

# BabelMeSH: Development of a Cross-Language Tool for MEDLINE/PubMed

Fang Liu, Paul Fontelo, Michael Ackerman

Office of High Performance Computing and Communications

National Library of Medicine, 8600 Rockville Pike, Bethesda, Maryland 20894

## Abstract

BabelMeSH is a cross-language tool for searching MEDLINE/PubMed. Queries can be submitted as single terms or complex phrases in French, Spanish and Portuguese. Citations will be sent to the user in English. It uses a smart parser interface with a medical terms database in MySQL. Preliminary evaluation using compound key words in foreign language medical journals showed an accuracy of 68%, 60% and 51% for French, Spanish and Portuguese, respectively. Development is continuing.

## Background

When searching the medical literature, it is often easier to use the language in which one is most accustomed to.<sup>1</sup> MeSH translations could be helpful to healthcare providers who are not too comfortable with English, in locating journal articles in MEDLINE/PubMed. Although foreign language journals are indexed in MEDLINE, citations are in English only. Using their own native language could help find certain concepts that are difficult to translate into English.

## Methods

BabelMeSH (<http://babelmesh.nlm.nih.gov>) is designed as a transparent cross-language interface. Users can submit medical terms in their native language (Spanish, French, or Portuguese). A parser translates the query into English using a multi-language MySQL database. BabelMeSH then sends the query to PubMed through E-Utilities and returns English citations to the user.

**Multi-language Database:** The major source for the Spanish and Portuguese databases is UMLS Metathesaurus. The sources for French were: French MeSH translation from National Institute of Health and Medical Research, France, and Unified Medical Lexicon for French from Publique-Hopitaux de Paris. The concept unique identifier (CUI) and its concept source were used to obtain translations from UMLS. Briefly, all the concepts in one foreign language and corresponding CUIs in UMLS were selected. If the English concept in UMLS from MeSH links to the same CUI, the English concept and the foreign language concept are paired. Otherwise, the English concept from another vocabulary was selected.

**Parser:** An auto-complete feature provides suggestions as the query is entered. The parser locates only words in the translation database and

deletes all others. The corresponding English translations are sent to PubMed through E-utilities (Fig. 1). For input with accented terms, the parser searches the accented terms column in the database. If no matches are found, it will transform the accented input to the unaccented format, and search the database again. BabelMeSH will suggest records that contain part of the non-English query, if no exact matches for some terms are found. The user may modify the suggestions before searching to PubMed.



Fig. 1. BabelMeSH (French) interface and result page

## Preliminary Evaluation and Discussion

We selected 45 articles in French, Spanish and Portuguese journals. The authors' English translation was used as the "gold standard". For each article, we combined all the foreign keywords as one query to BabelMeSH. Translation accuracy for French, Spanish and Portuguese were: 68.16±24.92%, 60.78±28.22% and 51.73±29.37% respectively. Common causes of errors were: inability to detect plural or singular terms, incorrect parsing and partial matches. Some complex phrases partially matched a record in database. If these were submitted as single terms or non-complex queries, BabelMeSH would have provided suggestions of partial matches. Future development will concentrate on these deficiencies.

## Conclusion

BabelMeSH provides an alternative resource for searching MEDLINE/PubMed to healthcare personnel for whose native language is not English. The auto-complete feature alerts users of medical terms in the database. We feel there is enough interest and progress to continue development.

## References:

1. Nelson SJ, et al. The MeSH translation maintenance system: structure, interface design, and implementation. *Medinfo*. 2004;11(Pt 1):67-9.