in:	
(in English)	(in Greek)	Greek	English
Government Departments	Υπουργεία	18	14
Universities	Πανεπιστήμια	21	20
Colleges	ΤΕΙ	14	8
Travel Agencies	Ταξιδιωτικά Γραφεία	39	4
Museum	Μουσεία	19	0
Transportation & Communication Services	Μέσα Μεταφοράς, Επικοινωνίες	12	7
Banks	Τράπεζες	28	13
Newspapers	Εφημερίδες	17	16
Television Stations	Τηλεόραση	12	3
Radio Stations	Ραδιόφωνο	37	7
Total / Σύνολο		217	92

The queries were submitted for search in the typical format of typing out the keyword separated by spaces. No advance search techniques were employed in order to simulate the input of a non-expert searcher. The ideal retrieval would be to get the website of that organization ranked first in the result set.

Table 2: Examples of queries

In Greek	Equivalent in English or transliterated
Υπουργείο Απασχόλησης και Κοινωνικής Προστασίας	Ministry of Employment and Social Protection
Υπουργείο Πολιτισμού	Ministry of Culture
Υπουργείο Μακεδονίας Θράκης	Ministry of Macedonia Thrace
Εθνική Τράπεζα της Ελλάδος	National Bank of Greece
Εμπορική Τράπεζα της Ελλάδος	Emporiki Bank
Δημοκρίτειο Πανεπιστήμιο Θράκης	Democritus University of Thrace
ΤΕΙ Καβάλας	Technological Education Institute (TEI) of Kavala
ΚΜ Ταξίδι & Τουρισμός	KM Travel and tourism
Τα Νέα	Nea [newspaper]
Νέα Ελληνική Τηλεόραση	NET [TV station]
Κόσμος	Kosmos [radio station]
Οργανισμός Αστικών Συγκοινωνιών Αθηνών(ΟΑΣΑ)	Athens Urban Transport Organization(OASA)

(On the left column of Table 2, the text in square brackets provides the subject category of the query.)

4.4. Evaluation Criteria

For every search we recorded the top ten results and their rank order. Then we evaluated whether the organization’s URL was found in the results set, and, if so, recorded the rank position and the number of times. The evaluation also counted whether there was an exact or partial match of the desired URL.

The score includes two components, the rank position, and the depth of the page as indicated in the URL. For example, if the correct URL were found in rank 1, then the score assigned was 100, if in rank two 90, and so on. If the URL were a partial match, that is, it came from a page in the website but not the top page, then, the score was adjusted depending on the depth of the page retrieved. The latter gives some credit for partial matches, assuming that the searcher will be able to identify that the returned result is related to the desired result. This way the search engine is penalized for the additional navigational effort that will be required by the user.

We also recorded the number of active and dead links returned in the top ten results. This provides an indication of the freshness of the search engine’s index.

5. Results

5.1. Qualitative aspects of searching

Table 3 presents how search engines respond to Greek queries that either have or do not have accent marks. It also shows whether the engines handle articles, prepositions, pronouns, etc.

Table 3: How search engines handle Greek accent marks

Search Engine	Greek with or without accents produce:	Handling of articles, prepositions, etc.	
		Greek	English
Anazitisis	different results	No	No
AnoKato	same results	No	Yes
Phantis	same results	Yes	Yes
Trinity	same results	Yes	Yes
Visto	same results	Yes	Yes
A9	different results	No	Yes
Altavista	different results	No	Yes
Google	different results	No	Yes
MSN	different results	No	Yes
Yahoo	different results	No	No

The five global search engines and one Greek return different results. The differences observed in the top ten results vary from providing totally different results, to having some small overlap in the results, but with differences in rank order.

5.2. Search Results by Rank Order

The 309 queries were submitted to each of the 10 search engines for a total of 3090 searches. Of those 276 queries or 2760 searches returned valid results, while 33 queries or 330 searches did not return any results at all.

Table 4: Rank distribution of all search results by search engine.

Greek / Global	Search Engines	Rank										Missed	Total Found	% Success Rate
		1	2	3	4	5	6	7	8	9	10			
Global	Google	199	11	8	1	2	3	1	1	1	1	81	228	73.79%
Global	Altavista	166	30	10	2	2	2	4	2	0	3	88	221	71.52%
Global	Yahoo	133	30	13	7	3	3	3	1	2	0	114	195	63.11%
Greek	Trinity	142	10	5	3	0	1	1	0	0	0	147	162	52.43%
Global	A9	106	17	11	3	4	5	3	2	1	0	157	152	49.19%
Global	MSN	101	18	12	6	5	3	3	1	1	0	159	150	48.54%
Greek	Visto	78	20	6	4	4	2	1	1	1	0	192	117	37.86%
Greek	Anokato	53	16	5	2	2	2	2	0	1	0	226	83	26.86%
Greek	Anazitisis	17	7	4	0	2	1	1	2	0	1	274	35	11.33%
Greek	Phantis	23	5	2	1	1	0	0	0	0	1	276	33	10.68%

Table 4 presents the rank distribution of the results for both the Greek and English queries by search engine. The table lists also the number of organizations missed by each engine, and their success rate. Of the organizations found it appears that most results were presented in the first three ranks. The global search engines have higher success rates, ranging from 48.54% to 73.79%, than the Greek engines which range from 10.68% to 52.43%. Google is the best performing global engine and Trinity is the best Greek engine.

The above results give an overall performance rate for the search engines but do not show how the engines respond to Greek or non-Greek queries. Table 5 and Table 6 present the rank distributions of the results by language. In Table 5 we see that AltaVista handles Greek queries better than all the other engines with a success rate of 72.81%, while Google follows closely with 70.96%, whereas MSN and A9 are fourth and fifth with 50.60% and 50.23% respectively. The best performance of the Greek engines was recorded by Trinity with 49.30%.

The rank distribution of the results from the queries in either English or in transliterated form is given in Table 6. These show mixed results, as we observe variations in performance for almost all the search engines. When compared to results from the Greek queries (Table 5) Google has increased its performance (80.43%), Yahoo!'s performance remained about the same (63.04%), whereas MSN, AltaVista, and A9 decreased theirs. Of the Greek

search engines Trinity's performance increased to 59.78%, whereas the performance of all other engines decreased.

Table 5: Rank distribution of results for Greek queries.

Greek / Global	Search Engines	Rank										Total Found	% Success Rate
		1	2	3	4	5	6	7	8	9	10		
Global	AltaVista	118	23	8	1	2	0	3	2	0	1	158	72.81%
Global	Google	131	10	5	1	2	2	1	1	0	1	154	70.96%
Global	Yahoo	104	12	8	5	3	1	2	0	2	0	137	63.13%
Global	MSN	79	9	8	6	3	2	1	1	1	0	110	50.69%
Global	A9	82	7	9	3	3	1	2	1	1	0	109	50.23%
Greek	Trinity	94	6	3	2	0	1	1	0	0	0	107	49.30%
Greek	Visto	63	16	4	3	4	0	1	1	1	0	93	42.86%
Greek	Anokato	32	11	3	1	2	2	1	0	1	0	53	24.42%
Greek	Anazitisis	12	5	3	0	2	1	1	2	0	1	27	12.44%
Greek	Phantis	18	1	1	1	1	0	0	0	0	1	23	10.59%
Total Queries												217	

Table 6: Rank distribution of results for English queries.

Greek / Global	Search Engines	Rank										Total Found	% Success Rate
		1	2	3	4	5	6	7	8	9	10		
Global	Google	68	1	3	0	0	1	0	0	1	0	74	80.43%
Global	AltaVista	48	7	2	1	0	2	1	0	0	2	63	68.48%
Global	Yahoo	29	18	5	2	0	2	1	1	0	0	58	63.04%
Greek	Trinity	48	4	2	1	0	0	0	0	0	0	55	59.78%
Global	A9	24	10	2	0	1	4	1	1	0	0	43	46.73%
Global	MSN	22	9	4	0	2	1	2	0	0	0	40	43.48%
Greek	Anokato	21	5	2	1	0	0	1	0	0	0	30	32.61%
Greek	Visto	15	4	2	1	0	2	0	0	0	0	24	26.09%
Greek	Phantis	5	4	1	0	0	0	0	0	0	0	10	10.87%
Greek	Anazitisis	5	2	1	0	0	0	0	0	0	0	8	8.69%
Total Queries												92	

An analysis of variance (ANOVA), using the SPSS statistical package, shows that there is a 100% significant difference in the mean performance of all 10 engines when the entire sample of all queries (Greek and English) is considered (Table 4). This applies to both groups of search engines, i.e., Greek and Global.

Significant difference is also found when the Greek queries are considered only (Table 5). More specifically, there is a 100% significant difference between the means of the Greek engines, whereas this is not as indicative as it ought to be among the global engines. When the English queries are considered, in Table 6, there is significant difference in the mean performance at the 97% level among the Global engines, whereas the same cannot be said with confidence for the Greek engines nor when all the engines are considered together. This is attributed to the many zero entries in the ranks between the 4th and 10th position that are recorded by the Greek engines when tackling English queries.

Tests on the paired differences show that in certain cases we can say with statistical certainty that some engines performed worse than all other engines, e.g., Anazitisis in Tables 4 and 5, A9 and MSN in table 4, and MSN in Table 6. The engines that performed better among the other

groups are the ones which are on top of the groups in the statistical analysis, i.e., Google and Trinity in Table 4, Trinity, AltaVista and Google in Table 5 and Google and Trinity performing better in Table 6.

5.3. Search results by subject category.

Using the method discussed in the section evaluation criteria all queries were scored and then grouped by category. This enables a finer evaluation of the performance of the search engines in the study. Table 7 shows the results of this evaluation grouped by language and by subject category. Based on the scoring the larger the number the better the performance of a search engine. Google from the global engines and Trinity from the Greek engines outperformed the other engines in their respective groups. But, this is not to say that Trinity's performance is

variation in rank amongst Greek and English, positions 1 and 7 respectively. Newspapers also ranged from rank 5 for Greek queries to rank 2 for English queries.

The statistical analysis of variance (ANOVA) of the results by subject category in the Greek queries (Table 7) shows a 100% significant difference in the mean performance of all engines, whereas in the English queries the difference is at the 95% level.

6. Active vs. Dead Links

The results returned by the search engines were also evaluated by measuring the percentage of active versus dead (non-active) links. Such evaluation measures the freshness of the index of the search engine and is an indicator for the levels of frustration the searcher would have to go. It is another level of evaluating the precision of

Table 7: Sum of the scores of the top ten results by subject category, language, and search engine.

search engines	Government Departments	Newspapers	Transportation & Comn. Services	Banks	Universities	Radio Stations	Colleges	Travel Agents	TV stations	Museums	
	Υπουργεία	Εφημερίδες	Μέσα Μεταφοράς & Επικοινωνιών	Τράπεζες	Πανεπιστήμια	Ραδιόφωνο	ΤΕΙ	Ταξιδιωτικά Γραφεία	Τηλεόραση	Μουσεία	
Greek - Ελληνικά											
Greek Ελληνικές	anazitis	407	63	321	702	776	529	243	1375	396	950
	anokato	767	502	200	580	1221	460	190	2268	634	200
	phantis	891	412	483	532	978	652	459	1065	502	80
	trinity	2338	1488	1429	2386	2614	271	1900	3641	806	672
	visto	1403	850	930	1340	2073	930	316	1650	400	300
Global Διεθνείς	a9	2206	1334	1030	2044	2304	1306	1478	2094	524	1304
	AltaVista	2145	1776	1311	2666	2385	2046	1493	3192	1049	1602
	google	2289	1866	1312	2953	3039	1841	1817	3100	1169	1712
	Msn	2206	1262	1030	2126	2418	1242	1430	2114	524	1260
	yahoo	1848	1435	1073	1985	1953	1519	1527	2827	818	1396
Totals:	16500	10988	9119	17314	19761	10796	10853	23326	6822	9476	
English - Αγγλικά											
Greek Ελληνικές	anazitis	413	897	170	381	958	180	106	149	7	
	anokato	865	1376	70	261	1371	642	0	290	0	
	phantis	437	554	90	495	1197	152	242	271	81	
	trinity	1528	1957	849	1519	2818	363	950	298	298	
	visto	300	1243	130	290	75	290	0	280	0	
Global Διεθνείς	a9	822	703	541	1244	2656	96	1091	135	172	
	AltaVista	1275	1109	494	1458	2844	136	1085	245	361	
	google	1609	1875	688	1650	2926	352	1169	281	181	
	Msn	822	688	531	1244	2391	96	1047	135	100	
	yahoo	1224	1057	495	1119	2785	130	1044	254	100	
Totals:	9295	11459	4058	9661	20021	2437	6734	2338	1300		

good. On the contrary when comparing the Greek and global engines the Greek engines failed miserably.

Based on the aggregate results for all search engines per category for Greek queries the coverage of the categories is in the following rank order: travel agencies, universities, banks, government departments, newspapers, colleges (TEI), radio stations, museums, transportation & communication services, TV stations.

Similarly, the aggregate results for all search engines for English queries show that the rank order of the coverage of the categories is: universities, newspapers, banks, government departments, colleges, transportation & communication services, travel agents, radio stations, and TV stations. Travel agencies is the category with most

the search and the cost to the user should they follow the dead links.

Table 8: Active vs. Dead Links for All Queries

All Queries 2006					
Search Engines		Active	%	Non Active	%
Global	Google	2927	96.00	122	4.00
Global	A9	2917	96.78	97	3.22
Global	Altavista	2910	95.57	135	4.43
Global	MSN	2862	95.27	142	4.73
Global	Yahoo	2651	91.35	251	8.65
Greek	Trinity	2572	94.49	150	5.51
Greek	Anazitis	2538	86.95	381	13.05
Greek	Phantis	2519	86.41	396	13.59
Greek	Visto	1745	74.80	588	25.20
Greek	Anokato	611	73.00	226	27.00

The results presented in Table 8 show the aggregate number of returned URLs for all 309 queries that were submitted to each search engine. These are further divided into those that were active and those that were non-active. Of the global search engines A9 had the highest percentage number of active links (96.78%), whereas Yahoo the highest percentage number of dead links (8.65%). Of the Greek search engines Trinity had the highest percentage of active links (94.49%) whereas AnoKato the highest percentage of dead links (27%).

These results despite being a good indication of the freshness of the index should not be considered in isolation. When compared to Table 4 which shows the overall success rate of the search engines one can see that for example, A9 has poor precision results successfully retrieving only 49.19% of the correct answers, while the dead links found in these results are 3.22%. Google, on the other hand, retrieves 73.79% of the correct results, while the non-active links are 4%. From a user's point of view getting higher precision in the top ten results is probably more valuable.

The performance of the Greek search engines with respect to active links in their result sets is disappointing. The dead links for all engines but Trinity (5.51%) range from 13.05% to 27%. Trinity is the best performing Greek search engine.

7. Conclusions

This study aimed at evaluating how search engines handle Greek language queries. The study evaluated ten search engines, five Greek and five global. Our results corroborate and extend the findings of [7]. The analysis shows that the global search engines ignore the characteristics of the Greek language, hence treating Greek queries differently. Despite this finding the performance of the global search engines outperforms that of the Greek engines. A set of 309 navigational queries was used in the evaluation. The rank distribution of all search results indicates that on average the search engines retrieved the desired document in the first three rank positions. However, the rate of success leaves much to be desired as the most successful engine, Google, was able to find the correct answer to only 73.91% of the English and 60.37% of the Greek queries. The engines seem to have poor coverage of the Greek web, and the results returned by the engines are different depending on how the searcher has typed the Greek query, e.g., with or without accents.

9. Appendix: List of search engines used in the study

Global Search Engines:

1. A9: <http://www.a9.com>
2. Google (<http://www.google.com.gr>)
3. Yahoo (<http://www.yahoo.com/>)
4. Altavista (<http://www.altavista.com>)
5. MSN (<http://www.msn.com>)

Therefore, the implications for Greek users are many as they need to be aware of the nuances to searching using Greek.

8. References

- [1] Alevizos T., Galiotou E., Skourlas C. (1988) Information retrieval and Greek-Latin text. International online information meeting. 12, London (06/12/1988). Learned Information Europe, Oxford, UK. pp. 791-801.
- [2] Bar-Ilan, J., Gutman, T. (2005) How do search engines respond to some non-English queries. *Journal of Information Science*, 31(1), Pages: 13-28, 2005.
- [3] Broder, A. (2002) A taxonomy of web search. *SIGIR Forum* 36, 2 (Sep. 2002), 3-10. <http://doi.acm.org/10.1145/792550.792552>
- [4] Efthimiadis, E.N. and Castillo, C. (2004) Charting the Greek Web. In: ASIST'04: American Society for Information Science and Technology Annual Conference. Providence, Rhode Island, November 13-18, 2004.
- [5] Hawking, D., Craswell, N., and Griffiths, K. Which search engine is best at finding airline site home pages? CMIS Technical Report 01/45 (March, 2001). http://es.csiro.au/pubs/craswell_tr01.pdf
- [6] Internet World Statistics. (2007) Greece: Internet Usage and Marketing report. Retrieved May 15, 2007, available at <http://www.internetworldstats.com/eu/gr.htm>.
- [7] Internet World Statistics. (2007) Internet World Users By Language (as of June 30, 2007). Retrieved September 5, 2007, available at <http://www.internetworldstats.com/stats7.htm>
- [8] Kalamboukis, T. Z. (1995). Suffix Stripping with Modern Greek. *Program*, 29(3), 313-321.
- [9] Lazarinis, F. (2007) Web Retrieval Systems and the Greek Language: do they have an understanding. *Journal of Information Science*, v5. doi:10.1177/0165551506076394
- [10] Mayer, T. (2005) Our Blog is Growing Up, And So Has Our Index. Yahoo! Search Blog, posted 8/8/05, retrieved May 15, 2007, <http://www.ysearchblog.com/archives/000172.html>
- [11] Moukdad, H. (2004) Lost In Cyberspace: How Do Search Engines Handle Arabic Queries? In: Access to Information: Technologies, Skills, and Socio-Political Context. University of Manitoba, Winnipeg, Manitoba. June 3 - 5, 2004. Proceedings Editors: H. Julien and S. Thompson. Available at: www.cais-acsi.ca/proceedings/2004/moukdad_2004.pdf
- [12] Moukdad, H., Cui, H. (2005) How do search engines handle Chinese queries? *Webology*, Volume 2, Number 3, October, 2005. <http://www.webology.ir/2005/v2n3/a17.html>

Greek Search Engines:

6. Anazitisis (<http://www.anazitisis.gr/>)
7. Ano-Kato (<http://www.ano-kato.com/>)
8. Phantis (<http://www.phantis.gr/>)
9. Trinity (<http://www.trinity.gr/>)
10. Visto (<http://www.visto.gr/>)