

The Newer, the Worse: the Status of Farsi Word Processing Softwares in Iran

Mortaza Kokabi

Associate-Professor, Dept. of Lib. & Inf. Science

Shaheed Chamran Univ., Ahwaz, Iran

Kokabi80@yahoo.com

In this paper, the problems of Farsi word processing software packages that are based on Word for Windows are enumerated. Being a non-roman script written from right to left and taken as Arabic, are two factors affecting the malfunctioning of Word for Windows. The problems such as: different keyboard layouts in different versions of Word; many problems encountered in dealing with digits; the problems concerning capital letters that sometimes do not attach when they must; and considering Farsi letters as misspelled by the spell checker are but some problems of the kind. Some suggestions are presented at the end.

Keywords: Farsi word processing software; Farsi; Arabic; non-roman languages

1. INTRODUCTION

The theme of ELPUB2005 is: From author to reader: challenges for the digital content chain. The theme considers the process of electronic publishing as a chain the first ring of which is a computer connected to a network, preferably Internet, and the final ring of which as the reader, that is, a person who is supposed to read the text sent. Discussing the digital content in that chain in the developed countries seems logical since there seems to be no problem concerning the actual readability of the texts sent in roman languages. But when the situation is considered in a developing country like Iran, it sometimes changes so drastically that conveying the message itself, not the content, is the main concern. The reason for this concern is firstly the official language of Iran, that is, Farsi and secondly, the word-processing software packages that are used for producing texts in Farsi. In other words, the major problem of electronic publishing in Iran lies mainly in the first ring of the chain.

Word for Windows is gradually substituting Zarnegar (gold writer) that is a word processor widely used by publishers of both scholarly journals and books in Iran. Word is much more powerful than Zarnegar, and has many great features such as compatibility with other Windows applications and especially with Word Wide Web (WWW) although it is poor in terms of Farsi. The main reason for this poorness is that Farsi versions of Word are in fact Arabic and some fonts in Word employ Arabic letters that are not used in Farsi. The process of substitution seems however, to be accelerating.

The main characteristic of Farsi however, that seems to be the most influential in word processors used for Farsi writing is its being a non-roman script that is written from right to left. The same characteristic exists in Arabic and that seems to be the reason why Arabic software packages are thought to be appropriate for typing Farsi. In the "Set Language" of "Language of Tools" option in the Standard Toolbar, there are sixteen Arabic languages beginning from Arabic Algeria to Arabic Yemen, but there is no Farsi (or Persian) language!

The two languages Arabic and Farsi however, are quite different. When the computer is started, one of the writings that is usually seen at the beginning is the phrase: Parsa 2001 Persian support for Arabic Windows (Millennium, 98SE, 98) and Microsoft office XP, 2000 and 97. That phrase indicates that the Arabic Windows is used, not the Farsi one, and the Persian support is employed for supporting Farsi. Being a non-roman script written from right to left and therefore mistaken as Arabic, are two factors affecting the malfunctioning of Word for Windows.

Newer versions of Word are sometimes so problematic that some users prefer earlier versions of Word and sometimes they prefer to return to Zarnegar! The following is but some of the problems that occur in (newer versions of) Word when it is used for producing a text in Farsi and especially a bilingual text in which the typist needs constantly to shift from Farsi to English and vice versa. It should be mentioned that there are two methods to shift from Farsi to English and vice versa in Word. The first is to click on a small square that is seen in the lower left hand side of the screen and showing letters "En", open it, and then change it to Farsi, after

which the letters “Fa” appear in the same square but the letters Ar appear in the Language Control box of the Formatting Toolbar indicating the employment of Saudi Arabia Arabic! The second method is to press simultaneously the right hand side keys “Ctrl + Shift” to change both the direction of typing and the fonts into Farsi, or the right hand side keys “Ctrl + Alt” to change only the fonts into Farsi, not the direction.

2. SOME PROBLEMS ENCOUNTERED IN FARSI WORD FOR WINDOWS

1. The keys in keyboard layouts of newer versions of Word act differently. In Word 78 the Farsi letter ﻑ is generated by pressing the D key while in XP 2003 version of Word, Shift + C generates the same letter. The typist, who has worked with the older versions, needs some time to work with the newer one to get used to the new keyboard layout!
2. Digits that are originally problematic in Farsi word processors produce special difficulties. In a Farsi text written from right to left, the numbers are written from left to right as in English. When numbers are typed in a Farsi text, since the direction must change, digits must be typed in special ways. For example, when typing a spanning date such as the date of publication of a multi-volume book, the date looks like: (1367-1365), that is, the first typed date that should be the first from left, comes last!
3. The same happens when typing dates divided by diagonal slashes. In the date: 1383/10/21 the number 21 (21) was typed first but came last! The cursor movement on such dates is strange; therefore correcting a wrong date is a cumbersome task! Highlighting a spanning date or dates with diagonal slashes is also difficult. The cursor movements are very unpredictable and when one thinks the highlighting is finished some part in the middle of the string of digits loses its highlights and the task must start from the beginning! Fig. 1 illustrates the case:

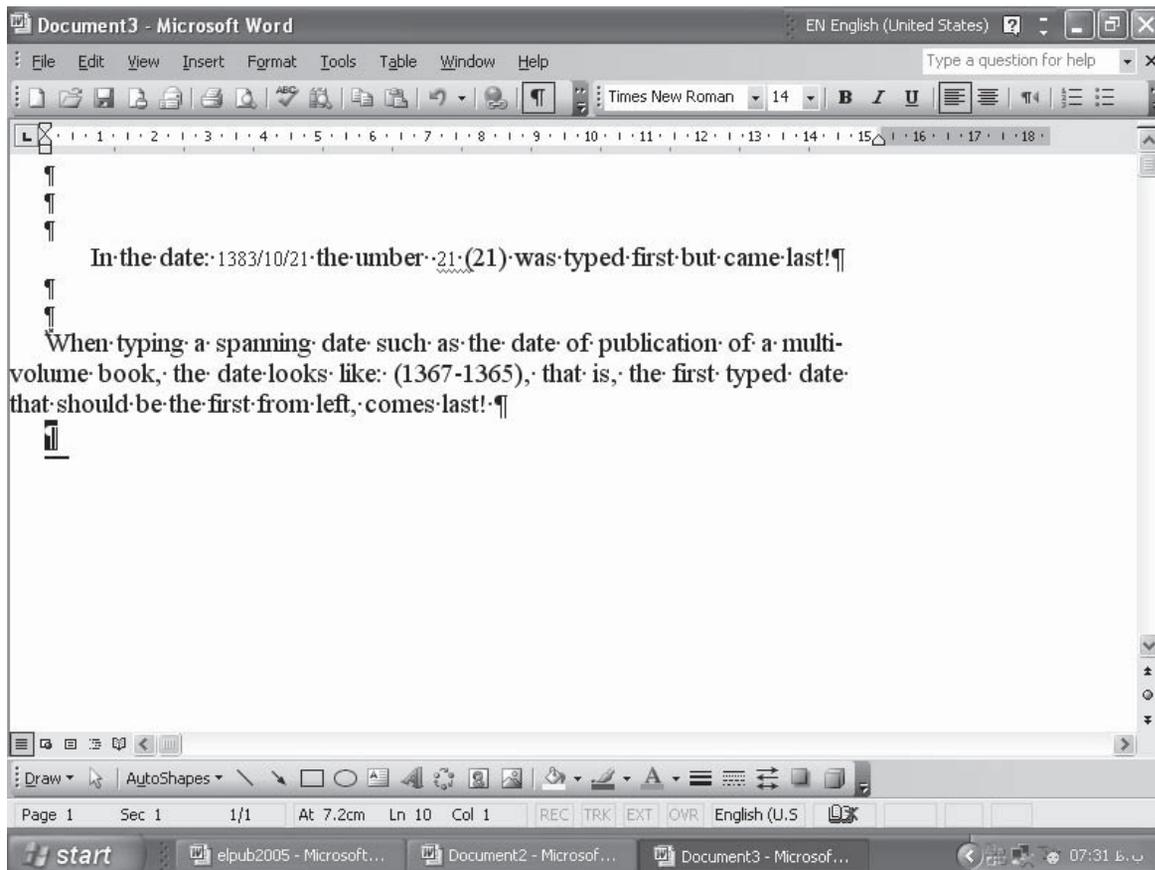


FIG. 1 SPANNING DATES AND DATE WITH DIAGONAL SLASHES SHOW SOME PROBLEMS

The problems concerning the numbers and digits do not end here.

In some XP versions of Word the numbers become Farsi in an English text and English in a Farsi text! Such texts produce unpleasant presentations when shown via Power Point. Footnote numbers are both in English and Farsi, either in the text or in the footnotes. A Farsi text with footnote numbers in English looks quite bizarre! Tables moved from statistical applications sometimes do not change into Farsi and a table with roman digits is seen in the middle of a Farsi text. Fig. 2 illustrates such a table:

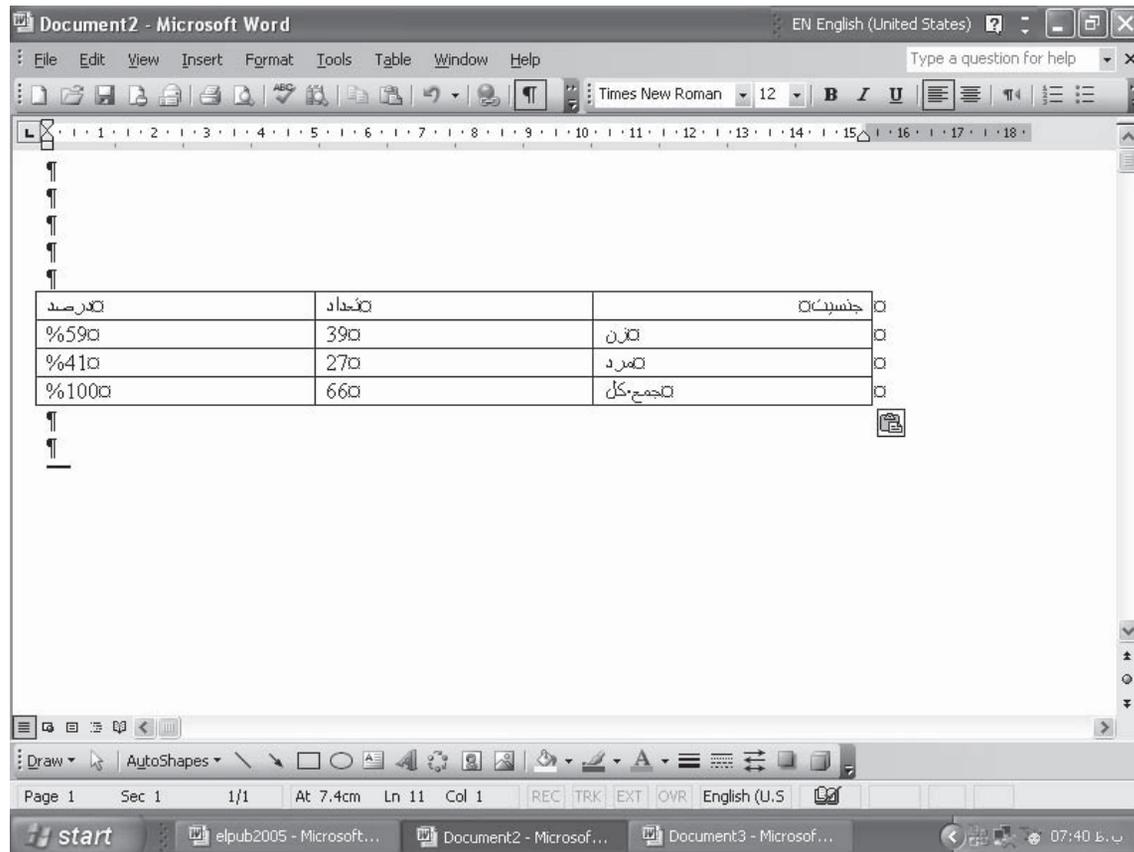


FIG. 2: TABLES IN FARSI WITH NUMBERS IN ENGLISH

- In a bilingual text of Farsi and English, correcting a misspelled Farsi word is awkward. The cursor movements are strange and the typist cannot be sure that the letter(s) typed are placed exactly where they should be. In such cases, one prefers often to delete the entire Farsi text and type it again! Correcting the mistakes of a Farsi text between two English texts is quite difficult. In example number 2 above, the author spent some minutes to place the cursor between 21 and “the number” and had to use the mouse to place the cursor between the two! When deleting a letter between an English text and Farsi one, one cannot be sure whether to delete the right letter.
- The capital letter **ي** in Farsi is sometimes connected to its preceding letter and sometimes not. In the word **فارسي** (Farsi) the underlined letter is the connected **ي** and in the word **بازي** (play) the underlined letter is non-connected capital letter **ي**. The same letter takes the form of a small **ي**, connected to the two letters before and after it. In the word **تغيير** (meaning, change) the underlined letters are two small letters **ي**. Since Farsi is a cursive language the connection and non-connection of letters is quite important. The small letter **ي** in some words typed in an older version of Word takes the form of a capital letter **ي** in some newer versions and becomes not only capital but also non-connected and produces a big spelling mistake!

The word **تغيير** for example, becomes **ت غ ي ي ر ي**. Fig. 3 presents an example:

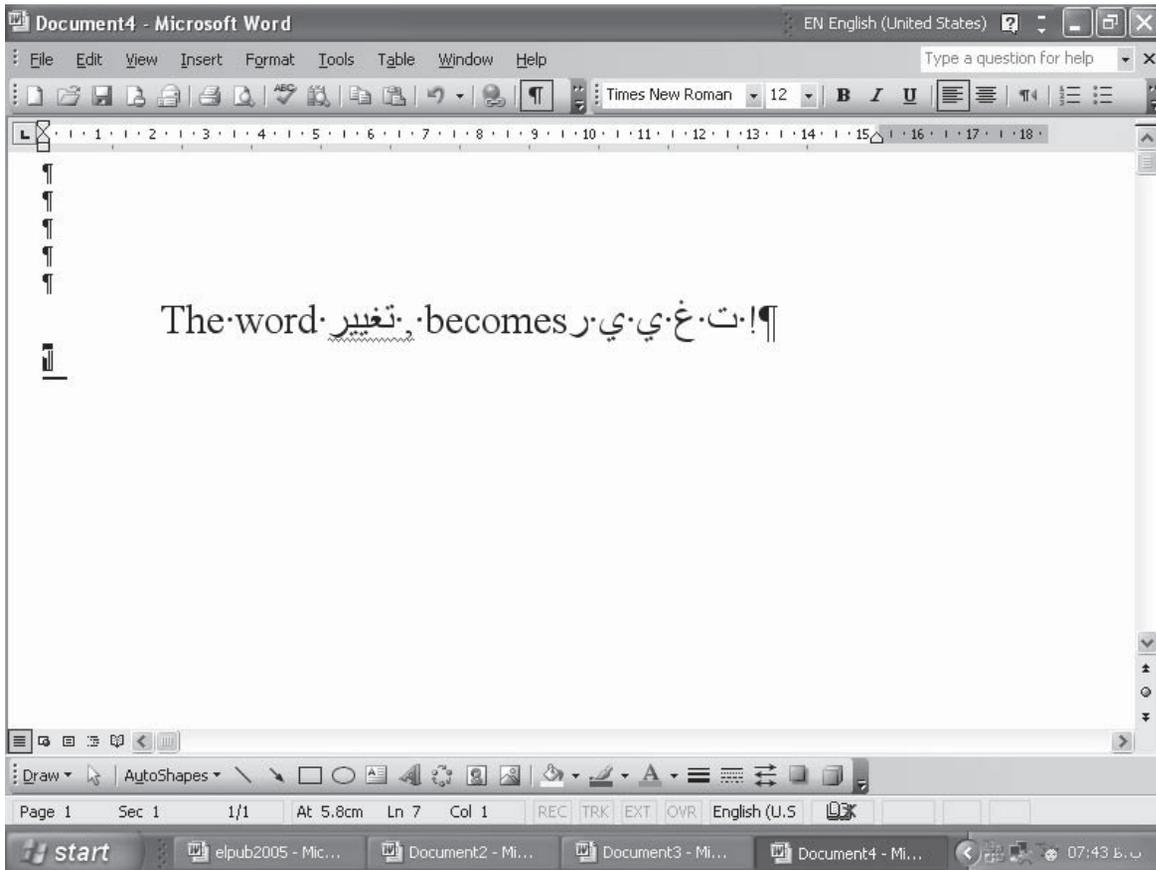


FIG. 3: A CASE INDICATING THE PROBLEM OF CURSIVE LETTERS IN FARSI SOFTWARES

6. When a file is sent as an attachment to an e-mail, the receiving computer must have the same fonts as the computer, which has produced the file. Otherwise the text file will probably have some problems.
7. The same happens when the receiving computer does not have the fonts that the transmitting computer has used to type the text sent. In such cases the text is really unreadable and should be sent either in English or via fax!
8. Since the Farsi versions of Word are in fact Arabic, the spell checker of these software packages considers the Farsi words as incorrect and underlines them in red. The problem here is not that the text is almost full of irritating words underlined in red, but that in lengthy texts, when the size of spell checking increases, Word automatically stops checking the spelling. In such cases, even the spelling of English words is not checked, the unpleasant fact the non-English speaking users like Iranians do not like very much. In the sentence: من در دانشگاه تدریس می کنم. (I teach in the university) Word considers the words: دانشگاه (meaning university) and می کنم as misspelled since they are Farsi words but considers the word تدریس as correct since it is Arabic. Fig. 4 gives an example of the case:

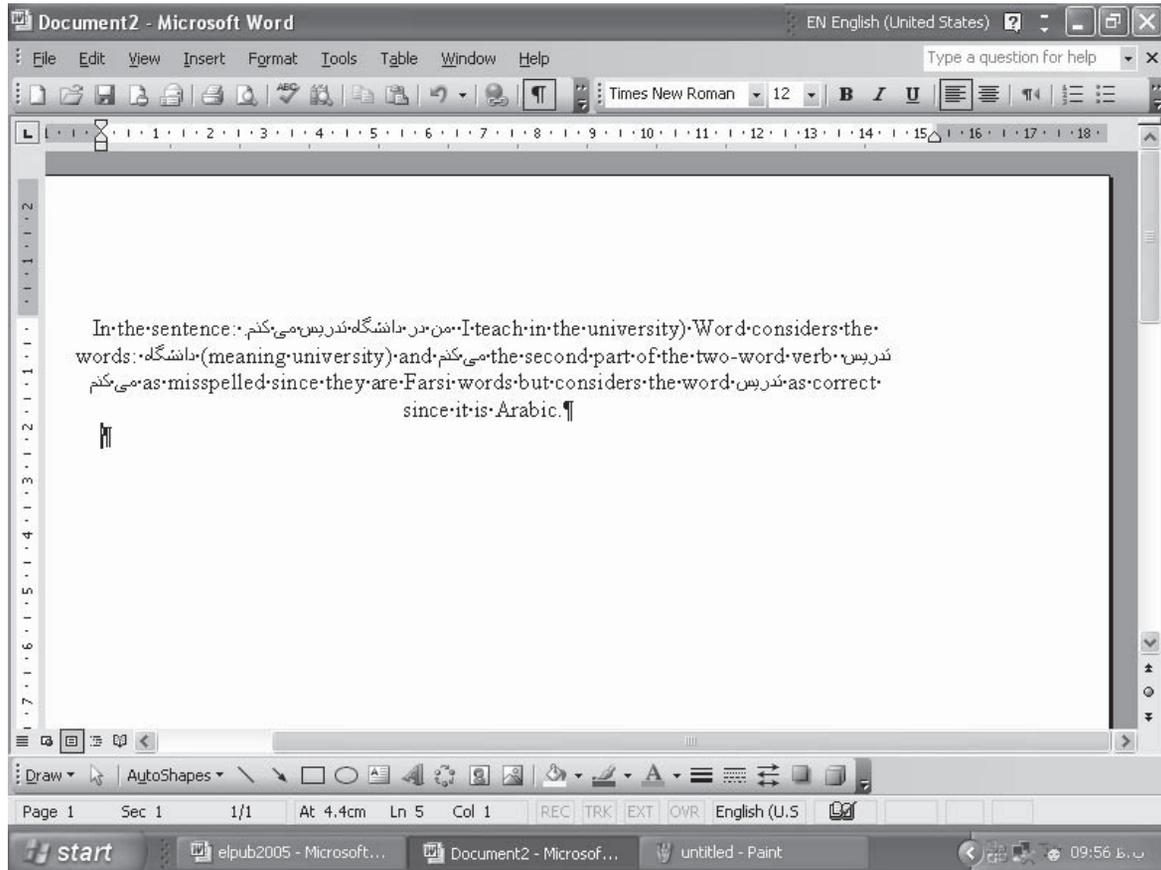


FIG. 4: AN EXAMPLE OF WRONG DOING OF ARABIC SPELL CHECKER WHEN DEALING WITH FARSI WORDS

9. Farsi maker software packages used in different versions of Word to make them Farsi are sometimes incompatible. Parsa 99 is a Farsi maker that is used to make earlier versions of Office Farsi. The newer versions employ other Farsi makers. The change in Farsi makers affects some other software. The Aryanpuor English-Farsi dictionary for example, cannot be read without Parsa 99. If another Farsi maker is employed along with the Aryanpuor dictionary, the fonts employed in Parsa 99 must be copied onto the computer to make that dictionary usable. A computer specialist must do that!

3. CONCLUSION AND SOME SUGGESTIONS

The above mentioned problems that are but some of the problems encountered in typing a bilingual text of Farsi and a roman alphabet text in Farsi Word for Windows can be categorized in two types. The first one is the problems of a non-roman cursive language written from right to left that seem to be the same for Arabic. The second one is considering Farsi as Arabic, or at least thinking that the latter can more or less handle the first. The radical solution for the first problem is developing software by computer companies such as Microsoft for special needs of languages such as Farsi and Arabic. And the only suggestion for the second problem is to develop a Farsi version of Word for Windows and not relying on Arabic ones. Iranian linguists' cooperation with the computer experts can do best.

4. REFERENCES

Asiasoft™ Farsi/Dari Support for Microsoft Windows 2000. Available at:
<http://www.liwal.com/windows/farsi%5C2k.htm>

Kokabi, Mortaza. The problems of desktop indexing of a book translated into a non-roman script: description of a real experience. In: Jan Engelen, Sely Maria de Souza Costa, & Ana Cristina Santos Moreira (eds.) Elpub2004: Building digital bridges: linking cultures, commerce and science: proceedings of the 8th ICC International Conference of Electronic Publishing.- Brasilia, Brazil: Department of Information Science and Documentation (CID), 2004.- 283-292

Linguist's Software: Fonts for the Whole World: LaserArabic & Farsi. Available at:
<http://www.linguistsoftware.com/la.htm#Testimonials>

Step-by-step learning of Word (in Farsi). Available at: <http://ourmind.netfirms.com/learning/word.htm>

What is a Farsimaker (in Farsi). Available at: http://www.bornaray.com/per_farsisaz.asp