THE PROBLEMS OF DESKTOP INDEXING OF A BOOK TRANSLATED INTO A NON-ROMAN SCRIPT: DESCRIPTION OF A REAL EXPERIENCE

MORTAZA KOKABI

Associate-Professor, Dept. of Lib. & Inf. Science, Shaheed Chamran Univ., Ahwaz, Iran
Kokabi80@yahoo.com

Zarnegar (gold writer) is a word processor widely used by publishers of both scholarly journals and books in Iran. Although it is gradually substituted by Word for Windows that is much more powerful than Zarnegar, the process seems to be slow and most Iranian publishers still prefer to receive manuscripts in Zarnegar than Word. There are many reasons for this preference: Word, though having many great features such as compatibility with other Windows applications and especially with Word Wide Web (WWW) is poor in terms of Farsi, the official language of Iran. The main reason for this poorness is that Farsi versions of Word are in fact Arabic versions. Zarnegar has been developed by Iranians familiar with Farsi script and acts more conveniently than Word in Farsi writing. The fonts employed in Zarnegar are mostly Farsi and seem to be more beautiful than Word due to the tradition of calligraphy among Iranians. Some fonts in Word employ Arabic letters that are not used in Farsi. This feature is of much dislike between Iranians using it. The author, when finished translating an English book into Farsi, confronted some problems when trying to prepare the index to the book. When providing an index to a translated book, the logical criterion to select the entries of the translated index is to base them on the entries of the original language index. Therefore, the entries of the original index were translated, and then rearranged in Farsi alphabet in another file When the task of allocating the page numbers to Farsi entries began by using the "Find" option under the "Edit" menu in Zarnegar, the different equivalents selected in each chapter for a single concept and not found totally by searching under the same term, showed up as the first major problem. The second major problem encountered was the similarity between the concepts used both in a very specific sense as well a general sense. The latter increased irrationally the number of pages containing some entries. Since the book translated was on the theoretical as well as practical aspects of library services, the word "use" was an example of problems caused by the similarity of words that have both specific and general meanings in library and information science context. Some suggestions are provided in the article as to how to start the task of translating the book by first translating the index into Farsi, providing an English-Farsi dictionary as well as a Farsi-English one for frequent reference to them in order to find selected equivalents, how to allocate page numbers while using the Farsi-English dictionary to provide the final Farsi index and how to overcome the problems mentioned above.

Keywords: Zarnegar, Word for Windows, desktop indexing, non-roman script, Farsi

INTRODUCTION

Zarnegar (gold writer) is a word processor widely used by publishers of both scholarly journals and books in Iran. Although it is gradually substituted by Word for Windows (hereafter called Word) that is much more powerful than Zarnegar, the process seems to be slow and most Iranian publishers still prefer to receive manuscripts in Zarnegar rather than Word. This is not to say that Iranian publishers accept only typed manuscripts, but a typed manuscript, especially when typed by a word processor gains more points when considered for publication than a

handwritten one. Faslname-ye Ketab (Book Quarterly), the journal published by the National Archive and Library of the Islamic Republic of Iran, even in its Fall 2003 issue, states that: "… Sending the diskette or file containing the article in Zarnegar or Word format will accelerate the process of publication"[1].

There are many reasons for this preference: Word, though having many great features such as compatibility with other Windows applications and especially with World Wide Web (WWW), is poor in terms of Farsi, the official language of Iran. The main reason for this poorness is that Farsi versions of Word are in fact Arabic versions.

THE SIMILARITIES AND DISSIMILARITIES OF FARSI AND ARABIC

Farsi and Arabic scripts and languages, although show much resemblance, have differences that make them somewhat incompatible. Farsi has four letters that Arabic does not, namely ?, (p) ?, (ch as in chat)?, (zh as in Zhivago) and ? (g as in gas) Arabic has seven more letters than Farsi. These are: ?, ?, ?,, ?, ?, the letter ? with a hamzih (will be siscussed later is some detail) on it, and the letter ? with two dots under it (the last two, are found in some fonts used in word but cannot be generated as single letters to be shown here). The marks of vocalization are employed more frequently in Arabic than in Farsi, the feature that makes Arabic to be more easily pronounced than Farsi. The easy pronunciation of Arabic is also due to the fact that each letter in Arabic alphabet has its own sound, but in Farsi, some different letters have similar sounds and that makes the writing of Farsi quite difficult. Letters ?, ?, ?, and ? in Arabic are pronounced in such a way that the listener familiar with the sounds, distinguishes them easily. These four letters are pronounced the same in Farsi, similar to pronunciation of ? in Arabic. The same is true for ? and ? in one hand and ? and ? on the other, in Farsi, that are pronounced similarly. The other examples of this kind are ? and ?, and ? and ?. Therefore sometimes even Iranians familiar with Farsi have difficulty distinguishing between them. The same is true for letters ?, ?, and ? that although are pronounced differently in Arabic, have the same pronunciation in Farsi. Arabic is in general, more routine than Farsi that has been affected by many other languages during its long history, and learning Arabic is much easier than Farsi.

THE PROBLEMS INVOLVED IN WRITING AND INDEXING FARSI

The following examples might illustrate the problems involved in writing and indexing Farsi:

Farsi is a cursive script: the word Farsi itself that is written as: ????? is made by connection of letters: ? ? ? ? ?. This feature gives the spaces between the words more meaning since gaps between the words play an important role in understanding Farsi. The word ????? is a single word that means, "Soldier". The same word, if written with a space in between, that is: ?? ??? means, "the open end" and with a slightly modified pronunciation means, "open-ended". Similarly, the word ?????? means, "twelve", while ?? ?? ?? means, "two out of ten" and ??? ??? means "drugged, medicated by drug". Word sometimes acts weakly in terms of these gaps: the word ??????? (meaning you have gone) must not have a space between the parts: ???? and ??. Word cannot generate this form and the operator must either type the word as: ?????? (absolutely without any space between the parts that makes the word unreadable, incomprehensible, and wrong) or ???? ?? (with a space between the parts). Both forms are incorrect, but the latter is more eligible than former. The third alternative is to press Shift + space bar to connect the two parts while the eligible form is maintained, but in even some 2000 versions of Word, this inserts a small vertical line between the parts, the fact that bothers the typist and in the printout the ultimate form is sometimes not desired. The same happens for the plural sign ?? in Farsi that is added to the end of some words. This ending can be attached to some letters before it and cannot, to some due to the

characteristics of the letters before ??. When this sign can be placed immediately after a letter without attaching to it, there is no problem, but when it should not be attached to a letter, unless the space is inserted, the parts connect to each other and make the word wrong. The correct form of the word: ???? ?? (fruits) is the same as typed but without any space between the parts. If the space is not inserted, the word becomes: ??????, since ? is of the kind of letters that attaches to the letters after it, and that is wrong. The situation becomes worse when the first part of such words type happens to be at the end of a line. The first part is placed at the end of the line and the second part at the beginning of the next line by Word. This form is both incorrect and irrational. The spaces between the words in Farsi has a great impact on indexing. The following example, taken from the book "Rules for filing catalog cards[2] explains the effect of the between-the-words spaces in indexing Farsi words:

?? ??????? ? ??? ?????? ????	?? ??????? ? ??? ?????? ????
?????	?? ?????
????? ????	?? ???
?????? ?? ????? ????? ???? ?????	?????
?????	????? ????
?? ?????	**??**???? ?? ????? ????? ???? ?????
????? ???	?????
?????? ????	????? ???
?? ???	?????? ????

FIGURE 1: A COMPARISON BETWEEN THE LETTER-BY-LETTER ARRANEMENT AND WORD-BY-WORD ARRANGEMENT OF FARSI ENTRIES

The two columns in fig. 1 show two different kinds of indexing in Farsi. The right column is called letter by letter indexing, since the spaces between the words are not considered. The left column is called word by word indexing, in which spaces between the words are important. The sixth and ninth entries in right column change place and become the second and third entries in the left column. Therefore selection of either type of indexing affects the order of entries. Farsi is quite irregular in terms of spaces between the words. Words like ???????? , ????? ????, ????? ????, ??????????, ?????????, ???????, ?????? are seen in both forms: the two parts attached or separated. Due to the irregularity of between-the-words spaces in Farsi, indexing of the first type, letter by letter is recommended. But there are two problems in this regard: the first is that this recommendation is suggested and partly observed only by Iranian librarians' and the second is that this type of indexing has its specific problems. The first three entries of the left column belong to names of three kinds flowers (the word ?? means flower). Adopting the word by word arrangement has the advantage of putting together related entries; therefore all entries belonging to the flowers are placed together. This is not always done in letter by letter indexing. The same three entries are scattered among other entries. The three entries are placed as first, sixth, and ninth in the right column, the one arranged in letter by letter arrangement. In a longer index, the problem becomes drastic!

Different forms of words written in Farsi in terms of spaces between the parts and also the effects of Arabic on Farsi, causes some problems in finding the entries to be included in an index when the "Find" option under the "Edit" menu is used. The Islamic Revolution with its religious nature strengthened the status of Arabic language in Iran and that has in turn strengthened the problems Farsi been tackling for centuries. The following are but some categories of the problems:

1. Some words are written in two forms due to employment of letters with different shapes and similar sounds: the words ???? and ???? (both taken from the French word “billet”, meaning ticket are an example of this kind. Some other examples are: ???? and ???? (room), ????? and ????? (battery), ??? and ??? (the name of a city in Iran). An example of this kind that causes much difficulty is the words: ?????? (meaning to perform, to act) and ?????? (meaning to put, to place in a physical manner). These two words are used very much in Farsi words as the second part of two-word verbs. They are wrongly used interchangeably.

2. Some words are written in two forms due to the repetition of a single letter: the words ???? and ????? (the boy name David), and ???? and ????? (a boy name) are some examples for this type.

3. Some words are written in two and even three forms; the original Arabic word being made Farsi: The Arabic word ????? (the instrument in which lamps were placed) is sometimes written as its Iranian form: ????? although the pronunciation is the same. The word is sometimes written as: ?????, something in between the Arabic and Farsi forms. The words ???? and ???? (life), ???? and ???? (some part of the property that is donated religiously), ???? and ???? (mirror) are of the same type. The third form is also sometimes seen.

4. The Words ???? and ????? (meaning Issac) are the same in terms of pronunciation but the latter is the Farsi form of the former that is Arabic. Some other examples of this kind are: ???? and ?????(a boy name, also meaning generous), and ?????? and ??????? (Ismaiil).

5. Some words are written in Farsi by using both the Arabic t letter ? and Farsi t letter ?. These are especially found in Arabic phrases sometimes used in Farsi. The words: ????? and ?????, ???? and ????, and ??? and ??? are some examples of this type.

6. The word ???? (God) is used in many Iranian compound names and phrases (in fact Arabic ones). The form is used in two alternative forms: ??? and ????. Personal names such as ?????? and ???????, ???? ??? and ???? ???? are written in both forms. Due to the religious ideas among Iranians, these names and phrases still occupy a large proportion in Iranian names.

7. There is a decorative ? (b) in Farsi that comes at the beginning of the present tense verbs. The verbs: ?????, ?????, ???? meaning from left to right I say, I listen, and I go, are but a few of examples of this verbs with decorative b. This ? is sometimes attached to the verb and sometimes not. Therefore, the above verbs might be written as ?? ????, ?? ????, and ?? ???. The same is true for adverbial ? that is added to the beginning of the adjectives to make adverbs. Words like ????? (hastily), ??????? (slowly), ??????? (nicely, beautifully), and ????? (slowly) are some examples of this kind of adverbs. These are sometimes written like: ?? ????, ?? ??????, ?? ?????? and ?? ???? alternatively.

8. The Arabic word ??? (meaning, son of) that is placed between the names of the father and son is sometimes written as ??. Therefore, ??? ??? ???? and ??? ?? ???? are the alternative forms of a single name pronounced similarly. Although this form of name is merely Arabic and rarely among Iranians, but when indexing a book that for any reason contains such names, these become problematic.

9. The Arabic words ????, ????., ???? , ????, ????, and ???? find their alternative Farsi forms as: ????, ????, ????; ????, and ????. In such words the final ? becomes the final ?, the pronunciation however is the same.

10. The most problematic letter in Farsi is the Arabic letter ? (hamzih). The problems caused by this letter are so numerous that need to be categorized here. Following is only three of the problems caused by hamzih:

a. Some words that have hamzih are written in various forms. The words ??????? , ??????? (responsibility); ????, ????(group); and ?????, ????? are only few of these words. Finding these words usually is a burden in even Farsi dictionaries, since there is no single rule as where to put these words in the alphabetical order (figure 2 below shall illustrate the issue in some derail).

b. The words ending with hamzih, when used in possessive case produce various forms. The word ?????? (beginning), for example, ends with hamzih. The possessive case ?????? ??? (beginning of the work) is also written as ?????? ???. When hamzih is placed on the unpronounced h in Farsi, the problem becomes worse. The word ???? (plan, map) ends with a letter ? (h) that, since is not pronounced is called unpronounced h. In the possessive case, the phrase "the plan of the house", finds three different forms: ???? ???? (without any hamzih), ????? ???? (with hamzih after unpronounced h), and ?????? ???? (with a ? instead of hamzih, with the same pronunciation). Three forms are used in modern Farsi alternatively.

c. Words with any form of hamzih are sometimes written in a form that is rather Farsi than Arabic. This is an inclination newly developed among Iranians who want to have a pure Farsi language. Therefore, words like ???? and ????? are also written as: ???? and ?????. And finally,

11. Foreign words that are written in Farsi find different forms. The word "theater" for example, finds three forms: ?????, ????? and ????. The three forms sounds similar, very much like the word in French.

HOW DOES ZARNEGAR DEAL WITH THE PROBLEMS OF WRITING AND INDEXING FARSI

Zarnegar has been developed by Iranians familiar with Farsi script and acts more conveniently than Word in Farsi writing. The fonts employed in Zarnegar are mostly Farsi and seem to be more beautiful than Word due to the tradition of calligraphy among Iranians. Some fonts in Word employ the seven Arabic letters that are not found and used in Farsi. This feature is of much dislike between Iranians using it. One of these letters, as mentioned above is ? with two dots under it that does not exist in Farsi, and its presence in a Farsi text is quite unpleasant. Zarnegar, however, is a DOS-based word processor, therefore not compatible with Word unless with the help of a conversion program developed in Iran for the purpose, that itself causes its specific problems. It works, however excellently in terms of writing Farsi. Although it cannot deal with all the problems mentioned above, but it has some solutions for some the problems. In the above examples, in case of ???? ?? for instance, when the first part that is ???? is typed, by pressing the space bar, although a space is placed between the parts, but after pressing the space bar for the second time to go to the next word, Zarnegar automatically takes the second part one space back and connects it to the first part in its correct Farsi form. Zarnegar has an excellent spell checker for Farsi and Word does not. This feature of Zarnegar is of much interest to people who use it although it is not as automatic and fast as the red underlines of the English spell checker of Word.

ZARNEGAR'S SPELL CHECKER AND ITS IMPORTANCE FOR INDEXING

Zarnegar's spell checker is very much similar to that of Word, but it has been modified to cater for at least some problems mentioned above. The first window of spell checker consists of subsidiary dictionary, matching, suggestion for wrong words, language of spell checking (Farsi or English), and continuation of former spell checking. Three of the features mentioned, that is, subsidiary dictionary, matching, and suggestion for wrong words can play an important role in solving some of the eleven types of problems mentioned before.

Subsidiary dictionary provides the possibility of having a specific dictionary for, and named, for specific purposes. This dictionary can be the basis of the terms that are going to include in the final index. The use of this subsidiary dictionary will be discussed in suggestions at the end of the paper.

Matching consists of six options as follows:

[ ] ...? = ...
[ ] ? = ?
[ ] ?? = 1
[ ] ? = ?
[ ] ? ? = ?
[ ] ?? = ?? (???)

As can be seen, of the six options three belong to hamzih, one belongs to ?? that is used for making nouns plural, and the other two belong to special forms of letter ?, the first letter of Arabic and Farsi alphabet. Activating first box from top, caters for problem 10b mentioned above. In that case, computer ignores the difference between ???? ???? and ????? ???? but ?????? ???? still remains a problem. Activating the second box from top, solves problem 10c, that is computer considers ???? and ???? as the same. Activating the bottom box makes computer ignore the difference between a word made by using plural sign ?? either connected or disconnected. In general, if one activates one or some of these boxes, computer ignores the differences and considers both forms as correct. Since one form must be selected to be included in the final index, it is suggested not to activate any of the boxes above, so that computer selects one form as the correct and the other as wrong, and maintains consistency.

Suggestion for wrong words is another option useful for indexing purposes. If this option is activated, Zarnegar suggests some words while checking the spelling. If one sticks to suggested words, one benefits from the advantages of a consistent, though not necessarily standard, text. This consistency is a great help in providing the index.

After selecting the required options, clicking on the bottom “check the spelling”, opens the next window. The options available in this window are: list of suggested words, replace with, replace, replace all, ignore, ignore all, and add to the dictionary. These features are all in Farsi and very similar to that of Word, except for provisions made for some cases of Farsi writing’s inconsistency.

The incompatibility of Zarnegar with Windows environment, the fact that does not let the texts prepared by it to be transmitted via internet, however, is its major disadvantage, the fact that takes its being under the umbrella of desktop publishing, under question.

ARRANGING THE ENTRIES AND PREPARING INDEX IN ZARNEGAR

Some 44 pages in Zarnegar’s manual have been allocated to arranging the entries of which, 10 pages belong to preparing the index[3]. Although some sophisticated methods have been presented both for arranging the entries and preparing index, surprisingly there is no solution provided for the problems mentioned above for difficulties in written words of Farsi. In Zarnegar’s index, it is possible to select the entries of the index while the text is typed. Even the sub-entries can be selected although the task is sometimes so tiring. That is why this option is rarely used by amateur typists such as the author who would like just to write by computer and not by hand. And as mentioned above, there is no solution provided for the 11 kinds of problem mentioned before.

The following example from page 339 of Zarnegar's manual illustrates some problems of inconsistency in arranging Farsi entries (page numbers have been deleted by the author from the example and the left column has been added to illustrate the difference):

??????.	??????.
???????.	?? (?????).
???? ?????? ?????	**???? ????.**
???? ? ????????	????? ??? ??.
???? ??? ?????? ?? ??? ?????? ?? ????????.	**??????? ????.**
???? ???? ????.	**???????**.
??????	??????.
???? ??????? ?????.	**?? ????.**
?????	????????.
??????? ? ??? ???? ?? ??????.	?????? ? ??????.
????? ???????.	??????.
?????? ??????	???????? ??????.
?????.	?????? ????? = ????? ?????.
??? (??????)	??????.
?? (?????).	**????.**
???? ????.	???????.
????? ??? ??.	???? ?????? ?????.
??????? ????.	???? ? ????????.
???????.	???? ??? ?????? **?? ??? ?????? ?? ????????.**
??????.	???? ???? ????
?? ????.	??????.
????????.	???? ??????? ??**???.**
?????? ? ??????	?????.
??????	??????? ? ??? ???? ?? ??????.
???????? ??????.	????? ???????.
?????? ????? = ????? ?????	?????? ??????
??????.	?????.
????.	??? (??????).

FIGURE 2 A COMPARISON BETWEEN LETTER-BY-LETTER ARRANGEMENT AND WORD-BY-WORD ARRANGEMENT, INDICATING THE PLACE OF WORDS WITH HAMZIH IN EACH CASE

The example above that probably illustrates a real index indicates the three problems usually existing in Farsi indexing. The right column is the real index adopted from Zarnegar's manual. It has used the letter-by-letter arrangement for the reason mentioned above. Although the order seems logical for Farsi entries, but the problem of scattering related entries is obvious. Highlighted entries in right column belong to "water", but there are five entries not related to water in between them. The left column has been added by the author to illustrate the difference in arrangement. The bold highlighted entries in the right column come almost at the beginning, since they are words having hamzih. The Iranian librarians do not agree with that arrangement and prefer the arrangement of the left column in which the words with hamzih come last. This order is recommended by "Rules for filing Persian catalog cards" mentioned above. It should be mentioned that the left column

order is not only used by librarians; but also some Farsi dictionaries use the same order to arrange their entries. As can be seen the words with hamzih are still problematic. And the last point in this regard is the superiority of Zarnegar over Word in using hamzih. Hamzih, when used in conjunction with the unpronounced h in possessive case must be placed over the unpronounced h; Zarnegar does that, but Word does not. That is why the phrase: ?????? ??????, generated by Word is like that. Hamzih (?) must be over ?, and the phrase in figure 2. is, in a sense, wrong.

PROBLEMS ENCOUNTERED IN A REAL EXPERIENCE

The author, when finished translating an English book into Farsi, confronted some problems when trying to prepare the index to the book. Some of these problems related to the structure of Farsi as mentioned above, affecting the arrangement of the index entries, and for which some relative solutions will be presented at the end. But there were other problems that needed to be solved.

When providing an index to a translated book, the logical criterion to select the entries of the translated index is to base them on the entries of the original language index. Therefore, the entries of the original index were translated, and then rearranged in Farsi alphabet in another file. A major advantage of doing that is the possibility of provision of a Farsi-English as well as an English-Farsi vocabulary list at the end of the translated book. This is an invaluable instrument especially in the fields not very known or very new. When the task of allocating the page numbers to Farsi entries began by using the "Find" option under the "Edit" menu in Zarnegar, the different equivalents selected in each chapter for a single concept and not found totally by searching under the same concept, showed up as the first major problem. This problem related to both forms and meanings. In other words, the Farsi equivalents of some English terms to be included in Farsi index, found were not the same both in form and in meaning.

The second major problem encountered was the similarity between the concepts used both in a very specific sense as well a general sense. The latter increased irrationally the number of pages containing some entries. Since the book translated was on the theoretical as well as practical concepts of library services, the word ??????? (use) was an example of problems caused by the similarity of words that have both specific and general meanings in library and information science context. The word "use" in library services has a specific meaning. The author found the Farsi equivalent ??????? as the best equivalent. Since the word ??????? was also used in cases where the words use, usage, apply, and other synonyms existed in the English text, the number of pages related to this term increased irrationally compared to that of the original text. The page numbers of Farsi equivalents for "utility", "information", and "relatedness" were also much more than the page numbers of the related terms in the English index. As a consequence, the Farsi index, except for some very specific terms, was almost useless!

SOME SUGGESTIONS AND SOLUTIONS

Based on what stated so far, some practical guidelines might be presented at the end to overcome the problem relatively:

1. Before any step taken in indexing Farsi entries, it is absolutely necessary to prepare a manual for writing Farsi. As was seen above, Farsi is quite irregular in terms of writing, both in terms of spaces between the words and the forms of some letters. Unless such a manual is prepared Farsi will continue to perform the same problems it has had so far. Although the Iranian Academy of Persian Language and Literature published a manual called "Dastur-e Khatt-e Farsi (Persian Orthography): Proposal"[4], it does not seem this 47-page manual be an absolute solution. There

is a 3-page list in this manual that deals with words with hamzih on different bases (letters). The forms suggested are sometimes the ones not accepted in modern Farsi. There are two main reasons for these probably unacceptable suggestions: the first is that, people in the Iranian Academy of Persian Language and Literature are influenced by Arabic language that, as mentioned before, has found superiority over the other foreign languages after the Islamic Revolution. For the same reason, the Farsi equivalent of the word "Idealism" for instance, has been selected as ?????????, the form that does not seem acceptable. To the best of the author's knowledge, the Farsi equivalent for idealism has so far been as ????????. The second reason is that since foreign words in Farsi are pronounced in French style rather than English style, hamzih has to be used almost frequently in Farsi equivalents of foreign non-Arabic-script words. As an example, the name of Christian month June, is pronounced in French style as: ????, needing a hamzih, while the English pronunciation of the same word in Farsi needs no hamzih. It is surprising that although the English language has found superiority over French in Iran, the pronunciation of foreign words is still in French style. There is also a 9-page alphabetical list at the end of this manual that lists the words that are acceptable in both forms. This irregularity is quite an obstacle in proper indexing of Farsi terms, and even the most sophisticated available indexing tools cannot help in this situation.

2. Sophisticated word processors such as Word for Windows should be developed for Farsi. Current Words used for Farsi writing are in fact English Word made Farsi by a Farsi maker. It is strongly recommended that a Farsi version of Word for Windows be developed in Microsoft Company in consultation with specialists in Farsi language and script.

3. Before starting the translation of the text, the original index should be translated. This provides an English-Farsi dictionary for frequent reference in order to find selected Farsi equivalents, either in form or in meaning. The English-Farsi dictionary can later produce a Farsi-English one The subsidiary dictionary as an option in Zarnegar's spell checker that was referred to above can be a good tool for the purpose. In case any change occurs in the dictionary, care should be taken as to transfer the change to the whole text being translated. When the task of translation of the text finalized, a simple copy-and-paste process can produce two more copies of the first version of the English-Farsi dictionary. Of three copies, one can remain as the English-Farsi dictionary; the second, with some modifications as the Farsi-English dictionary; and the third, with English equivalents deleted and page numbers added as the final Farsi index.

4. The text size of the Farsi translation of the text should be determined before starting the translation. The size can be obtained either by asking the publisher one usually works with, or gaining information as to what the common size acceptable to most publishers is. In this way, there will be no future change in size to change the place thereby, change the page numbers related to concepts of the index. Some other information as to which pages, that is, recto or verso start the first page of a chapter should also be known.

5. All the chapters or parts of the text should be produced in a single file, provided the software used have enough capacity to include the text without any malfunction. And finally,

6. To prevent the general concepts to be mixed with synonymous ones, using techniques employed in Zarnegar seems to be useful. As mentioned before, Zarnegar employs techniques by which one can highlight the index terms and their subsets to include them in the index. Although the techniques seem useful, the human role should not be ignored. This is human intelligence that decides which terms are specific, worth including in the index and which ones are not. The dictionary mentioned in the suggestion number three could be useful in this regard. The techniques should be simpler than what they are now. Simpler techniques such as adding a specific sign to the term, and specific signs to relate the subsets to the main concept need to be developed.

NOTES AND REFERENCES

All the references are in Farsi.

1. National Library of the Islamic Republic of Iran. Faslname-ye Ketab (Book Quarterly), 14(3), fall 2003, p. 4.
2. Farhad Vaziri. *Rules for filing Persian catalog cards*. Tehran: National Library of the Islamic Republic of Iran, 1995, p.22.
3. SinaSoft. *Zarnegar*. Tehran: SinaSoft, 1992, p. 320-344.
4. Iranian Academy of Persian Language and Literature. *Dastur-e Khatt-e Farsi (Persian Orthography): proposal*. Tehran: Iranian Academy of Persian Language and Literature, 2000.