Syntactic differences of adverbials and attributives in Chinese-English code-switching

Lin Wang [a], Haitao Liu [a,b,*]

[a] School of Foreign Languages, Zhejiang Gongshang University, Hangzhou, 310018, China
[b] Department of Linguistics, Zhejiang University, Hangzhou, 310058, China

A B S T R A C T

Based on a Chinese-English code-mixed treebank, this paper investigates the effect of code-switching on dependency distance and dependency direction of two major grammatical relations (adverbials and attributives). It was found that (1) mixed dependencies generally present longer dependency distances than monolingual ones, especially the adverbial dependencies with preposition or adverb dependents and attributive dependencies with pronoun dependents; (2) mixed adverbial dependencies with English verb heads and Chinese adverb and noun dependents and mixed attributive dependencies with Chinese noun heads and English noun dependents present shorter dependency distances than monolingual ones; (3) word order differences are largely found in adverbial dependencies with adverb, noun or preposition dependents and potentially in attributives with preposition dependents. These findings suggest that: (1) code-switching is constrained by the grammars involved in it; (2) the syntactic properties of dependents mainly determine the word order of adverbials and attributives.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Code-switching (hereafter CS) involves both language comprehension and production, and may happen within a sentence (intra-sentential CS) or between sentences (inter-sentential CS). Based on the naturalistic Chinese-English code-switching corpus, this paper quantitatively investigates the syntactic differences between highly-frequent monolingual and mixed syntactic adverbial and attributive relations.

In recent decades, studies on code-switching have covered various languages, integrating effort from various fields like syntax, psycholinguistics and sociolinguistics (Chan, 2009; Bullock and Toribio, 2009). Syntactically, researchers are mainly interested in syntactic properties of two languages influencing the code-switching. Some researchers have proposed models or constraints for code-switching, such as the Free Morpheme and Equivalence Constraints (Sankoff and Poplack, 1981), the Phrase Structure Congruence Constraint (Woolford, 1983), the Government Constraint (Di Sciullo et al., 1986), the Functional Head Constraint and the Word Order Integrity Corollary (Belazi et al., 1994), and the Matrix Language Frame Model (Myers-Scotton, 2002). Some researchers propose that there are no specific constraints for code-switching and that "pure" languages and mixed languages are governed by the same constraints or principles underlying universal grammar—The Null Theory

---

* Corresponding author. Department of Linguistics, School of International Studies, Zhejiang University, No. 866 Yuhangtang Road, 310058 Hangzhou, China.
E-mail address: lhtzju@gmail.com (H. Liu).

(Mahootian, 1993; MacSwan, 1999, 2000; Chan, 2003). Other scholars proposed that there are no constraints on code-switching because it is highly variable (Bokamba, 1989; Gardner-Chloros and Edwards, 2004). From the perspective of psycholinguistics, researchers have studied the brain responses to the production of language switches (Jackson et al., 2001) and the comprehension of language switches (Moreno et al., 2002; Proverbio et al., 2004), the relations between CS and aphasia or dementia (Hyltenstam, 1995; Myers-Scotton and Jake, 2000), and the relations between code-switching and the mental lexicon (Wei, 2009). In addition, various sociolinguistic models and theories have been developed to explain the social motivations or functions of conversational code-switching (Gumperz, 1982), the Markedness Model in Myers-Scotton (1993, 2006) and the social network study on CS in Milroy and Wei (1995). However, very few studies have provided a corpus-based, comparative syntactic and cognitive analysis on CS between two typologically distant languages.

To some extent, the diversity and variation of the code-switching appear to defy any economical and universal syntactic account (Chan, 2009). Hudson (2010) indicated that the positive side of the structuralist approach is its focusing on the complex internal language structure, "but it had negative effect of encouraging linguists to ignore everything outside language, and in particular, to assume that language is organized differently from everything else in our minds" (Hudson, 2010:107). Instead, "Language is a part of general cognition" (Hudson, 2010:108) and "knowledge of language is knowledge" (Goldberg, 1995:5). Word Grammar, a cognitive linguistic theory developed primarily by Richard Hudson, aims to marry "the general insights of cognitive science into how our minds work with the enormous amounts of detail that linguists analyze" (Hudson, 2010:108). As a result, Word Grammar is an effective linguistic theory to study language structures and how they are organized in our mind. Therefore it is adopted in our study.

Two important syntactic properties in Word Grammar—the Dependency Distance (hereafter DD) and the Dependency Direction—are investigated in our study. It is found in cognitive science and linguistics (Hudson, 1995; Hiranuma, 1999; Gibson, 1998, 2000; Temperley, 2007; Liu, 2008; Gildea and Temperley, 2010) that the dependency distance, that is, the linear distance between a word (dependent) and its head in a sentence can measure syntactic complexity and language processing difficulty. On one hand, the mean dependency distance (hereafter MDD) can be used to measure the parsing difficulty of different sentence structures, such as subject-extracted vs. object-extracted structures, and center-embedded vs. right-dependent sentences (Hsiao and Gibson, 2003; Liu, 2008). As a result of the pressure for least effort, in natural languages, there seems to exist a tendency to minimize the dependency distance in language production (Hudson, 1995; Gibson, 1998; 2000; Liu, 2008) and comprehension (Hawkins, 2004; Temperley, 2007). Properties of dependency relations, the type of relationships and their dependency direction are possible factors to influence the dependency distance. Jiang and Liu (2015) found that human languages have a tendency to minimize the dependency distance and that the constantly high percentage of adjacent dependencies is one of the most contributing factors. On the other hand, MDD may serve as a metric to compare the processing difficulty of different languages. Liu (2008) found that different languages have different MDDs, and Chinese is assumed as difficult to process because of the length of its MDD. Recently, MDD has been used in **the syntactic and cognitive studies on code-switching**. By measuring the dependency distance of German-English monolingual and code-mixed discourse corpus, Eppler (2010) found that longer distance facilitates the code-switching and proposed the Distance Hypothesis:"code-switching is more likely in syntactic relations with long dependency distance" (Eppler, 2011:145).

Another syntactic metric—the dependency direction—can be used as an indicator in language typology (Liu, 2010). The dependency direction is "head-initial" when the head precedes the dependent, and "head-final" when the head follows the dependent. Whether the syntactic properties of the head or the dependent determine the word order of the dependencies is of great interest to many linguists. Mahootian (1993), MacSwan (1999) suggest that heads determine the order of their complements, and Chan (2008) proposed that the lexical heads may not determine the word order of their complements in code-switching, but functional categories must. Findings on complements push researchers to probe further how word order is determined in mixed dependencies involving adjuncts, which will be pursued in this paper.

Based on a Chinese-English code-mixed treebank, Wang and Liu (2013) reported that, compared with monolingual Chinese and English corpora, the mixed corpus presents many syntactic differences such as differences in dependency distance and dependency direction. There are two types of dependencies in the mixed corpus: monolingual and mixed dependencies, and mixed dependencies present longer dependency distances than monolingual ones. Major grammatical relations (subject, object, attributive and adverbial) contribute to the differences in dependency distances. It is the distributions of major grammatical relations with different dependency directions in monolingual and mixed dependencies that cause the word-order difference.

Grammatical relations reveal the syntactic features of code-switching. According to Word Grammar, "syntactic structure is the network of dependencies between words which satisfies their mutual needs", and "syntax is the area of grammar which holds all the information about dependency needs" (Hudson, 2010: 154). A dependency, generally a traditional syntactic relation, is an unequal relation between the dependent and its head. Analysis into major syntactic relations between words from the same language and/or different languages may display more clearly the syntactic differences caused by Chinese-English code-switching. Based on a Chinese-English code-mixed treebank of 19 766 word tokens, this paper investigates adverbial and attributive dependency relations, because firstly, adverbials (18.44%) and attributives (12.04%) constitute the largest proportions of all syntactic dependency relations; secondly, compared with subject and object, adverbial and attributive are more syntactically complicated in Chinese and English, both of which are SVO languages. In brief, this paper mainly aims to answer 3 questions: