Some Saliant as well as Divergent and Convergent Linguistic Developments in Chinese— A Big Data and Trans-Millennial Approach

Benjamin K. T'SOU 鄒嘉彥

City University of Hong Kong / Hong Kong University of Science and Technology

In the last few decades surrounding the emergence of the new millennium, the Pan-Chinese speech communities have witnessed epochal changes. With the additional and rapid technological developments in the information age, we could now also catch sight of the concomitant new developments of lexical and grammatical variations which may be related to the emergence of tangible and intangible variegated cultural artifacts and salient events. In this paper, we propose to review some latitudinal and longitudinal developments by means of LIVAC*, a 700-million-character Chinese corpus that draws on media materials from six Chinese speech communities since 1995 and covers a critical period in contemporary Chinese history. It provides an uncommon window to explore the mutual relationships between language and the broader societal and cultural contexts in which they are embedded so as to deepen our understanding of the Chinese language and of its diverse communities of speakers.

Our focus will range from the rise of tri-syllabicity and portmanteau-like words, linguistic enrichment in quantitative and qualitative terms, resulting from indigenous and cross-regional enrichment to innovative manipulation of grammatical as well as pragmatic devices such as light verbs and metaphors, and alternate modes of obtaining information. The rise of contrastive convergent and divergent trends also invites explanations in the larger context of the Chinese society.

*LIVAC: https://en.wikipedia.org/wiki/LIVAC_Synchronous_Corpus

References:

Dong, X. F. (2011). *Lexicalization:The Origin and Evolution of Chinese Disyllabic Word (Revised Edition)* [词汇化:汉语双音词的衍生和发展（修订本）]. Beijing: The Commercial Press

Lü, S. X. (1963). Xiandai Hanyu Shuangyinjie Wenti Chutan [现代汉语双音节问题初探 An enquiry to the question of disyllabification in Chinese]. *Zhongguo Yuwen [*中国语文*]*, 1, 11 – 23.

Ke, J. Y. 2007. A cross-linguistic quantitative study of homophony. *Journal of Quantitative Linguistics* Vol.13, pp. 129-159.

Tsou, B. K. 1972. Reordering in Diachronic Syntax, *Papers from the Eighth Chicago Linguistics Society Meetings,* ed. P. Peranteau et. al. Chicago, 1972, pp.591-612.

Tsou, B. K. (2016a). Sociolinguistic Aspects of the Chinese Language. Oxford Bibliographies in Chinese Studies, Oxford University Press.

Tsou, B. K. (2016b). Lexical Variation of Dialectal Origin in Modern Standard Chinese. *The Encyclopedia of Chinese Language and Linguistics*. R. Sybesma, W. Behr, Y. Gu, Z. Handel and C.-T. J. Huang. (Eds.). Leiden: Brill.

Tsou, B.K. & Ji, Y. X. [2022] Contrastive simplification of near-synonym MSC compounds in Mandarin and Cantonese: A study on morpho-syllabic differentiation in dialectal developments. (IACL -28)

Tsou, B. K. & Kwong, O.Y. (2015a). LIVAC as a Monitoring Corpus for Tracking Trends Beyond Linguistics [從 LIVAC 追蹤語料庫探索泛華語地區語言以外的演變]. *Journal of Chinese Linguistics Monograph Series.* 25 (2015/12/01), 447 – 471.

Tsou, B. K. & Kwong, O. (2015b). Some Quantitative and Qualitative Characteristic Features of the Chinese Language. *The Oxford Handbook of Chinese Language and Linguistics*. Sun, C.fF and Wang, W.(Eds). Oxford University Press. Tsou, B. K. & Yip, K. F. (2020). A corpus-based comparative study of light verbs in three Chinese speech communities. In *Proceedings of the*

34th Pacific Asia Conference on Language, Information and Computation, PACLIC (pp. 302-311).

Yip, K.F. & Tsou, B.K. (2022). MSC variations in metaphorization among Pan-Chinese speech communities (IACL -28)