

Abstract submitted to NACCL 28

**Is China Entering WTO or 世界贸易组织  
--A Corpus-based Study of English Acronyms in Chinese Newspapers**

Key words: language contact, code-switching

Hai Hu  
Indiana University Bloomington  
huhai@indiana.edu

In recent years, there has been a discussion in China about whether commonly used English borrowings (e.g. WTO, GDP) should be allowed in Chinese texts (Hu 2002; Yu & Zhu 2003; among others). Despite the intensifying debate on “Englishization” and the “purity of Chinese”, few studies have systematically investigated the type and degree of direct English borrowing into Chinese. In this study, using corpora from newspaper sources, I attempt to answer the following questions: 1) for all instances of a foreign concept (e.g. *World Trade Organization*), what percentage of it is expressed in English (WTO), and in Chinese (世界贸易组织), and 2) what factors are at play in language users’ choice between the English and Chinese terms? This study uses two datasets to calculate the exact percentage of English usage for every concept, and linear models to quantitatively determine the role of possible predictors.

Ten concepts (WTO, CEO, etc.) are examined in the first dataset, which consists of three Chinese newspapers from year 2005 to 2014. Table 1 illustrates the percentage of English borrowings for four concepts from one newspaper, *Southern Metropolis Daily* (南方都市报).

Concept	Dependent Variable	Predictor 1:	Predictor 2:	Predictor 3:
	<i>PercentOfEn</i>	“ <i>entrenchment</i> ”: frequency	“ <i>lg economy</i> ”: NumOfChar in CHN	“ <i>first appearance</i> ”
WTO	54%	9690	5 (世贸组织/世界贸易组织)	CHN=EN
CEO	68%	9717	5 (首席执行官)	CHN
CPI	95%	19257	8 (居民消费价格指数)	CHN
EMBA	97%	2293	12 (高级管理人员工商管理硕士)	EN

Table 1 Four concepts from the newspaper *Southern Metropolis Daily*

Three factors in Table 1 are hypothesized to predict the percentage of English borrowings in a linear mixed-effects model (following Zenner et al. 2014). The level of “*entrenchment*”, operationalized as the frequency of the concept in the corpora (English and Chinese taken together), predicts that the more entrenched a concept is, the more likely it will be used in the native form (i.e. Chinese). “*Language economy*”, operationalized as the number of characters in the Chinese translation, predicts that the more concise the Chinese translation is, the more likely the Chinese form will be used. “*First appearance*” notes which of the two terms appeared *first* in Chinese newspapers, the prediction being that if the English term appeared earlier than its Chinese counterpart, a higher *PercentOfEn* would be expected.

The results of the first dataset show that different concepts have a varying *PercentOfEn*, ranging from 2% to 98%. One unexpected but interesting finding is that *proper nouns* (e.g. WTO, OPEC and IMF) all have a lower *PercentOfEn*. Crucially, the linear mixed-effects model demonstrates that the three predictors can predict *PercentOfEn* in the first dataset (all  $p < .05$ ), confirming the hypothesis. That is, how much English is used for one concept can be predicted by 1) how often the concept occurs, 2) how long the Chinese translation is, and 3) whether the English or Chinese form first appeared in the newspaper.

In order to confirm the results, a second dataset with 31 concepts was used. This dataset contains news articles of *Xinhua News Agency*, obtained through the CCL corpus of Peking University. A multiple linear regression shows that the same three predictors can also predict the *PercentOfEn* (all  $p < .05$ ).

In sum, not all concepts have a very high percentage of English borrowings. Language users choose the English term either because the Chinese translation is too long, or because the concept is not yet well entrenched in the Chinese lexicon, i.e. it may still need more time to “settle in”. This result demonstrates the process of acceptance of new concepts into a language, and how loanwords and the native terms are competing for acceptance. It also predicts that with deeper entrenchment and the newly created shorter translations (such as 美职篮 for NBA), some Chinese terms are likely to win out in the future.