

## Néologisme en chinois : une méthode basée sur le « clustering » pour faciliter l'introduction de termes occidentaux vers le chinois

Julien Roche, Zhenling Zhang, Jia Yangli

Equipe Condillac « Terminologie & Ontologie » – Université Savoie Mont-Blanc (France)  
Knowledge Engineering & Terminology Research Centre – University of Liaocheng (China)

*Dans cet article, nous proposons une méthode pour faciliter la traduction de termes occidentaux vers le chinois avec une approche dite de “clustering”, en tenant compte de l’aspect phonétique et sémantique des mots. On détaillera tout d’abord les trois principales manières de transcrire en chinois des termes étrangers. Puis nous exposerons une méthode dont le but est de sélectionner des mots chinois qui auront la meilleure probabilité d’entrer dans la composition du nouveau terme. Enfin nous présenterons à travers différents exemples de transcription, un programme informatique dédié à l’application de cette méthode.*

*Mots clés : Mots d’emprunt, néologisme, clustering, terme, distance Levensthein*

Edward Sapir: Languages, like cultures, are rarely sufficient unto themselves [1]. Some contacts among people, such as intertribal trade, cultural and spiritual interchanges, war, mixed immigration, etc. inevitably lead to some kind of linguistic inter influencing. The simplest kind of influence that one language may exert on another is the “borrowing” of words [1].

Dans l’histoire récente de la Chine, de nombreux termes occidentaux ont été introduits en chinois et continuent de l’être. Même s’il s’agit de transcription dans un système d’écriture différent, cela relève du néologisme puisqu’il faut trouver à partir de caractères chinois existants la combinaison idéale à même de représenter le nouveau concept. Par exemple « hacker » est introduit en chinois par la combinaison de deux caractères existants 黑(hei) et 客(ke), qui ont respectivement le sens ici de « illégal » et de « visiteur ». Ces deux caractères assemblés se rapprochent tout aussi bien du concept que de la prononciation.

Avec le développement économique de la Chine, on ne peut ignorer l’importance du marché chinois pour les entreprises étrangères. La justesse du nom chinois d’une marque ou d’un produit étranger constitue à lui seul un élément marketing fondamental. Par exemple « Coca-cola » est traduit par « 可口可乐 » (ke kou ke le) chaque caractère est non seulement porteur d’un son mais aussi d’un sens. De nombreux caractères possèdent le même son mais peuvent avoir un sens très différent. C’est à quoi nous nous intéressons dans cet article : comment trouver les caractères chinois qui une fois assemblés représenteront au mieux le sens et la prononciation du terme étranger.

Nous détaillerons dans une première partie les principes utilisés pour introduire un terme étranger en chinois : la transcription purement phonétique, la transcription sémantique et enfin une combinaison des deux. La difficulté est d’obtenir la combinaison optimale.

C’est pourquoi nous proposerons dans une deuxième partie, une méthode pour faciliter la création d’un terme chinois à partir de son équivalent étranger. Par exemple le terme « yacht » est traduit en chinois par 游艇 (you ting) le premier caractère signifie : « voyager » et le deuxième : « petit bateau ». On a ici une bonne combinaison entre le sens et la phonétique du mot. En utilisant une approche par « clustering », c’est-à-dire une méthode qui consiste à regrouper des éléments qui partagent des similitudes selon des critères, nous voulons proposer

une sélection de mots chinois existants (un ou plusieurs caractères) qui auront la meilleure probabilité d'entrer dans la composition du nouveau terme. Les deux critères utilisés pour le « clustering » sont la distance sémantique et phonétique. Pour chaque mot chinois existants dans un dictionnaire donné, est attaché une distance phonétique avec le terme occidental que l'on veut introduire. Cette distance phonétique est calculée en utilisant l'algorithme « distance Levenshtein ». Quant à la distance sémantique, elle est calculée à partir d'une liste préférentielle de significations que l'on souhaite faire apparaître dans le nouveau terme. Les mots chinois possédant les distances les plus courtes sont mis en évidence selon le principe de transcription que l'on souhaite mettre en avant : une transcription phonétique, une transcription sémantique ou bien une combinaison des deux.

La dernière partie de cet article sera consacrée à la présentation d'une application dédiée à la mise en pratique de cette méthode à travers différents exemples de transcription.

### **Bibliography**

- [1] Edward Sapir. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and Company, 1921.
- [2] Viviane Alleton. *Chinese Terminologies: On Preconceptions*. In *New Terms for New Ideas*, ed. Lackner et al., pp.15-34.
- [3] Shi Youwei. *Chinese Loanwords*. Beijing: The Commercial Press, 2000.
- [4] Gao Mingkai, Liu Zhengyan, Mai Yongqian, Shi Youwei. *A Dictionary of Chinese Loanwords and Hybrid Words in Chinese*. [Z].Shanghai: Shanghai Lexicographical Publishing House, 1984.
- [5] DR. Bourigault, Ch. Jacquemin, Marie-Claude L'Homme, *Recent Advances in Computational Terminology*. John Benjamins Publishing Company, 2001.
- [6] Wang Li. *A Draft History of the Chinese Language*. Beijing: Zhong Hua Book Company, 1980.3.
- [7] Feng Zhiwei. *The Semantic Loanword and Phonetic Loanword in Chinese language*. *Aspects of Foreign Words/Loanwords in the Word's language*, Proceedings for 11th International Symposium, The National Institute for Japanese Language, p200-229, 2004, Tokyo.
- [8] Shen Guowei. *New Er Ya*. Shanghai: Shanghai Lexicographical Publishing House, 2011.4
- [9] Peter Newmark. *Pragmatic translation and literalism*. *TTR : traduction, terminologie, rédaction*, vol. 1, n° 2, 1988, p. 133-145. URI: <http://id.erudit.org/iderudit/037027ar>.
- [10] Dictionary Editing Room of the Linguistics Institute of Chinese Academy of Social Sciences. *Modern Chinese Dictionary*. 2012.6.
- [11] Black, Paul E., ed. (14 August 2008), "Levenshtein distance", U.S. National Institute of Standards and Technology