摘 要

本研究基于"中文十亿字标注语料库"，从跨变体的视角出发，考察了影响中国大陆汉语、中国台湾汉语和新加坡汉语分析型致使构式中"使""让""令"交替的句法、语义、语言变体等因素。本研究首先从子语料库（六百万字左右）中分别自动抽取出"使""令""让"的索引行，并人工排除了非致使用法，然后从清理结果中随机选择了 3070 条（占有效致使用法的 30%）索引行，对其进行了 27 个自变量的标注。接着，利用条件推断决策树分析了各个自变量间的交互作用，并使用随机森林模型得出了自变量的重要性排序。对于随机森林判定为具有显著效应的五个自变量，本研究进一步采用多项逻辑斯蒂回归分析考察了其影响的方向和强度。

统计分析结果显示，1）"令"构式中被使者的形式通常是非定指的"人"字，被使者常充当感事，结果谓词倾向于使用形容词，"令"构式常用于关系从句，这说明["令人"+结果]呈现出词汇化倾向；2）相较于"令"，"使"和"让"都更倾向于用系动词和及物动词作为结果谓词；3）然而，"使""让"对被使者语义角色的偏好有所不同，当被使者是受事或受益（害）者时，"使"的概率显著提高，而"让"更倾向于充当施事的被使者；4）随机森林将"语言变体"评估为最重要的自变量，而多项逻辑斯蒂回归模型也进一步显示，"语言变体"与"小句类型"之间存在显著的交互效应：中国大陆汉语中"使""让"在不同小句类型中的概率差异更大，中国台湾汉语和新加坡汉语中"小句类型"对"使""让"概率的影响较小，且"让"的概率总是高于"使"，而三个语言变体在"令"的用法上未表现出显著差异。

本文的结论是，汉语分析型致使构式中"使""让""令"的交替受句法、语义等语言内部因素和变体这一语言外部因素的共同影响，且表现出了显著的跨变体变异。本文体现了认知社会语言学将语言内部和外部因素整合在同一框架内进行分析的优势，并在更多数据和参项的基础上丰富了汉语分析型致使构式与汉语变体变异的研究。

关键词：认知社会语言学；分析型致使构式；构式交替；汉语变异；多因素分析

Abstract

This study examines the onomasiological choice of three near-synonymous Chinese causative markers, i.e. *shi*, *rang* and *ling* and its lectal variation in Mainland Chinese (MC), Taiwan Chinese (TC) and Singapore Chinese (SC) using the 'Tagged Chinese Gigaword Corpus'. All the observations with *shi*, *rang* and *ling* were automatically retrieved from a sub-corpus (around six million words) and then were manually checked to avoid spurious hits. We randomly selected 3070 observations (30% of all valid observations) and annotated them with 27 predictors. Then, we used conditional inference trees and random forests to disentangle the interactions between predictors and to obtain an importance ranking of all the predictors. Finally, a multinomial logistic regression model was built with the significant predictors evaluated by the random forest.

Our analyses show that: 1) the CAUSEE of *ling*-constructions tends to be an experiencer with the form of *ren*, followed by adjectives, and *ling*-constructions significantly appear more in a relative clause, showing a lexicalization tendency; 2) both *shi* and *rang* are more favored when the effected predicate is a transitive verb or a copula; 3) the probability of *rang* significantly increases when CAUSEE is an agent, whereas *shi* is more favored when CAUSEE is a theme or a patient; 4) VARIETY is evaluated by the random forest to be the most imporant predictor. The multinomial logistic regression further detects significant interaction between VARIETY and CLAUSETYPE: The influence of CLAUSETYPE is stronger in MC, with *shi* and *rang* favoring different clause types, whereas in TC and SC *rang* is always more favored than *shi*. The three varieties do not show significant variation on the usage of *ling*.

To conclude, the choice of *shi*, *ling* and *rang* is jointly affected by multiple factors, and significant lectal variation has been found. This study not only shows the benefit of Cognitive Sociolinguistics for synthetically analyzing language internal and external factors, but also contributes to the research on Chinese analytic causative constructions and variation in Chinese varieties by analyzing large-scale data and incorporating

considerable predictors.