

A comprehensive corpus of French loanwords into Vietnamese: 533 nouns that are still in use

This corpus is a selection of 533 Vietnamese nouns of French origin, based on a corpus of currently 1038 words, compiled on the basis of various sources. It provides the data of our conference paper on tonal, syllabic and segmental aspects of French loanwords into Vietnamese (Scholvin & Meinschaefer, 2018). Corpora from Barker (1969), Huynh (2010) and V. K. Nguyễn (2013) were taken as a starting point. Informal interviews with Vietnamese informants helped to expand the corpus. The informants are native speakers of Vietnamese living in Germany who have learned Vietnamese in Vietnam as a first language and acquired German in their adult life as a second language. Although they do not have any knowledge of French, they are aware of the French origin of the words they mentioned. For all 533 selected nouns, it has been checked that they are still in use, drawing on native informants' judgments as well as on word frequency and use in the World Wide Web and in a Vietnamese dictionary (Bùi et al., 2003). Concerning the pronunciation of loanwords in the corpus, the phonetic transcriptions of the Vietnamese loanwords were initially generated automatically on the basis of the orthographic representation (Kirby, 2008) and then checked with reference to native informants' pronunciation. Phonetic transcriptions of the French source words are based on the standard hexagonal pronunciation as may be found in common dictionaries (Rey-Debove & Rey, 2013). We publish this corpus as we want to be transparent about our findings and share the data that we collected.

Barker, M. E. (1969). The Phonological Adaption of French Loanwords in Vietnamese. *Mon-Khmer Studies Journal*, 3, 138–147.

Bùi, K. V., Chu, B. T., Đào, T., Tuệ, H., Hành, H. V., Lê, K. C., ... Lộc, V. (2003). *Từ Điển Tiếng Việt. Viện Ngôn Ngữ Học*. Hà Nội - Đà Nẵng: Nhà Xuất Bản Đà Nẵng.

Huynh, S. (2010). *Les mécanismes d'intégration des mots d'emprunt français en vietnamien*. Paris: Harmattan.

Kirby, J. (2008). vPhon: a vietnamese phonetizer.

Nguyễn, V. K. (2013). *Từ ngoại lai trong tiếng Việt*. Thành Phố Hồ Chí Minh: Nhà xuất bản Tổng hợp Thành phố Hồ Chí Minh.

Rey-Debove, J., & Rey, A. (Eds.). (2013). *Le nouveau Petit Robert. Dictionnaire alphabétique et analogique de la langue française. Version électronique*.

Scholvin, V., & Meinschaefer, J. (2018). The integration of French loanwords into Vietnamese: A corpus-based analysis of tonal, syllabic and segmental aspects. In *International Conference on Austro-Asiatic Linguistics (ICAAL) 7*.