GLOBALDIX
a Unique Multilingual Dictionary
for the WorldWide Market

Mika Herpiö

Mika Herpiö has a master's degree from the Helsinki University of Technology. He is the director of business development and a partner in Kielikone Oy. mika@kielikone.fi www.kielikone.fi

Kielikone Oy is the leading language technology company in Finland, with more than 10 years of experience in electronic dictionaries. We started with diskette dictionaries in the late 1980s, and our latest products run on Linux and WAP telephones. Besides dictionaries we offer terminology management software, machine translation, morphological analysis and parsing software. Our focus is on products for corporate customers. We publish electronically works from more than ten publishing houses in Finland, Sweden and the UK*. They are all available in the MOT Dictionary Bookshelf, where users can select quality dictionaries and use them through a single, common user interface. MOT is available through varied platforms, such as Windows, Mac and Unix (Linux, Sun Solaris, etc), over the Internet and intranets, on mobile phones via WAP/GPRS service, and soon through Nokia Communicator memory cards as well.

Currently we are developing a new multilingual electronic dictionary based on the K Dictionaries semi-bilingual Password series. The first version, including well over twenty languages, will be launched this fall. It will be available both as part of the MOT Dictionary Bookshelf and as a stand-alone product named *GlobalDix*, that will be a shrink-wrapped software package for individual users, as well as for all the above-mentioned platforms.

All the multilingual dictionary solutions with which we are familiar list an item and its translations on a single line – just like database spreadsheet-tables filled with words from various languages. The problem with this kind of approach is, of course, the many grades of synonymy that words have in each language and the varying degrees of correspondence that a word and its translation(s) may have. For exmple, the structure of an English-German dictionary is not the same as that of an English-French one. Since a word may have several meanings, the structure of an article in the dictionary depends on whether the same translation can be used for all the meanings or whether a different translation is needed for each meaning.

There is an analogy between *GlobalDix* and multilingual terminology databases. Multilingual terminology databases are handy, useful and precise tools, since each entry is based on a concept that is common to all the languages covered. Password dictionaries are also based on descriptions of concepts and have a common set of definitions that are translated into various languages. As we learn from terminology, a well-defined concept can work as an anchor that binds together words from different languages. The English explanation in Password dictionaries may work as a similar anchor that connects words in many languages. Since most of the different Password versions are based on a fairly similar monolingual lexical core, in many cases the user can find how this meaning is translated into various languages. Combining these dictionaries together into a single software product enables cross-searches from any one language into any other language. For example, searching for the Finnish *erinomainen* would bring up its English definitions together with their translations in the other Password language versions based on the same core, and the translaions of the word in the Password versions not sharing the same core.

The first version of *GlobalDix* will contain the following language versions from the Kernerman Semi-Bilingual Dictionaries series: Chinese (traditional and simplified), Czech, Dutch, English, Estonian, Finnish, French, German, Hungarian, Icelandic, Indonesian, Italian, Japanese, Korean, Latvian, Lithuanian, Norwegian, Polish, Portuguese (Portugal and Brazil), Russian, Slovak, Spanish, Swedish and Turkish. It will cover a total of approximately 1.5 million headwords (all languages combined). The user is expected to have a basic knowledge of English, so that he/she can browse through the possible different meanings and choose the right one. The solution is aimed at helping the average person in her/his communicative needs, mainly reading foreign-language newspapers, websites, magazines and books. Another special application of *GlobalDix* is cross-language information retrieval (CLIR), where it will be possible to search for information in any of the languages covered through the user's mother tongue.

* WSOY, Gummerus and others, including over 20 titles for Finnish users. Gustava: Engelsk ordbok – Swedish-English-Swedish dictionary. HarperCollins: Collins English Dictionary, Collins English Thesaurus, Collins French Dictionary, Collins German Dictionary, Collins Italian Dictionary and Collins Spanish Dictionary.