

# Expressing conditionality in Mandarin

## A corpus-based study of *rúguǒ* and *zhǐyào*

Weiyang Chen and Jacqueline Evers-Vermeul

Zhejiang University, China / Utrecht University, The Netherlands

After a broad overview of Mandarin Chinese conditionality marking, this paper presents a corpus-based analysis of two conditional connectives, *rúguǒ* and *zhǐyào* (both translatable as ‘if’), from a syntactic and a cognitive perspective. We examine their use in narrative and informative texts along four parameters: clause order, position of the connective within the clause, domain, and counterfactuality. For all parameters, the two connectives displayed robust profiles across genres. Both connectives preferred an antecedent-consequent clause order. They displayed flexibility in their position, behaving like adverbs, with *rúguǒ* showing a stronger preference for the pre-subject position than *zhǐyào*. In terms of domains, *zhǐyào* has a stronger preference for content conditionals than *rúguǒ*, which is also frequently used in the epistemic domain. In our data, only *rúguǒ* was used meta-metaphorically and in counterfactuals. We argue that both connectives can be translated with ‘if’, but *zhǐyào* also matches ‘so/as long as’.

**Keywords:** conditionals, conjunctions, counterfactuals, mental spaces, domain

### 1. Introduction

Conditionality, as exemplified in the *if-then*-clauses in (1) and (2), is very frequent in everyday language (Austin 1961/1979; Haiman 1978; Stalnaker 1981; Comrie 1986; van der Auwera 1986; Sweetser 1990; Wierzbicka 1997; Dancygier 1998; Dancygier & Sweetser 2005; Gauker 2005; Ippolito 2013; Douven 2016). It reflects the ability to reason about alternative possibilities, make inferences, imagine possible correlations between situations, and understand how the world would change if certain correlations were different. Conditional reasoning is a central part of human thinking, because much of our knowledge is conditional (Johnson-Laird & Byrne 2002, 2010). Wierzbicka (1997) claims that the IF-relation is a conceptual primitive. Understanding exactly how conditionals work “provides basic insights

into the cognitive processes, linguistic competence, and inferential strategies of human beings” (Ferguson, Reilly, ter Meulen & Traugott 1986, 3).

- (1) If it rains, they’ll cancel the game. (Dancygier & Sweetser 2005, 34)
- (2) If you are hungry, there are biscuits on the sideboard.<sup>1</sup>  
(Dancygier & Sweetser 2005, 40)

Conditionals such as (1) are the most common type of conditionals, and are known as hypothetical or indicative conditionals: the *if*-clause expresses the antecedent that will lead to the event expressed in the consequent. The conditional in (2) is called a ‘biscuit conditional’ (Siegel 2006) or ‘relevance conditional’ (Bhatt and Pancheva 2005): the *if*-clause applies to the illocutionary act performed in uttering the main clause, rather than to its propositional content.

It is often claimed that the prominence of conditionality is reflected in the fact that many, if not all, languages have lexical and/or grammatical resources to mark conditionality (Comrie 1986; Ferguson et al. 1986; Wierzbicka 1997). Languages can mark conditionals with syntactic structures such as word order. For instance, the conditionality of both the English example in (3) and the Dutch example in (4) is indicated by V1, the finite verb in first position; by contrast, regular constative clauses show subject-verb-object in English, and verb second in Dutch. Alternatively, languages may use tense to mark counterfactual conditionals, for example the past tense in sentence (5). Moreover, languages employ specific conjunctions such as *if* in English in (6) and *als* ‘if’ in Dutch in (7).

- (3) Had he been ill, he would not have gone to the party.
- (4) Is hij ziek, dan gaat hij niet naar het feest.  
is he sick then go he not to the party  
“If he is ill, then he will not go to the party.”
- (5) *If* I were you, I would not go to the party.
- (6) *If* he is ill, he will not go to the party.
- (7) *Als* hij ziek is, dan gaat hij niet naar het feest.  
if he sick is then go he not to the party  
“If he is ill, he will not go to the party.”

In the first part of this paper (Section 2) we will present an overview of the ways in which Mandarin Chinese marks conditionality. This issue has been addressed

1. Austin’s example is: “There are biscuits on the sideboard if you want them” (1961/1979, 210). From this, it can be inferred that ‘There are biscuits on the sideboard whether you want them or not’, and that anyway ‘There are biscuits on the sideboard.’

before, but information is scattered over a variety of publications. In the second part of the paper (Sections 3 to 6), we will zoom in on two of the most frequent conditionality markers in Chinese: *rúguǒ* ‘if’ and *zhíyào* ‘if, so/as long as’. Because it is not clear whether and how these two markers differ exactly, we have conducted a corpus-based study, in which we examine the use of *rúguǒ* and *zhíyào* in narrative and informative contexts from a syntactic and a cognitive perspective. The paper ends with a conclusion in Section 7.

## 2. Expressing conditionality in Chinese

Like English *if*, German *wenn* and *falls*, and Dutch *als*, Chinese has specific linguistic means to mark conditionality (Chao 1968, 1976; Li & Thompson 1981; Wang et al. 1994; Cheng & Huang 1996; Chierchia 2000; Xing 2001; Shen 2003, 2008; Bhatt & Pancheva 2005; Wang 2010a, 2010b). In the following subsections, we will first introduce a variety of conditional markers in spoken Chinese (Section 2.1), and illustrate that Chinese has the linguistic means to introduce unconditionals (*no matter what/whether*), necessary conditionals (*only if*), as well as sufficient conditionals (*if*) (Section 2.2). Then, we will zoom in on other linguistic means for marking conditionality (Section 2.3), and on the marking of counterfactuals (Section 2.4).

### 2.1 Conditionals in spoken Chinese

Chao (1968, 116) lists a variety of conjunctions in spoken Chinese that can all be translated with ‘if’: *yàoshi*, *jiǎrú*, *ruòshi* (monosyllabic synonyms *rú* and *ruò* limited to Classical Chinese), *tǎngruò*, *jiǎruò*, *jiǎshǐ*, *tǎngshǐ*, and *shèruò* in approximately descending order of frequency of occurrence. See some examples in (8)–(10), taken from Chao (1968, 104 and 116).

- (8) nǐ *yàoshi* kànbùqǐ            tā, tā jiù kànbùqǐ            nǐ  
 you if      look down upon 3SG 3SG then look down upon you  
 “If you look down upon him, he will look down upon you.”
- (9) *yàoshi* nǐ búhuì,      wǒ kěyǐ jiāo nǐ  
 if      you not know I      can teach you  
 “If you don’t know how, I can teach you.”

- (10) *tāng ruò nǐ zǎo jǐng gào le tāmen, nà cì chū shì huò zhě jiù néng*  
 if you early warn PAR 3PL that accident perhaps then can  
*bì miǎn le*  
 avoid PAR  
 “If you had warned them early, that accident could perhaps have been avoided.”

After checking the frequency of occurrence of Chao’s ‘if’-word list in the Corpus Query System developed under State Language Commission, we find that the order accords with his observation, except for the pair *jiǎ ruò* and *jiǎ shǐ* and the two most frequent conjunctions in terms of total frequency, *rú guǒ* and *zhǐ yào*, which were not mentioned in Chao’s list.<sup>2</sup>

**Table 1.** Frequency of conditional conjunctions in the Corpus Query System

Conjunction	Total frequency of use	Frequency of conjunction use
<i>rú guǒ</i>	12340	12331
<i>zhǐ yào</i>	3965	435
<i>yào shì</i>	1818	1809
<i>jiǎ rú</i>	730	670
<i>ruò shì</i>	347	326
<i>tāng ruò</i>	286	286
<i>jiǎ ruò</i>	97	89
<i>jiǎ shǐ</i>	155	150
<i>tāng shǐ</i>	43	43
<i>shè ruò</i>	19	19

Chao (1968, 116) claims that the order from *yào shì* to *shè ruò* (in Table 1) also represents the likelihood of the supposition, so that the later ones are more apt to go with suppositions contrary to fact. However, there is doubt whether these expressions by themselves can distinguish the degree of likelihood (Wang 2010b).

## 2.2 Unconditionals, and necessary and sufficient conditionals in Chinese

Based on the type of conditionality, conjunctions can be categorized into markers for unconditionals (*no matter what/whether*), necessary conditionals (*only if*), and sufficient conditionals (*if*). All three kinds of conditionals can be demonstrated

2. <http://202.114.40.175:8080/cqs/index.gsp>

by the italicized expressions in (11), taken from Mao Zedong's speech entitled *Serving the People*.

- (11) *yīnwèi wǒmen shì wèi rénmin fúwù de, suǒyì, wǒmen rúguǒ yǒu quēdiǎn, jiù bú pà bierén pīpíng zhǐchū. Bùguǎn shì shénme rén, shéi xiàng wǒmen zhǐchū dōu xíng. Zhǐyào nǐ shuō de duì, wǒmen jiù gǎizhèng. Nǐ shuō de bànfǎ duì rénmin yǒu hǎochù, wǒmen jiù zhào nǐ de bàn.*  
 “If we have shortcomings, then we are not afraid to have them pointed out and to be criticized, because we serve the people. Anyone, no matter who, may point out our shortcomings. As long as your criticism is right, we will correct them. If what you propose benefits the people, then we will act upon it.”  
 (*Serving the People*, by Mao Zedong, 1944)

For unconditionals, there are *wúlùn* ‘no matter whether/what’ and *búlùn* ‘regardless of’ with a quantifier-like particle *dōu* ‘all’ in the consequent clause, as illustrated in (12) and (13) respectively.<sup>3</sup> Necessary conditionals are connected with *zhǐyǒu... cái* ‘only if... then’, as in (14). The antecedent *p* necessitates the consequent *q*, and is the sole necessary condition of *q*; there is no consequent without the antecedent. As a result of Occidental influence, a relatively new word *chúfēi* ‘unless’ can also be used to express a necessary condition, as in (15). What looks or sounds like a necessary condition is commonly understood as also sufficient (Chao 1976, 257). According to Xing (2001), necessary conditionals are harsher and less tolerant than sufficient conditionals.

- (12) *wúlùn shì dàiyī xuéshēng, háishì dàsì xuéshēng, dōu yīng*  
 no matter whether be freshmen or seniors all should  
*rènzhēn xuéxí*  
 hard study  
 “No matter whether you are freshmen, or seniors, you should study hard.”
- (13) *búlùn xīlà dáchéng hé zhǒng xiéyì, dōu jiāng yǒulì*  
 regardless of Greece reach what kind agreement all will benefit  
*ōuyuán*  
 Euro  
 “Regardless of what agreement Greece will reach, it will benefit Euro.”

3. Unless stated otherwise, quoted examples are from the corpus studied in this paper, and developed by the Center for Chinese Linguistics of Peking University (CCL), taken from [http://ccl.pku.edu.cn:8080/ccl\\_corpus/](http://ccl.pku.edu.cn:8080/ccl_corpus/).

- (14) *zhīyǒu niánmǎn shíbā suì, cái yǒu gōngmín xuǎnjǔquán*  
 only if age eighteen years then get citizen franchise  
 “Only if people reach the age of eighteen, then they can be granted franchise.”
- (15) *chúfēi nǐ qù, fǒuzé wǒ bú qù*  
 unless you go otherwise I not go  
 “Unless you go, I won’t go.”

Chinese also has connectives expressing sufficient conditionality. This includes the two markers that are central to this paper. In (16), the conjunction *jiǎrú* ‘if’ marks the antecedent in the first clause.

- (16) *jiǎrú xiàyǔ, wǒmen jiù zài wūli chī fàn*  
 if rain we then at indoors eat food  
 “If it rains, we’ll eat indoors.” (Li & Thompson 1981, 632)

### 2.3 Other linguistic markers expressing conditionality

Example (16) displays a case of double marking of conditionality: a conjunction *jiǎrú* ‘if’ marks the antecedent in the first clause, and an adverb *jiù* ‘then’ marks the consequent in the second clause. In this instance of double marking, the connectives work in pairs to mark both the protasis (*p*) or antecedent, and the apodosis (*q*) or consequent, just like English *if...then...* However, in Chinese it is also common to mark just one of the two clauses. In (17), there is no *jiù* in the consequent when another adverb *huì* ‘will’ fills the position. In (18), there is no conjunction in the antecedent, but a *jiù* in the consequent. Chao (1976, 256) mentions that very frequently Chinese let the *if-then* relation be gathered from *jiù* ‘then’.

- (17) *nǐ rúguǒ gēn le wǒ, jiānglái huì chī hěn duō kǔ de*  
 you rúguǒ follow PAR I future will eat very many hardship PAR  
 “If you marry me, you will have many hardships!” (CCL)
- (18) *nǐ dǎdiànhuà gěi ta, wǒ jiù búyòng xiě xìn le*  
 you telephone to 3SG I then not need write letter PAR  
 “If you telephone to him, I won’t need to write.” (Chao 1968, 116)

If used at all, English *then* normally precedes the subject (Chao 1968, 114). The consequent marker *jiù*, however, never precedes the subject when it co-occurs with antecedent markers, as is illustrated in (16) and (18).<sup>4</sup> Indicating a sufficient

4. A sentence like *tāmen dōushì běifāng rén, jiù wǒ shì nánfāng rén* ‘They are all Northerners, only I am a Southerner’ (Tian 2006, 20) is a different case where *jiù* ‘only’ is a typical adverb, restricting the range or scope.

condition, it must occur after subjects, or before predicates in Chinese clauses without an overt subject. The restricted position of the adverb *jiù* matches its part of speech, modifying the predicate and staying closer to the predicate (see Tian 2006 for a detailed account of the origin of *jiù* and its historical development).

A conditional clause can occur without an ‘if’-word by merely having negators such as *bù* ‘not’ in one or both clauses, as in (19), or with no syntactic change at all, as in (20).<sup>5</sup>

- (19) nǐ *bù* lái(,) wǒ *bú* qù  
 you not come I not go  
 “If you don’t come, I don’t go.” (Chao 1968, 116)

- (20) bàba qù, wǒ gēn tā qù  
 father go I follow 3SG go  
 “If father goes, I’ll go with him.” (Li & Thompson 1981, 633)

Li and Thompson (1981) claim that this kind of clausal dependence is established by the speaker’s intention. In such cases, the type of coherence relation between the two clauses is not signaled explicitly and must be inferred by the hearer from his/her knowledge of the situation and of what has been said to that point (1981, 641).

Amazed by the latter type of examples provided by his colleagues, Comrie (1986) regarded Mandarin Chinese as a typological exception that does not have a formally identifiable syntactic construction whose basic function is to encode conditionals. He discusses sentence (21) in which no word order, connective, or verb tense indicates conditionality or any other coherence relation.

- (21) zhāngsān hē jiǔ, wǒ mà tā  
 zhangsan drink wine I scold 3SG  
 “Zhangsan drinks wine. I scold him.” (Comrie 1986, 82)

Comrie claims that the relation between the two clauses can be interpreted as temporal, causal or conditional, i.e. ‘When/Because/If Zhangsan drinks wine, I scold him.’ From the ambiguity in the interpretation of the two sentences in simple parataxis, he concludes that Chinese has no clear prototypical conditional construction: although there are some particles translatable as ‘if’, most conditionals are in principle

5. Bare conditionals with *wh*-binding have received some attention in study (Cheng & Huang 1996; Chierchia 2000; Bhatt & Pancheva 2005). Bare conditionals are characterized by the presence of one or more *shéi* ‘who’ (*wh*-word in Chinese) in the antecedent clause, which have to be matched by an equal *shéi* in the consequent. The adverb *jiù* ‘then’ may be optionally present in the consequent clause. e.g. *shéi xiān jìnlái, wǒ xiān dā shéi* ‘who first enter, I first hit who’ meaning ‘If X enters first, I hit X first’. Bhatt and Pancheva (2005, 680) claim that it is the syntactic dependency between the operator and the variables that it binds that make bare conditionals conditional. This so-called bare conditional with *wh*-binding is not within the scope of our study.

ambiguous and are interpreted as conditional only from the context. Of his three interpretations, we judge that the most prominent reading is the temporal one. Causal is a possible alternative, but the conditional interpretation is only marginal.<sup>6</sup>

However, we agree with Comrie that, on his view of “extreme context dependence”, though the relationship between sentences is not made explicit, the context allows the speakers to juxtapose sentences without confusing listeners. In contrast to English that has an elaborate system of overt grammatical categories, Chinese is “a context-dependent language”, as Shen argues in his recent work (2016, 160). Potential ambiguities are usually resolved by either the linguistic or the situational context.

Spoken Chinese displays another marker of conditionality. As (22) illustrates, a clitic *de huà* ‘the matter of, in the event that’ can be used at the end of the antecedent clause. The prevalence of *de huà* and its compatibility with conditional conjunctions can be explained by observing that the “condition expressions serve as topics in Chinese” (Chao 1968, 85). Chao further suggests that a conditional clause can be regarded as a clause subject. When it co-occurs with a conditional conjunction, its role as a topic marker is evident, introducing a topic to comment on. Therefore, after most conditional clauses *de huà* can be added, as in (23). When *de huà* occurs at the sentence-final position, it often indicates an afterthought, as in (24).

- (22) nǐ yǒu qián      *de huà*,      jiù bù huì xiàng wǒ jiè      qián le  
 you have money in the event that then not will from I    borrow money PAR  
 “If you had money, you wouldn’t have to borrow money from me.”  
 (Li & Thompson 1981, 634)

- (23) yàoshi bùkěn      *de huà*, nà jiù suàn le  
 if    unwilling in the event that that then let go PAR  
 “If (it is a matter of) his not willing, then let it go.”      (Chao 1968, 118)

- (24) qián      béng cún      le, yàoshi jiù      yào yòng *de huà*  
 money not    deposit PAR if      soon need use    a question of  
 “The money does not need to be deposited, if it’s to be used right away.”  
 (Chao 1968, 133)

We should like to say, in concluding this section, that Chinese may not have one favored or prototypical conditional construction, but it does have various means

6. Conditional and temporal relations are sometimes quite similar (cf. the German connective *wenn* that can mark both types of relation). The same perfective particle *-le* can be used in both conditional and temporal relations between clauses. Thus when an aged parent says *Wǒ sile sāngshì cóngjiǎn* ‘When I die, the funeral should be simple’, it is a time relation; while if a young husband says *Wǒ sile nǐ dīnghǎo zài jià* ‘If I die, you’d better marry again’, it is a conditional (Chao 1968, 117).



to mark conditionality. This is in line with Wierzbicka's (1997, 25) claim that "IF is a universal human concept lexicalized in all languages."

#### 2.4 Counterfactuals in Chinese

There has been some discussion whether Chinese makes a distinction between conditionals and counterfactuals. Conditionals allow us to imagine that something might happen that we think can happen (e.g., *If I drink coffee, I won't sleep well*), whereas counterfactuals allow us to imagine that something might happen that we think cannot or know did not happen (e.g., *If I had drunk coffee, I wouldn't have slept well*). Bloom (1981) points out the absence of counterfactuals in Chinese, and hypothesizes a cognitive difference between English and Chinese speakers with regard to counterfactual thinking. Similarly, Comrie (1986) claims that Chinese displays no distinction in the degree of hypotheticality. It will be seen that their conclusions are mainly based on the fact that Chinese does not have tense marking.

Li and Thompson (1981, 647) mention that in a Chinese conversation, addressees infer the exact type of conditional message "from the proposition in the second clause, and from their knowledge of the world, and of the context in which the sentence is being used." Chinese does not mark counterfactuality with backshifting in tense like *if* (V-past, or V-pluperfect) in English. English [*if...will*]-sentences and [*if* (V-past)...*would*]-sentences are translated into Chinese with no indication of tense difference unless there is an explicit time adverbial (e.g. tomorrow, in the future, at that time, earlier, then). Comrie gave two examples, (25) and (26), and noted that the same Chinese sentence in (27) could be a translation of either of these sentences in English.

(25) *If you kiss me, I will buy you a beer.*

(26) *If you kissed me, I would buy you a beer.*

(27) *rúguǒ nǐ qīn wǒ, wǒ jiù mǎi jiǔ gěi nǐ*  
 if you kiss I I then buy beer to you  
 "If you kiss/kissed me, I will/would buy you a beer." (Comrie 1986, 90)

From the context-free Chinese version in (27) we cannot tell whether it expresses real possibility or counterfactuality. But a lack of tense marking does not mean a lack of distinction between real possibility and counterfactuality. As to the ways to express counterfactual thinking, Chinese is no exception. It has other linguistic means for "counterfactual talk and thought", and counterfactual meaning can be clearly conveyed.

Chen (1988) shows that counterfactual messages are indicated in Chinese by five types of linguistic devices. For example, counterfactuality can be derived if the sentence includes a temporal reference, or time adverbial, such as *zǎo* ‘early’ in (28), *nàshíhòu* ‘at that time’, or *zuótiān* ‘yesterday’. Connectives *yàobúshì*, *ruòbúshì* ‘if not’, *rúguǒ búshì* ‘if not’, or a compound *búshì* ‘not’, all containing a negator *bù* ‘not’ can also express counterfactuality, as (29) illustrates for *yàobúshì*. The particle *le* in (30) indicates counterfactuality as well. Chou (2000) mentions a similar list of Chinese counterfactuality markers, but adds the rhetorical interrogative. This construction is unique to Chinese and is shown in (31).

- (28) *rúguǒ nǐ zǎo chūfā, jiù búhuì wù le huǒchē*  
 if you early leave then will not miss PAR train  
 “If you had left half an hour earlier, you would not have missed the train.”
- (29) *yàobúshì zánmen ràng tā dāng le jīnglǐ, fàndiàn de shēngyì*  
 if not we let 3SG become PAR manager restaurant PAR business  
*jiù búhuì zhème hónguǒ*  
 then will not so good  
 “If he had not been promoted to be the manager, the restaurant business would not have been so good.” (CCL)
- (30) *nǐ qù le, shìqíng jiù búhuì dào zhège dìbù le*  
 you go PAR thing then will not reach this situation PAR  
 “If you had been there, it would not have been like this.”
- (31) *yàobúshì wǒ, tā huì yǒu jīntiān?*  
 if not I 3SG can have today  
 “If it were not for me, could he be what he is now?” (Chou 2000, 62)

To conclude, speakers of English and Chinese are equally able to reason counterfactually in their native languages (Au 1983, 1984). Chinese is no exception to Wierzbicka’s (1997) claims that all languages have a class of “counterfactuals” and that IF...WOULD is a universal human concept, lexicalized and grammaticalized in all languages, like IF.

### 3. *Rúguǒ* versus *zhǐyào*

Given the variety of Chinese conditional markers, it is interesting to examine whether seemingly similar markers actually display similar syntactic distributions and semantic profiles or not. In this paper, we therefore zoom in on *rúguǒ* and *zhǐyào*, the conjunctions with the highest total frequency of use (12,340 and 3,965 respectively, see Table 1).

Logically, both *rúguǒ* and *zhǐyào* mark a sufficient condition in the antecedent, indicating that when *p* occurs, *q* occurs. In the constructed example in (32), ‘raining’ is a sufficient, but not a necessary condition for the ground to get wet; without rain, the ground may also get wet. Similarly, *rúguǒ* in (33) indicates that ‘raining’ is a sufficient condition for the game to be cancelled. Without rain, the game may also be cancelled, for other reasons.

(32) *zhǐyào xiàyǔ, dì jiù huì shī*  
 if rain ground then will wet  
 “If it rains, the ground will get wet.”

(33) *rúguǒ míngtiān xiàyǔ, bǐsài jiù qǔxiāo*  
 if tomorrow rain game then cancel  
 “If it rains tomorrow, the game will be cancelled.”

In spite of the fact that both *rúguǒ* and *zhǐyào* mark a sufficient condition, there is as yet no general agreement as to what are the exact similarities and differences between these two markers. Xing (2001, 106–114) devotes a section to a discussion of these two conjunctions. Here, we will review previous literature in order to find whether there are differences between these markers in terms of morphology and syntax (Section 3.1) and/or in terms of semantics (Section 3.2).

### 3.1 Morphology and syntax of *rúguǒ* and *zhǐyào*

*Rúguǒ* is the most commonly used protasis (*p*) marker, occurring 12,331 times as a conditional conjunction. *Rúguǒ* is consistent in this profile, as its conjunction use makes up 99.9% of the total frequency of use. As Table 1 shows, the conjunction use of *zhǐyào* takes up only about 11% of its total frequency (435 out of 3,965 instances). This difference between the two markers can be related to the fact that *rúguǒ* has a higher degree of grammaticalization than *zhǐyào*. As a combination of *rú* ‘like’ and *guǒ* ‘result’, the sufficient marker did not come into use in *Dream of the Red Chamber* (1791), but later developed into a fixed grammaticalized element.

*Zhǐyào* arose out of the combination *zhǐ* ‘only’ and *yào* ‘want/need’. Unlike *rúguǒ*, *zhǐyào* has retained its original use next to its conjunction use, and has not completely lost the meaning or autonomy of each individual morpheme. For example, in (34) *zhǐyào* functions as a verb phrase. Putting the conjunction *rúguǒ* in this position would result in an ungrammatical clause, as is illustrated in (35), because it would result in clause where *wǒ* ‘I’ is not accompanied by a verb.

(34) *wǒ zhǐyào nǐ tóngyì jiù hǎo le*  
 I only want you agree then good PAR  
 “I only want you to agree. Then it will be good.”

- (35) \*wǒ *rúguǒ* nǐ tóngyì jiù hǎo le  
 I if you agree then good PAR  
 \*‘I if you agree. Then it will be good.’

Note that the ungrammaticality in (35) is not due to the fact that *rúguǒ* occurs after the subject, as both *rúguǒ* and *zhǐyào* can appear before or after the clausal subject, no matter whether the subjects in the *p* or *q* are the same or different. Compare, for instance, (36) and (37), which are both grammatical. The two syntactic positions (i.e., before and after the subject) cannot always be distinguished, given that Chinese is a pro-drop language and therefore allows its subject to be ‘left out’, as in (38).

- (36) *rúguǒ* nǐ zài nàlǐ, jiù gènghǎo le  
 if you be there then better PAR  
 ‘If you are there, it will be better.’
- (37) nǐ *rúguǒ* zài nàlǐ, jiù gènghǎo le  
 you if be there then better PAR  
 ‘If you are there, it will be better.’
- (38) *rúguǒ* zài nàlǐ, jiù gènghǎo le  
 if be there then better PAR  
 ‘If (someone) is there, it will be better.’

The variation in position makes researchers wonder about the syntactic nature of Chinese conjunctions:

Chinese conjunctions are hardly distinguishable from prepositions or adverbs. The conjunction-like words occupy the typical position of adverbs, namely, between the subject and the verb. In fact, the status of the Chinese conjunction is so uncertain that Dragunov did not even recognize it as a separate class.

(Chao 1968, 790)

According to Wang et al. (1994), the difference in position might result in a difference in scope of the conjunction. For example, in (39), *zhǐyào* can restrict *nǐ* ‘you’, *dào Shànghǎi* ‘to Shanghai’, or *qù yí tàng* ‘go once’, depending on the stress of the tone.

- (39) *zhǐyào* nǐ dào Shànghǎi qù yí tàng, zhè jiàn shì jiù jiějué le  
 if you to Shanghai go one CL this CL thing then solve PAR  
 ‘If you go to Shanghai once, this problem will be solved.’
- (40) nǐ *zhǐyào* dào Shànghǎi qù yí tàng, zhè jiàn shì jiù jiějué le  
 you if to Shanghai go one CL this CL thing then solve PAR  
 ‘If you go to Shanghai once, this problem will be solved.’

(Wang et al. 1994, 103)

If the stress is on *nǐ*, it means “we don’t need anybody else”; only “you” can realize the consequent of “solving the problem”. If the stress is on *dào Shànghǎi*, it means “no other place is needed”; only “Shanghai” meets the condition and causes the consequent. If the stress is on *qù yí tàng*, it means “only once is enough” and “there is no need to go again”. However in (40), *zhǐyào* follows *nǐ*, so *nǐ* is not within the scope of the conjunction, and only the last two interpretations remain possible. In written text, the ambiguity is hardly perceived. Therefore, the English translations of (39) and (40) do not vary in spite of the position change of *zhǐyào*.<sup>7</sup>

The position of *rúguǒ* and *zhǐyào* can be alternated in yet another way, namely by adjusting the order of the clauses that are connected by the conjunction. Normally, the antecedent *p* precedes the consequent *q* in conditionals. Conditions are like topics, and topics go before the comment (Chao 1968; Haiman 1978). But in some cases, this clause order can be reversed, as (41) illustrates for *zhǐyào*.

- (41) *nǐ bǎ tā dài zǒu ba, zhǐyào nǐ hǎo hǎo de dài tā*  
 you PAR 3SG take away EX if you good good PAR treat 3SG  
 “Take her away, if you treat her well!” (Thunderstorm, by Cao Yu)

This example is a non-initial conditional, where *p* follows *q* and the speaker makes a request “you take her away”. As Ford and Thompson (1986, 368) point out, when the speaker either proposes an action or makes a request in the main clause,

7. In spoken Chinese, connectives like *rúguǒ* and *zhǐyào* can even occur in clause-final position, as in (i) and (ii). This presents a sharp difference with English conjunctions, which only occur clause-initially. The clause-final position alters the nature of the sentence. For example, the Chinese sentence in (iii) is regarded as a complex sentence in which the condition acts as the subject, but the English translation is a compound sentence that includes a subordinate and a main clause.

- (i) *nǐ lái, zhǐyào*  
 you come only want  
 “Only if you come.”
- (ii) *nǐ néng lái de huà, rúguǒ*  
 you can come in the event of if  
 “If you can come (it will be good).”
- (iii) *éluósī de jiǔ rúguǒ xiàng zhōngguó shìchǎng shàng de jiǔ yīyàng*  
 Russia PAR alcohol if resemble China market on PAR alcohol same  
*duō jiù hǎo le*  
 many then good PAR  
 “If alcohol in Russia were as readily available as in the Chinese market, it would be good.”

non-initial conditionals can express the speaker's respect for, or deference to, the authority of the interlocutor.

### 3.2 Semantics of *rúguǒ* and *zhǐyào*

In the West, the 'if'-conditional is the core of many studies on conditionality. Chinese studies put forward *rúguǒ* as the most prominent marker, but they label in a slightly different way to introduce semantic distinctions between markers: instead of using the term *conditional* as a cover term for all kinds of 'if'-relations, the Chinese literature makes a distinction between the subtypes *hypothetical*, *conditional* and *counterfactual* markers, where it seems that the subtype *conditional* refers to those relations that indicate a high degree of certainty that the consequent will happen.

Before the 1950s, linguists saw no need to distinguish between *zhǐyào* and *rúguǒ*, and both were treated as markers of the category of hypotheticals (Lü 1942; Wang 1943, 1944). After the 1950s, there has been a tendency to classify *zhǐyào* into a different category as a conditional rather than a hypothetical marker (Wang et al. 1994; Xing 2001). Both hypotheticals and conditionals express a causal relationship to be realized (Wang 2012, 122). Despite their logical resemblance, they are not completely interchangeable. If we interpret Xing (2001, 106–114) correctly, the most important difference between the markers has to do with the hypothetical nature of the antecedent: as (42) illustrates, *zhǐyào* cannot be replaced by *rúguǒ* when the hypothesized condition is factual. Similarly, Xing (2001) thinks *wǒ xīshēng* 'I die' in (43) is a pure hypothesis, and can therefore only go with *rúguǒ*, not *zhǐyào*. With *zhǐyào* in the sentence, the event in the consequent is highly likely to happen.

(42) *zhǐyào*/\**rúguǒ* wǒ yǒu yì kǒu qì zài, cǐ chóu bùnéng bú bào  
if I have one CL breath have this hatred cannot not revenge  
“If I live, revenge must be taken.” (Xing 2001, 109)

(43) *rúguǒ*/\**zhǐyào* wǒ xīshēng le, qǐng bǎ zhè fèn cáiliào zhuǎnjiāo gěi  
if I die PAR please PAR this CL document forward to  
shàngjí  
superior  
“If I die, please pass on the document to my superior.” (Xing 2001, 108)

However, this classification has been disputed. Wang (2012) mentions the lack of evidence for the distinction between hypotheticals and conditionals, and points out that the subcategory of hypotheticals is not consistent with the overall system of classifying complex sentences: conditional relations, like causal and contrastive

ones, concern the relationship between two clauses, but hypotheticals concern the relation between what is stated in the antecedent clause and reality. Wang (2012) therefore proposes to treat *rúguǒ* as a typical conditional (i.e., in the general sense of the term).

Still, there are other differences between *rúguǒ* and *zhǐyào*. For instance, Xing (2001, 113) argues that the coverage, field or scope of *rúguǒ* is bigger than that of *zhǐyào*, which is why only the former can occur in (44) and (45).

- (44) *rúguǒ* shuō yào yǒu wèntí, zhè cái shì zuì gēnběn de wèntí  
 if say want exist problem this then be most fundamental PAR problem  
 “If there is a problem, this is the most fundamental problem.”  
 (Xing 2001, 113)

- (45) *rúguǒ* tā bù kěn lái, wǒ gāi zěnmē bàn ne  
 if 3SG not willing come I shall what do EX-Q  
 “If he is not willing to come, what shall I do?”  
 (Xing 2001, 113)

It seems hard to pinpoint the exact difference between *rúguǒ* and *zhǐyào*. According to Wang (2012), Chinese grammarians focus on logic and semantic dimensions to classify conditionals, while their Western colleagues use formal, semantic and cognitive dimensions. An exception is the work of Xing (2001), who lists six types of conditionals from a functional perspective: inference, reaction, question, imperative, comment and verification. However, this list does not seem useful for discriminating between *rúguǒ* and *zhǐyào*, because some of them seem to focus on form instead of function (cf. question, imperative), while others overlap in function (e.g. reaction and comment).

Wang (2012) advocates the collection of empirical data and the extension with a cognitive approach, stating that Dancygier and Sweetser’s (2005) work on conditionals might facilitate the study of Chinese conditionals. Except for preliminary work done by Shen (2003, 2008), this recommendation has not been put to practice yet. In the remainder of this paper, we will therefore add to the body of evidence on *rúguǒ* and *zhǐyào*, by presenting the results of a corpus-based study on the syntax and semantics of the two markers. We will look at these markers from a cognitive perspective, using the categories proposed by Dancygier and Sweetser (2005). Before we present the methodology (Section 5) and results of our corpus-based study (Section 6), we will introduce these cognitive categories, illustrating their applicability in the study of Chinese conditionals (Section 4).

#### 4. A cognitive approach to conditionals

A more sophisticated theoretical framework is needed to further explore the uses of Chinese conditionals. Working with English conditionals, Dancygier and Sweetser claim that the job of *if* in a conditional construction is to prompt the set-up of a mental space. They consider Mental Spaces Theory as an “economical and elegant” way to “attribute some of the functional diversity to a few specific parameters of interpretation” (2005: 16). Within cognitive linguistics, conditionality is a way of mapping human cognition and the construal of conditional relations between events onto language. According to them, conditionals can function in six domains, with the content, epistemic, and speech-act domains being the most commonly found in English conditionals. Sweetser (1990) elucidates the functioning of conditional (and other) connectives in these three domains. Dancygier and Sweetser (2005) supplement three meta-type domains: metalinguistic, meta-metaphorical and meta-spatial. In the next subsections, we will discuss these six domains in more detail, and illustrate how they can be applied to Chinese conditionals.

##### 4.1 Content domain

An example of a content conditional is shown in (46). The speaker is using the *if*-clause to set up a mental space wherein he ties his handkerchief on the cut, and then predicts that it will stick to the cut in that envisioned situation. In other words, the speaker is talking about a possible attempt at bandaging, and the conditional construction marks his representation of a contingent relationship between that portrayed event and its predicted result (Dancygier & Sweetser 2005). A Chinese example of a content conditional from our corpus study is provided in (47).

(46) If I tie my handkerchief on it [a cut], it'll stick.

(Dancygier & Sweetser 2005, 16)

(47) *zhìyào zhuānxīnzhìzhì, tòng xià gōngfú, jiānchíbúxiè de nǚli, jiù*  
*if concentrated work hard persevere PAR hard then*  
*yídìng huì yǒu shōuhuò*  
*definitely will have harvest*

“If you are concentrated, work hard and persevere in your efforts, you will definitely succeed.”



## 4.2 Epistemic domain

In an epistemic conditional, reasoning can go from effect to likely cause, which is a frequent case, and vice versa: from cause to likely effect, in which it follows the direction of a content causal contingency (Dancygier and Sweetser 2005). In (48), this epistemic reasoning goes from effect to cause, and can be paraphrased as “If I know the fact in *p*, then I conclude *q*.” In other words, the speaker’s knowledge that the typing happened is a precondition for her conclusion about the loving. In the Chinese Example (49), the knowledge of his understanding the unspoken is a precondition for the speaker’s conclusion about the sensitivity.

(48) If he typed her thesis, he loves her. (Dancygier & Sweetser 2005, 117)

(49) *zhǐyào tā tīng míngbái le, huò shì kàntòu yányǔ zhōng de*  
 if 3SG listen clear PAR or be see through conversation in PAR  
*ànshì, tā de sīxiǎng shì xīnì de*  
 implication 3SG PAR mind be sensitive PAR  
 “If he understood it, or got what was implied from the conversation, he was sensitive.”

## 4.3 Speech-act domain

Speech-act conditionals have been noticed as a special and interesting class for some time (van der Auwera 1986; Sweetser 1990). The speaker can use an *if*-clause to set up a speech-act space, and then utter a speech act that is taken to be effective within that space. For example, the biscuit conditional in (2) can be paraphrased as “if you are hungry as stated in *p*, then let us consider that I perform this speech act of offering by saying *q*.” An example of a Chinese speech-act conditional is provided in (50). Xing’s (2001) example in (45) can also be classified as a speech-act conditional.

(50) *zhǐyào nǐ gǎn xìngqù, wǒ jiù jiǎng gè gùshi*  
 if you feel interest I then tell CL story  
 “If you are interested, I will tell you a story.” (Shen 2008, 121)

## 4.4 Metalinguistic domain

Metalinguage is a language for talking about the object language (Johnson-Laird & Byrne, 2002, 653). This definition helps us to see a metalinguistic space as one consisting of a pairing between a content space and a language or code space, as in (51). On the basis of the correlation between language and labels, the speaker uses the language being spoken as a basis for predicting the choice of labels.

Metalinguistic clauses focus on a chosen expression, often by repeating it, as in (52), in which *yíchǎn* ‘legacy’ is repeated in *jīngshén yíchǎn* ‘spiritual legacy’. Similarly, Xing’s (2001) example in (44) can be classified as a metalinguistic conditional, in which *wèntí* ‘problem’ is repeated in *gēnběn de wèntí* ‘fundamental problem’.

- (51) If we were speaking Spanish, he would be your uncle.  
(Dancygier & Sweetser 2005, 127)

- (52) *rúguǒ* (shuō) tā cóng fùqīn nàlì jìchéng le shénme yíchǎn, nàjiùshì  
if (say) 3SG from father there inherit PAR what legacy then be  
jìchéng le wèi kēxué ér xiànshēn de jīngshén yíchǎn  
inherit PAR for science thus sacrifice PAR spirit legacy  
“If (we say) she has inherited any legacy from her father, then she has  
inherited the spiritual legacy to sacrifice for the sake of science.”  
(Shen 2008, 120)

#### 4.5 Meta-metaphorical domain

Meta-metaphorical conditionals express a relationship between metaphorical mappings. For instance, (53) establishes and develops a metaphorical relationship between two domains: bridges and horses. It can be read as “If you accept the possibility of metaphorically talking about bridges as if they were horses by agreeing to call the Golden Gate the thoroughbred of bridges, you will also accept calling the Bay Bridge a workhorse” (Dancygier and Sweetser 2005, 132). A similar metaphorical example in Chinese is given in (54).

- (53) If the beautiful Golden Gate is the thoroughbred of bridges, the Bay Bridge is the workhorse.  
(Dancygier & Sweetser 2005, 132)

- (54) *rúguǒ gōngzuò shì lèqù, nàme rénshēng jiù shì tiāntáng*  
if work be joy then life then be paradise  
“If work is joy, then life is paradise.”  
(CCL)

#### 4.6 Metaspatial domain

In metaspatial conditionals, the protasis seems to be setting up a background mental space that does not fit smoothly into one of the categories mentioned so far. The example in (55) might be paraphrased as “if your parents would give a name like Utah to your sister, then I’d like to know what name they would give you.” The Chinese example in (56) is also a metaspatial conditional. According to

Shen (2008), the relationship between you and me in the real world is projected to the relationship between the cowboy and the loom girl in the legend.<sup>8</sup>

- (55) “Utah’s my sister. She can do better than me when I was her age, but my form is improving. I’m definitely better than Rhiannon. Want to see?”  
 “Not today,” I said. “If Utah’s your sister, are you Wyoming or Nevada?”  
 (Dancygier & Sweetser 2005, 137)

- (56) *rúguǒ nǐ shì niúláng, wǒ jiù shì zhīnǚ*  
 if you be cowboy I then be loomgirl  
 “If you are the cowboy, then I am the loom girl.” (Shen 2008, 120)

## 5. A corpus-based analysis of *rúguǒ* and *zhīyào*

In Section 4, we have shown that a cognitive approach to Chinese conditionals seems promising. So far, however, we have only supplied anecdotal examples to illustrate the relevance of the six domains for the Chinese language. Such a qualitative approach is valuable in itself, because the categories for classification must first be identified before they can be quantified (Schmied 1993; McEnery & Wilson 2001). However, inferential statistics are necessary to enable researchers to generalize conclusions beyond sporadic observations (Núñez 2007; McEnery & Hardie 2012).

In this paper, we therefore examine the use of *rúguǒ* and *zhīyào* in a natural language corpus. Statistical analyses cannot help us find absolute differences between the two connectives (i.e. whether certain uses are grammatical or not), but they can facilitate discerning relative differences (e.g. whether these connectives have different preferences in terms of, for instance, syntactic position or domain).

### 5.1 Corpus description

The samples are collected from a corpus developed by the Center for Chinese Linguistics of Peking University (CCL, [http://ccl.pku.edu.cn:8080/ccl\\_corpus/](http://ccl.pku.edu.cn:8080/ccl_corpus/)). It is a raw corpus without segmentation and tagging, including both Classical Chinese and modern Chinese. CCL contains 477 million Chinese characters (1.06 GB) and offers a keyword-in context function for inspecting the context of a given keyword through its web interface.

8. Niulang and Zhinü are two legendary figures in Chinese folklore. They are a couple: the former is a cowboy, and the latter is a daughter from the Heavenly Emperor, named loom girl.

## 5.2 Sample

The data we collected are from the modern Chinese part of the corpus. We looked at written texts, in which the sentences are in deliberately planned form, the only exceptions being realistic dialogues in drama and fiction and texts of actual speech recorded for scientific or legal purposes. Two genres of the corpus were analyzed: narrative and informative, because previous research has found that the profile of conjunctions may show genre-sensitivity (Li, Evers-Vermeul & Sanders 2016).

For the narrative corpus, we searched the connectives in the category of novels, and got 1,557 fragments containing *rúguǒ*, and 1,488 fragments for *zhǐyào*. A random selection of 50 fragments was carried out at intervals of every 30 fragments. If fragments were incomplete or showed non-connective uses of the selected words (cf. the verbal use of *zhǐyào* discussed in Section 3.1), we excluded them and searched for the next suitable one.

For the informative fragments, we looked into the first 10,000 occurrences containing the two connectives out of the categories of encyclopedia and news reports. We applied the same standard for the inclusion of suitable fragments. For each connective, we picked 25 from each of these two categories. In total, 100 fragments were randomly selected for *rúguǒ* and *zhǐyào* respectively.

## 5.3 Analytical methods

In order to systematically measure the distribution and profile of both connectives, we tabulated all the conditionals in the selected fragments according to four parameters: clause order, position of the connective, domain, and counterfactuality.

### 5.3.1 *Clause order and position of the connective*

The antecedent *p* often precedes the consequent *q*, but, as we discussed in Section 3.1, this clause order can sometimes be reversed. We checked whether *rúguǒ* and *zhǐyào* display different profiles in their preference for the *p-q* order versus the reversed order.

In addition, we compared the flexibility of the position of the two connectives under investigation. In our analyses we distinguished three categories: with the conjunction preceding the subject (pre-subject, as in (39)), following the subject but before the predicate (pre-predicate, as in (40)), or in a sentence with no subject. The latter case makes it impossible to establish whether the conjunction is in pre-subject or pre-predicate position (see Example (38)).

### 5.3.2 Domain

The conditional relation expressed in each fragment was analyzed in terms of the six domains distinguished by Dancygier and Sweetser (2005): the content, epistemic, speech-act, metalinguistic, meta-metaphorical and metaspatial domains (see Section 4). Based on Sweetser's (1990) Domain Theory, a number of studies have been carried out on causal connectives in English, other European languages, and Chinese (e.g. Sanders & Sweetser 2009; Evers-Vermeul, Degand, Fagard & Mortier 2011; Li, Evers-Vermeul & Sanders 2013, 2016), but no systematic corpus-based investigation has been done so far to explore domain differences in the use of Chinese conditionals.

In line with the method used by Li et al. (2013), we devised a paraphrase test to increase the reliability of the domain analysis (see Table 2).

**Table 2.** Paraphrases used to distinguish the six domains\*

Domain	Paraphrase
Content	in the mental space set up by $p$ , the fact that $p$ leads to the fact that $q$
Epistemic	in the mental space set up by $p$ , the fact that $p$ leads to the conclusion/claim/ inference that $q$
Speech act	in the mental space set up by $p$ , the fact that $p$ leads to the question/advice/ command that $q$
Metalinguistic	there is a pairing between $p$ and $q$ involving a content space and a language space
Meta-metaphorical	there is a relationship between $p$ and $q$ involving metaphorical mapping
Metaspatial	there is a relationship between $p$ and $q$ involving setting up different types of space

\*  $p$  and  $q$  correspond to the propositions expressed in the antecedent and the consequent

In order to ensure the reliability of the domain analysis, we adopted the strategy of what Spooren and Degand (2010) call two-coders-discuss: two coders analyzed the corpus independently of each other, and afterwards discussed the differences. Parameters like clause order and position of the connective can easily be classified in an objective way, which is why an analysis by a single coder sufficed.

### 5.3.3 Counterfactuality

Of all the fragments that designated a content domain, we determined whether *rúguǒ* and *zhǐyào* expressed real possibility or counterfactuality. For example, from the context of (57), we could judge that there is a real possibility for the addressee to move the leg and get to work. Hence, (57) can be labeled as an indicative conditional. Similarly, we could infer that at the moment of the utterance in (58), *wǒ mǔqīn* 'my mother' did not know this yet, so this is not a real fact that

has already happened. The counterfactuality is expressed with a particle *le* and the connective *rúguǒ*.

(57) *zhìyào* nǐ kěn dòng yí dòng nǐ de tuǐ, nǐ búhuì bù fādá de  
 if you willing move one move you PAR leg you cannot not rich PAR  
 “If you are willing to work hard, you will become rich.” (CCL)

(58) wǒ mǔqīn *rúguǒ* zhīdào le zhè jiàn shì, tā yíding hèn wǒ  
 I mother if know PAR this CL thing 3SG definitely hate I  
 “If my mother knew this, she would hate me.” (*Thunderstorm*, by Cao Yu)

## 6. Results

To establish inter-rater agreement, we first looked at Cohen’s Kappa of the codings before discussion. The result of the two-coders-discuss was good (Cohen’s Kappa = .63). Then, we performed general loglinear analyses to find out whether the use of *rúguǒ* and *zhìyào* differed in terms of the four parameters: clause order (Section 6.1), position of the connectives (Section 6.2), domain (Section 6.3), and hypotheticality (Section 6.4). This type of statistical analysis allowed us to also check whether the connective profiles varied with genre.

### 6.1 Clause order

Table 3 shows the distribution of the connectives over the two clause orders.

**Table 3.** Clause order of *rúguǒ* and *zhìyào* in two genres

Genre	<i>rúguǒ</i> <i>p, q</i>	<i>q, rúguǒ</i> <i>p</i>	<i>zhìyào</i> <i>p, q</i>	<i>q, zhìyào</i> <i>p</i>	Total
Narrative	48	2	46	4	100
Informative	50	0	50	0	100
Total	98	2	96	4	200

Across genres, the connectives did not differ in their preferred clause order ( $\chi^2(1) = 1.35, p = .25$ ), as the overwhelming majority of the Chinese conditionals obeyed the topic-comment order (cf. Chao 1968; Haiman 1978): *p* first, *q* second. However, clause order itself displayed genre sensitivity ( $\chi^2(1) = 6.85, p = .009$ ): the reversed order (*q, p*) occurred with both connectives, but only in narratives. In the informative genre, we did not find any examples of reversed order either for *rúguǒ* or *zhìyào*. The sentences in (59) and (60) present examples of the reversed order for each of the connectives.

- (59) yǒude shíhòu yě juéde duì tā bù shífēn gōngjìng sìde, rúguǒ rénmen  
 some time too feel to 3SG not very respectful seem if people  
 jiào tā “Wáng Dé”  
 call 3SG “Wang De”  
 “Sometimes he felt not respected, if people called him *Wang De*.”  
 (Old Zhang’s Philosophy, by Lao She)
- (60) shénme wǒ dōu dāyìng nǐ。 zhǐyào nǐ zhēnde ài wǒ  
 whatever I all agree you as long as you really love I  
 “I can promise everything. As long as you love me truly”  
 (San jin chun qiu, by Feng Xiangguang)

In (60), a conditional remark is added as an afterthought, or a compromise for what is promised or predicted. In spoken language, it would be marked by a faster tempo in speech. In Europeanized Chinese, *if*-clauses after the principal clause are not limited to afterthoughts exclusively. But instances like (59) sound foreign in a written text without the change of tempo.

## 6.2 Position of the connectives

Table 4 presents the connective frequency per position.

Table 4. Position of *rúguǒ* and *zhǐyào* in two genres

position	<i>rúguǒ</i>		<i>zhǐyào</i>		Total
	narrative	informative	narrative	informative	
pre-subject	33	22	26	9	90
no subject	14	28	21	36	99
pre-predicate	3	0	3	5	11
Total	50	50	50	50	200

Although both connectives favor the pre-subject over the pre-predicate position, the connective *rúguǒ* has a stronger preference for the pre-subject position than *zhǐyào* ( $\chi^2(2) = 6.40, p = .04$ ). The latter connective occurs more often in clauses without a linguistically expressed subject (i.e. in clauses where it cannot be determined whether the connective precedes or follows the subject), and slightly more in clauses with the connective clearly in pre-predicate position. This positioning pattern of the connectives is consistent across genres, even though there is a main effect of genre ( $\chi^2(2) = 17.07, p < .001$ ): in informative texts, there are more clauses without an overt subject than in narrative texts. An example of *rúguǒ* in pre-predicate position can be seen in (61), while (62) presents an instance of

*zhǐyào* without an overt subject. The translation of (62) has to supply a subject as required by English syntax.

- (61) wǒmen *rúguǒ* róngrěn le zhèyàng de chǒulèi, háiyou shénme bùnéng  
 we if tolerate PAR such PAR ugliness else what cannot  
 róngrěn de  
 tolerate EX  
 “If we tolerate such ugliness, what else can’t we tolerate?”

(Bai Hui, by Zhang Wei)

- (62) *zhǐyào* dú diǎn shū, shí diǎn zì, jiù kěyǐ xiě wénzhāng  
 as long as read a little book know some word then can write article  
 “As long as one is literate and reads a little, he can write” (CCL)

### 6.3 Domain distribution

Table 5 displays the distribution over domains. First, we found a main effect of genre ( $\chi^2(3) = 23.54, p < .001$ ), caused by the fact that speech acts are more frequent in narrative texts. Second, there was a main effect of connective ( $\chi^2(3) = 12.01, p < .007$ ).<sup>9</sup> Although both connectives frequently occur in all three basic domains (i.e. content, epistemic, speech act), they both prefer the content domain, which expresses a causal relationship in the possible world. Across genres, however, the connective *zhǐyào* has a stronger preference for content conditionals than *rúguǒ*, which is also frequently used in the epistemic domain. Apparently, *zhǐyào* is preferred in describing a prediction or result based on the condition in the protasis material.

Table 5. Domain distribution of *rúguǒ* and *zhǐyào* in two genres

domain	<i>rúguǒ</i>		<i>zhǐyào</i>		Total
	narrative	informative	narrative	informative	
content	19	28	28	39	114
epistemic	14	17	9	9	49
speech act	17	3	13	2	35
metalinguistic	0	0	0	0	0
meta-metaphorical	0	2	0	0	2
metaspatial	0	0	0	0	0
Total	50	50	50	50	200

9. In the statistical analyses, we collapsed the three meta-types in order to avoid loss of power.



In our data, *rúguǒ* is the only connective that is used meta-metaphorically (see the two instances in (63) and (64)). However, from the small difference in frequency in the meta-metaphorical domain, we cannot tell whether this indicates an absolute or a relative distinction between *rúguǒ* and *zhǐyào*, because the meta-types of use occur hardly at all.

- (63) *rúguǒ* bǎ yǔzhòu bǐyù chéng wúbiānwújì de hǎiyáng de  
 if PAR universe compare to infinite PAR ocean in the event  
 huà, nàme, yínhéxì zhǐ shì dàhǎi zhōng de yí gè hěn xiǎo hěn xiǎo  
 that then galaxy only be ocean in PAR one CL very small very small  
 de xiǎo dǎo  
 PAR small island  
 “If the universe can be metaphorically compared to an infinite ocean, then the galaxy is only a tiny, tiny island in the ocean.” (CCL)

- (64) *rúguǒ* bǎ dìfāng xìqǔ hé guānzhòng bǐyù chéng yú hé shuǐ de  
 if PAR local opera and audience compare to fish and water PAR  
 guānxi, xìqǔ de xiànzhuàng yǒudiǎn hézhézhǐfù de  
 relationship opera PAR status quo somewhat fish in waterless ditch PAR  
 wèidào  
 feeling  
 “If the relationship between the local opera and audience is metaphorically compared to that between fish and water, the current situation of the local opera is like the fish in a waterless ditch.” (CCL)

If we replace *rúguǒ* in these examples with *zhǐyào*, the sentences would not sound as natural, because *zhǐyào* assumes a higher likelihood of the consequence to actually happen (see also Section 6.4). The connective *zhǐyào* is not compatible with the weak *de huà* ‘in the event that’ in (63) and the uncertain *yǒudiǎn* ‘somewhat, probably’ in (64). By contrast, *rúguǒ* is not constrained by these epistemic markers in the consequent *q*.

The absence of the metalinguistic and the metaspatial type of conditional in our corpus does not mean that these are never used in daily language (compare the *rúguǒ* examples in Section 4.4 and 4.6). It just happened that our relatively small corpus did not include such cases.

#### 6.4 Counterfactuality

Within the content domain, we further investigated whether the two markers under investigation expressed real possibility or counterfactuality (see Table 6).

Table 6. Counterfactuality of *rúguǒ* and *zhǐyào*

counterfactuality	<i>rúguǒ</i>		<i>zhǐyào</i>		Total
	narrative	informative	narrative	informative	
indicative	16	29	28	39	112
counterfactual	3	0	0	0	3
Total	19	29	28	39	115

As in other languages, counterfactual talk does occur in Chinese discourse, though the frequency in our data is quite low. The fact that all three instances occur in the narrative genre results in a main effect of genre ( $\chi^2(1) = 4.45, p = .03$ ). All three counterfactuals go with *rúguǒ*, which results in a main effect of connective ( $\chi^2(1) = 10.34, p = .001$ ). The three instances are given in (65)–(67).

- (65) *zhǐyǒu* mǎdá shēng shǐzhōngrúyī de hōnglōng hōnglōng xiǎng zhe,  
 only engine sound ceaselessly PAR roaring roaring sound ASP  
*rúguǒ* méiyǒu zhè shēngyīn, wǒ jiù gèng nánshòu le  
 if not this sound I then more upset PAR  
 “Only the engine was roaring all the time. If there had not been this sound, I  
 would have felt more uneasy.” (Good Luck, by Ye Nan)

- (66) tā shuō, “kuīle shì zài qù jiē nǐ de lùshàng chū de  
 3SG say fortunately be on go pick up you PAR way happen PAR  
 shì, *rúguǒ* shì zài huílái de lùshàng, lián nǐ yě dājìnqù le”  
 accident if be on back PAR way including you too involve PAR  
 “He said: “Thank God the accident happened on the way to pick you up. If it  
 had happened after we picked you up, you would have been involved in the  
 accident too.” (Sunset on the River Hudson, by Tian Xiaofei)

- (67) tā zuì tònghèn zìjǐ de yòu shǒu. Zhè zhī shǒu *rúguǒ* zǎodiǎn kǔnbǎng  
 3SG most hate elf PAR right hand this CL hand if earlier bind  
 yīxià yěxǔ jiù méiyǒu hòulái de guài shì le  
 done perhaps then not later PAR weird thing PAR  
 “He hated his own right hand most. If this hand had been bound earlier,  
 maybe the strange thing would not have happened.” (Bai Hui, by Zhang Wei)

In expressing counterfactuality, *zhǐyào* sounds infelicitous. At least in our corpus, there is no case of *zhǐyào* expressing counterfactuality. This is in line with previous observations that *zhǐyào* carries a stronger tone and more powerfully emphasizes the likely realization of the consequent based on the antecedent, while *rúguǒ* is not so affirmative in its tone (cf. Wang et al. 1994; Xing 2001). The less determinate *rúguǒ* can go together with *yěxǔ* ‘maybe’, as (67) indicates, but *zhǐyào* cannot

(see (68)). The marker *zhǐyào*, however does go well with *yídìng* ‘definitely’, *bìdìng* ‘necessarily’, and *bì* ‘definitely’, as (69) illustrates for *yídìng*.

- (68) *#zhǐyào zǎodiǎn chūfā, yěxǔ jiù búhuì chídào*  
 as long as early leave maybe then will not late  
 “#As long as you start early, maybe you won’t be late.”
- (69) *zhǐyào zǎodiǎn chūfā, jiù yídìng búhuì chídào*  
 as long as early leave then definitely will not late  
 “As long as you start early, you definitely won’t be late.”

A further look at the ending tone of the consequent affirms the varying strength of the tone they carry: the ending tone of *zhǐyào* is often exclamatory, as in (70). Contrarily, *rúguǒ* sounds much weaker in such an expression of strong determination. Likewise, the question marker *ne*, which often goes together with *rúguǒ*, does not match the ending tone of *zhǐyào*, as (71) shows. If we use the English ‘if’ to translate *rúguǒ*, we might, at this point turn to ‘so/as long as’ for an exact equivalent of *zhǐyào*.

- (70) *zhǐyào/?rúguǒ néng qīnzì kàn shàng yì yǎn, sǐ yě bì yǎn le*  
 if can in person look up one eye die too close eye EX  
 “If I could see it once, I would rather die!”
- (71) *rúguǒ /\*zhǐyào wǒ dédào le Xiùxiu de xiāoxi, wǒ gāi zěnme bàn*  
 if / as long as I get PAR Xiuxiu PAR news I should what do  
 ne  
 EX-Q  
 “What should I do if /\*as long as I heard from Xiuxiu?”

Native speakers can immediately decide that *zhǐyào* does not fit in the context of (71). We do not mean that *zhǐyào* cannot occur in any question consequent (cf. a rhetorical question like *Zhǐyào shì fāguāng de dōu shì jīnzi ma?* ‘Are those that glitter all gold?’), but it definitely cannot be used in a *ne*-ending question that expresses a doubt. If we use the two connective clauses in isolation, that is, if we leave out the consequent clause and only keep the antecedent as in (72) and (73), this incompatibility of *zhǐyào* and *ne* is more evident. Both connective clauses can be used independently without the consequent, but one expresses an ‘only if’ wish, and the other expresses a ‘what if’ doubt. This is why (72) and (73) need different translations for the connectives.

- (72) *zhǐyào néng qīnzì kàn shàng yì yǎn*  
 Only if I can see it once!
- (73) *rúguǒ wǒ dédào le Xiùxiu de xiāoxi ne*  
 What if I heard from Xiuxiu?

## 7. Conclusion

The current paper set out to present an overview of conditional marking in Mandarin Chinese. On the basis of a literature review, we were able to show that – despite the absence of tense marking – Chinese has a variety of markers to indicate unconditionals, necessary conditionals, sufficient conditionals, and counterfactuals. Hence, Chinese is no exception to Wierzbicka's (1997, 25) claims that IF is a universal human concept lexicalized in all languages and that all languages have a class of counterfactuals.

The literature review has also shown that the exact profiles and distribution preferences for *rúguǒ* and *zhǐyào*, which both express a causal and implication relation in a possible world, are not clear. We therefore conducted a corpus-based analysis of these markers of sufficient conditionality, classifying their use along syntactic parameters, and – following Wang's (2012) recommendation – cognitive parameters.

Across genres, the connectives did not differ in their preferred clause order. The overwhelming majority of the Chinese conditionals obeyed the topic-comment order: *p* first, *q* second (Chao 1968; Haiman 1978).

The position of the connectives within the clause showed greater variation than their English counterparts, *rúguǒ* and *zhǐyào* being able to both precede and follow their subject. In spoken discourse, they can even appear at the final position of a clause. Although both connectives favor the pre-subject over the pre-predicate position, the connective *rúguǒ* has a stronger preference for the pre-subject position than *zhǐyào*. The latter connective occurs more often in clauses without a linguistically expressed subject (i.e. in clauses it cannot be determined whether the connective precedes or follows the subject), and slightly more in clauses with the connective clearly in pre-predicate position.

Analyzing the connectives from a cognitive perspective, we looked at their distribution over domains (Sweetser 1990; Dancygier & Sweetser 2005). Although both connectives frequently occur in all three basic domains (i.e. content, epistemic, speech act) and have a preference for the content domain, the connective *zhǐyào* has a stronger preference for content conditionals than *rúguǒ*, which is also frequently used in the epistemic domain. This means that in these domains, the connectives show a relative and not an absolute difference. Due to the small number of the three meta-types of conditionals, no strong conclusions could be drawn about these types of conditionals.

It appears, however, that an absolute difference between the connectives can be found in terms of counterfactuality, which is only allowed in combination with *rúguǒ*. The zero-occurrence of counterfactual use of *zhǐyào* in our data can be explained by the fact that *zhǐyào* displays a much stronger predictive link between

the occurrence of the antecedent and the realization of its consequent than *rúguǒ*. When used to introduce a subordinate conditional clause, *rúguǒ* is equivalent to ‘if’. By contrast, *zhǐyào* does not always simply seem to mean ‘if’, as is often thought, but finds an important English equivalent in ‘so/as long as’. The three instances of ‘so long as’ in (76), taken from a speech of President Obama, should therefore all be translated into *zhǐyào*. Only then it generates the empowering effect it has in the English language.

- (74) Your future is in your hands. Your life is what you make of it. And nothing – absolutely nothing – is beyond your reach. *So long as* you’re willing to dream big. *So long as* you’re willing to work hard. *So long as* you’re willing to stay focused on your education.

In this paper, we have shown the added value of conducting corpus-based studies on actual language use. The application of inferential statistics – even on a relatively small scale, and with only two genres – allowed us to detect subtle relative differences between the connectives that would otherwise have gone unnoticed. Future studies could benefit from statistics on a larger scale, for instance by applying collocation analyses (Gries, 2013) to the study of Mandarin Chinese conditionals (cf. the recommendations in Wei, Speelman & Evers-Vermeul, *submitted*). This would be a way to test whether *rúguǒ* and *zhǐyào* also display differences in their preference for co-occurrence with other linguistic elements, such as particles (e.g. *le*, *ne*) and perspective markers (e.g. *yěxǔ* ‘maybe’, *yídìng* ‘definitely’), and in other genres (e.g. argumentative texts). These improvements in both corpora and methods open up new opportunities for linguistic research at the discourse level, thereby allowing us to enhance our understanding of human language.

## Acknowledgements

This paper was first presented at the North American Conference on Chinese Linguistics (NACCL-27) at UCLA. We owe a debt of gratitude to Dr. Gang Chen for being a terrific coder and colleague. Thanks to Professor Bernard Comrie for a face-to-face discussion about the article at University of California, Santa Barbara. Our thanks also go to President Emeritus of the California Institute of Integral Studies (CIIS) Joseph Subbiondo and Professor Chaofen Sun (Stanford University) for their invaluable advice. We are particularly grateful to the general editor of CLD for his assistance and the anonymous reviewers for their generous comments and suggestions. Any mistakes or shortcomings remain the sole responsibility of the authors.

## References

- Au, T. 1983. Chinese and English counterfactuals: The Sapir-Whorf hypothesis revisited. *Cognition* 15: 155–187.
- Au, T. 1984. Counterfactuals: In reply to Alfred Bloom. *Cognition* 17: 289–302.
- Austin, J. L. 1961/1979. Ifs and cans. In J. O. Urmson, and G. J. Warnock (eds.), *Philosophical Papers* (3rd edition). 205–232. Oxford: Oxford University Press.
- Bhatt, R., and R. Pancheva. 2005. Conditionals. In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*. 638–687. Malden, MA/Oxford: Blackwell.
- Bloom, A. 1981. *The Linguistic Shaping of Thought*. Hillsdale, NJ: Lawrence Erlbaum.
- Chao, Y. R. 1968. *A Grammar of Spoken Chinese*. Berkeley/Los Angeles: University of California Press.
- Chao, Y. R. 1976. *Aspects of Chinese Sociolinguistics*. Stanford: Stanford University Press.
- Chen, G. 1988. Ying han jiashe tiaojianju bijiao [A formal description of Chinese expressions of hypothetical conditions in comparison with English]. *Foreign Language Teaching and Research* 73 (1): 10–18.
- Cheng, L., and J. Huang. 1996. Two types of donkey sentences. *Natural Language Semantics* 4: 121–163.
- Chierchia, G. 2000. Chinese conditionals and the theory of conditionals. *Journal of East Asian Linguistics* 9: 1–54.
- Chou, C. 2000. Chinese speakers' acquisition of English conditionals: Acquisition order and L1 transfer effects. *Second Language Studies* 19 (1): 57–98.
- Comrie, B. 1986. Conditionals: A typology. In E. C. Traugott, A. ter Meulen, J. S. Reilly, and C. A. Ferguson (eds.), *On Conditionals*. 77–99. Cambridge: Cambridge University Press.
- Dancygier, B. 1998. *Conditionals and Prediction*. Cambridge: Cambridge University Press.
- Dancygier, B., and E. Sweetser. 2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge University Press.
- Douven, I. 2016. *The Epistemology of Indicative Conditionals*. Cambridge: Cambridge University Press.
- Evers-Vermeul, J., L. Degand, B. Fagard, and L. Mortier. 2011. Historical and comparative perspectives on subjectification: A corpus-based analysis of Dutch and French causal connectives. *Linguistics* 49 (2): 445–478.
- Ferguson, C. A., J. S. Reilly, A. ter Meulen, & E. C. Traugott. 1986. Overview. In E. C. Traugott, A. ter Meulen, J. S. Reilly, and C. A. Ferguson (eds.), *On Conditionals*. 3–20. Cambridge: Cambridge University Press.
- Ford, C. E., and S. A. Thompson. 1986. Conditionals in Discourse. In E. C. Traugott, A. ter Meulen, J. S. Reilly, and C. A. Ferguson (eds.), *On Conditionals*. 353–372. Cambridge: Cambridge University Press.
- Gauker, C. 2005. *Conditionals in Context*. Cambridge, MA: MIT Press.
- Gries, S. 2013. 50-something years of work on collocations: What is or should be next ... *International Journal of Corpus Linguistics* 18 (1): 137–165.
- Haiman, J. 1978. Conditionals are topics. *Language* 54: 564–589.
- Ippolito, M. 2013. *Subjunctive Conditionals: A Linguistic Analysis*. Cambridge, MA: MIT Press.
- Johnson-Laird, P. N., and R. M. J. Byrne. 2002. Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review* 109 (4): 646–678.

- Johnson-Laird, P. N., and R. M. J. Byrne. 2010. Conditionals and possibilities. In M. Oaksford and N. Chater (eds.), *Cognition and Conditionals*. 55–68. New York: Oxford University Press.
- Li, C. N., and S. A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California.
- Li, F., J. Evers-Vermeul, and T. J. M. Sanders. 2013. Subjectivity and result marking in Mandarin: A corpus-based investigation. *Chinese Language and Discourse* 4: 74–119.
- Li, F., Evers-Vermeul, J., and T. J. M. Sanders. 2016. On the subjectivity of Mandarin reason connectives: Robust profiles or genre-sensitivity? In N. Stukker, W. Spooren, and G. Steen (eds.), *Genre in Language, Discourse and Cognition*. 15–49. Berlin/New York: De Gruyter Mouton.
- Lü, S. 1942. *Zhongguo wenfa yaolüe*. [Essentials of Chinese Grammar]. Beijing: Commercial Press.
- McEnery, T., and A. Hardie. 2012. *Corpus Linguistics*. New York: Cambridge University Press.
- McEnery, T., and A. Wilson. 2001. *Corpus Linguistics (2nd edition)*. Edinburgh: Edinburgh University Press.
- Núñez, R. 2007. Inferential statistics in the context of empirical cognitive linguistics. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, and M. J. Spivey (eds.), *Methods in Cognitive Linguistics*. 87–118. Amsterdam: John Benjamins.
- Sanders, T., and E. Sweetser (eds). 2009. *Causal Categories in Discourse and Cognition*. Berlin: Mouton de Gruyter.
- Schmied, J. 1993. Qualitative and quantitative research approaches to English relative constructions. In C. Souter, and E. Atwell (eds.), *Corpus Based Computational Linguistics*. 85–96. Amsterdam: Rodopi.
- Shen, J. 2003. Fujū san yu: Xing, zhi, yan [Compound sentences in three conceptual domains: Acting, knowing and uttering]. *Chinese Language* 294 (3): 195–204.
- Shen, J. 2008. Fuci he lianci de yuanyu yongfa [Meta-linguistic usage of adverbs and conjunctions]. *Research on Chinese as a Second Language* 5: 113–125.
- Shen, J. 2016. *Mingci he dongci* [Nouns and Verbs]. Beijing: Commercial Press.
- Siegel, M. E. A. 2006. Biscuit conditionals: Quantification over potential literal acts. *Linguistics and Philosophy* 29: 167–203.
- Spooren, W., and L. Degand. 2010. Coding coherence relations: Reliability and validity. *Corpus Linguistics and Linguistic Theory* 6 (2): 241–266.
- Stalnaker, R. C. 1981. A theory of conditionals. In W. L. Harper, R. Stalnaker, and G. Pearce (eds.), *Ifs: Conditionals, Beliefs, Decisions, Chance, and Time*. 41–55. Dordrecht: D. Reidel Publishing Company.
- Sweetser, E. E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.
- Tian, Y. 2006. *Pingding fuci “jiu” de qiyi xianxiang* [Ambiguities of the evaluative adverb *jiu*]. Beijing: Peking University Press.
- van der Auwera, J. 1986. Conditionals and speech acts. In E. C. Traugott, A. ter Meulen, J. S. Reilly, and C. A. Ferguson (eds.), *On Conditionals*. 197–214. Cambridge: Cambridge University Press.
- Wang, C. 2010a. Hanyu tiaojianju xiaoju jian de yuxu leixing [Ordering patterns of clauses in conditional complex sentences in Chinese]. *Shijie hanyu jiaoxue* [Chinese Teaching in the World] 24 (4): 468–482.
- Wang, C. 2010b. Jiashe xing dengji yu hanyu tiaojianju [Degree of hypotheticality and Chinese conditional sentences]. *Hanyu xuebao* [Chinese Linguistics] 32 (4): 59–69.

- Wang, C. 2012. Tiaojianju fenlei de guoneiwai bijiao [A comparison of conditional categorization in China and West], *Dongfang yuyan xue* [Oriental Linguistics] 11: 115–128.
- Wang, L. 1943, 1944. *Zhongguo xiandai yufa* [Modern Chinese Grammar]. Shanghai: Commercial Press.
- Wang, W., X. Zhang, M. Lu, and H. Cheng. 1994. *Xiandai hanyu fuju xinjie* [A New Study of Complex Sentences in Modern Chinese]. Shanghai: Huadong shifan daxue chubanshe.
- Wei, Y., Speelman, D., and J. Evers-Vermeul. (submitted). Collocation analyses at the discourse level using Chinese corpora. Submitted for publication.
- Wierzbicka, A. 1997. Conditionals and counterfactuals: Conceptual primitives and linguistic universals. In A. Athanasiadou and R. Dirven (eds), *On Conditionals Again*. 15–60. Amsterdam/Philadelphia: John Benjamins.
- Xing, F. 2001. *Hanyu fuju yanjiu* [A Study of Chinese Complex Sentences]. Beijing: Commercial Press.

### *Authors' addresses*

Weiyang Chen  
School of International Studies  
Zhejiang University  
Hangzhou, 310058  
P.R. China  
chenweiyang@zju.edu.cn

Jacqueline Evers-Vermeul  
Utrecht Institute of Linguistics OTS  
Trans 10  
Utrecht University  
3512 JK Utrecht  
The Netherlands  
j.evers@uu.nl