Petek Kurtböke, Monash University

The Impact of Corpus Planning on Bilingual Dictionaries

Abstract

This paper examines the impact of language planning on lexico- graphical practices. Dictionaries have always been amongst the main agents of language planning in establishing the standard forms, meanings, pronunciations and spellings. While mono- lingual dictionaries, generally compiled by government agencies, have dictated the 'norm', the impact of language planning on bilingual dictionaries, generally compiled by individuals or private bodies, has been, in Tietze's terms, *schizoglossic*. In this article, specific reference is made to the Language Planning activities in Turkey, and the impact of these activities on Turkish-English/English-Turkish bilingual dictionaries.

1. Introduction

Language planning has been defined as an organised activity to regulate language issues in a society (de Vries 1991). Such language issues may include "the creation of new alphabets, the codification of morphology, standardization, the development of 'plain language', spelling reform, language maintenance, and the elimination of gender discrimination in language" (Clyne 1992). A common model used to describe the activities of language planners has been proposed by Haugen (1966, 1983, 1987):

Table 1

	Form (policy planning)	Function (language cultivation)
Society (status planning)	1] Selection a. Problem identification b. Allocation of norms	3] Implementation a. Correction procedures b. Evaluation
Language (corpus planning)	2] Codification a. Graphization b. Grammatication c. Lexication	4] Elaboration a. Terminological modernization b. Stylistic development

Status planning, which is usually the agenda of politicians and bureaucrats, involves developing a national identity and language spread at national and international levels. *Corpus planning*, on the other hand, is usually the agenda of linguists, lexicographers and experts alike who intend "...(i) to give the language a terminology for scientific and technical purposes; (ii) to resolve normative/structural questions of correctness, efficiency, and stylistic levels; and/or (iii) to support an ideological cause by eliminating sexist, racist, or militaristic elements in the language" (Clyne 1992).

Selection refers to the decision of replacing one language with another or one variety with another when there are two conflicting norms. Examples of this can be seen in the replacement of Arabic with Hebrew in Israel, and the urban with rural dialects in Norway. The selected code may enjoy a higher status in the society. *Codification* involves giving a form to the norm chosen and "the typical product of all codification has been a prescriptive orthography, grammar, and dictionary" (Haugen 1983). Selection and codification will be successful only if they are followed by *implementation* and *elaboration. Implementation* usually takes the form of books, pamphlets, newspapers and textbooks. *Elaboration*, on the other hand, is "the continued implementation of a norm to meet the functions of a modern world" (Haugen 1983). The language form which has been selected and codified may be spread by individuals (e.g. a writer), a government agency or an institution.

2. Language planning in Turkey

Status planning and corpus planning went hand in hand in Turkey, especially in the early years of language planning. Nationalism was "the central pillar of Kemalist ideology" and created a strong "demand for the purification of the Ottoman language by replacing its foreign elements with genuine Turkish words..." (Heyd 1954:19). The selection of Turkish as the norm, and the recodification of the language using the Latin alphabet announced the start of an official language policy in 1928. The implementation and elaboration would be carried out by a semi-official agency, the Turkish Linguistic Society (TDK), which was founded in 1932. TDK's activities included the elimination of Arabic and Persian elements in the language and the creation of a new Turkish lexicon based on a number of processes: a) collecting words from old Turkish texts (*tarama*) b) collecting words from folk vernacular (*derleme*) c) deriving new words using Turkish roots and derivational morphemes (*türetme*) c) compounding (*birleştirme*) (Doğançay-Aktuna 1995). After 1935 "many

long lists and entire books full of proposed new vocables for all the walks of life were published" (Tietze 1962) as well as a monolingual dictionary (*Türkçe Sözlük*) in 1944, which has been an influential agent in establishing the new 'norm' (Heyd 1954:40).

While Turkish language planning is commonly labelled as a successful movement (Fishman 1971), it did not enjoy a smooth transition. Tietze (1962) referred to its linguistic consequences as *schizoglossia*. That is, the public had to read the new vocabulary proposed in the official word lists, when for example reading a newspaper article, and refer to the notes in brackets and/or annotations to translate constantly from the new Turkish into Ottoman. Tietze found the situation problematic for the lexicographer:

> ...the Turkish language, both as a system, as a set of rules for communication, and as a social institution, with the relative distribution of the areas of application of these rules, is at present in a stage of transition. However, the evolution is not so straight and steady as one might conclude. A continual tampering with the language, although it was a historical necessity, has created contrasting attitudes of reaction toward stability and change... This state of the language, with its conflicting concepts, creates serious problems for the lexicographer. Which of these concepts is he going to follow? Can he combine more than one concept or all concepts? (Tietze 1962).

Tietze, himself, was involved in the revision of the *Redhouse* which was the most-widely used Turkish-English dictionary during the time of the language reform in Turkey. Starting from 1600's to early 1900's, most bilingual dictionaries of Turkish had been compiled by missionaries and/or interpreters at the Ottoman Court (Kurtböke 1994), and the *Redhouse* was no exception. It was compiled by Sir James Redhouse who was both affiliated with the Near East Mission of the United Church and an interpreter to the Grand Vizier. He had already published smaller Turkish-English/English-Turkish dictionaries in 1856, 1857 and 1861 and 1884. The complete dictionary appeared in 1890 and has been widely used since. The 1890 edition naturally included Arabic and Persian vocabulary which was then in current use.

Before the language reform it was the tendency of dictionary-makers to include, apart from the Turkish words and commonly used borrowings, all nouns and adjectives found in an Arabic and/or Persian dictionary. In fact, Sir James Redhouse had included "the entire noun and adjective material of the Arabic *Kamus* and of Persian *Burhan-i Kati* without having to bother to find out which of these words, or which of their various meanings, had ever actually been used in Turkish; he was justified in doing this because all this material was, potentially, at least, a

part of the Ottoman language" (Tietze 1962). Language reform broke this tradition and brought challenging questions to the lexicographer: "How much of the loan vocabulary is to be taken into the dictionary?" and "What to do with newly coined words and expressions?" (Tietze 1962).

The official decision to exclude the majority of loanwords from the Turkish lexicon left the lexicographers on the Redhouse committee with concern. They felt "nothing of value should be dropped" from the text of the original *Redhouse* dictionary. There was even regret that some entries had been omitted to make room for new material (Avery 1974). Since the subsequent editions of the Redhouse dictionary "stood firmly on the shoulders" of the 1890 edition, it required considerable effort to discard materials which were so well attested by James Redhouse's earlier work (Tietze 1962). This was a difficult decision to make and led to the new tendency of listing both the 'old' and the 'new' together regardless of their current status in the language, including both the Arabic and Latin script, thus combining the 'historical' and the 'period' dictionaries in one (Zgusta 1991, Kurtböke 1996).

The Redhouse dictionary, as stated by the editor, reflected the "American point of view of its editors" (Avery 1974). The other dictionary which was on the market during the high time of corpus planning reflected the British point of view and could not escape schizoglossia either. In his preface to the first edition of *The Oxford Turkish-English Dictionary* in 1947, Hony stated that most existing dictionaries of Turkish-English had derived from the *Redhouse* whose content was out of date due to the massive changes occuring in the Turkish language. The editor recognised that the "purification was badly needed and had brought the written language into line with the spoken." However, he blamed the reformers for being "carried away by their zeal" and "for introducing a great numbers of obsolete Turkish words, and what is worse..." for inventing "hundreds of new words" which he hardly regarded "as an improvement of the language." Also in the introduction to the 1957 edition he stated that invented words and meanings had not really caught on and those few which had, were "positively harmful" and had impoverished the language. Amongst those "senseless innovations" as he called them, are today's firmly established *sağlık, savaş, çeviri* and so on. As in the case of *Redhouse English-Turkish Dictionary*, the tendency to list the 'old' and the 'new' continued also in *The Oxford English-Turkish Dictionary* which was first published in 1952 and revised in 1978 and still included a considerable number of Ottoman equivalents. Interestingly, an important recommendation made by Tietze in 1960 seems to have gone unnoticed in Turkish-English dictionaries,

that is, "the lexicographer who includes 'old' material (from 1300 to 1935) will have to be very careful to provide each word with an appropriate label: *learned, obsolete, archaic, only known to the older generation*, etc."

As can be seen from the brief history above, massive changes in the lexicon resulted in long strings of words appearing in bilingual dictionaries. Attitudes towards the official language planning agency varied and since the individual lexicographer thought loan vocabulary was of value, he had to make room for both the 'old' and the 'new' in the dictionary. Tietze (1962), in fact, pointed to the lack of objective criteria in letting the new vocabulary into dictionaries: "if the criterion is the degree of reception and assimilation by the language, how can this be measured?" Frequency counts of the 'old' and 'new' words have not been widely available in Turkey so far although İmer (1994) reports that while in 1930 the percentage of Turkish words in the language of the press was 35%, today it has reached 70%, and the remaining 30% includes not only Persian and Arabic words but also borrowings from European languages.

3. Conclusion

When placed in a Language Planning framework, such as the one proposed by Haugen (1987), and when considered in relation to such activities in many other countries, the changes which have taken place in the Turkish lexicon may seem less dramatic than the way they were described by Tietze over 30 years ago. However, the individual lexicographer's attitude to the changes introduced by the corpus planners has long determined the number and the choice of the 'old' words to be taken into bilingual dictionaries. Since objective measurements such as frequency counts have not been available, what the lexicographer believed was of value entered the dictionary. It should also be considered that the lexicographers involved, in the works mentioned, have lived through the era of language planning themselves and the vocabulary they regarded as valuable was not necessarily used by the younger generation. Understandably, a certain extent of *schizoglossia* still exists and before embarking on a new bilingual dictionary project, lexicographers should objectively measure the impact of corpus planning on the current Turkish lexicon.

References

Avery, R. 1974 Preface. *Redhouse Turkish-English Dictionary*. Redhouse Press, İstanbul.

Cobarrubias and J. Fishman (eds) 1983. *Progress in Language Planning: International Perspectives*. Mouton de Gruyter.

Clyne, M. 1992. Language Planning. In W. Bright (ed) *Encyclopedia of Linguistics*. Vol 1. Oxford University Press. pp. 84–86.

de Vries, J. 1991. Towards a Sociology of Language Planning. In Marshall (ed). pp. 37–52.

Doğançay-Aktuna, S. 1995. An Evaluation of the Turkish Language Reform After 60 Years. *Language Problems and Language Planning*. 19:3. pp. 221–249.

Fishman, J. 1971. The Impact of Nationalism on Language Planning. In J. Rubin and B. Jernudd (eds) *Can Language Be Planned?* University Press of Hawaii, Honolulu. pp. 3–20.

Haugen, E. 1966. Language *Conflict and Language Planning*. Harvard University Press, Cambridge, Mass.

Haugen, E. 1983. The Implementation of Corpus Planning: Theory and Practice. In Cobarrubias and Fishman (eds). pp. 269–289.

Haugen, E. 1987. *The Blessings of Babel: Bilingualism and Language Planning*. Mouton de Gruyter.

Heyd, U. 1954. *Language Reform in Turkey*. The Institute of Oriental Studies, Jerusalem.

Hony, H.C. and F. İz 1978. *The Oxford Turkish-English Dictionary*. Oxford University Press.

İmer, K. 1994. Toplumbilim Işığında Türk Dil Devriminin Değerlendirilmesi. In *Uygulamalı Dilbilim açısından Türkçenin Görünümü*. Dil Derneği, Ankara. pp. 17–20.

Kurtböke, P. 1994. 17th Century Italian-Turkish Dictionaries. In *Euralex '94 Proceedings*, University of Amsterdam, Amsterdam. pp. 536–541.

Kurtböke, P. 1996. Türk Sözlükçülüğünün Sorunları 2. Kuram 10. İstanbul.

Marshall, D. (ed) 1991. *Language Planning*. John Benjamins, Amsterdam.

Tietze, A. 1962. *Problems of Turkish Lexicography*. In Householder, F.W. and S. Saporta (eds) Problems in Lexicography. Indiana University Press, Bloomington. pp. 263–272.

Zgusta, L. 1991. The Polysemy of "history". *Lexicographica 7*. pp. 1–10.