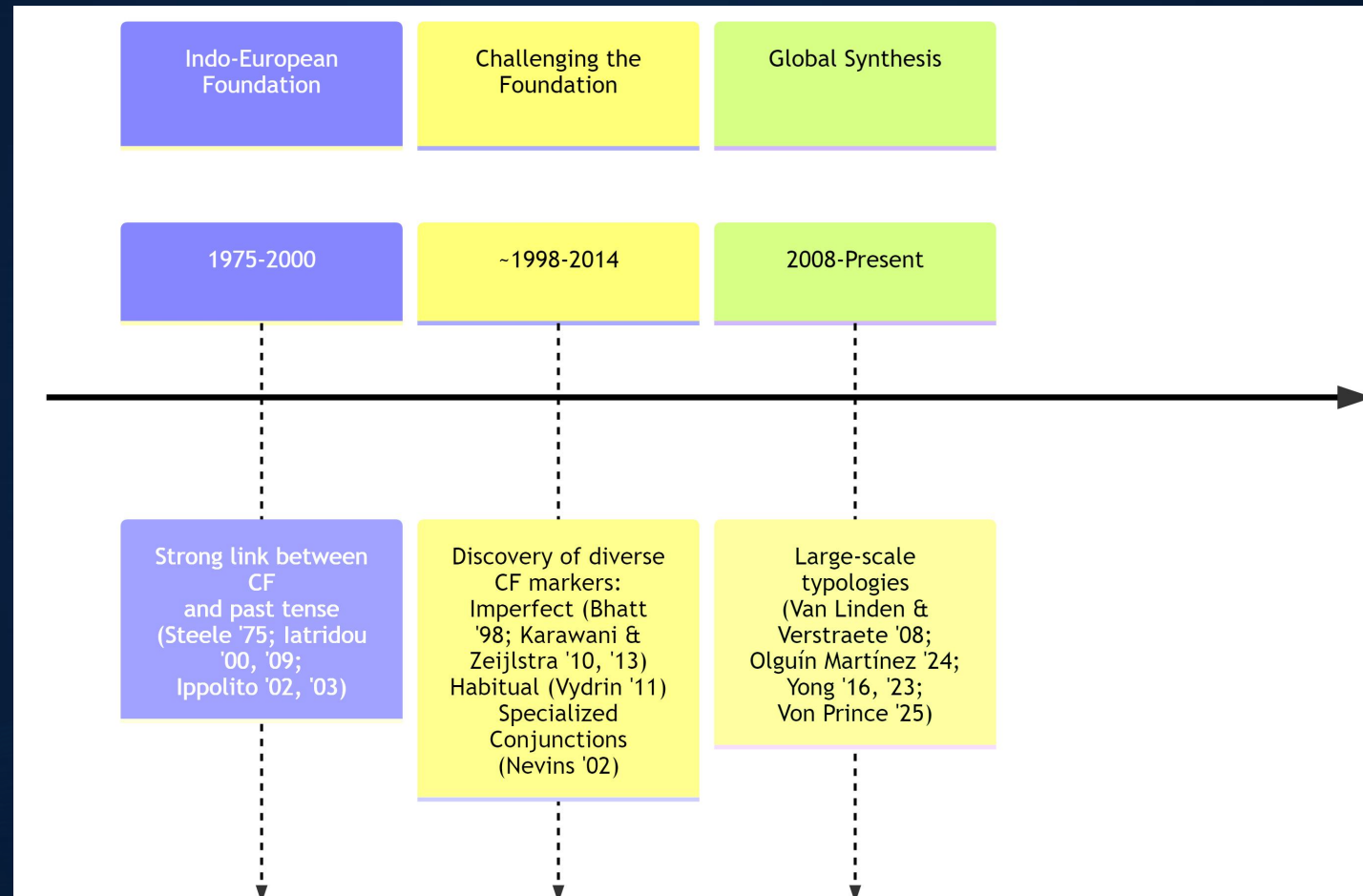




# **Unveiling Counterfactuality in Mandarin Chinese: A Comparative Study with English and its Typological Implications**

# Introduction

Figure 1. The Typological Expansion of Counterfactual Marking



# Introduction

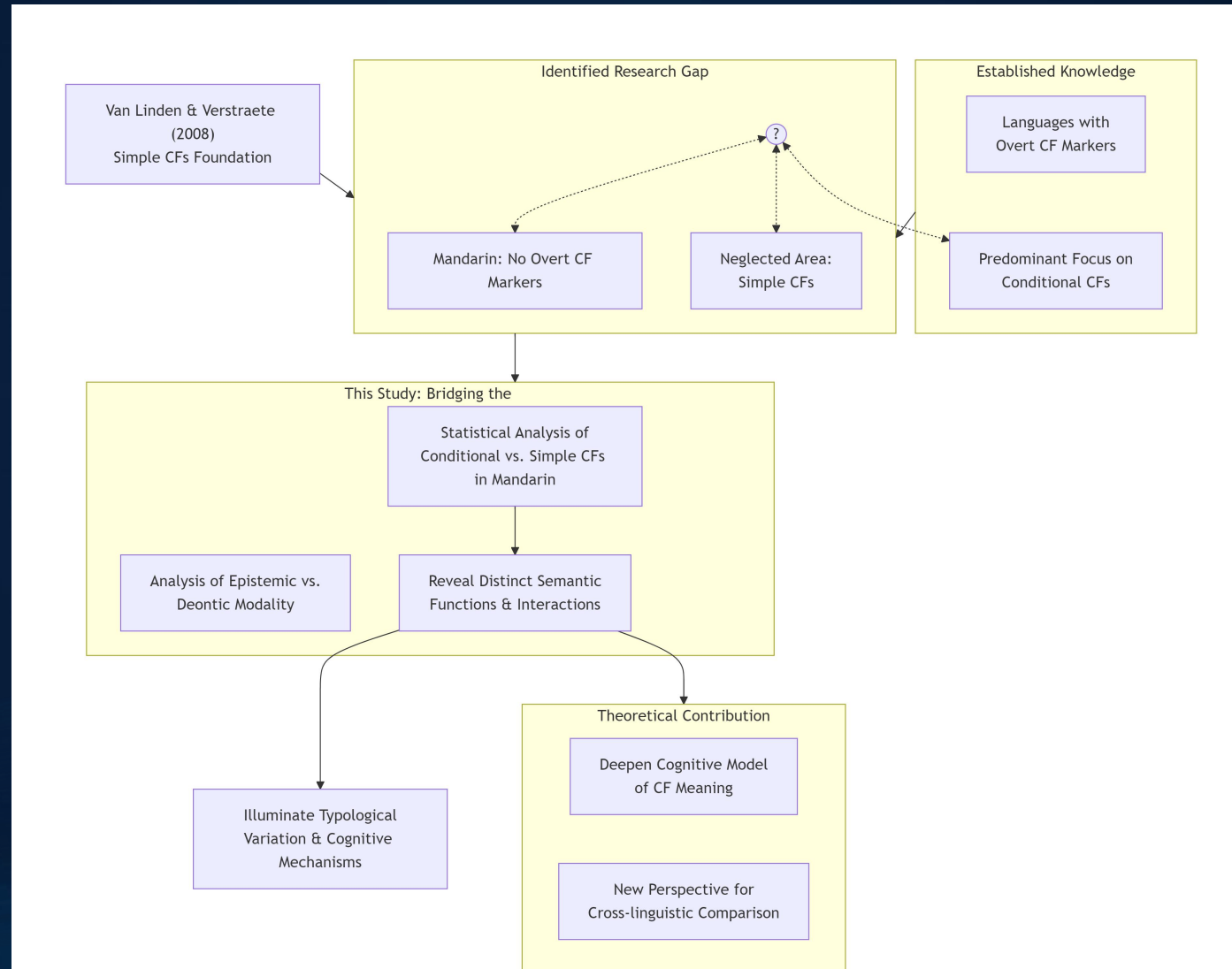
Table 1. Key Cross-Linguistic Challenges to the Indo-European Bias

| Language / Language Group | Key CF Marker(s) Identified    | Key Scholar(s) & Year             | Core Insight                                                 |
|---------------------------|--------------------------------|-----------------------------------|--------------------------------------------------------------|
| Indo-Aryan                | Imperfect                      | Bhatt (1998)                      | Extended CF marking beyond simple past tense.                |
| Iranian                   | Habitual / Dedicated Markers   | Vydrin (2011)                     | Demonstrated diversity even within a single language family. |
| Slavic/Romance            | "Fake" Imperfective            | Grønn (2008)                      | Highlighted the critical role of aspect.                     |
| Palestinian Arabic        | Imperfect                      | Karawani & Zeijlstra (2010, 2013) | Provided a detailed formal analysis of a non-past marker.    |
| Zulu, French, Persian     | Past, Perfective, Imperfective | Bjorkman & Halpert (2012)         | Cataloged multiple strategies across disparate languages.    |
| Mandarin (Tenseless)      | Specialized Conjunctions       | Nevins (2002)                     | Proved tenseless languages have systematic CF strategies.    |
| Turkana, Yoruba, Nahuatl  | Various (Mood, Particles)      | Lazard (2006)                     | Expanded the typology to include diverse, global languages.  |



# Outlines

Figure 2. Mapping the Research Gap and Objectives



# Research Questions

1. **Feature Identification:** What constellation of linguistic features triggers counterfactual interpretations in Mandarin Chinese, a language without dedicated CF markers?
2. **Internal Typology:** What are the systematic differences in the expression of counterfactuality between conditional and non-conditional in Mandarin?
3. **Underlying Mechanism:** What are the underlying pragmatic and semantic mechanisms—specifically, how do epistemic and deontic modalities interact with syntactic structures—that enable the construction of CF meaning in Mandarin?
4. **Cross-Linguistic Significance:** What does this reveal about universal patterns in the linguistic encoding of alternative worlds?

# Definitions:

(1) Iquito (Zaparoan: Peru) (I-wen Lai, 2009: (2), (4), (7))

|    |                                                                                                  |                      |                |              |
|----|--------------------------------------------------------------------------------------------------|----------------------|----------------|--------------|
| a. | Nu= <b>ti</b> =huárataamicaáca                                                                   | raati-aa-cura        | iina           | ampisitaaja, |
|    | 3S= <b>CF</b> =two_days_away                                                                     | drink-PFV-RPST       | DET            | medicine     |
|    | jaa                                                                                              | nu= <b>ti</b> =ácari | anaji-qui-ø.   |              |
|    | already                                                                                          | 3SG= <b>CF</b> =now  | recover-PFV-ZT |              |
|    | 'If he had taken the medicine the day before yesterday, he would have already recovered by now.' |                      |                |              |

(1a) is incompatible with the following sentence to cancel the counterfactually:

|    |                               |                        |         |      |
|----|-------------------------------|------------------------|---------|------|
| b. | Ca=quí=nacusi-i-ø             | nu=raati-ø-cura        | cuuta   | nuu. |
|    | NEG=1SG=know-IPFV-ZT          | 3SG=drink-GNR.PFV-RPST | perhaps | 3SG  |
|    | 'I don't know if he took it.' |                        |         |      |

(2) Mandarin (Sino-Tibetan: China)

a. 他要是在那儿住过，就会认识那儿的人了。

ta yaoshi zai nar zhu guo jiu

3SG if-be at there live PTCL, then  
hui renshi nar de ren le.  
MDL know there PTCL people MDL  
'If he had ever lived there, he would have known the people there.'

(2a) is compatible with the (2b) and (2c) to either cancel or reinforce the counterfactually:

b. 其实他确实在那儿住过。

qishi ta qushi zai nar zhu guo  
in\_fact he actually at there live PTCL  
'In fact, he has ever lived there.'

c. 显然他没在那儿住过。

xianran ta mei zai nar zhu guo  
obviously he NEG at there live PTCL  
'Obviously, he has never lived there.'

# Definitions

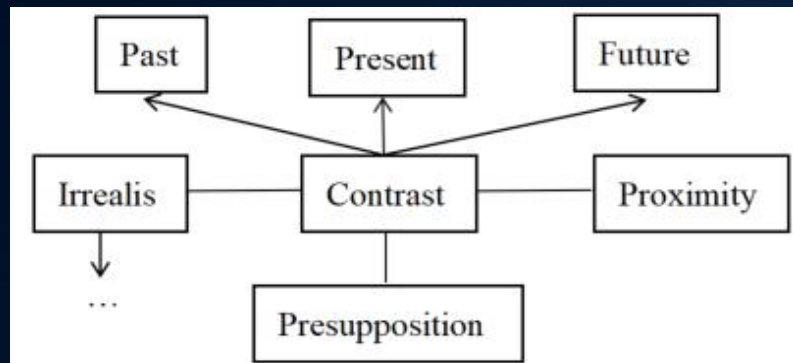
Table 2. The Binary features of CF-related Constructions of their Markedness in English

| Construction Type                                           | Irealis | Contrast | Presupposition | Proximity | Example                                                                                                           | Marked |
|-------------------------------------------------------------|---------|----------|----------------|-----------|-------------------------------------------------------------------------------------------------------------------|--------|
| Canonical CF Conditional<br>(protasis alone/apodosis alone) | +       | +        | +              | +         | "If I had known, I would've helped."<br>"What if you had taken that job offer?"<br>"I would not have done that. " | Y      |
| Future Less Vivid                                           | +       | +        | +/-            | +         | "If it rained tomorrow, I would not go."                                                                          | Y      |
| "As if" Comparison                                          | +       | +        | +              | +         | "He acted as if he were rich. "                                                                                   | Y      |
| Optative Wish                                               | +       | +        | +/-            | +         | "I wish I had bought a new car."<br>"I wish I would pass the exam next year."                                     | Y      |
| Mistaken Identity                                           | +       | +        | +              | +         | "Imagine you had wings."                                                                                          | Y      |
| Weak Necessity                                              | +       | +        | +/-            | +         | "You ought to/should have washed your clothes."<br>"You ought to/should wash your clothes. "                      | N      |
| Failed Attempt                                              | +       | +        | +              | +         | "He almost won (but didn't)."                                                                                     | N      |
| Rhetorical Question                                         | -       | +        | +              | +         | "Isn't it shameful to waste so much food?"                                                                        | N      |
| Past Intention                                              | +       | +        | +              | -         | "He wanted to go inside, but was afraid."                                                                         | N      |

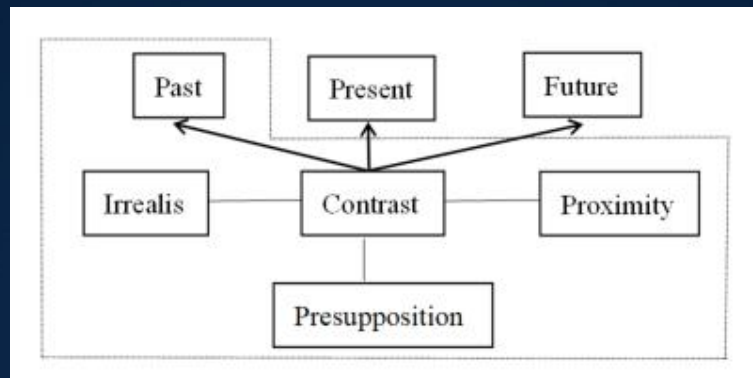


# Definitions:

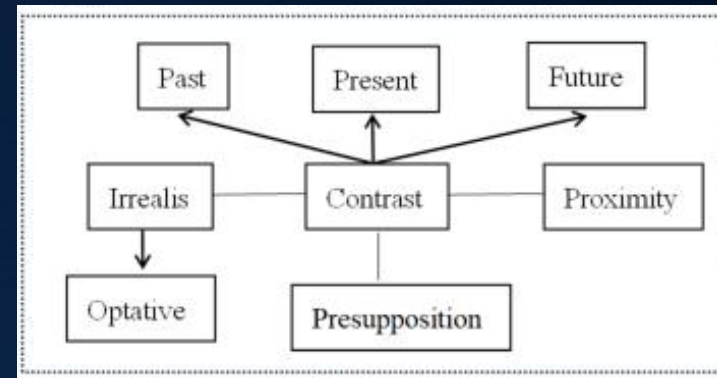
Figure 3. Semantic map of CFs



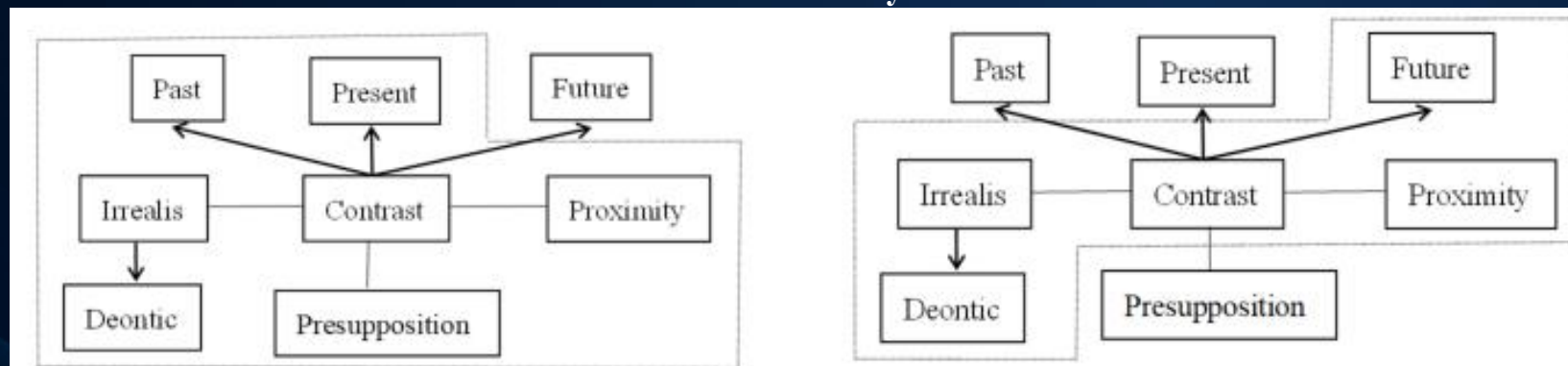
failed-attempt



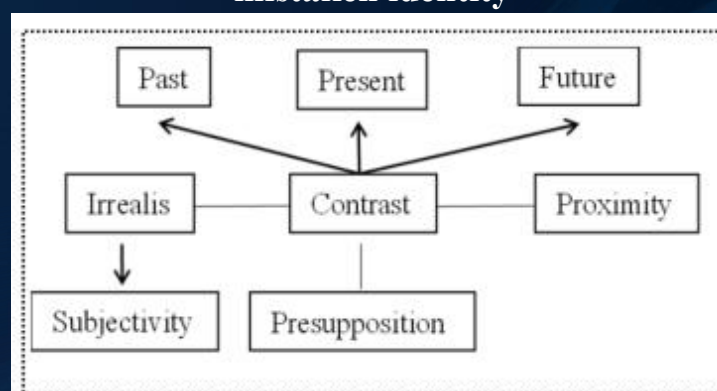
optative wish



weak necessity



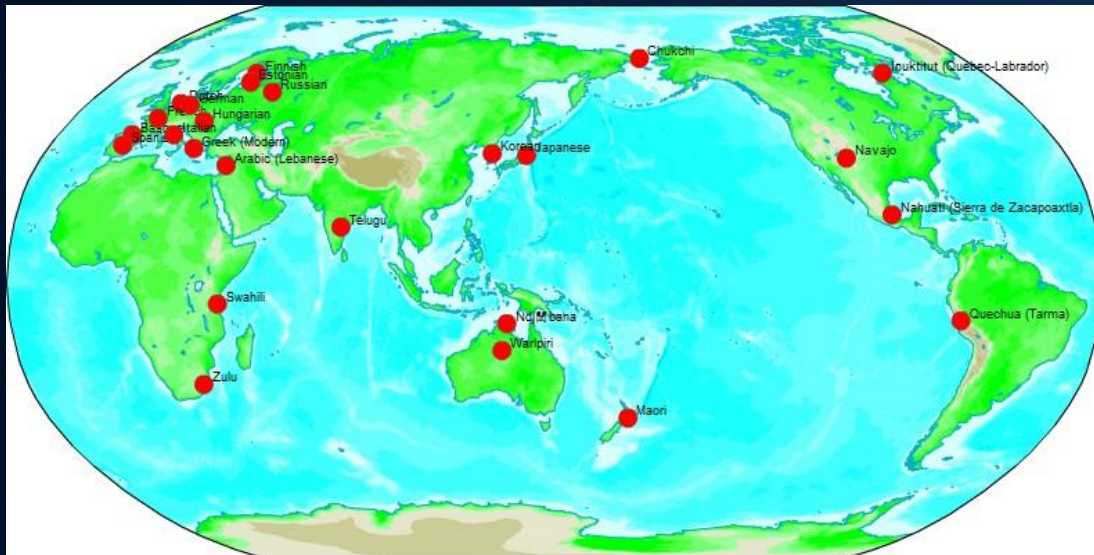
mistaken identity



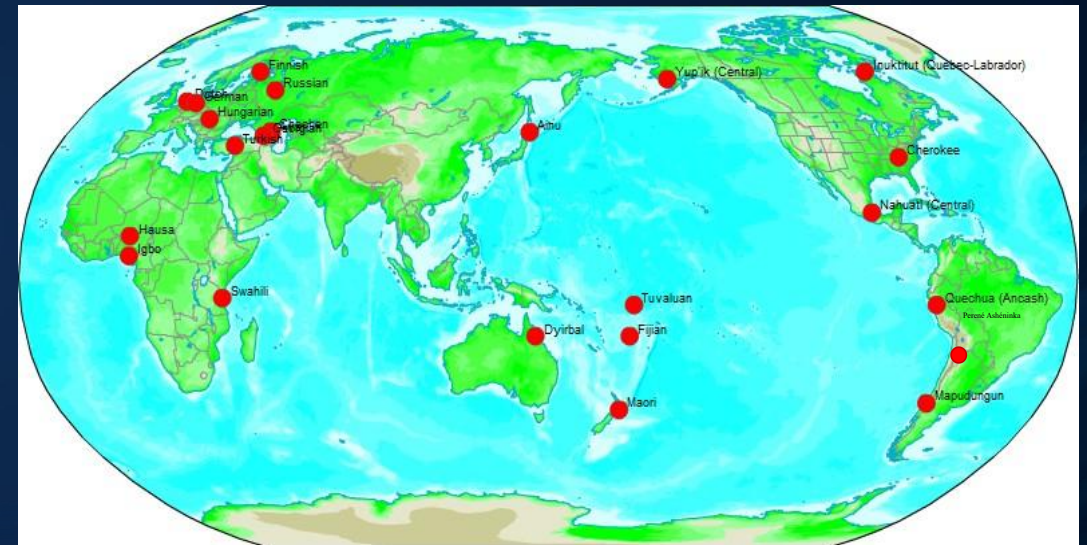


# Definitions

**Figure 4. The distribution of shared markings between weak modality and CFs**



**Figure 5. The distribution of shared markings between past intention and CF**



# Definitions

A common scenario occurs when futurity intervenes: weak irrealis typically cancels presupposition, while strong irrealis weakens it, giving rise to the typical FLV. This explains why many marked CFs lack a falsity presupposition, a phenomenon previously subsumed under Von Stechow & Iatridou (2023)'s concept of "X-marking". Certain languages derive CFs from past subjunctive meanings, which originally do not encode proximity. As a result, markers extended from past subjunctive often lose proximity in grammaticalization, leading to constructions like "past intention" that are widely attested crosslinguistically but lack proximity features. In some cases, languages even extend conditional meanings beyond the irrealis domain into interrogatives, giving rise to rhetorical questions. This analysis reveals a hierarchical sequence of semantic components:

Contrast > Irrealis > Proximity > Presupposition

Components higher in the hierarchy constitute core CF semantics, while lower components are more prone to weakening or loss.

## Definitions

A construction in Mandarin is considered a counterfactual if it exhibits systematic consistency in its formal marking with the prototypical CF semantic bundle ([+Irrealis, +Contrast, +Proximity, +Presupposition]), even if some features (like Proximity or Presupposition) are contextually variable.

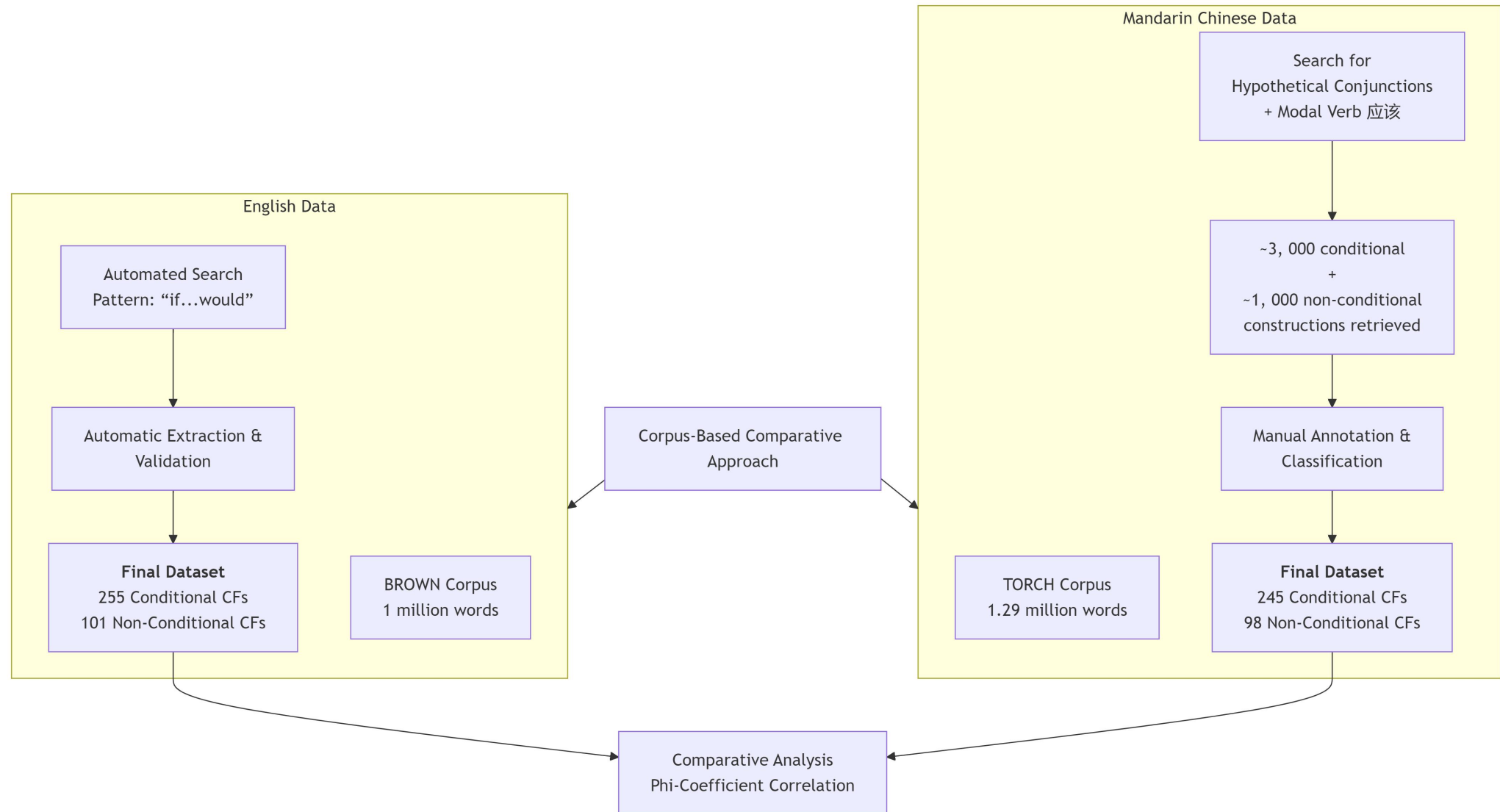


# Methodology : the Corpora

Table 3. Basic Statistics of the Copora

| Corpus                                                                    | Size               | Register                                                                                                                                                                                                                                                                                                                      |
|---------------------------------------------------------------------------|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Texts of Recent Chinese (Brown family, 2009, 2013 summer edition) (TORCH) | 1.29 million words | Press: Reportage<br>Press: Editorial<br>Press: Reviews<br>Religion<br>Skills and Hobbies<br>Popular Lore<br>Belles Lettres, Biography, Essays<br>Miscellaneous: Government and House Organs<br>Learned<br>Fiction: General<br>Fiction: Mystery<br>Fiction: Science Fiction<br>Fiction: Adventure<br>Fiction: Romance<br>Humor |
| Brown University Standard Corpus of Present-Day American English (BROWN)  | 1 million words    |                                                                                                                                                                                                                                                                                                                               |

# Methodology : the Corpora

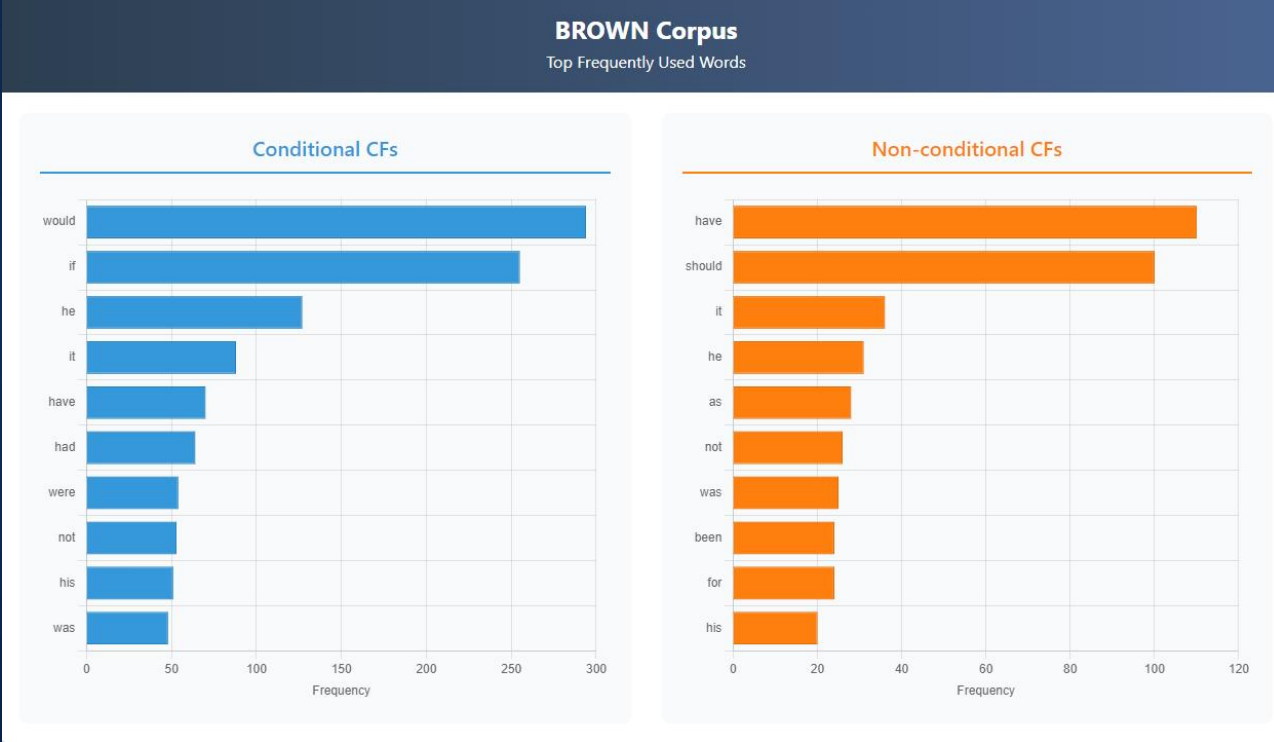


# Methodology : Frequency Test

Figure 6. Top Frequently Used Word Categories of Chinese Conditional vs Non-Conditional CFs



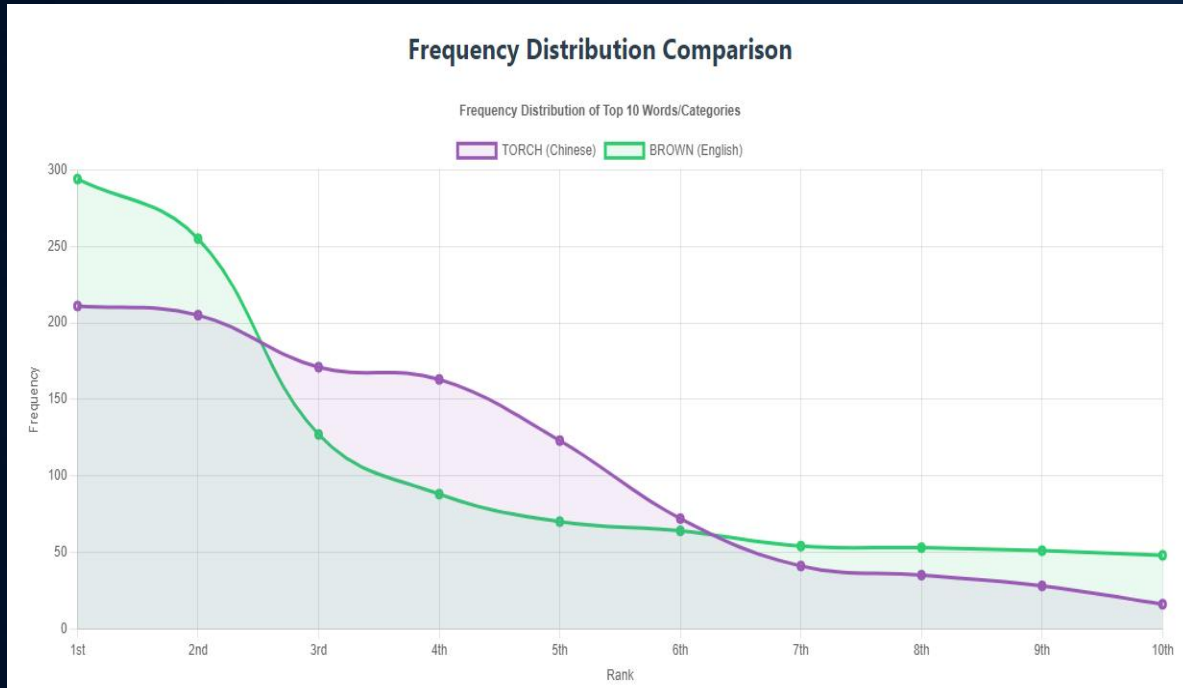
Figure 7. Top Frequently Used Word Categories of English Conditional vs Non-Conditional CFs





# Methodology : Frequency Test

Figure 8. Frequency Distribution Comparison



TORCH variance: 5910 - words spread more evenly"

BROWN variance: 8120 - concentrated in few words"

Chinese shows gradual frequency decline, English shows sharp drop-off.

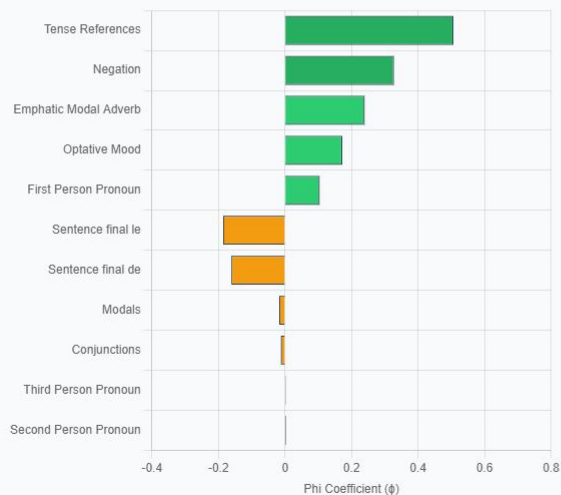
# Methodology : Correlation Analysis



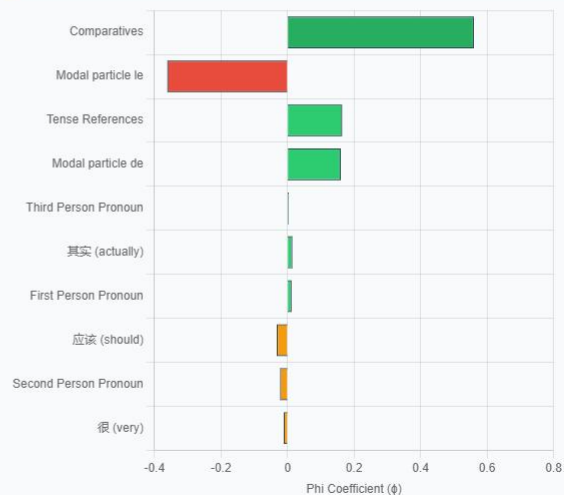
## TORCH Corpus (Chinese)

Phi Coefficient Correlation of Counterfactual Features

### Conditional CFs



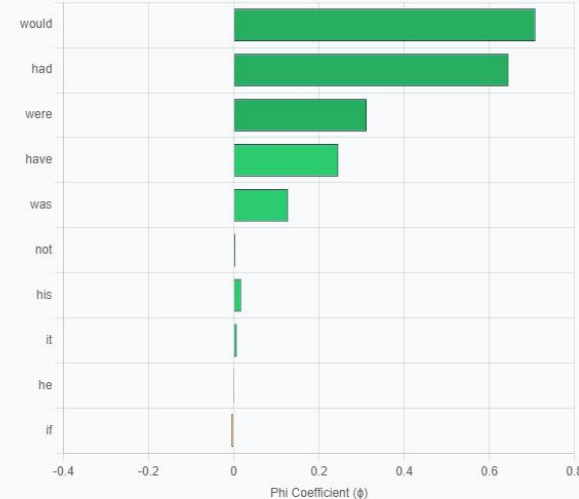
### Non-conditional CFs



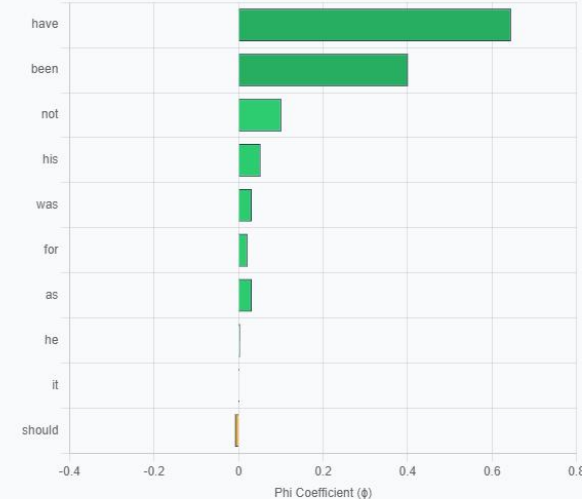
## BROWN Corpus (English)

Phi Coefficient Correlation of Counterfactual Features

### Conditional CFs



### Non-conditional CFs



We used phi coefficient correlation to measure association between linguistic features and counterfactual constructions at the sentence level. This approach counts feature presence per sentence (not total frequency), preventing distortion from repeated usage within single sentences.

# Findings: the Role of Modality in Chinese CFs

## Conditional CFs → Epistemic in Protasis: An Embedded □P

|                                                                                                      |      |             |       |      |       |      |        |
|------------------------------------------------------------------------------------------------------|------|-------------|-------|------|-------|------|--------|
| (4a) 如果有电，灯就亮了。                                                                                      |      |             |       |      |       |      |        |
| Rúguǒ                                                                                                | yǒu  | diàn,       | dēng  | jiù  | liàng | le.  |        |
| if                                                                                                   | have | electricity | light | will | on    | PTCL |        |
| If there was electricity, the light would be on.                                                     |      |             |       |      |       |      |        |
| (CF, with implicit EM certainty in A.) (□A) ∧ (A → B)                                                |      |             |       |      |       |      |        |
| (4b) 如果有电，灯就亮了；如果没电，灯就不会亮。                                                                           |      |             |       |      |       |      |        |
| Rúguǒ                                                                                                | yǒu  | diàn,       | dēng  | jiù  | liàng | le;  |        |
| if                                                                                                   | have | electricity | light | will | on    | PTCL |        |
| rúguǒ                                                                                                | méi  | diàn,       | dēng  | jiù  | bù    | huì  | liàng. |
| if                                                                                                   | NEG  | electricity | light | will | NEG   | will | on     |
| If there is electricity, the light will be on; if there is no electricity, the light will not be on. |      |             |       |      |       |      |        |
| (Non-CF, with implicit EM possibility in A.) (◇A) ∧ (A → B) ∨ ((◇¬A) ∧ (¬A → ¬B))                    |      |             |       |      |       |      |        |

In standard conditionals, multiple possible worlds are considered where the premise may or may not hold, reflecting possibility. However, in counterfactual conditionals like example (4a) '如果有电，灯就亮了' (interpreted as 'If there WAS electricity, the light WOULD be on'), the logic shifts to consider only worlds where the premise holds, creating an implicit certainty operator (Embedded □P) that activates counterfactual meaning. This is facilitated by linguistic elements like the sentence-final particle 'le,' which helps establish this certainty. In contrast, example (4b) '如果有电，灯就亮了；如果没电，灯就不会亮' explicitly states both possibilities, canceling the certainty operator through negation and resulting in a standard hypothetical interpretation rather than a counterfactual one.



## Findings: the Role of Modality in Chinese CFs

## Conditional CFs $\longrightarrow$ Epistemic in Protasis: An Embedded $\Box P$

(5) 如果必须修改任务时, 一般也应经上级批准。

|                   |                            |                  |                 |                     |                    |
|-------------------|----------------------------|------------------|-----------------|---------------------|--------------------|
| Rúguǒ<br>if       | <b>bìxū</b><br><b>must</b> | xiūgǎi<br>revise | rènwù<br>task   | shí,<br>time        |                    |
| yībān<br>normally | yě<br>also                 | yīng<br>should   | jīng<br>through | shàngjí<br>superior | pīzhǔn.<br>approve |

If the task **needs to** be modified, it generally should be approved by a superior.

$$(A(\Box P)) \wedge (A \rightarrow B)$$

(6) 但有一点可以肯定：如果申花上诉国际足联，那么后者将会要求双方提供签订的工作合同。

|         |               |           |          |                 |         |          |
|---------|---------------|-----------|----------|-----------------|---------|----------|
| Dàn     | yǒu           | yīdiǎn    | kěyǐ     | <b>kěndìng:</b> | rúguǒ   | Shēnhuā  |
| but     | have          | one_point | could    | <b>sure</b>     | if      | Shenhua  |
| shàngsù | guójì zúlián, |           | nàme     | hòuzhě          | jiāng   | huì      |
| appeal  | FIFA          |           | then     | latter          | will    | will     |
| yāoqiú  | shuāngfāng    | tígōng    | qiāndìng | de              | gōngzuò | hétóng.  |
| require | both          | provide   | signed   | PTCL            | work    | contract |

But one thing is **certain**: if Shenhua appeals to the FIFA, the latter will require both parties to provide the signed employment contract.

$$\Box(A \rightarrow B)$$

The implicit  $\Box P$  in A is often embedded within connectives (e.g., (4a)) and functions as an external modal of A, distinct from the internal modality of A in (5) or the external modality at the clause level in (6):

# Findings: the Role of Modality in Chinese CFs

## Conditional CFs → Epistemic in Protasis: An Embedded □P

(7) 如果有可能,就不要打翻牛奶瓶。

|       |      |             |        |     |      |
|-------|------|-------------|--------|-----|------|
| Rúguǒ | yǒu  | kěnéng,     | jiù    | bù  | yào  |
| if    | have | probability | then   | NEG | will |
| dǎ    | fān  | niú'nǎi     | píng.  |     |      |
| knock | over | milk        | bottle |     |      |

If **possible**, don't knock over the milk bottle.

(8) 如果当初我真的跑去南京读大学, 那我肯定会哭死的。

|       |           |            |        |     |     |         |       |            |
|-------|-----------|------------|--------|-----|-----|---------|-------|------------|
| Rúguǒ | dāngchū   | wǒ         | zhēnde | pǎo | qù  | Nánjīng | dú    | dàxué,     |
| if    | back_then | I          | indeed | run | to  | Nanjing | study | university |
| nà    | wǒ        | kěndìng    | huì    | kū  | sǐ  | de.     |       |            |
| then  | I         | definitely | will   | cry | die | PTCL    |       |            |

If I had **really** gone to Nanjing to study at university back then, I would definitely have cried to death.

- To show uncertainty or a simple "maybe," phrases like "rúguǒ... yǒu kěnéng" (如果...有可能, "if possible") are used. This creates an open hypothesis.
- To show certainty—specifically in hypothetical or "what if" scenarios that aren't true—speakers add words like "zhēn" (真, "really") as in "rúguǒ... zhēn de" (如果...真的, "if it were really the case"). This is common in counterfactual conditionals.



# Findings: the Role of Modality in Chinese CFs

## Conditional CFs → Epistemic in Protasis: An Embedded □P

(9) 如果不是这个非常时期，我们一定会很轻松的聊上几天。

|                                                                                                        |      |       |          |       |      |                 |         |        |  |
|--------------------------------------------------------------------------------------------------------|------|-------|----------|-------|------|-----------------|---------|--------|--|
| Rúguǒ                                                                                                  |      | búshì |          | zhègè |      | fēicháng shíqī, |         | wǒ men |  |
| if                                                                                                     |      | NEG   |          | this  |      | abnormal        |         | 1PL    |  |
| yíding                                                                                                 | huì  | hěn   | qīngsōng | de    | liáo | shàng           | huì     | tiān.  |  |
| definitely                                                                                             | will | very  | relaxed  | PTCL  | chat | PTCL            | a while | chat   |  |
| If it weren't for this extraordinary period, we would definitely be able to chat casually for a while. |      |       |          |       |      |                 |         |        |  |

(10a) 要是你有任何问题，你都可以来找我。

|                                                       |     |      |       |          |     |     |      |      |      |     |
|-------------------------------------------------------|-----|------|-------|----------|-----|-----|------|------|------|-----|
| Yàoshì                                                | nǐ  | yǒu  | rènhe | wèntí,   | nǐ  | dōu | kěyǐ | lái  | zhǎo | wǒ. |
| if                                                    | you | have | any   | question | you | all | can  | come | find | me  |
| If you have any questions, you can always come to me. |     |      |       |          |     |     |      |      |      |     |

(10b\*) 要不是你有任何问题，你都可以来找我。

|                                                                          |     |      |       |          |     |     |      |      |      |     |
|--------------------------------------------------------------------------|-----|------|-------|----------|-----|-----|------|------|------|-----|
| Yàobushì                                                                 | nǐ  | yǒu  | rènhe | wèntí,   | nǐ  | dōu | kěyǐ | lái  | zhǎo | wǒ. |
| if-not                                                                   | you | have | any   | question | you | all | can  | come | find | me  |
| If it weren't for you having any questions, you could always come to me. |     |      |       |          |     |     |      |      |      |     |

As mentioned earlier, negation demonstrates a relatively close correlation with Mandarin CFs. However, negation in Mandarin encompasses several distinct forms, such as 没 (méi), 没有 (méiyǒu), 不 (bù), and 不是 (búshì), e. Our analysis of the corpus reveals that the type of negation most closely associated with counterfactuality is often factive negation *búshì* (不是, "not\_be"), typically occurring in the initial position of the sentence, where it negates an assumed or presupposed fact. The negation marker *búshì* (不是, "not\_be") in such constructions typically functions to constrain potential outcomes, thereby contributing to the expression of EM certainty as in (9), where the negation not only negates a specific alternative but also reinforces the certainty of the ensuing assertion. The factive operator inherently presupposes the truth of the embedded proposition, thereby interfering with the licensing of NPIs. This blocking effect is clearly observable in examples such as (10a) and (10b)



# Findings: the Role of Modality in Chinese CFs

## Conditional CFs → Epistemic in Protasis: An Embedded □P

(11) 就算是你昨天提醒了我，我也不会改变主意。

|            |     |           |         |        |     |    |
|------------|-----|-----------|---------|--------|-----|----|
| Jiùsuànshì | nǐ  | zuótiān   | tíxǐng  | le     | wǒ, | wǒ |
| even_if    | you | yesterday | remind  | PTCL   | me  | I  |
| yě         | bù  | huì       | gǎibiàn | zhǔyì. |     |    |
| also       | NEG | will      | change  | idea   |     |    |

Even if you had reminded me yesterday, I still wouldn't have changed my mind.

(12) 但使龙城飞将在，不教胡马度阴山。（《唐诗·出塞》）

|         |           |          |       |     |      |       |       |              |
|---------|-----------|----------|-------|-----|------|-------|-------|--------------|
| Dànshǐ  | Lóngchéng | fēijiàng | zài,  | bù  | jiāo | Húmǎ  | dù    | Yīnshān.     |
| if_only | Long city | hero     | exist | NEG | let  | enemy | cross | Yin mountain |

If only the great general of Longcheng, Li Guang, were still alive, the Xiongnu cavalry would never have been allowed to cross the Yin Mountains. (Poems in Tangy Dynasty, 725-750 CE)

(13) 早知道这样，我就不来了。

|            |          |    |      |     |      |      |
|------------|----------|----|------|-----|------|------|
| Zǎozhīdào  | zhèyàng, | wǒ | jiù  | bù  | lái  | le.  |
| know_early | this     | I  | then | NEG | come | PTCL |

If I had known it earlier, I would not have come.

(14) 微夫人之力，不及此。（《左传》）

|            |          |      |       |     |       |      |
|------------|----------|------|-------|-----|-------|------|
| Wēi        | fūren    | zhī  | lì,   | bù  | jí    | cǐ.  |
| if_not_for | that_man | PTCL | force | NEG | reach | this |

If it were not for the strength of that person, we would not have achieved this. (Zuo Zhuan, 4<sup>th</sup>-3<sup>rd</sup> BCE)

In Mandarin Chinese, the implicit certainty (□P) within a counterfactual protasis can be explicitly signaled by specific elements in the C0 position.

# Findings: the Role of Modality in Chinese CFs

## Conditional CFs → Epistemic in Protasis: An Embedded □P

Table 3. Correlation test of □P and CFs in Chinese

|                                                                                                 |                                | Non-CF | CF  | Total | φ≈0.239 indicating a weak to moderate positive correlation. |
|-------------------------------------------------------------------------------------------------|--------------------------------|--------|-----|-------|-------------------------------------------------------------|
| <i>zhēnde</i> (真的, "indeed")<br><i>guǒzhēn</i> (果真, "indeed")<br><i>dàngzhēn</i> (当真, "indeed") | with <b>Emphatic Modals</b>    | 14     | 28  | 42    |                                                             |
|                                                                                                 | without <b>Emphatic Modals</b> | 2770   | 217 | 2987  |                                                             |
| Total                                                                                           |                                | 2784   | 245 | 3029  |                                                             |

|                             |                                 | Non-CF | CF  | Total | φ≈0.547 indicating a moderate to strong positive correlation. |
|-----------------------------|---------------------------------|--------|-----|-------|---------------------------------------------------------------|
| <i>búshì</i> (不是, "not_be") | with <b>Factive Negation</b>    | 3      | 81  | 84    |                                                               |
|                             | without <b>Factive Negation</b> | 2781   | 164 | 2945  |                                                               |
| Total                       |                                 | 2784   | 245 | 3029  |                                                               |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                | Non-CF | CF  | Total | φ≈0.5060 indicating a moderate positive correlation |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|--------|-----|-------|-----------------------------------------------------|
| zǎo(xiē/diǎnr) (早(些/早/点儿), "earlier/early/a bit earlier")<br>dāngshí (当时, "at that time")<br>dāngchū (当初, "in the beginning/originally")<br>céng (曾, "once/ever")<br>shìxiān (事先, "in advance")<br>guòqù (过去, "in the past")<br>qùnián (去年, "last year")<br>yǐqián (以前, "before/in the past")<br>cóngxiǎo (从小, "since childhood")<br>le1 (了 1, "perfective marker")<br>yǐ (已, "already")<br>xiànzài (现在, "now")<br>jīntiān (今天, "today")<br>zhìjīn (至今, "up to now") | with <b>past references</b>    | 5      | 72  | 84    |                                                     |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | without <b>past references</b> | 2779   | 173 | 2945  |                                                     |
| Total                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                | 2784   | 245 | 3029  |                                                     |

## Findings: the Role of Modality in Chinese CFs

Conditional CFs → Epistemic in Apodosis: Unrelated to Counterfactuality

Table 4. Correlation test of EMs in Apodosis and CFs in Chinese

|                                                                                                                                                                                                                                                                                 |                                                                                       | Non-CF | CF  | Total | $\phi \approx 0.0170$ indicating no correlation. |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------|-----|-------|--------------------------------------------------|
| <b>kěnéng</b> (可能, "maybe")<br><b>néng</b> (能, "can")<br><b>yěxǔ</b> (也许, "perhaps")<br><b>bìxū</b> (必须, "must")<br><b>yīdìng</b> (一定, "certainly")<br><b>bìrán</b> (必然, "inevitably")<br><b>huì</b> (会, "will")<br><b>yīnggāi</b> (应该, "should") <b>zhǔn</b> (准, "definitely") | with<br><b>May:</b> $\Diamond P$<br><b>Must:</b> $\Box P$<br><b>Will:</b> $\Box P$    | 2454   | 211 | 2654  |                                                  |
|                                                                                                                                                                                                                                                                                 | without<br><b>May:</b> $\Diamond P$<br><b>Must:</b> $\Box P$<br><b>Will:</b> $\Box P$ | 330    | 34  | 375   |                                                  |
| <b>Total</b>                                                                                                                                                                                                                                                                    |                                                                                       | 2784   | 245 | 3029  |                                                  |



# Findings: the Role of Modality in Chinese CFs

## Conditional CFs → Ambiguous Cases: sentence final *le* (了) and *de* (的)

|                                                  |      |             |       |        |       |             |
|--------------------------------------------------|------|-------------|-------|--------|-------|-------------|
| (4a) 如果有电，灯就亮了。                                  |      |             |       |        |       |             |
| Rúguǒ                                            | yǒu  | diàn,       | dēng  | jiù    | liàng | le.         |
| if                                               | have | electricity | light | will   | on    | <b>PTCL</b> |
| If there was electricity, the light would be on. |      |             |       |        |       |             |
|                                                  |      |             |       |        |       |             |
| (4c) 如果有电，灯就会亮的。                                 |      |             |       |        |       |             |
| Rúguǒ                                            | yǒu  | diàn,       | dēng  | jiùhuì | liàng | de.         |
| if                                               | have | electricity | light | will   | on    | <b>PTCL</b> |
| If there is electricity, the light will be on.   |      |             |       |        |       |             |

$PERF(le)=\lambda P\lambda t.\exists e[P(e)\wedge Completed(e)\wedge \tau(e)\subseteq t]$

$FUT(le)=\lambda P\lambda t.\exists e[P(e)\wedge Initiation(e)\wedge t<\tau(e)]$

$STAT(de)=\lambda P\lambda t.\exists s[P(s)\wedge State(s)\wedge t\subseteq \tau(s)]$

$PAST(de)=\lambda P\lambda t.\exists e[P(e)\wedge Dynamic(e)\wedge \tau(e)\subset t]$

|                                  |        |     |       |                                                            |
|----------------------------------|--------|-----|-------|------------------------------------------------------------|
|                                  | Non-CF | CF  | Total | $\phi\approx-0.186$ indicating a weak negative correlation |
| with sentence final <i>le</i>    | 1346   | 35  | 1381  |                                                            |
| without sentence final <i>le</i> | 1438   | 210 | 1648  |                                                            |
| Total                            | 2784   | 245 | 3029  |                                                            |
|                                  | Non-CF | CF  | Total | $\phi\approx-0.161$ indicating a weak negative correlation |
| with sentence final <i>de</i>    | 948    | 16  | 964   |                                                            |
| without sentence final <i>de</i> | 1836   | 229 | 2065  |                                                            |
| Total                            | 2784   | 245 | 3029  |                                                            |

## Findings: the Role of Modality in Chinese CFs

Non-Conditional CFs → Ambiguous Cases: sentence final le (了) and de (的)

(5a) 他应该昨天来。

|                                     |         |           |      |
|-------------------------------------|---------|-----------|------|
| Tā                                  | yīnggāi | zuótiān   | lái. |
| He                                  | should  | yesterday | come |
| He should have come yesterday. (CF) |         |           |      |

(5b) 他应该昨天来了。

|                                       |         |           |      |      |
|---------------------------------------|---------|-----------|------|------|
| Tā                                    | yīnggāi | zuótiān   | lái  | le.  |
| He                                    | should  | yesterday | come | PTCL |
| He must have come yesterday. (Non-CF) |         |           |      |      |

$$\text{PERF}(le) = \lambda P \lambda t. \exists e [P(e) \wedge \text{Completed}(e) \wedge \tau(e) \subseteq t]$$

# Findings: the Role of Modality in Chinese CFs

Non-Conditional CFs → Ambiguous Cases: sentence final le (了) and de (的)

(6) 他应该昨天来的。

|    |         |           |      |      |
|----|---------|-----------|------|------|
| Tā | yīnggāi | zuótiān   | lái  | de.  |
| He | should  | yesterday | come | PTCL |

He must have come yesterday. (Non-CF)

He should have come yesterday. (CF)

(6a) 他应该是昨天来的。

|    |         |     |           |      |      |
|----|---------|-----|-----------|------|------|
| Tā | yīnggāi | shi | zuotian   | lái  | de.  |
| H  | should  | be  | yesterday | come | PTCL |
| e  |         |     |           |      |      |

He must have come yesterday. (Non-CF)

(6b) 他是应该昨天来的。

|    |     |         |           |      |      |
|----|-----|---------|-----------|------|------|
| Tā | shi | yīnggāi | zuotian   | lái  | de.  |
| He | be  | should  | yesterday | come | PTCL |

He should have come yesterday. (CF)

PAST (de)

[CP Tā [ModP yīnggāi [TP shì  
[zuótiān lái] de]]]

$\square$ should[ $\exists e \wedge \text{lái}(e, \text{tā}, t)$ ]

STAT (de)

[CP Tā [TopP shì [ModP yīnggāi  
[TP zuótiān lái]]] de]

$\square$ should[zuótiān-lái(tā)]  $\wedge$   
static marker



## Discussions: Realis vs. Irrealis in Chinese CFs

The expression of counterfactuality in Mandarin Chinese reveals a fundamental split between Conditional CFs (like 'if...then...') and Deontic CFs (like 'should have...').

In **Conditional** CFs: Weak irrealis conjunctions like *rúguǒ* (如果, "if") don't inherently create 'closeness' to reality. To build this, the language strengthens the premise internally using realis elements. Examples: Factive negation, Perfective *le* (PERF(*le*)), Past-anchoring *de* (PAST(*de*)), and emphatic modals. Function: These elements act like a  $\Box$ P operator, adding epistemic certainty to the hypothetical premise, making it feel closer to a reality that didn't happen.

In **Deontic** CFs: The modal operator like *yīnggāi* (应该, "should") has a built-in ordering source that already ranks worlds against an ideal norm. This inherent 'closeness' to an ideal world favors irrealis elements. Examples: Bare verb forms or Stative *de* (STAT(*de*)). Function: These forms align with the hypothetical, normative nature of the obligation, contrasting the ideal world with the real one.

Key Insight: Conditional CFs in Mandarin are built by internally strengthening a premise with realis markers, while Deontic CFs leverage the inherent irrealis "closeness" of modal operators."

# Discussions: Premise Strengthening in Conditional CFs

"Since 'if' is weak, Mandarin (and other languages) use various strategies to strengthen the premise for a counterfactual reading.

## 1. Aspectual & Emphatic Markers:

PERF(le) and PAST(de) complete the action, grounding the hypothesis.

Emphatic modals like zhēnde (真的, "really") boost the premise's plausibility.

## 2. Syntactic Reinforcement & Lexicalization:

This often involves elements in the C<sup>0</sup> position (sentence-initial).

Factive Negation: Sentence-initial negation frames the premise as a given fact. Combination with Strong Irrealis. Lexicalized CF Conjunctions: In many languages, these strengthening strategies fossilize into dedicated CF worlds.

## Discussions: Premise Strengthening in Conditional CFs

Table 3. Languages with specialized CF conjunctions

| Languages            | CF conjunctions                                                                                                                                         |
|----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| Eton                 | <i>bɛn</i>                                                                                                                                              |
| Polish               | <i>gby</i> = if+subjunctive                                                                                                                             |
| Hausa                | <i>baicin/dàʔ/da-ace</i> = if+negation/earlier/if+subjunctive ‘say’                                                                                     |
| Tagalog              | <i>kundi</i> = if+negation                                                                                                                              |
| Madarin              | <i>yaobushi</i> = if+negation                                                                                                                           |
| Ashéninka Perené     | <i>airorika</i> = if+negation                                                                                                                           |
| Chol                 | <i>machik</i> = negation+irrealis                                                                                                                       |
| Hebrew               | <i>ilu</i>                                                                                                                                              |
| Tamashek             | <i>ənǎɛdr/enǎɛkk</i>                                                                                                                                    |
| Maltese              | <i>kieku</i>                                                                                                                                            |
| Modern Arabic        | <i>law</i>                                                                                                                                              |
| Palestinian Arabic   | <i>law</i>                                                                                                                                              |
| Bulgarian            | <i>da</i>                                                                                                                                               |
| Welsh                | <i>pe(d)</i>                                                                                                                                            |
| Nahuatl              | <i>ma iga</i>                                                                                                                                           |
| Sierra Popoluca      | <i>meʔgaʂ</i>                                                                                                                                           |
| Hindi                | <i>kadācit/kahi</i>                                                                                                                                     |
| Lango                | <i>kónō</i>                                                                                                                                             |
| Maʔdi                | <i>kēsú</i>                                                                                                                                             |
| Savosavo             | <i>monei</i>                                                                                                                                            |
| Haya                 | <i>kubá</i>                                                                                                                                             |
| Samoan               | <i>ana</i>                                                                                                                                              |
| Georgian             | <i>rom</i>                                                                                                                                              |
| Greek                | <i>av/εαν...θα</i>                                                                                                                                      |
| Izi                  | <i>óómé/óbùlé</i>                                                                                                                                       |
| Huave                | <i>pot</i>                                                                                                                                              |
| Cambodian            | <i>oy-tae</i>                                                                                                                                           |
| Chinantec(Sochiapan) | <i>laʒjuah</i> <sup>21</sup> = evidential+‘say’<br><i>sáʒjmí</i> <sup>21</sup> = perhaps+terminative adverb<br><i>laʒjuah</i> <sup>21</sup> = assume+if |

|     |                                |                                                                |               |                        |          |                       |
|-----|--------------------------------|----------------------------------------------------------------|---------------|------------------------|----------|-----------------------|
| (7) | Polish (Indo-European: Poland) |                                                                |               | (Xrakovskij, 2005: 58) |          |                       |
|     | a.                             | <b>Jeśli-by-ś</b>                                              | jutro         | śpiewał,               | to       | ja                    |
|     |                                | <b>if-SBJV-2SG</b>                                             | tomorrow      | sing.IPRF.SG           | then     | I                     |
|     |                                | by-m                                                           | tańczył.      |                        |          |                       |
|     |                                | would-1SG                                                      | dance.IPRF.SG |                        |          |                       |
|     |                                | 'If you sang tomorrow, then I would dance.'                    |               |                        |          |                       |
|     | b.                             | <b>Gdy-by</b>                                                  | już           | wrócił,                | to       | spotkali-by-śmy       |
|     |                                | <b>if-SBJV</b>                                                 | already       | come_back.PRF.3SG.M    | then     | meet.PRF.PL-would-1PL |
|     |                                | się                                                            | z             | nim                    | jutro.   |                       |
|     |                                | REFL                                                           | with          | him                    | tomorrow |                       |
|     |                                | 'If he had returned already, then we would meet him tomorrow.' |               |                        |          |                       |

|     |                                                                    |     |              |           |                                     |     |      |
|-----|--------------------------------------------------------------------|-----|--------------|-----------|-------------------------------------|-----|------|
| (8) | Hausa (Afro-Asiatic: Nigeria, Niger)                               |     |              |           | (Smirnova & Dobronravin, 2005: 435) |     |      |
|     | <b>Da-ace</b>                                                      | ba  | ta           | dauko     | shi                                 | ba  | da   |
|     | <b>if-SBJV</b>                                                     | NEG | PM.3SG.F.PRF | carry_out | OP.3SG.M                            | NEG | CONJ |
|     | watakila.                                                          |     |              |           |                                     |     |      |
|     | probably                                                           |     |              |           |                                     |     |      |
|     | 'If you had come five minutes earlier, you would have found meat.' |     |              |           |                                     |     |      |



## Discussions: Modal Operators in Deontic CFs & Typological Outlook

"Deontic CFs operate differently. The 'closeness' is managed by the modal operator itself.

The modal (e.g., yīnggāi / "should") projects and ranks worlds based on norms. Its closeness function is context-activated. Comparatives like gèng (更, "more") can trigger this, highlighting the gap between the ideal CF world and reality. The Tense-Irrealis Puzzle: Deontic CFs often feel past-oriented but reject overt past markers like PERF(1e). Instead, they use irrealis forms (bare verbs, STAT(de)). The 'past' component is often covert, implied by context. This creates a surface illusion where only irrealis is visible, but the meaning relies on a hidden 'past irrealis' activation.

|     |                                    |                                                                  |              |                                  |          |
|-----|------------------------------------|------------------------------------------------------------------|--------------|----------------------------------|----------|
| (9) | Gooniyandi (Australian: Australia) |                                                                  |              | (McGregor, 1990:548-549)         |          |
|     | a.                                 | Maa                                                              | thoowoorndoo | waj- <b>ja-aladdi</b>            | marlami. |
|     |                                    | meat                                                             | rotten       | throw- <b>SBJV-IRR.1SG.N.CLF</b> | nothing  |
|     |                                    | ‘I should have thrown out the rotten meat, but I didn’t bother.’ |              |                                  |          |
|     | b.                                 | Jail-nhingi                                                      | yoowooloo    | dirib- <b>a-ayoondi.</b>         |          |
|     |                                    | something-ABL                                                    | man          | enter- <b>SBJV-IRR</b>           |          |
|     |                                    | ‘He should go inside (but he doesn’t/won’t).’                    |              |                                  |          |

## Conclusions

English employs a 'Main Trigger' Mechanism. It relies on a small set of highly grammaticalized, explicit markers like "would" and "if...had". This results in a centralized and systematic encoding of CFs.

Mandarin Chinese, by contrast, utilizes a 'Gestalt Feature' Mechanism. In the absence of dedicated markers, CF meaning emerges from a dispersed set of linguistic cues—conjunctions, modals, particles, and contextual inferences—that combine to form a holistic, pragmatic interpretation.

## Conclusions

This distinction is further refined within Mandarin itself, where two distinct paths to achieving 'closeness' to reality are observed:

In Conditional CFs, weak irrealis markers like *rúguǒ* (如果) require internal premise strengthening. This is achieved through realis elements like factive negation, perfective *le*, and past-anchoring *de*, which activate an embedded  $\Box P$  operator to establish closeness.

In Deontic CFs, the modal operator *yīnggāi* (应该) has closeness built-in through its normative base. This favors irrealis elements like the stative *de* and bare verb forms, contrasting the ideal world with reality.



## Conclusions

Future research should explore this 'Gestalt vs. Main Trigger' distinction across a wider range of languages to deepen our understanding of how syntax, morphology, and pragmatics interact to shape our imagination of alternative worlds.

**Thank you.**