

Electronic Applications of the Semi-Bilingual Dictionary

Ilan J. Kernerman

K Dictionaries Ltd

Nahum 10 Tel Aviv 63503 Israel

ijk@kdictionaries.com

Abstract

As dictionaries get integrated in daily life, their role and functioning transform. Increasingly, nowadays, dictionary use takes place in computerized environments. The change from printed to electronic media affects what a dictionary is, or may be, and how it will be used. This paper discusses some aspects of the ensuing differences with reference to electronic applications of (Kernerman) semi-bilingual English learner dictionaries. From simple PC products through scholarly research and multimedia CD-ROMs, current directions include versions online and in e-learning, various hand-held devices and cellular phones, multilingual lexicography, etc, opening new prospects for future developments.

1 Introduction

The first major wave of dictionaries in electronic applications began in the 1980s, especially in the form of hand-held devices, diskettes and CD-ROM, and some multilingual versions. What these initial applications generally had in common was a fairly straightforward conversion of content from printed format to electronic media without significant transformation of the content itself. The main differences concerned exploiting the great calculation power and speed of computers, and their relative compactness and user-friendliness. Basically, instead of sifting through pages and crowded text to find an entry, one could type it in and receive immediately the same information as in the book, use sophisticated search and retrieval facilities, and have access to supplementary features (such as grammatical information) via given keys. Less attention went into revitalizing the lexicographic content and presentation in accordance with the new computerized environments.

In a sense, lexicography and dictionary compilation were much earlier in “understanding” computers – adapting to suit the computer world and applying its new potential for their own requirements – such as through corpora or desktop publishing, than did dictionary products per se. As an analogy, a good computerized dictionary would resemble filming a fine theater performance of a classic play and showing it on screen, as occurred in the early days with cinema shows which flatly copied a scene from the stage. Later on, gradually, movies actually began to incorporate their unique cinematographic qualities into a new up-to-date production of the hundreds-year-old play, thus developing a novel artistic and cultural stream in its own right while moving on from cinema to television to video etc.

2 The New Wave

The first PC versions of the Password semi-bilingual Czech and Estonian dictionaries did little more than reflect the printed edition transferred onto computer¹. (A semi-bilingual

dictionary consists of the headword, definition(s) and example(s) of use, and translation for each meaning.) Their importance lies in this seemingly simple technical accomplishment, to convert the dictionary from paper to electronic media, so commonly taken for granted now.

The Estonian semi-bilingual version soon had another role in a project held under the auspices of the European Union's Fourth Framework Programme COPERNICUS 1994, which financed language engineering research, reusable language resources and some pilot applications [Langemets 1997]. The project called GLOSSER was designed to support the process of reading (and learning to read) in a foreign language. The computer was used as a reference tool, designed to help to find the lemma and the right sense in the dictionary. For its morphological analysis GLOSSER had access to the Xerox Part-Of-Speech Disambiguator for the English language, and the guidelines used for the encoding scheme of the dictionary were tagged in SGML. GLOSSER examined sentences word by word, and had only a word-based access to the dictionary, although multiword expressions appeared in the dictionary as well, usually at the end of an entry according to a key word in the multiword expression. The semi-bilingual dictionary was particularly useful for this program because it combines monolingual and bilingual dictionary cores in a single body.

3 The Next Wave

The next step can be attributed to Pass-Q-Word, launched in 1997². This was a sophisticated multimedia product, derived from the Password semi-bilingual English learner's dictionary, while utilizing the interactive nature of the computer. To begin with (in certain ways, as with GLOSSER), the dictionary text was broken down into its constituents, then recomposed with the aim of suiting specifically the requirements and capacities of the computer and the user. Each and every item in the English-English text was morphologically analyzed, so that any word used in the definitions or examples could be hyperlinked to an appropriate headword. Users could access the dictionary through either their native language or English, aided by extensive search options, and any information given was retained for repeated use. Headwords had vocal pronunciation, as well as self-recording and playback abilities to compare the user's pronunciation with the default pronunciation. The computer enabled an even deeper intervention into the very heart of the semi-bilingual lexicographic concept, by letting the users activate the translation at will, so that the dictionary could function as either monolingual or bilingual. The presentation was very colorful, clear and attractive, both for the background screen and for highlighting different components of the entry, such as distinguishing the definition by a certain color and the examples of use by another. The visual additions included illustrations supplementing entries, and crossword puzzles for word exercises and games, both directly linked to the dictionary core. The dictionary included the usual editing and printing functions, and was compatible with other programs [Stock 1997].

Pass-Q-Word was a ground-breaker with state-of-the-art features, but with drawbacks too, namely of being complex, expensive and slow to manufacture, and possibly hitting the market ahead of its time. The high cost of developing and manufacturing each new language version stranded it at a premature stage with its initial French and Italian versions only³. However, the experience gained in creating it proved useful later. The electronic version of the lower-level Passport dictionary appeared in 1999⁴, containing many of the important features mentioned above, though simpler, friendlier-to-use, and somewhat more down to

earth [Siani 1999; Cohen 1999; Gambacciani 1999]. While the Passport electronic dictionary can be a stand-alone product in its own right, it can easily function within a wider context and assume a secondary role as may be appropriate. For example, it is possible to open a file inside the dictionary, and to hyperlink a word from the file into the dictionary. This can be compared to reading a document in one's computer or on the Internet, and pointing at a word in order to receive its translation or other lexical information⁵.

4 The Now Wave

Pass-Q-Word, Passport, and other such new electronic versions of printed dictionaries that appeared in recent years are the start of a significant breakthrough away from the traditional type of dictionary we have been accustomed to, especially in pedagogical lexicography. For instance, when utilizing the miraculous hyperlink capacity, so that each and every item in the text can be transferred to its appropriate headword, we are no longer concerned with a limited defining vocabulary in the same sense this has been dealt with extensively in 20th century lexicography. This and other tools permit instant and endless navigation throughout the dictionary, and unlimited search options, with appreciable influence on morphology, phraseology, etc. For example, the lexicographer (and user) is altogether differently preoccupied by whether "touch wood" is a derivative of "touch" or of "wood". Another example: the vocal pronunciation facilities carry the potential to sooner or later do without phonetic transcription in learners' dictionaries.

From another direction, the vast changes in many daily aspects of our lives – regarding information, technology, communication, globalization, and so on – on the one hand continuously increase the need for and use of dictionaries, while on the other hand lead dictionaries to their integration into numerous other products and services – such as word-processors, language courses, information services online or by mobile phone – and to gradual loss of the dictionary's independent status [Levine 2001]. The following are examples of present-day applications of the semi-bilingual dictionary, which hint at some such possible new horizons.

- A hand-held device, integrating the semi-bilingual Korean version with comprehensive English-Korean and Korean-English dictionaries⁶.
- An e-learning course available on CD and online, incorporating a pre-intermediate semi-bilingual dictionary. Every word in the 12-level course can be referred to the dictionary, and of course every word within the dictionary is hyperlinked to its headword⁷.
- Dictionary programs based on various semi-bilingual versions developed for mobile hand-held devices, PDAs and personal computers, corporate servers, etc. Through OEM arrangements with hardware manufacturers, the dictionary can be built-in with a computer operating system like Windows program, anti-virus software, etc⁸.
- An innovative multilingual dictionary incorporating twenty language versions from the semi-bilingual series⁹. Apart from its unprecedented number of languages, it is unique in sharing more or less the same English-English lexicographic core. The implication is that every sense of a multi-meaning headword has its own information (definition, example, etc) with the specific translation to each language. Thus, through the intermediary of the English entry, all the languages are interconnected, with a large range of language-pairs and

translations [Herpiö 2001]. It is made for CD-ROM, for online use on the Internet and intranets, and for cellular telephone memory cards.

5 Conclusion

The future may bring with it dictionaries incorporated in any word-based activity, and the rapid changes in technology and our daily lives may well make this paper seem archaic before long. The electronic dictionaries described here are quite likely a temporary phase that will give way to more superior solutions – whatever these may be. New dictionaries will continue to suit themselves to their new media, in order to best serve their compilers' purposes and users' needs.

Acknowledgement

I am grateful to Raphael Gefen for his valuable comments and help with this paper, and would like to thank also the Euralex referees for their inspiring assessment reports on my draft proposal.

Endnotes

- 1 The semi-bilingual Czech version *Password PC Anglický výkladový slovník Anglicko-český a česko-anglický slovník Elektronická kniha pro PC* was published on a 5.4" three-diskette-set by Mladá Fronta and Swing, Praha, 1993. The Estonian *Password Inglise-Eesti Seletav Sõnaraamat* was published on a 3.5" four-disk-set by TEA and Festart, Tallinn, 1996.
- 2 *Pass-Q-Word* was created by Password Publishers and The Q Group in 1996.
- 3 The French and Italian *Pass-Q-Word* were prepared but not published.
- 4 The first version of *Passport Electronic Dictionary* for speakers of Italian by Password Publishers and McGraw-Hill Libri Italia was based on *Passport English Learner's Dictionary* and published as part of a package including the textbook *Grammar in Action* (1999).
- 5 The *Babylon* dictionaries which are available on the Internet are installed in the user's computer, and show the translation when pointing at a word with the mouse cursor.
- 6 The electronic hand-held device by Sharp (*PW-4100S*) includes YBM Si-sa's *Si-sa Elite English-English-Korean Dictionary*, *Si-sa Elite English-Korean Dictionary*, *Si-sa Elite Korean-English Dictionary*, and TOEIC exams (2000).
- 7 The *OK English Learner's Dictionary* is part of Edusoft's *English+ Millennium* on CD and the *English Discoveries* online (2001).
- 8 The *Socrat* series by Arsenal has nine language versions of *Password* for OEM, PDA, etc (2002).
- 9 The first version of *GlobalDix* by Kielikone on CD and online brings together 20 language versions from the *Kernerman Semi-Bilingual Dictionaries* series (2002).

References

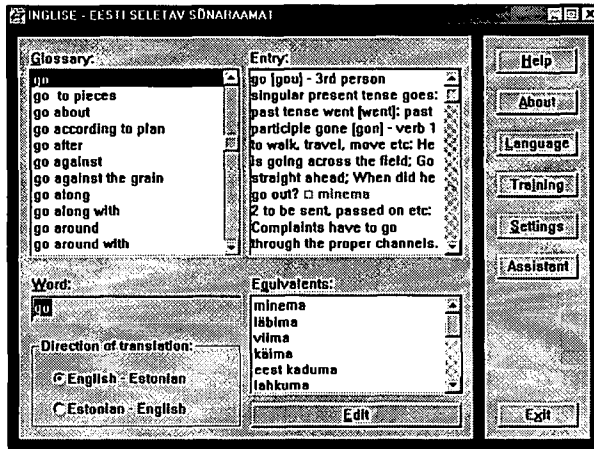
- [Cohen 1999] Cohen, H., 1999. Designing Passport for Computer, in: *Kernerman Dictionary News*, 7, Password Publishers, Israel.
- [Gambacciani 1999] Gambacciani, M., 1999. Passport Dizionario Multimediale Inglese-Italiano, in: *Kernerman Dictionary News*, 7, Password Publishers, Israel.
- [Herpiö 2001] Herpiö, M., 2001. GlobalDix: a Unique Multilingual Dictionary for the Worldwide Market, in: *Kernerman Dictionary News*, 9, K Dictionaries, Israel.
- [Langemets 1997] Langemets, M., 1997. GLOSSER as a Practical Application of the Semi-Bilingual Dictionary Concept, in: *Kernerman Dictionary News*, 6, Password Publishers, Israel.
- [Levine 2001] Levine, C.M., 2001. The Coming Boom in English Lexicography: Some Thoughts about the World Wide Web (Part 1), in: *Kernerman Dictionary News*, 9, K Dictionaries, Israel.

[Siani 1999] Siani, A., 1999. Computerizing Passport: Print to Screen, in: *Kernerman Dictionary News*, 7, Password Publishers, Israel.

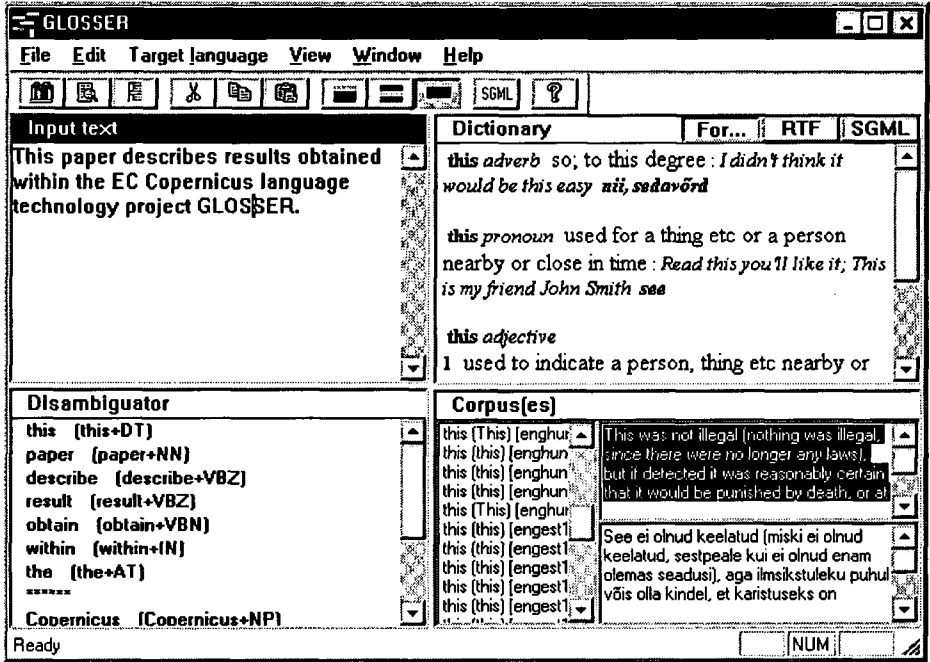
[Stock 1997] Stock, R., 1997. The Innovative Pass-Q-Word, in: *Kernerman Dictionary News*, 6, Password Publishers, Israel.

Appendices

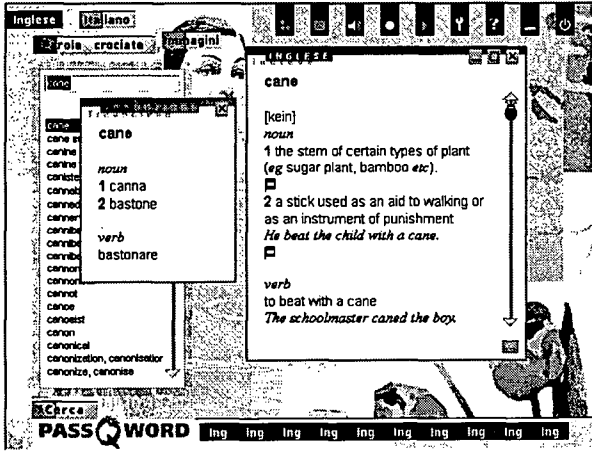
1 Screen from *Password Inglise-Eesti Seletav Sõnaraamat*, Festart and TEA, 1996:



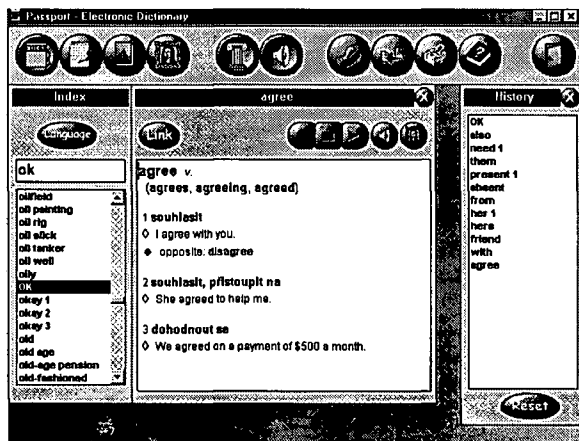
2 GLOSSER, main window:



3 Screen from Pass-Q-Word English/Italian, Password Publishers and The Q Group, 1997:



- 4 Screen from *Passport Electronic Dictionary English/Czech*, Password Publishers and Fragment, 1999:



- 5 Screen from *GlobalDix* (pre-release beta version), Kielikone and K Dictionaries, 2002:

