

Predicting the Sentential Positions of PPs in the History of Mandarin: A Statistical Study

Old Chinese allows adverbial preposition phrases to appear in postverbal position, as in 齊宣王見孟子於雪宮 ('The king Xuan of Qi met Mengzi at Snow Palace'). On the other hand, texts of Middle Chinese and later stages of the history reveal that adverbial PPs tend to appear in preverbal position 於獵場見齊莊 ('At the hunting place met Qizhuang'). Among previous studies (Peyraube 1994, Sun 1996, Zhang 2002, and Liu 2004), Zhang (2002) in particular has argued that the iconic association between the syntactic positions and the meanings of PPs, the monosyllabicity of matrix verbs, and the syntactic complexity of VPs are factors in the positional shift of adverbial PPs from the postverbal to the preverbal position. However, this work does not account for how the change actually took place.

This study proposes a mixed effects model that predicts the sentential positions of adverbial PPs, building on a statistical model that has been developed to predict speakers' choice in different syntactic constructions concerning the English dative alternation (Bresnan et al. 2005). I argue that the relative weights of the driving forces that contribute to the positions of adverbial PPs vary in each historical stage, rather than being a limited set of factors with constant significance, as implicit in previous studies. By doing so, this study will show a dynamic picture of interaction between multiple factors, which yield syntactic variation.

First, I conducted a corpus study to analyze multiple factors which are related to the sentential position of adverbial PPs. I looked at 5000-character-blocks of texts from the following works: *Mencius* (孟子, 4th century BC), *The Historical Records* (史記, 2nd-1st centuries BC), *Shi shuo xin yu* (世說新語, 5th century) and *Zutangji* (祖堂集, 10th century). I annotated data in terms of the semantic type of PP, the existence of a postverbal direct object, the monosyllabicity of the matrix verb, and co-occurrence with a resultative complement or adjunct phrase, shown to be relevant by Zhang (2002). Further, I annotated new factors, whether PPs bear aspectual roles or not (Tenny 1994), the number of postverbal XPs (Economy within VP; Fang 2006, Peck and Sells 2006), the total number of VPs, and the historical period of a given datum.

All these factors were then used in the mixed effects logistic regression model. The preverbal position or postverbal position is modeled as a function of the presence of a PP denoting the endpoint, the presence of an object in postverbal structure, the number of syllables in a main verb, + the presence of resultative complement, the presence of durative adverbial phrase in the postverbal structure, the number of XP within VP, the number of VP within VP', the historical period to which a given sentence belongs. The model evaluates each factor's strength as a predictor of the response values, i.e. whether in diachronic syntactic variation, the position of an adverbial PP is realized as [PP VP] or [VP PP]. The conditional probability of a response is systematically linked to a linear combination of fixed cross-group explanatory variables introduced above, and also to a randomly varying normally distributed group effect in relation to the writing styles of authors and the lexical bias of verbs.

The resultant model successfully demonstrates the multi-variable nature of determining the sentential position of adverbial PPs. It also correctly classifies 95% of the data overall. The aspectual role of PP, Economy within VP and the monosyllabicity of the matrix verb were found to be significant predictors. In contrast, iconic association, the co-occurrence of resultative complements and durative adverbial PPs within postverbal structure were not at all predictive factors, revealing that these factors are not driving forces, but later development after the syntactic change took place. In addition, the weight of the aspectual role of PP was relatively consistent throughout all historical stages, whereas the weight of Economy within VP, and the monosyllabicity of the verb become heavier towards the Middle Chinese period, suggesting that the degree of contribution of some factors may change in different historical periods.