

Annotated database of conventional euphemistic expressions in Chinese: explanatory notes

The doctoral research project on Chinese euphemisms „Conceptualisations of Death, Illness, and the Body in Chinese Euphemisms: A Survey of Sensitive Vocabulary in Modern Chinese Lexicography“ is aimed to contribute to the universal typology of euphemistic expressions. Since Chinese is greatly divergent in its typological characteristics from languages traditionally involved in the studies of sensitive vocabulary, the inclusion of linguistic evidence of Mandarin Chinese into the general discussion on euphemisms can enhance our understanding of how languages deal with taboo entities.

Present research treats euphemization and dysphemization as processes of conceptualization, focusing on the semantic means of euphemistic formation, namely metaphorical and metonymical extensions. Euphemisms related to three major domains Death (Dying and the Deceased, Funerals and Burials, Afterlife, Old Age and Aging), Disease (Physical and Mental Illnesses, Physical and Mental Disabilities) and Body (Sex and Sexual Desire, Non-Marital Sex, Genitals and Other Body Parts, Prostitution, Excretion, Menstruation) constitute the core of the present project. In order to collect and assess this sort of data, an annotated index of euphemistic expressions was compiled through qualitative analysis of existing common and specialized Chinese monolingual lexicography.

Why dictionaries?

The project is focused on conventional euphemisms found in common lexis. Therefore, the priority in the selection of data is given to conventionalised expressions that were or are used by a relatively large number of speakers in both

formal and informal communicative situations and codified in common lexicon. Well-established, well-entrenched, or even clichéd euphemisms are considered here conventional, as opposed to unconventional euphemisms creatively coined by a certain author in a certain text and/or strongly associated with those. The novel euphemisms might be perceived by speakers as unusual, too specific or too individual.

Which dictionaries?

The major specialised dictionaries of Chinese ‘tactful expressions’ *wěiwǎnyǔ* 委婉语 were taken as the main source of the databank, namely:

- *Hanyu Weiwanyu Cidian* 《汉语委婉语词典》 by Zhang Gonggui (1996)
- *Qianci, Jingci Wanci Cidian* 《谦词敬词婉词词典》 by Hong Chengyu (2010)
- *Tongyong Weiwanyu Cidian* 《通用委婉语词典》 by Zhu Jingsong (2018).

I also retrieved the material of contemporary Chinese dictionaries of common vocabulary, including:

- *Hanyu Dacidian* 《汉语大词典》 (HDC, 2010)
- *Guoyi Cidian* 《重編國語辭典修訂本》 (GC, 2015)
- *Xiandai Hanyu Guifan Cidian* 《现代汉语规范词典》 (GF, 2014)
- *Xiandai Hanyu Cidian* 《现代汉语词典》 (XHC, 2016)

Only dictionaries?

I also included several new euphemisms common in digital communication which I came across accidentally, when I had personal interviews with native speakers or read both formal and informal blogs (including fora for young mothers, quora-like Q&A posts on *Zhihu* platform etc.) These are currently not codified in dictionaries in most of cases. However, their “commonness”, their prevalence in the speech of native speakers might suggest that they are in the process of fossilization into common vocabulary. These novel entries were tested in terms of their occurrence in the largest Mandarin Chinese text corpora available: CCL created by the Center of Chinese Linguistics at Peking University, with ca. 470 million characters (http://ccl.pku.edu.cn:8080/ccl_corpus); and BCC, created by the Beijing Language and Culture University, comprising ca. 15 billion characters (<http://bcc.blcu.edu.cn>).

How exactly were these sources processed?

First, entries were manually retrieved from the specialized dictionaries. Zhang (1996) and Hong (2010) organized their entries thematically, thus, it was unproblematic to work with these lexicographic sources and find the expressions related to the sensitive domains Death, Illness, Sex, Genitalia, Excretion etc. However, Zhu’s dictionary (2018) does not organise its entries by theme, so retrievals were made manually by analysing the whole corpus of the dictionary.

Second, I dissected the lemma presentation and additional information found in the lexica for each entry, including search options, article structure, various data types, language usage guidance (prescriptive/descriptive), synonyms and equivalents, cultural information, diasystematic labelling, etymology, example sentences, collocations and phrases, idioms, and cross-references. I was

particularly focused on establishing the pragmatic and variational markedness of an expression, including pragmatic characteristics of the denotata and of the language user (is it said about people of a specific age? A specific gender? Is it common among speakers of a specific profession? Social group? Religious group?), obsolete status of entries (diachronic markedness)¹, regional status (diatopic markedness) and stylistic status (aka diaphasic markedness, if the expression belongs to a certain register, either formal or informal; or has additional evaluative meanings: jocular, ironic, playful). Several pragmatic features could be combined in a single term, cf. euphemism ‘bury jade and inter incense’ *zàngyù máixiāng* 葬玉埋香 for funeral or death of a beautiful (appearance) young (age) woman (gender) of virtue (moral standards).

Third, this information was expanded and cross-checked with the help of major common lexica of Chinese language. E.g., euphemism ‘not good’ *bùxíng* 不行 denoting a person’s death is marked as such (婉词, 指人死亡) in GF (2014), HDC (2010), GC (2015), and XDC (2016), but has not been found in any of the specialized dictionaries of tactful expressions.

Forth, in order to identify the word motivation and the mechanism of formation, I conducted the morphemic analysis of each of the component in the entry.

And finally, the mechanism of formation was discovered and further elaborated: either formal or semantic. If the euphemism was built by a metaphoric extension,

¹ It is important to underline that ‘obsolete’ status of an entry is recognized even if the Chinese word is still in use in a different, non-euphemistic meaning, cf. euphemistic meaning of the term ‘leak air’ *xìèqì* 泄气 is flatulence, farting (放屁), however, nowadays it is used in its literal meaning ‘lose air’ = ‘deflate’ (of a balloon or of a car tire) and metaphorically (but still non-euphemistically) ‘feel discouraged’, ‘lose hope’, ‘get disappointed’ (XHC 2016).

the metaphorical identification procedure (MIP) was applied, centred on differentiation of the so-called “basic” meanings of linguistic expressions (its concreteness, relatedness to bodily action, precision, and sustainability).

Data presentation

ID and Target Domains

Each entry was assigned to one of the three main domains: Death (D), Illness and Deficiency (ID), and Body (B).

Entries which belong to domains Death and Body have additional letters in their descriptors: DD for Death: Dying And The Deceased; DF for Death: Funerals And Burials; DA for Death: Afterlife; DO for Death: Old Age And Ageing; BS for Body: Sexual Desire And Sex; BN for Body: Non-Marital Sex; BG for Body: Genitals And Other Body Parts; BP for Body: Prostitution; BE for Body: Excretion; BM for Body: Menstruation.

Pinyin

All entries within a certain subdomain are organised “alphabetically” based on pinyin – the standard romanization for Mandarin Chinese. To avoid the segmentation problem caused by undefined „word“ boundaries in Chinese, each pinyin syllable associated with a character is separated from another one by spaces.

Translation

Each entry is translated into English in two ways—literally (word-for-word or morpheme-for-morpheme) and idiomatically. This allows us to understand what kind of figurative language (over 90% of entries are products of semantic shifts) is deployed in order to convey its euphemistic meaning. Meaning in Chinese is given based on the lexicographic data retrieved from Chinese dictionaries of common vocabulary and specialized dictionaries of tactful expressions.

Mechanism of euphemisation

Each database entry has been classified according to its corresponding mechanism of euphemistic formation: semantic or non-semantic (formal, including borrowing, reduplication, abbreviation and deletion, phonemic and graphic modifications etc.).

Elaborations on metaphors include the relevant source and target domains; if it is a metonymy or reversal, I explicitly state which vehicle and tenor are involved. If there are several possible motivations, I specify those in “remarks” section, cf. several motivations of the metaphoric euphemism ‘orioles and swallows’ *yīngyàn* 莺燕 for prostitutes: both birds are ‘spring birds’ (Zhang 1996), and the spring season has a sexual connotation in Chinese. On the other hand, orioles are said to be good at singing and swallows are said to be good at dancing (GC 2015), thus referencing singers, dancers, and prostitutes.

The latest updated version from 04 July 2022 contains minor corrections related to the formal presentation of the data: entries related to the same subdomain were sorted in alphabetical order (pinyin romanization); translation of several Chinese words in remarks section was revised and its formatting unified.

1. GC 2015. Ministry of Education, Republic of China 中华民国教育部 (Ed.) (2015). *Chongbian Guoyu Cidian Xiudingben* 重編國語辭典修訂本 [A Revised Dictionary of the National Language]. Retrieved September 1, 2021, from: <http://dict.revisEd.moe.edu.tw/>
2. GF 2014. Li, Xingjian 李行健 (Ed.) (2014). *Xiandai Hanyu Guifan Cidian* 现代汉语规范词典 [A Standard Dictionary of Modern Chinese] (3rd Edition). Waiyu Jiaoxue yu Yanjiu Chunbanshe.
3. HDC 2010. Luo, Zhufeng 罗竹风 (Ed.) (1993). *Hanyu Da Cidian* 汉语大词典 [Comprehensive Chinese Word Dictionary] (CD-ROM Edition, ISBN 962-07-0255-7, Pleco version). Shanghai Lexicographical Publishing House. Supplemented by: Office of Chinese Dictionary Compilation (Ed.) 汉语大词典编纂处 (2010). *Hanyu Da Cidian Dingbu* 汉语大词典订补 [Supplementary Volume of Comprehensive Chinese Word Dictionary]. Shanghai Lexicographical Publishing House.
4. Hong, Chengyu 洪成玉 (2010). *Qianci Jingci Wanci Cidian* 谦词敬词婉词词典 [A Dictionary of Self-Deprecatory, Honorific and Tactful Expressions] (Extended edition). Commercial Press.
5. XHC 2016. Chinese Academy of Social Sciences, Research Institute of Language Department of Dictionary 中国社会科学院语言研究所词典编辑室 (Ed.) (2016). *Xiandai Hanyu Cidian* 现代汉语词典 [Contemporary Chinese Dictionary] (7th edition). Commercial Press.

6. Zhang, Gonggui 张拱贵 (1996). *Hanyu Weiwanyu Cidian* 汉语委婉语词典 [A Dictionary of Chinese Tactful Expressions]. Beijing Language and Culture University Press.
7. Zhu, Jingsong 朱景松 (Ed.) (2018). *Tongyong Weiwanyu Cidian* 通用委婉语词典 [Dictionary of Common Tactful Expressions]. Yuwen Chubanshe.