What if Chinese had Linguistic Markers for Counterfactual Conditionals? Language and Thought Revisited

Gary Feng (garyfeng@duke.edu)
Department of Psychology: Social and Health Sciences, Box 90085,
Duke University, Durham, NC 27708, USA

Li Yi (ly3@duke.edu)
Department of Psychology: Social and Health Sciences, Box 90085,
Duke University, Durham, NC 27708, USA

Abstract

This study critically examines the widely held belief that the Chinese language provides no consistent linguistic forms for expressing counterfactual ideas (Bloom, 1981). In Study 1, native speakers of Chinese identified counterfactual sentences from a large corpus of Chinese texts. A number of syntactic and/or lexical forms are found to be highly predictive of counterfactuality. Aspect modification and other markers emerge as potentially universal linguistic mechanisms for counterfactual marking. Study 2 demonstrated that these linguistic markers significantly influence readers' interpretation of whether a sentence is factual or counterfactual, even after semantic and pragmatic information was controlled. Study 3 showed that the counterfactual markers influence on-line sentence comprehension in a self-paced reading paradigm. Taken together, the data reject Bloom's conjuncture regarding Chinese counterfactuals and reveal cross-linguistic similarities and differences in linguistic representation of counterfactual thoughts.

Introduction

Counterfactual (CF) thinking, as hard as it might sound, is a ubiquitous and essential part of everyday life. For instance, regrets – wishing something hadn't had happened – are demonstratively automatic, irrepressible responses to unfavorable outcomes. Other examples include children's pretend play, understanding false beliefs, lying, and movie watching. While in some of these cases CF thinking is arguably non-linguistic, language may mediate CF thoughts and is unmistakably involved whenever a CF idea is shared among people.

Communicating a CF idea, however, can be tricky, because it involves two pieces of information: (a) the message that describes a state of affair and (b) the fact that it is false. Language should do a good job for (a), but how to convey the meta-message regarding the fictional nature of the message can be an interesting engineering problem. Sometimes the context tells it all, as when we watch an impersonation of President Bush. However, it seems more economical achieve both (a) and (b) with language, particularly if speakers of the language routinely communicate CF ideas. And once accepted and formalized, these linguistic markings of CF may in turn facilitate counterfactual reasoning.

Bloom (1981, 1984) took this idea a step further and entertained the possibility of a language that:

> … has no distinct lexical, grammatical, or intonational device to signal entry into the counterfactual realm, to indicate explicitly that the events referred to have definitely not occurred and are being discussed for the purpose only of exploring the might-have-been or the might-be. [Bloom, 1981, p.16]

Bloom further reasoned that because speaker of that language

> "… have not been led by their language to construct schemas specific to counterfactual speech and thought … they would typically do so [counterfactual reasoning] less directly, with a greater investment of cognitive effort and hence less naturally than their English-speaking counterparts. [p. 22]"

Bloom identified Chinese as a living example of such a language, and conducted experiments to show deficiencies in counterfactual reasoning among Chinese speakers. Bloom's linguistic determinism was met with strong criticisms from linguists and philosophers. Empirical support for Bloom's claim is weak at the best. With the exception of Bloom's original studies (1981; 1984), most studies reported similar performances between Chinese- and English-speakers in counterfactual reasoning tasks (Au, 1983; Liu, 1985; Hsu et al., 2004; Wu, 1994; Yeh & Gentner, 2005). Despite all these criticisms, Bloom's work captured the imagination of the public and scholars alike, in part because no one can explain how Chinese speakers could reason counterfactually without a language that supports it.

The present study considers a different hypothetical – that the Chinese language has consistent linguistic markers for counterfactuals that are comparable to those in English. If true, it explains the equal performance of Chinese- and English-speakers. It also shifts our attentions from simple language differences to potentially universal principles in linguistic expressions of thoughts.

Marking counterfactuals

Can a language communicate counterfactual thoughts without ANY consistent linguistic marking? In other words, is it possible to encode and decode the meta-message (that the message is CF) linguistically without agreeing on a common code book? Here we define *linguistic marking* as linguistic forms that are highly predictive of a CF interpretation. Expressed in conditional probability, the probability of a sentence being CF is much higher in the presence of this marker then not. It should be noted that we are not arguing for logically necessary or sufficient markers of CF, nor do we think they exist. From the cognitive processing perspective, the high conditional probability serves as a signal for the listener to disregard the literal interpretation of the sentence and to commit to a CF one. The decision making, however, is based on an interaction between the CF markers and other variables such as semantics and contexts.

We argue that the answer to the above question is generally *No*, with some potential exceptions. The reason CF marking has to be explicit and consistent is that the meta-message is in general not available in the message itself and cannot be computed from the linguistic message. Try stripping out the tense and aspect marking from the following sentence:

> If you had done your homework, you wouldn't have failed the class.

The bare bones of the sentence, "if you do your homework, you do not fail the class," tell nothing about the missed homework and its consequence. This is a crucial difference between CFs and metaphors. Although both require non-literal interpretations, metaphors can often be understood via featural comparisons or other local computations, and thus can function without explicit linguistic markings.

There are, however, apparent exceptions, as in "if he has money, I am Bill Gates." The absurdity, and thus the CF interpretation, of the above sentence only arise from the shared knowledge that the speaker is (probably) not Bill Gates. Conceptually, marking the CF-ness with intentional violations of semantic knowledge is an example of linguistic marking.

In principle, any linguistic devices (e.g., intonation) can potentially be exploited to mark CF, as long as the listener and speaker agree on a common codec. The English expression "if pigs fly" and the Chinese equivalent "unless the sun comes from the west" strongly signal counterfactuality because of violations of common semantics. Violating discourse/pragmatic principles is another way of marking, as in Bloom's study, "Everyone knows X is not Y. But let's suppose X is Y, then …" There are obvious limitations with semantic or discourse level markings, and they are not discussed here in the interest of space.

Two arguably more effective strategies are lexical and syntactic marking. Compared to the neutral "if", words such as "suppose" or "pretend" are strong invitations for CF thinking. English, as well as many other languages, primarily use syntax to differentiate regular "Open" conditionals and CF conditionals. In English, the subjunctive mood of the main verb strongly suggests a CF reading. Special structures such as "had A, then B" provide unequivocal evidence for a CF interpretation. Lexical and syntactic marking does not depend on shared semantic or pragmatic knowledge and thus can be used with much greater precision and to address a broader audience.

According to Bloom, lexical and syntactic marking is not available in the Chinese language. Bloom (1981) suggested that Chinese speakers could reason counterfactually, but only by engaging in an odd style of argument ("Everyone knows X is not Y. But let's suppose X is Y, then …") that is unwieldy and imprecise.

The elusive Chinese CF markers

Bloom's assertion was based more on intuitions than serious linguistic research. In fact, despite the popularity of the topic, we were only able to find a handful of linguistic inquires on this topic (e.g., Chao, 1968; Jiang, 2000; Wu, 1999). Bloom's conjecture may "feel right" to naïve Chinese speakers because such discussions almost certainly arise in the context of comparing Chinese and English. Being an isolate language, Chinese has essentially no verb inflections that correspond to the "subjunctive mood" in English. Focusing exclusively on syntactic marking in Chinese is a biased approach.

Yeh and Gentner (2005) distinguished semantically "transparent" and "non-transparent" counterfactuals. The former contains direct contradictions with the listener's world knowledge (e.g., "if pigs can fly…"), whereas the truth value of the non-transparent ones is ambiguous from the sentence itself. The authors imply that counterfactuals can be reliably marked by exploiting common grounding, and world knowledge.

There are also attempts to identify lexical markers of CF in Chinese. Chao speculated that the many "if-words" in Chinese may serve to signal different degrees of certainty, and some could be used exclusively for counterfactual propositions (Chao, 1968). It is interesting to note that Bloom himself used word equivalents of "suppose" instead of the plain "if". Chao's claim, however, appeared to be based on intuitions in classic Chinese rather than modern, vernacular Chinese. Recently, Jiang (2000) conducted an insightful linguistic analysis of potential markers of counterfactuality in Chinese, only to reject all of them on the logical ground that each form could potentially be used in non-counterfactual contexts. In doing so, however, Jiang (2000) might have raised the bar too high. By the same criterion, English would not have any consistent linguistic markers for counterfactuals because the usual marking

device, the subjunctive mood, is also used in other situations, too (Lycan, 2001).

In an interesting study by Wu (1994), she asked native Chinese speakers to identify counterfactual sentences from newspapers and looked for recurring linguistic forms in the sample. Although she was able to identify some interesting patterns, these are not markers by our definition because Wu's approach yields the probability of using a linguistic form if the sentence is counterfactual, whereas we argue the conditional probability in the opposite direction should be used.

Linguistic marking and cognitive performance

To summarize, we hypothesize that linguistic marking of CF has to be explicit and consistent in order to communicate CF ideas reliably. This marking may happened at different levels of linguistic analysis, including but not limited to lexical, syntactic, semantic, and discourse/pragmatic levels. In addition, we define linguistic markers with a statistical analogy, i.e., linguistic forms that predict the CF interpretation. These markers play an important role in CF sentence processing, but only in conjunction with other sources of information.

We report three studies to test these hypotheses. Study 1 aims to establish that there are consistent linguistic markers for CF in Chinese. To this end, we searched through a large corpus of Chinese sentences and asked native Chinese speakers to classify them into Open conditional and CF conditions. It is shown that there are distinct lexical and syntactic forms in Chinese that strongly predict the counterfactual interpretation of a sentence. Study 2 controls semantic/discourse factors and demonstrates that these markers are still informative. Together, Studies 1 and 2 strongly refute Bloom's (1981) basic premise that Chinese does not provide linguistic signals of counterfactuality.

Study 3 asks a different question – if linguistic markers are strongly predictive of a counterfactual interpretation of the sentence, do readers engage in a different process as soon as they see the markers? A self-paced reading task was used to investigate the time course of marker processing.

Study 1: Corpus Study

The purpose of this study is to identify linguistic markers of Chinese counterfactuals. We define a marker as a consistent linguistic form that strongly predicts a counterfactual interpretation of the sentence.

Methods

The study was based on a list of potential counterfactual markers in Chinese compiled from Wu (1994) and other prior research. They fell in several categories, see Table 1.

Using those markers as keywords, a naïve Chinese native speaker was asked to find 200 sentences containing those markers, either from an open-access Chinese text database (Peking University) or from the internet search engines. Two other native Chinese speakers were asked to judge whether each sentence was a counterfactual or an open conditional one. Extensive training was provided prior to the work so that the coders were confident and consistent in making the categorization.

Findings and discussion

Inter-coder reliability

The inter-coder reliability was calculated for the two raters. They agreed with each other 86% of the time. After controlling the agreements that would occur by chance, they agreed with each other 73% of the time (Krippendorff's alpha=0.73).

Strength of the markers

The percentage of sentences judged as counterfactual (out of 200) was calculated for each marker. Table 1 shows the average percentages of the two raters.

Table 1: Percentages of counterfactual responses for each marker

category	marker	% CF
Temporal reference	早(early)	83%
Aspect marker	了 (perfect/perfective marker)	21%
Negators:	要不是 (had it not been the case)	91%
	没 (did not)	14%
	要不然 (had it not been the case)	43%
Predicates	就好了(would have been great if only)	55%
	还以为… (had thought)	91%
	原来应该 (should have been)	92%
Others	…的话 (in the case)	9%
	真的 (really)	10%

Several observations are in order. (1) A number of lexical and syntactic markers are strongly associated with the CF reading of sentences (over 90%). We do not have the baseline statistics of the overall percentage of CF sentences in Chinese, making it hard to interpret the ones with lower scores. (2) As the gloss English translations suggest, each of the markers are used in different and limited contexts. We did not find a structure as productive as the English subjunctive mood. (3) Except for the first two rows, the rest markers are words or lexicalized phrases. (4) The "temporal reference" and "aspect marker" categories (Wu, 1994) are the only syntactic markers we were able to identify. They are productive and are often combined with other markers and/or semantics to mark the CF-ness. For instance, 了 is not a strong marker by itself, but it is often obligatory at sentence end to change the aspect to perfective.

Study 2: Sentence Comprehension

In real language use, semantic and linguistic cues work together. Yeh and Gentner (2005) suggest that Chinese readers have trouble with non-transparent counterfactuals that do not violate semantic knowledge. Study 2 examines whether those counterfactual markers identified in Study 1 are informative above and beyond contextual effects.

Methods

Participants

The English portion of the study was conducted in the United States. Thirty native-speaking undergraduate students at Duke University participated in the English part of the study, where they received course credit for participating. The Chinese data were collected in Beijing, China. Thirty paid undergraduate students at Peking University participated.

Materials

48 sentence frames were made so that the antecedents and the consequents had no logical connection. In other words, knowing the antecedent tells nothing about the consequent.. There were 4 conditions for English material: open conditional in present tense, counterfactual in present tense, open conditional in past tense, and counterfactual in past tense. Sentences are the same across conditions except for the tense and CF markers.

Similarly, unpredictive sentence frames were created in Chinese and potential markers are inserted in the appropriate places in the sentence. There were 6 conditions for Chinese material: open conditional, open conditional with marker ("吧/ba") at the end, counterfactual with aspect marker ("了/le"), counterfactual with temporal reference, counterfactual with negator 1 ("要不是/yao4 bu4 shi4"), and counterfactual with negator 2 ("没/mei2"). Sentences were the same across conditions except for the markers.

Procedure

Participants were asked to read a sentence, and then judge whether a statement was true or false based on their understanding of the sentence they read before. They identified the likelihood of the statements being true on a scale of 0-100%..

There was an additional task in the Chinese test. In this forced choice task, participants were asked to read a dialog with a blank in it, and then to choose one sentence from a pair of sentences to fill in the blank. There were 2 kinds of settings of the dialog: open setting where the correct answer is an open conditional, and counterfactual setting where the correct answer is a counterfactual sentence. The pair of sentences always consisted of an open conditional and a counterfactual conditional.

Results

English: Counterfactual rating

Figure 1 shows the percentage of English sentences judged to be counterfactual for each condition. Having past tense subjunctive is a strong signal that the sentence should be read as a counterfactual. However, the data on present tense subjunctive (such as "If Michael did not play Basketball, he would play tennis") is ambiguous. One possibility is that the syntactic form for present tense subjunctive is similar to that of the simple past tense and readers might have confused the two. Alternatively, like Chinese speakers, English readers may need both the semantic cue and the subjunctive mood cue to trigger a counterfactual reading.

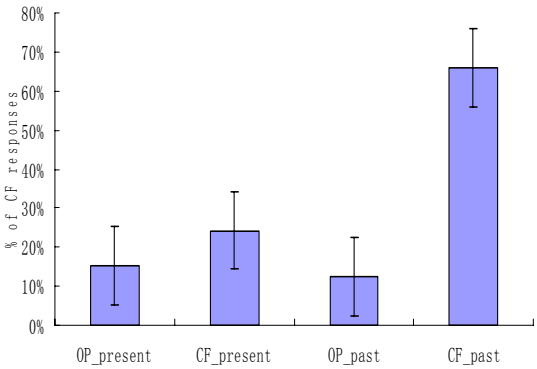

Figure 1: Percentage of judging English sentences as counterfactual in each condition in Study 2

Chinese: Counterfactual rating

Figure 2 shows the percentages of counterfactual responses for each condition. The data are based on identical sentence frames that only differ by the markers. Even after semantic information was controlled, lexical (negator NG1) and syntactic (Aspect Marker (AM) and Temporal Reference (TR)) markers significantly increased readers' chance to read them as counterfactuals.

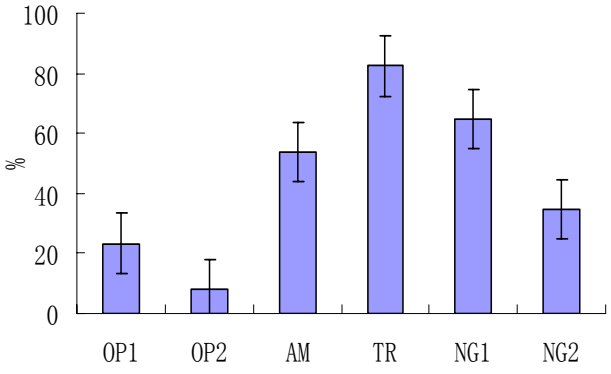

Figure 2: Percentage of judging Chinese sentences as counterfactual in each condition in Study 2

Chinese: Forced-choice

As a measure of the strength of the marker, the Phi coefficients were calculated for the data from this task. Phi is derived from Chi-square and is a measure of the associations between settings and sentence types. Table 2 shows phi values of each pair of comparison. Consistent with other tasks, having a lexical marker (the negators) or a syntactic marker (Aspect marker (AM) or Temporal reference (TR)) in the sentence increase the readers' CF interpretation of the sentence.

Table 2: Phi values of each pair of comparisons

comparisons	phi values
OP2-AM	0.89**
OP1-TR	0.74**
OP1-NG1	0.53**
OP1-NG2	0.31*

Note: *p< .001, *p< .0001.

Study 3: Self-paced Reading

Study 3 investigates the time course of the effect of counterfactual markers on on-line sentence comprehension. It is predicted that there is an increased cognitive load in doing counterfactual readings reasoning, and it will take more time.

The processing load increase may begin as soon as the reader identifies the marker for counterfactual conditionals. The immediacy of the effect of the marker should be a function of the predictive strength of each marker, particularly in the Chinese study. On the other hand, a reader may wait until the end of the sentence to integrate information.

A caveat with the self-paced reading paradigm is that it may not be sensitive enough to identify a strategy change in reading, and participants' normal reading processes may be altered by the one-word-at-a-time reading method.

Methods

Participants

Thirty undergraduate students at Duke University participated in the English part of the study. The Chinese study was conducted in Beijing, China. Thirty undergraduate students at Peking University participated in the Chinese part of the study. They received payment for participating.

Materials & procedure

The same English and Chinese material in Study 2 was used in Study 3. Using a self-paced reading paradigm, Chinese and English speakers were asked to read sentences word-by-word on a computer screen. Reading time of each word was recorded. Participants were asked to read a sentence with self-paced reading paradigm, and then judge whether a following statement was true or false according to the meaning of that sentence.

Results

English: Mean word reading time

Figure 3 shows the mean reading time of each word in English sentences in each condition. ANOVA showed that English speakers' mean word reading time was significantly longer for sentences in counterfactual condition in past tense than in open condition in present tense ($F(3) = 6.10$, $p < 0.05$).

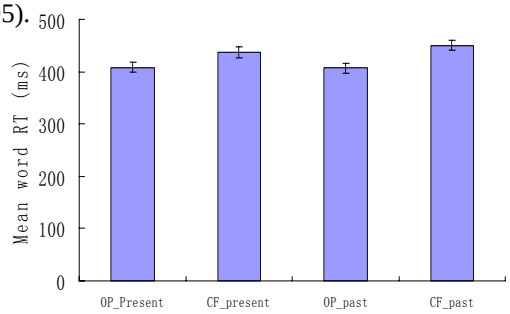

Figure 3. Mean reading time of each word in English sentences by condition

English: Reading time by region

English sentences were divided into several regions, and the reading time per word was averaged by region for each condition. Figure 4 shows the mean reading time by region in English sentences in each condition. ANOVA did not show any significant differences.

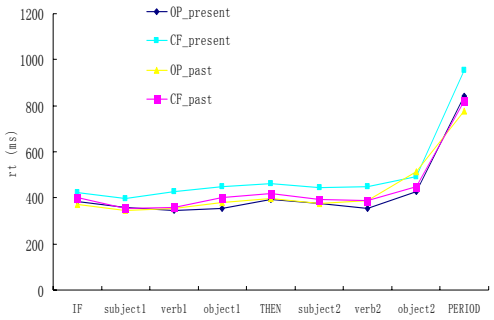

Figure 4. Mean reading time by region for English sentences in each condition

Chinese: Mean word reading time

Figure 5 shows the mean reading time of each word in Chinese sentences in each condition. ANOVA F1 analysis showed there is no difference with mean word reading time across conditions in Chinese.

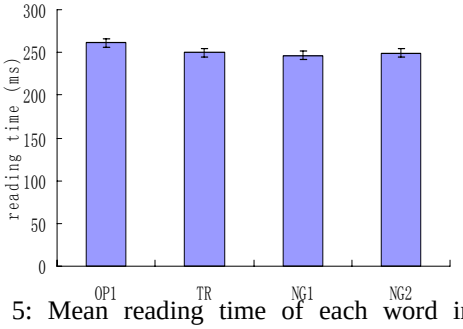

Figure 5: Mean reading time of each word in Chinese sentences by condition

Chinese: Reading time by region

Chinese sentences were divided into several regions, and the reading time was averaged by region for each condition. Figure 6 shows the mean reading time by region in Chinese sentences in each condition. ANOVA F1 showed that the time spent at the end of the sentence was significantly longer for sentences in the counterfactual condition with temporal reference than in the open condition ($F(1, 29)=10.05$ $p<0.005$).

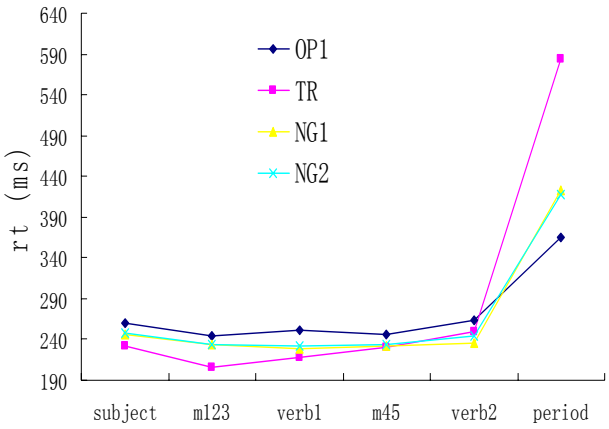

Figure 6. Mean reading time by region for Chinese sentences in each condition

General Discussion

Several key findings emerge form the three studies reported here. First, it is possible to identify linguistic forms (lexical or syntactic) that reliably predict a counterfactual reading of a Chinese sentence. These forms are shown to be informative in naturally occurring sentences in which they work with contextual cues to highlight counterfactuality. They also independently signal a counterfactual reading when contextual information is controlled. Contrary to Bloom's (1981) assertion, the Chinese language provides both lexical and syntactical devices to mark counterfactuals.

This nicely explains the puzzling finding in the past two decades that Chinese speakers' counterfactual reasoning ability is on par with that of English speakers. No substantial differences should be expected if both languages mark counterfactuality in the language itself. For Bloom (1981), our finding suggests that the logic of the study was false, and the cross-language differences he reported were most likely due to translation and other technical problems, as pointed out by Au (1983). The criticism applies equally to any study that accepted the false premise by Bloom.

Secondly, the study began to uncover some potential language-universals. For example, few people would have guessed before this study that Chinese uses the temporal (tense) and aspect markers to signify a counterfactual interpretation. The analogy with the English subjunctive mood – which modifies the tense and aspect of the main verbs – is obvious. More research, particularly cross-linguistic, is needed to identify linguistic and cognitive universals.

Last but not least, it is clear that counterfactual markers affect how people understand counterfactual conditionals. When there is no other useful information, readers of Chinese can rely solely on the linguistic markers to solve the problem. Message from the self-paced reading stud is less clear-cut, but overall reading time was longer, and for the Chinese a significant sentence-end wrap up effect is observed. We are now in the process of conducting eye movement experiments, where reader can move their eyes at will and re-read sentences if necessary. The eye movement technology is expect to yield rich information about the time course of counterfactual processing, particularly the immediacy of the marker effect.

Acknowledgments

This study was partially supported by a summer research grant from the Asian/Pacific Studies Institute (APSI) at Duke University. The authors also wish to acknowledge Dr. Xiaolin Zhou and Dr. Xiangzhi Meng for assistance with conducting the experiments in China.

References

Au, T. K (1983) Chinese and English counterfactuals: The Sapir-Whorf hypothesis revisited. Cognition, 15, 155-187.

Au, T. K. (1984) Counterfactuals: In reply to Alfred Bloom,*Cognition*, 17, 289-302.

Bloom, A. H. (1981) *The Linguistic Shaping of Thought: A Study in the Impact of Language on Thinking in China and the West.* Hillsdale. Erlbaum Associates.

Chao, Yuen Ren. (1968). *A Grammar of Spoken Chinese*. Berkeley & Los Angeles: University of California Press.

Jiang, Y. (2000). Counterfactual interpretations of Chinese conditionals. *[Studies and Explorations on Syntax (Chinese)]*, 10, 257-279.

Liu, L. G (1985). Reasoning counterfactually in Chinese: Are there any obstacles? *Cognition*, 21, 239-270.

Yeh, D. & Gentner, D. (2005). Reasoning counterfactually in Chinese: Picking up the pieces. *Proceedings of the Twenty-seventh Annual Meeting of the Cognitive Science Society*, pp. 2410-2415.

Wu (1994) *If Triangle Were Circles,…" – A Study of Counterfactuals in Chinese and English.* The Crane Publishing Co., Ltd, TW.