A Multivariate Quantitative Study on English Modal Construction from a Variationist Linguistics Perspective 一 A Case Study of *must, have to, have got to*

Abstract: From a corpus-based variationist linguistics perspective, this study explores the major factors influencing the choice of "must", "have to" and "have got to" , these three constructions and its diachronic evolution in American English, by means of COHA corpus data and R software. After the annotation of nine predicative factors including tense, genre, year, etc, Conditional Inference Tree modal and Conditional Random Forest modal are used to analyze the data. The results show that: The frequency of "must" is declining from 1810 to 2009 in American English; Although there is still a high frequency of "have got to" in the "novel" genre, the frequency of "have got to" is decreasing in the diachronic evolution. The frequency of "have to" is on the rise. The main factors influencing the selection of these three constructions include "tense", "genre" and "year". In addition, the influence of other predictive variables selected in this study on the selection of three construction variants is not obvious. Through further analysis on the standardized frequency distribution of "must" and "have to" from 1820 to 2019 in American English, it can be concluded that when expressing the meaning of "necessary to do sth.", there was a trend that "must" was gradually replaced by "have to".
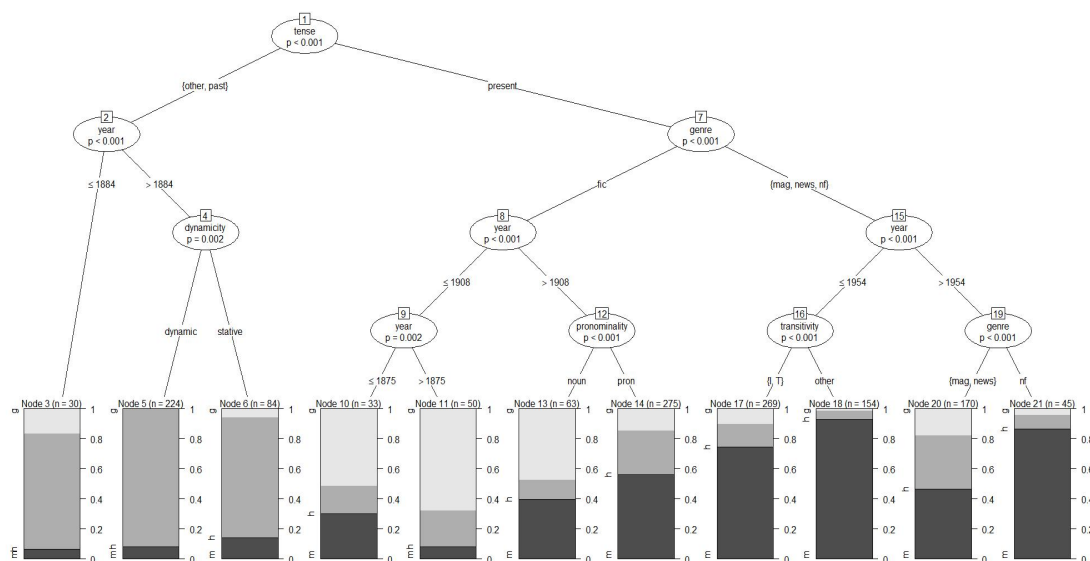

Figure 1 Conditional Inference Tree model on predictive variables influencing the choice of *must, have to* and *have got*
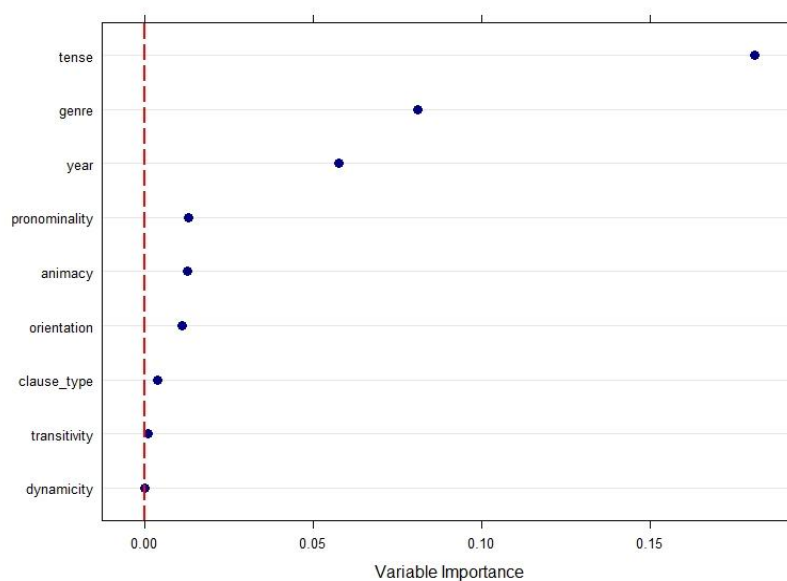
Figure 2 Conditional Random Forest analysis on predictive variables influencing the choice of *must, have to* and *have got to*

Bibliography

BIBER D, JOHANSSON S, LEECH G, CONRAD S, FINEGAN E. Longman Grammar of Spoken and Written English [M]. London: Longman, 1999.

BOD R, HAY J, JANNEDY S. Probabilistic linguistics [M]. Cambridge, MA.: The MIT Press, 2003.

BRESNAN J. Is syntactic knowledge probabilistic? Experiments with the English dative alternation [A]//FEATHERSTON S, STERNEFELD W. (eds.) Linguistics in Search of its Evidential Base [M]. Berlin: Mouton de Gruyter, 2007, 75-96.

CAPPELLE B, DEPRAETERE I, LESUISSE M. The necessity modals *have to*, *must, need to* and *should*: using n-grams to help identify common and distinct semantic and pragmatic aspects [J]. Constructions and Frames, 2019, 11(2):220-243.

CLOSE J, AARTS B. Current change in the modal system of English [C]//LENKER U, HUBER J, MAILHAMMAR R. English historical linguistics 2008. Amsterdam: John Benjamins, 2010: 165-182.

DEPRAETERE I, REED S. The handbook of English linguistics [M]. Oxford: Blackwell, 2006.

GRIES S. Multifactorial analysis in corpus linguistics: a study of particle placement

KRUG M G. Emerging English modals: a corpus-based study of grammaticalization [M].

LEECH G. Modality on the move: the English modal auxiliaries 1961-1992 [C]// FECCHINETTI R, KRUG M G, PALMER F R. Modality in contemporary English. Berlin: Mouton de Gruyter, 2003: 223-240.

PEREK F. Argument structure in usage-based construction grammar [M]. Amsterdam: John Benjamins, 2015.

QUIRK R, GREENBAUM S, LEECH G, et al. A comprehensive grammar of the English language [M]. London: Longman, 1985.

SMITH N. Changes in modals and semi-modals of strong obligation an epistemic necessity in recent British English [A]//FECCHINETTI R, KRUG M G, PALMER F R. (eds.) Modality in contemporary English [M]. Berlin: Mouton de Gruyter, 2003.

SZMRECSANYI B, GRAFMILLER J, HELLER, B, RÖTHLISBERGER M. Around the world in three alternations: Modeling syntactic variation in varieties of English [J]. English World-Wide, 2016(37): 109-137.