
Au-delà des caractères simples : analyse des combinaisons lexicales spécialisées autour du yangsheng en didactique de la traduction humaine et assistée par l'IA

Pascale Chazaly Elbaz^{*1,2}

¹Institut de Management et de Communication Interculturels – Institut Français de Recherche sur l'Asie de l'Est (IFRAE, CNRS/INALCO) – France

²IFRAE – CNRS : UMR8043, CNRS – France

Résumé

La langue spécialisée est constituée de cristallisations linguistiques et sociologiques qui se laissent appréhender sous des formes diverses, depuis certains stéréotypes lexicaux jusqu'à des habitus discursifs (Elbaz & Miao, 2022). L'étude de ces cristallisations s'avère complexe en chinois moderne, où le monosyllabisme des morphèmes est équilibré par le disyllabisme et, plus rarement, le polysyllabisme des mots (Drocourt, 2022). Aux monosyllabes "héritaires", qui représentent le fonds lexical du chinois moderne, viennent s'ajouter les disyllabes, constituant la grande majorité du lexique, les polysyllabes, moins nombreux, et les expressions phraséologiques, ou "formules". Lors de ces combinaisons, les morphèmes conservent leur forme phonologique (Packard, 2015) mais perdent leur polysémie. Pour aborder l'apprentissage de ces différentes unités et combinaisons lexicales en cours de traduction, nous avons choisi une méthode exploratoire nous permettant de distinguer les acceptations d'un caractère dans les différentes combinaisons lexicales qu'il forme, comprendre comment l'ajout d'un deuxième caractère réduit la polysémie et, à plus grande échelle, observer la variation de ses cooccurrences. Nous travaillons sur le terme *yāngshēng*, formé de deux monosyllabes "héritaires", au même titre que *wúwéi* (non-agir) ou *xiūxíng* (s'exercer dans la pratique) (Sun, 2006). A la fois théorisation d'une hygiène de vie et culture de soi, le *yāngshēng* consiste à nourrir en soi le principe vital par des exercices appropriés ; à entretenir la vie et à préserver sa santé (Ricci, 2006). Nous explorons les multiples unités lexicales autour de ce terme, à savoir les caractères individuels *yāng* et *shēng* ; les combinaisons lexicales formées par chacun des caractères, - *bǎoyāng*, *tiáoyāng* - et - *shēngmìng*, *shēngcún* - ; les cooccurrences du terme. En plus d'une attention particulière portée à la stabilisation des formes langagières, nous nous penchons sur la nature grammaticale de ces formes et de leurs composants et sur les différentes combinaisons possibles. Ainsi, si la forme *yāngshēng* s'est imposée dans la langue, le terme *shéngyāng* existe également. Nous nous appuyons sur la lexicologie explicative et combinatoire (Mel'čuk, 2012, 2023) et sur la didactique de la langue spécialisée (Cavalla et al., 2020 ; Nation, 2022 ; Han, 2022). Enfin, nous testons différents moteurs de traduction automatique afin de mesurer leur efficacité pour traduire ces différentes unités lexicales en à partir des recherches en linguistique computationnelle (Wang & al., 2018 ; Han & al. 2021) et de nos propres recherches (Elbaz, 2023).

*Intervenant