

Recall of Synonyms in Keyword-Based Information Retrieval in RICeST

Shapour Reza Berenjian¹, Ali Reza Berenjian²

¹ Regional Information Center for Science and Technology (RICeST), Shiraz, Iran

² University of Applied Science and Technology, Tehran, Iran.

Abstract: This study was conducted to develop and introduce the outline of the software for ‘Automatic Calling of Synonyms in Keyword-based Information Retrieval in RICeST’. The software functions such that a synonym is extracted by providing every given entity. In case the entered keyword or its synonym is not meant by the operator, it could be changed or modified to do a new search. In case of variable spelling, the operator may alter the spelling and do a new search. This would help avoid information loss and elimination of relevant documents so that irrelevant documents are not retrieved. As some linguists contend, although two words may be synonymous in a chain of words, they may vary in meaning in different discursive situations after adopting new syntactic roles. Still, the present study hinges on the premise that synonyms are the words with the same meaning in a chain of words. This system was designed and implemented in the Regional Centre for Information Science and Technology (RICeST). This article reports part of the research project carried out in RICeST. This system is used on the RICeST intranet and website. It is recommended that further studies be conducted to modify the software so that [Shapour Reza Berenjian, Ali Reza Berenjian. **Recall of Synonyms in Keyword-Based Information Retrieval in RICeST**. *Rep Opin* 2016;8(7):29-38]. ISSN 1553-9873 (print); ISSN 2375-7205 (online). <http://www.sciencepub.net/report>. 6. doi: [10.7537/marsroj080716.06](https://doi.org/10.7537/marsroj080716.06).

Keywords: Synonyms, Information retrieval, RICeST

1. Introduction

An internal search engine is one of the most frequently-used parts of a website. A built-in search facility is highly important in a website as it helps reduce the time to access the content. Since most of the leading websites use data banks to store the content and information, it has become a necessity to be able to search through data banks. With the rapid development of computer sciences in the world and their introduction into the third world countries, languages have tended to adapt to these technologies. It is further complicated when scripts have to be adapted to computer technologies in addition to language. Farsi language and script needs to be modified by computer scientists and linguists in order to be adapted to these technologies. One of the procedures to respond to these needs is to develop software to modify Farsi language and script computationally.

As some linguists contend, although two words may prove synonymous in a chain of words, they may have varying meanings in another chain of words or when adopting different grammatical roles. Still, we assume the words as synonyms when they have the same meanings in a chain of words. It should be noted that synonyms play an important role in understanding the meaning of a word. Thus, knowing the synonyms of a word may help decrease disagreements in choosing keywords so that it will facilitate the selection of appropriate keywords.

1.1 Statement of the problem

Synonymy

Modarresi (2006) considers synonymy as the equivalent of ‘sameness of meaning’. She writes, ‘words with the same meanings are called synonyms. Synonymy is possible only when an element could be replaced with another element in discourse without changing the meaning.’ According to philologists, synonyms are the words with the same meaning as other words or phrases (Bahar, 1990).

The word synonym is derived from the Greek *sunónomos*, which literally means ‘with the same meaning’. Synonyms are two or more words with the same or similar meanings (e.g. heart and cardio). It is also referred to as two or more words denoting the same meaning (Dehkhoda, 1992).

Some totally-different vocabulary words may have identical meanings as follows (words in Farsi):

Elm & Danesh, Nour & Roshanaie, Khaneh & Sara, Jameh & Lebas, Sabr & Shakibaie, Solh & Ashti, Dasht & Biaban, Khoob & Nik, Gham & Andooh, Ghamgin & Andohnak (Anvari & Ahmadi Givi, 2009).

Palmer believes that two words are synonymous when they share the same antonym (as cited in Safavi, 1981). Synonyms are words with different lexical structures but with identical meaning in usage. When words denote the same object, they are said to be identical in a less technical sense. In such cases, there could be stylistic differences (Falk, 1992). Synonyms may have similar, equivalent or identical meanings. Under certain circumstances, a word may be more suitable than its equivalent, or two words may only be synonymous in certain senses (Richard, 1993).

We tend to express the same idea in various ways. The more we are talented in writing, the more we are comfortable with diversity of expression. This makes it difficult for the information retrievers to guess the likely phrasing of sentences by authors. We need specific retrieval systems that are capable of adapting ideas but not words. Besides, various words may be used in different disciplines to denote the same entity. In an accurate compliance system, we may lose important information unless we use synonyms. A natural language processing system should be able to extend the query automatically using the name of the locations and synonyms (Mehrad & Falahati Fomani, 2005).

Some linguists believe that it is rare or impossible for two words to have identical meanings. The reason for the lack of absolutely-identical synonyms is the inconsistency among all sematic features in the words. Thus, we consider the words as synonymous when they belong to the same grammatical category and sematic field. In this regard, two or more words are synonymous when they belong to a similar semantic network and grammatical category (Khoda Parasti, 1997). Two words are synonymous when they share a set of features. Synonyms may be considered as lexical interpretations. It is difficult to find two absolutely synonymous words. When two words share the same sense, one may adopt an extra, distinguishing sematic feature to diverge from another (Falk, 1992).

Modarresi (2006) writes, 'one should bear in mind that two words are synonymous when they share the same semantic features. Synonyms can be thought of as lexical interpretations. It is difficult to find two absolutely identical synonyms.'

Considering the above discussion, absolute synonymy is rare. In this regard, even the words with quite similar senses diverge in their subtle pragmatic aspects. For example, the words in each pair of square brackets are close in meaning: {Sabet, Paydar, Paya, Bargharar, Bar Jay, Bar Makan} and {Hadaf, Amaj Gah, Barjas, Neshan Gah}. A glance at a Farsi dictionary will reveal a plethora of such lexical groups with the same conceptual relations. Every Farsi speaker understands that the words *Hadaf* and *Amaj Gah* vary in their usage. A study of the content and features of words demonstrates that no two words bear exactly the same meaning because they may slightly vary in terms of content and semantic loading. It seems that linguistic features do not allow for completely similar meanings between two words. Synonyms may only gain identical senses in certain collocations in a chain of discourse (Modarresi, 2006). Thus, synonymy is relative between two words so that no two words are absolutely synonymous.

Khoda Parasti (1997) argues that an individual phenomenon does not matter in synonyms. Rather we have to deal with a set of phenomena amounting to twelve categories. These twelve comparative categories include stylistic/sematic/figurative/virtual/syntactic/functional/total/approximate/absolute/complete/affective and explanatory each of which accommodating subcategories originated from them based on specific criteria. In sum, synonymy, homonymy (homophony and homographs), polysemy and hyponymy could be studied in a network of conceptual relations.

Absolute synonymy

It is often said that synonyms have the same meaning. However, such definition is against the economy of language. Absolute synonymy is rare. Rather, there is near or partial synonymy relations among some words. Even the words with quite similar senses diverge in their subtle pragmatic aspects. For example, the words in the following sets are near synonyms: {Sabet, Paydar, Paya, Bargharar, Bar Jay, Bar Makan}, {Hadaf, AmajGah, Barjas, Neshan Gah} and {Bartari, Ollov, E'tela, Rojhan, Balatari, A'laie, Arfa'ee} (words are in Farsi). A glance at a Farsi dictionary demonstrates the broadness of such conceptual relationship in Farsi language. Every Farsi speaker well knows the usage of these words. Thus, the differences in the meanings of synonyms arise from dialectic, stylistic (literary, non-literary, etc.), categorical (written, spoken, formal, informal) and diachronic uses. These differences are typically associated with usage and pragmatics. Still, if we discard pragmatic differences as potential differentiating factors, we may find ample synonyms in any language (Gholamali Zadeh, 1995). According to Modarresi (2006), types of synonymy are as follows:

A) Dialectic synonymy: synonyms may belong to a different dialect such as *Madandarin* Khorasani dialect vs. *Namadari* in Tehrani dialect.

B) Stylistic synonymy: two words may prove synonymous in two different/specific stylistic conditions. For example, the Farsi words *Zan*, *Khanom*, *Hamsar*, *Zaifeh* and *Madar-e Bachehhaare* synonymous in different cultural and spoken styles.

C) Non-affective synonymy: sometimes, two words are synonymous but carry variable affective loading. Still, they bear the same semantic content such as *Doust* and *Yar*.

D) Contextual synonymy: some synonyms suffer collocative limitations. That is, they may only collocate with certain words. For example, *Sara* and *Khaneh* have the same meaning, but *Khaneh* collocates with *Marizto* form the compound word *Mariz Khaneh* (hospital) while *Sara* cannot make such a compound word with *Mariz*. The same is true with

Fased and *Kharab* as the latter tends to collocate with *Divar* but not the former.

E) General vs. particular synonymy: in a synonym pair, a word may have a more general and inclusive sense comparing with its counterpart. Thus, they may be synonymous in certain uses. For example, in the synonym pair *Mashin* and *Automobile*, the former is more general and inclusive than the latter as it may denote any mechanical device.

F) Homonymy: words with the same pronunciation and spelling but different meanings are homonyms such as the word *Shiir* in Farsi – meaning a lion, a faucet and milk – where both the spelling and pronunciation are the same. One should note that homophony and homographs are examples of partial homonymy.

G) Homographs (partial homonymy): two words with the same spelling but different pronunciations are homographs such as *Shour* and *Dour* that have similar spellings but different pronunciations.

1.1.1 Causes of synonymy in words

Bahar (1990) believes that some words are meant to embellish the language and create balance and equilibrium. Although the use of synonyms was not still in vogue in the fourth and first half of fifth century AH, it is still traceable in the sermons and preface to chapters as well as in the praises and hyperboles. Although juxtaposition of synonyms was not necessary in prose, it was crucial in poetry to create balance and rhyme. Thus, the use of synonyms was introduced into prose through poetry.

The causes of synonymy are the diversity of expression and crystallization of actualized tropes such as *Gardidan* (to gyrate) that means *Gardesh* (circular movement) literally but has gained a figurative sense of *Sirurat* (transformation) figuratively. *Nemoodan* means *to present* and *to offer* literally but has gained a figurative sense of *to do* over time.

From the fourth century AH onward, synonyms have increasingly found their way into writings. Particularly, Arabic synonyms are used more frequently. From the sixth century AH, the use of synonyms, puns and Arabic words have increased in Farsi writings, which will be discussed in the following sections.

In the fourth century AH, religious terms or the vocabulary with no Farsi equivalent, obsolete and poetic words as well as formal Arabic vocabulary or Arabic synonyms were used in Farsi prose. However, these words do not exceed five in one hundred words providing that it is a translation from Arabic. The number of such words does not exceed three in one hundred words in the texts written directly in Farsi, cited or translated from Farsi or Pahlavi sources (Bahar, 1990).

Many factors contribute to the creation of new synonyms some of which are as follows:

Firstly, different varieties of language account for the creation of new synonyms so that a special relationship is created between two varieties of language. Example: *Gorikhtan* (to flee) and *Jim Shodan* (to bug out).

The second cause of synonymy is the assignment of the specific to general such as *Helleh* (garment) that is a kind of garment but also generally referred to as all types of garments.

Third is the meaning assignment that creates synonymy among words such as *Kawthar* that is a spring in the paradise but also used to name any spring.

Fourth, synecdoche may also contribute to the creation of synonyms. For example, *Azal* refers to the beginning of creation but is generally referred to any kind of commencement.

Fifth, changes over the course of time may contribute to the creation of new synonyms. For instance, the meaning of *Shookh* (dirty) has altered to humorous over time.

Sixth, the use of rhetorical devices and figures of speech such as metaphor and irony helps juxtapose some vocabulary as synonyms.

Seventh, the borrowing of vocabulary from other languages and coining their equivalents are another source of new synonyms in Farsi. Nowadays, the majority of synonyms created by the Academy of Persian Language and Literature fall in this category. Still, most of these words are of foreign origin with synonyms in their native language. According to Modarresi (2008), the differences among synonyms arise from dialectic, stylistic (literary, non-literary, etc.), categorical (written, spoken, formal, informal), diachronic (non-synchrony), contextual (having different collocations) and affective loading. Thus, semanticists deny the existence of absolute synonymy among words.

The majority of Arabic words borrowed into Farsi have had Farsi equivalents or fallen into similar semantic domains. Most often, semantic differences that functioned as the distinguishing factor in Arabic language did not matter in Farsi so that the words appeared as synonymous by the elimination of distinguishing factors. Example: *Ghana*=*Servat* (wealth), *Gham*=*Hozn* (sorrow)

The majority of Arabic words inflected variously in Arabic language were treated as synonyms. Such words usually share the same stem and fall into the same semantic field. Example: *Naf'e*=*Manfe'at* (benefit), *Manhoos*=*Nahs* (ill-omen)

The majority of the words borrowed from Arabic have synonyms in Farsi. Thus, they form synonym sets comprising Arabic and Farsi words. Such

synonyms are numerous in Farsi. Example: *Elm*=*Danesh* (knowledge), *Hers*=*Aaz* (greed)

Synonyms are created either advertently or inadvertently in the process of word coinage in Farsi where Arabic words play a role. The morphology of the coined words is such that they may fall into the following categories.

1. The same (Arabic) stems and variable (Farsi) suffixes: normally, suffixes help create synonyms. Example: *Gham+gin* (sad), *Gham+nak* (tragic); *Hozn+alood* (sorrowful), *Hozn+angiz* (mournful)

2. Different (Arabic) stems and the same (Farsi) suffixes: in such constructions, the stems are synonymous. Example: *Vahshat+nak* (dreadful), *Howl+nak* (frightening), *Vahm+nak* (awful); *Gham+alood* (sad), *Hozn+alood* (sorrowful)

3. Using the words synonymous to the Arabic word and affixing mechanism in Farsi such as *Montaghem=Entegham Ju* (revenge seeker); *Amigh=Omgh Dar* (deep); *Makhdoush=Khadsheh Dar* (defaced)

4. Applying the Arabic morphological rules and fabrication of words such as *Jeddiyat* (seriousness); *Bashariyat* (humankind); *Khejalat* (shame); *Khajel* (shameful)

•
reating synonyms from a pair of Farsi-Arabic or Arabic-Arabic antonyms using Farsi prefixes such as *Abad#Kharab* (built#destroyed); *Na Abad=Kharab* (destroyed); *Na Ashna=Gharib* (stranger)

•
ending Arabic words into Farsi vocabulary through the elimination of inconsistent and bizarre elements: normally, such words have synonyms in either Farsi or Arabic. Example: *Marjo'e=Marjo'ee* (returned); *Kolliya=Kolli* (all/total) (Khoda Parasti, 1997)

In the vocabulary network, the words are categorized into cognitively synonymous sets. These sets are called synonym sets. Thus, a synonym set comprises a set of words of the same grammatical category that can replace one another in a given context. Every synonym set denotes a different concept. The members of the set represent the given concept in different texts. Synonym sets are related together via words and semantic-conceptual relations. The relations in the Word Net may exist among words of the same category (e.g. synonymy, antonymy, hyponymy, meronymy) or of different categories (e.g. features and derivatives) (Davari Ardakani & Fakourian, 2013).

1.1.2 How the search engine works

RICeST search engine is unilateral and keyword-based functioning according to the system of bag of words. Despite the relative success of this system, two major semantic concerns are involved: semantic ambiguity and synonymy. When the phrase entered

into the search engine involves a word with semantic ambiguity, irrelevant documents containing the given word are retrieved in addition to the relevant documents. With regard to synonymy, it is always likely that the engine fails to retrieve many relevant documents due to the lack of the entered keywords in the documents (Manns, 2000). A user does not necessarily intend to retrieve the documents containing the exact keywords; rather, he often prefers to retrieve documents containing the words with his intended meanings. Indeed, he looks for the documents containing his intended meaning but not the keywords (Ahmadi Nasab, 2012). It should be noted that synonyms play an important role in understanding the meaning of a word. Therefore, knowing the synonyms of a given word may help decrease diversity in selecting keywords and facilitate appropriate keyword selection.

If a word with semantic ambiguity is searched in a query, the retrieved documents may contain the keyword but not the intended concepts and content. On the contrary, a document may not be retrieved in a query despite the existence of synonyms in the document. This is because synonyms were not used in the query. Most of the existing approaches to demystifying the meanings of words and/or identifying synonyms draw upon lexical categories such as in Word Net (Mehrad & Falahati Fomani, 2005). A word is tagged with one of the concepts it represents (such as a group of synonyms in Word Net) in order to disambiguate it. The concept is identified by using the context in which it appears. Moreover, one may also use synonyms. To this end, extended query technique is used whereby synonyms are added to the query (Voorhees, 1994).

1.2 Literature review

Although thorough definitions were provided for synonymy in previous sections, a few more definitions by predecessors are offered below.

In the past, various definitions were offered for synonymy in Persian literature. For example, Fakhr ad-Din ar-Razi (606 AH) observes that synonymy is the consecutive use of individual words denoting the same sense. According to philologists, a synonym is a word that is similar to another in terms of semantics (Bahar, 1990). Ahmadi Givi and Anvari (2009) contend that words may bear three types of relationships:

1. Lexically different but semantically similar: such words are synonyms.
2. Semantically different but lexically similar: these are analogous.
3. Not only different both semantically and lexically but also opposing semantically: these are antonyms.

Mehrad and Falahati (2005) write

Synonymy is one of the most well-known semantic relations that justifies the existence of glossaries and crossword puzzles in magazines and newspapers. Synonymy has a simple but tricky definition: different lexical elements with the same meaning. This definition, however, does not make it clear what we mean when we talk of the same meaning. We may respond using the concept of interchangeability: two words are synonymous when we can replace one for another in a sentence without changing the meaning or acceptability of the sentence. No doubt, real synonymy will be rare once interchangeability is meant in any given linguistic context. This is because we often find it impossible to interchange the synonymous words in a sentence. Thus, we arrive at a weaker version of synonymy definition: two words are synonymous when we can replace one for another in some linguistic contexts.

Sibawayh asserts that synonymy consists of two different words with the same meaning. According to Modarresi (2008), the words or phrases with the same sense are referred to as synonymous. Synonymy is only possible when an element of discourse could be replaced with another without a change in the meaning of the discourse. One should note that two words are considered synonymous when they share a set of similar features. In many dictionaries, synonyms are cited to further illustrate the meaning of a word. The following dictionaries are published in Iran on synonyms:

1. A Comprehensive Dictionary of Persian Synonyms and Antonyms (Khoda Parasti, 1997)

2. Synonyms and Idioms Dictionary (Mohammad Shad)

3. The researcher has recently been informed that Mr. Ali Purhosseini – a researcher with the research panel of Astan-e Quds-e Razavi – has compiled a Dictionary of Persian Synonyms and Antonyms in 6 volumes sponsored by Faragostar Mohaghegh Research Institute. By compiling this dictionary, Mr. Ali Purhosseini aimed to help accelerate the acquisition and retrieval of information. In the preface, he does not limit the application of the dictionary content to information storage and retrieval. Rather he asserts that the dictionary may help in the compilation of Thesauri, choosing headlines, broadening a writer's/speaker's vocabulary to add diversity to texts or speech. The dictionary involves 24921 entries, 41826 synonyms, 6781 antonyms and 250,000 words. Hopefully, the dictionary will be published and distributed as soon as possible (aqlibrary, 2014).

2. Moreover, JafarSadeghi (2011) compiled the Dictionary of Analogous, Synonymous and Antonymous Words in Farsi, which comprises vulgar words and idioms, new and common vocabulary as

well as contemporary Farsi works. This dictionary is a supplement to the Dictionary of Slang Words by Jamal Zadeh. The dictionary is intended to provide the scholars of Persian prose with an invaluable source aiming to eliminate lexical ambiguities and compile a special dictionary of the Farsi words commonly used in contemporary Persian prose. Some of the features of this dictionary include:

A) Citing the pronunciation of words and some phrases

B) In each entry, the word in parentheses indicates alternative pronunciation.

C) In the words with two different pronunciations, the contemporary pronunciation is favored.

D) Silent *ha* at the end of the words is considered as a diacritic.

E) The meaning of each word is illustrated based on its contemporary sense.

F) Examples are cited for appropriate usage of each entity.

G) The origin of non-Farsi and non-Arabic words is cited.

Most of the dictionary writers such as Dehkhoda, Moin, Amid, Nafisi, Aryanpur, Bateni and Golestani have carried out studies on synonyms. They tended to cite synonyms in their dictionaries. Still, they did not compile a specialized dictionary of synonyms. Retrieval systems in information search have recently come into vogue. However, there is no discussion on retrieval systems except for in the books on information system. During a search, the user usually enters a keyword in order to retrieve the documents containing a given concept but not necessarily the keyword. To help achieve this goal, some retrieval systems use thesauri. A thesaurus comprises a set of controlled vocabulary illustrating the relations between contents and concepts. A common approach to disambiguation is to use a lexical network. That is, the intended keyword may be tagged with its different senses to be disambiguated, and extended query could be used to solve the issue of synonymy (Ahmadi Nasab, 2012).

To the best of the author's knowledge, no study has already been conducted in Iran on *Automatic Calling of Synonyms in Keyword-based Information Retrieval*.

2. Discussion

2.1 Implication

Identification of synonymous words and/or semantic differences among words helps us clarify and understand the senses of words. Synonyms may be used to determine the exact meaning of the keywords in keyword-based retrieval of information and journal articles, which helps with disambiguation

and retrieval of relevant documents. This research project may also be applied to the following situations.

1. Presenting Farsi synonyms to the users
2. Providing students of Farsi language with synonyms
3. Users' selection of synonyms to increase the accuracy of retrieved documents
4. Increasing the accuracy of search through choosing appropriate keywords

Method

A number of 3500 word entries and their synonyms as cited in *A Comprehensive Dictionary of Persian Synonyms and Antonyms* were selected and typed into the computer. Subsequently, a software developer used the RICEst program to develop and implement the software in cooperation with the department of systems design and operation. Although some of the entries in the dictionary were not common in contemporary Farsi language, they were extracted and analyzed along with their synonyms in order to obtain the rules required for this project.

System testing

Conducted in a real environment, system testing proved successful. Since we have developed a special algorithm, it stands out from other existing algorithms. This system can provide immediate answers in high volumes in the shortest time (Figure 1).

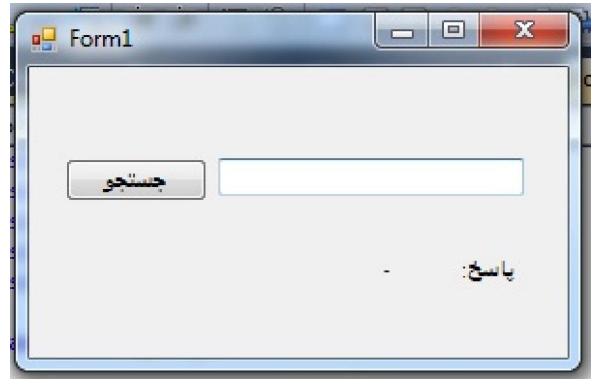


Figure 1. Search window (before installation on RICEst)

The user can enter the keyword into the search box and click the search tab. The search result will be shown on the bottom of the page (Figure 4 & 5). The software was initially tested by the researcher and software developer. Then it was piloted on E-Journal Search section of the RICEst website. Figures 6-9 illustrate various stages of the process.

Input

A sheet was created in Excel to add the entries in columns (Figure 2). The words were entered individually at first and collectively in later stages such as Khahar – Abji/ Hamshireh (sister).

Row	Column A	Column B	Column C
5292		خنده	خورد فرورس
5293		فانشنگی	خورد فرورس
5294		خود سرانه	خورد گمانه
5295		اسنداد	خورد گمانگی
5296		خودسر	خورد گمانه
5297		القدر	خورد گمانی
5298	نوزامند	ای تواری	خورد گمانا
5299	نوزامندی	ای تواری	خورد گمانی
5300	مناکف	گمانا	خورد گمانی
5301		مکشایم	خورد گمانا
5302		ممانون	خورد گمانی
5303	غیر	آشنا	خورد گمانی
5304		مهارگی	خورد گمانی
5305	مهر	آفتاب	خورد گمانی
5306		مجره	خورد گمانی
5307		مژده	خورد گمانی
5308		سرف کردن	خورد گمانی
5309		خوراکی	خورد گمانی
5310		خوراک	خورد گمانی
5311		آهنر	خورد گمانی
5312	مهر	آفتاب	خورد گمانی
5313		حمام	خورد گمانی
5314	ناخوشی	خورد	خورد گمانی
5315		اسکنول	خورد گمانی
5316	فخر شایند	نایبند	خورد گمانی
5317	بد اخلاقی	نیک خلقی	خورد گمانی
5318	بد شناسی	خوش شناسی	خورد گمانی
5319	بد صدا	خوش آواز	خورد گمانی
5320	بد هیول	خوش قامت	خورد گمانی
5321	مهر بار	ساده	خورد گمانی
5322	بد سفین	خوش کلام	خورد گمانی
5323	بد خط	خوش نویس	خورد گمانی
5324	بد خلق	نیک خلقی	خورد گمانی
5325	بد خو	مهربان	خورد گمانی
5326	بد خوراک	شکام پرست	خورد گمانی
5327		آفتاب آفتاب	خورد گمانی
5328	بد خیم	ای خنجر	خورد گمانی
5329	بد جاس	نیک ذات	خورد گمانی

Figure 2. A sample input datasheet in Excel

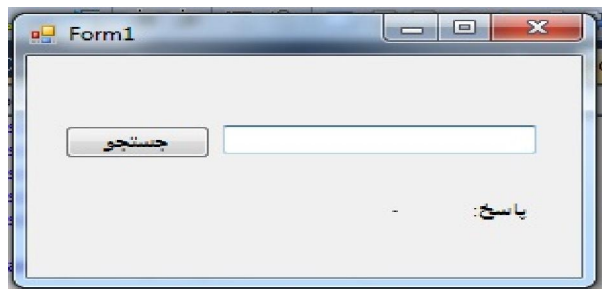


Figure 3. A sample search window (before installation on RICeST)

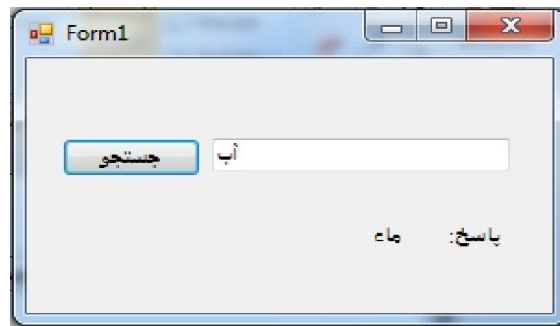


Figure 5. Aab (آب) search window (before installation on RICeST)

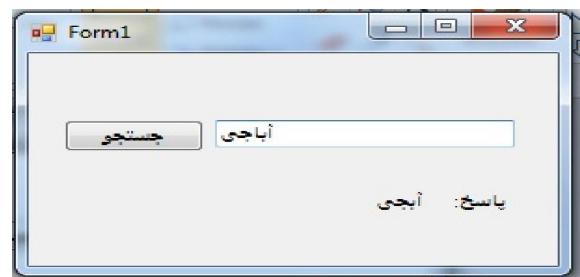


Figure 4. Abaji (آباجی) search window (before installation on RICeST)

The software was tested by checking various keyword searches before being installed on RICeST. Examples are illustrated in Figure 4 and 5. The word *Abaji* (آباجی) is searched in Figure 4, and the search result is *Abji* (آبجی). In Figure 5, the word *Aab* is searched, and the result is *Ma'* (ماء) as its synonym. These tests were conducted to ensure the accuracy of the software. Then the software was temporarily installed on the E-Journal Search section of the RICeST website (Figure 6).



Figure 6. Search window (after installation on RICeST)

As shown in Figure 6, the search window pops up with empty fields. Next to the third field on the left, a small square box appears under the operators *AND/OR*. The phrase in red print beside the box reads 'search together with the keyword synonym'. The user may check the box if he wishes to search the keyword together with its synonym. The search result contains

the keyword and its synonym. For example, the word *Aab* in figure 7 is searched together with its synonym. The result is illustrated in Figure 8 where *Ma'* is also illustrated as the synonym of the keyword. The result of searching the word *Abkhori* (trough/canikin) is illustrated in Figure 9.

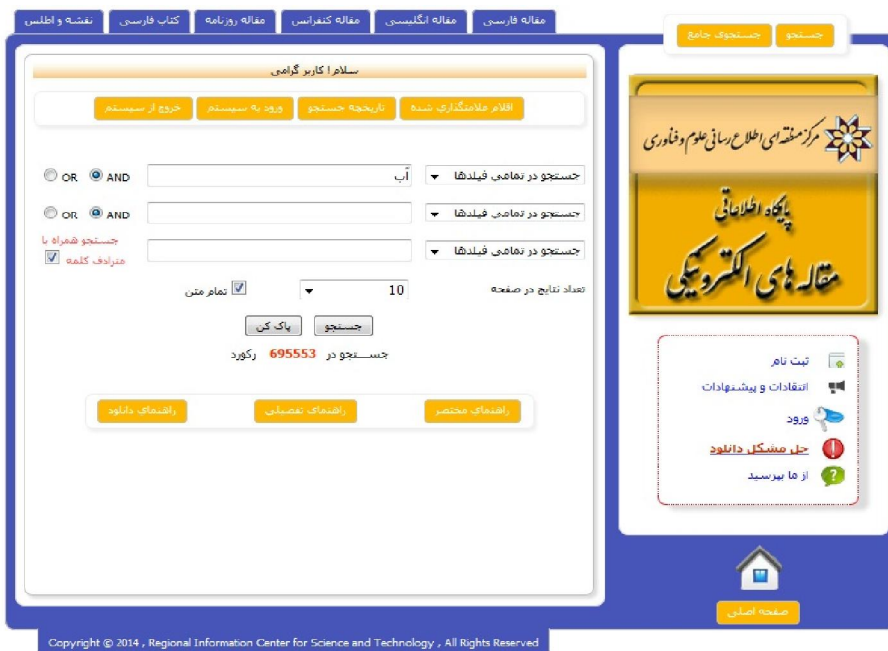


Figure 7. Search window for *Aab* and its synonym in RICeST



Figure 8. Results of searching *Aab* and its synonym in RICeST



Figure 9. Results of searching *Aabkhor* and its synonym in RICEST

Programming language

In this project, Visual Studio in the form of C# was used to develop the software due to its compatibility with RICEST software. In this system, the user may search a keyword together with its synonym one at a time.

2.2 Advantages of the development of Automatic Calling of Synonyms

This system may have multiple uses as follows:

- To increase the keywords to be searched
- To be used in information storage and retrieval
- To be used in language teaching
- To be used in compiling thesauri
- To broaden the repertoire of vocabulary in writers, poets, lecturers and orators
- To be used in adding lexical diversity to writings and lectures

3. Conclusion

This system was designed and implemented in the Regional Centre for Information Science and Technology (RICEST). This article reports part of the research project carried out in RICEST. This system is used on the RICEST intranet and website. It is recommended that further studies be conducted to modify the software so that:

A) The user may see more than one synonym of the same keyword.

B) The user may see the antonym(s) of a given keyword in addition to its synonym(s).

C) The software could be written using other programming languages.

D) The software could support other languages such as English, French, etc.

E) Due to a lack of research on synonyms, except for the present one, further studies may be conducted on larger corpora to determine the status of synonyms in language.

References

1. Ahmadi Givi, H., & Anvari, H. (2009). *Farsi grammar 1*. Tehran: Fatemi Cultural Institute.
2. Ahmadi Nasab, F. (2012). A semantic study of document retrieval in the Regional Centre for Information Science and Technology and developing a semantic model to improve information retrieval. Shiraz, Research project.
3. Anvari, H., & Ahmadi Givi, H. (2006). *Farsi grammar 2*. Tehran: Fatemi Cultural Institute.
4. ar-Razi, F.D. (1938). *Nahayat al-Ijaz fi Derayat al-E'jaz*. Cairo.
5. Bahar, M.T. (1990). *Stylistics*. Tehran: Amir Kabir Publications.
6. DavariArdakani, N., & Fakourian, N. (2013).

- Investigating morphological semantics and its application in lexical networks. *Second National Conference on Teaching Farsi Language and Linguistics*, Shiraz, Iran.
7. Dehkhoda, A.A. (1992). *Lexicon*. Tehran: Tehran University Publications.
 8. Falk, J.S. (1992). *Linguistics and Language: A Survey of Basic Concepts and Applications*. Kh. Gholamali Zadeh (Trans.). Mashhad: Astan-e Quds-e Razavi.
 9. Gharib, A.A., & Bahar, M. et al. (1992). *A grammar of Farsi language (five masters)*. Tehran: Vazheh Publications.
 10. Gholamali Zadeh, Kh. (1995). *Structure of Farsi language*. Tehran: Ehya-e Ketab Publications.
 11. Jafar Sadeghi, H. (2011). *Dictionary of Analogous, Synonymous and Antonymous Words in Farsi*. Tehran: A'lami Publications.
 12. Khoda Parasti, F. (1997). *A Comprehensive Dictionary of Persian Synonyms and Antonyms*. Shiraz: Danesh Nameh Fars Publications.
 13. Mehrad, J., & Falahati Fomani, M.R. (2005). *Semantics and information retrieval*. Mashhad: Computer Library Publications.
 14. Modarresi, F. (2008). *From syllable to sentence*. Tehran: Chapar Publications.
 15. Reza Khah, Z., & Kazemian, M. (2012). *A dictionary of synonyms*. Tehran: Research Center for Humanities and Cultural Studies.
 16. Richards, J.C., Platt, J., & Weber, H. (1993). *Longman dictionary of applied linguistics*. H. Vosooghi, Mirhosseini, A.A. (Trans). Tehran: Center for Book Translation & Publication.
 17. Safavi, K. (2007). *An introduction to semantics*. Tehran: Pezhvak Keyhan Publications.
 18. Voorhees, E. (1994). Query expansion using lexical-semantic relations. In Proceedings of the Seventeenth Annual International Conference on Research and Development in Information Retrieval, page 61-69.
 19. <http://aqlibrary.ir/blog/book> (1393).

7/25/2016