

The Factuality Status of Chinese Necessity Modals

Exploring the Distribution Via Corpus-Based Approach

Carlotta Sparvoli

Alma Mater, Università di Bologna, Italia

Abstract This paper is intended to test the deontic vs anankastic hypothesis outlined by Sparvoli 2012. The stipulation is that, in past contexts, deontic modals trigger a counterfactual inference, while anankastic modals (here called ‘goal-oriented modals’) either trigger an actuality entailment effects (‘only possibility’ modals) or a generic non-factual reading (‘mere necessity’ modals). The result of this corpus-based study conducted in a Chinese-English parallel corpus confirm the crucial role played by the deontic vs goal-oriented contrast in the marking of factuality in Chinese and shows that the factuality value decreases across a cline from goal-oriented to deontic modals.

Keywords Actuality entailment. Counterfactuality. Deontic modality. Goal-oriented modality.

Summary 1 Introduction. – 2 Background. – 2.1 The Deontic vs Anankastic Contrast. – 2.2 Modals and Factuality. – 2.3 Counterfactuality and Temporal Orientation. – 2.4 Counterfactuality in Chinese. – 3 Hypothesis and Prediction. – 3.1 Anankastic Strength and Actuality Entailment. – 3.2 The Working Hypothesis. – 3.3 The Prediction. – 4 The Method. – 5 The Study. – 5.1 Keyword 1. *Should Have*. – 5.1.1 Past Counterfactual of Wish. – 5.1.2 Past Counterfactual of Reprimand. – 5.2 Keyword 2. *Had to*. – 5.2.1 Temporal Feature Bleach in Embedded Position. – 5.2.2 Unexpected Data. Backshift in First-Person Narrative. – 5.3 Distribution of the 要 *yào* Tokens. – 6 Conclusion.

1 Introduction

The framework here adopted relies on the differentiation between deontic and anankastic modalities. Based on von Wright (1963), this theory postulates that modals pertaining to duty and necessity are distributed within a semantic domain having two poles (Sparvoli 2012): namely, the *deontic*, which expresses an obligation (ancient Greek *déon*) and is related to a moral duty, grounded on a principle, as in (1a); and the *anankastic* (from *anánkē*, literally ‘rope, wire’), which indicates a practical necessity, linked to a specific purpose, as in (1b).

1. a. ‘We should be modest and prudent’
(translated from Alleton 1984, 200)
[deontic]
- b. *To get to the station you have to take bus 66.*
(Van der Auwera, Plungian 1998, 80)
[anankastic]

Anchored in the notion of ‘inevitability’, the anankastic expresses what ‘cannot be done otherwise’ and makes it possible to establish a unique and consistent class for expressions which are commonly related to different modalities, such as the necessity depending on natural law, circumstances or a given goal (or wish). Rough equivalents of the anankastic modality are found in the “participant-external non-deontic” (Van der Auwera, Plungian 1998), the “goal-oriented or teleological” (von Fintel, Iatridou 2007) and in the “neutral” or “circumstantial dynamic” modality (Palmer 1990). Importantly, the anankastic domain includes markers of different binding force, ranging from weak to strong anankastic modals (as ‘must’ and ‘cannot but’, respectively) (Sparvoli 2012).

Along these lines, this paper focuses on the factuality¹ reading triggered by Chinese modals in past contexts. The working hypothesis is that (i) deontic modals such as 应该 *yīnggāi* ‘should’ yields counterfactuality, that is, they trigger the inference that “the speaker believes a certain proposition not to hold” (Iatridou 2000, 231) and such meaning is understood via an inference; (ii) the strongest anankastic modals, such as 不得不 *bùdébù* ‘cannot but’ or 只好 *zhǐhǎo* ‘can only’, trigger an uncancellable inference that the event took place in the actual world, therefore they are implicative, yield actuality entailments (Bhatt 1999; Hacquard 2006) and have a factual reading; (iii) 必须 *bìxū* ‘have to’ preferably gets a factual interpretation; (iv) weaker anankastic modals, such as 得 *děi* and 要 *yào* ‘must’, have a distribution similar to imper-

¹ For an account of equivalent labels of ‘factuality’, such as ‘actuality’, see Giannakidou, Mari 2016, 82.

fective modals in French or Italian, thus they are not implicative and are compatible with both counterfactual and factual interpretations.

This hypothesis, already outlined in Sparvoli (2012), will be explored through a corpus-based study. To facilitate the identification of Chinese modals in past contexts, we selected the most prominent English (counter)factual necessity markers, respectively, *should have* and *had to*, to then identify the Chinese equivalent in the bilingual token thus retrieved. We browsed two subsets of the *E-C English-Chinese Parallel Concordancer*, published by the Hong Kong Institute of Education,² namely, the *E-C English Novels* (0.807 million words) and the *E-C Chinese Novels* (0.181 million words). In total, we processed 795 tokens and manually tagged the valid ones (527) against five types of eventualities (counterfactual, factual, habitual, non-factual in matrix position, non-factual embedded). Finally, we filtered the tokens including modal markers (387) for analysing their distribution across those types of eventualities.

§§ 2, 3, 4 and 5 illustrate, respectively, the theoretical framework, the prediction, the method and the study. The results show that the factuality reading of Chinese modals of duty and necessity is gradient: it extends from a unique factual reading for strong anankastic modals to a unique counterfactual reading for the deontic. Between these two poles are located the weaker anankastic modals, which can also have habitual reading and thus have a similar distribution of the imperfective form of the Italian *dovere*.

2 Background

2.1 The Deontic vs Anankastic Contrast

Though interchangeable in a positive context, the classification into deontic or anankastic modality is based on the different interaction with negation (Sparvoli 2012). Namely, the negation of a prominent³ deontic marker produces a Prohibition, like ‘should not’, while the negation of the anankastic produces an Exemption, like ‘don’t have to’, ‘need not’. In other words, deontic modals scope over negation, while anankastic modals scope under negation (Lü [1942] 1944). In Chinese, the categorisation into either one of these two modalities, though expressed in different terminology, is already found in the modality in-

² Further details on the corpus are provided in the Bibliography.

³ The underlying principle of the concept of “modal prominence” (Li 2004, 176) is that the different modal meanings of polysemous markers can be ranked into four categories: namely, prominent markers (that is, prototypical, as for 应该 *yīnggāi* in the deontic and epistemic modalities); frequent but non-prominent; non-frequent; not used.

vestigation prior to 1949 (Sparvoli 2012). In this literature, the prominent markers of these two modalities are the deontic (应)该/当 (*yīng*) *gāi/dāng* ‘should’ (2a), and the anankastic 必须 *bìxū* ‘must’ and 得 *děi* ‘have to’ (2b); the latter two are positive polarity items, negated via suppletive forms expressing Exemption, like 不必 *búbì*, 无需 *wúxū* ‘don’t have to’ or 不用 *bùyòng*, 甭 *béng* ‘need not’.⁴ The classification of 要 *yào* is more difficult, since it can have the meaning of 必要 *bìyào* ‘must’, 需要 *xūyào* ‘need’, 想要 *xiǎngyào* ‘would like to’, 快要 *kuàiyào* ‘is going to’, or 将要 *jiāngyào* ‘will’ (Li 2004, 162). In a normative context, following von Wright, who “classified ‘must’ as anankastic but ‘must not’ as deontic” (1963, VIII-2, 157), we labelled 要 *yào* as a weak anankastic and 不要 *búyào* as a deontic. It must be noted that, in this corpus-based study (see Chart 1),⁵ 要 *yào* also occurs as a dynamic marker, indicating some “necessity internal to a participant engaged in the state of affairs” (Van der Auwera, Plungian 1998, 80), as in (2c).⁶

2. a. 我们应⁴该/应当谦虚谨慎。 (Alleton 1984, 200) [deontic]
wǒmen yīnggāi/yīngdāng qiānxū jǐnshèn
 we should be.modest be.prudent
 ‘We should be modest and prudent’.
- b. 去火车站得⁴坐第六六路公共汽车。 (Li 2004, 107) [anankastic]
qù huǒchē-zhàn dēi zuò dìliùliù lù gònggōngqìchē
 go train-station have.to sit 66 CLF bus
 ‘To get to the station you have to take bus 66’. (Van der Auwera, Plungian 1998, 80)
- c. 鲍里斯每晚要⁴睡十个小时才能正常活动。 (Li 2004, 107) [dynamic]
Bàolǐsī měi wǎn yào shuì shí ge xiǎoshí
 Boris every night need sleep ten CLF hour
cái néng zhèngcháng huódòng
 then can normally function
 ‘Boris needs to sleep ten hours every night for him to function properly’. (Van der Auwera, Plungian 1998, 80)

⁴ Concerning the status of 得 *děi*, Lü Shuxiang clarified that the negative form of 得 *děi* is 不用 *bùyòng*, 甭 *béng* ‘need not’: “[*děi* 得] 表示否定用‘不用、甭、不能用’不得” (1984, 143). In other words, in Chinese linguistics prior to 1949, the homograph 得 is considered to have three distinct forms, *dé*, *de* and *děi*, wherein the latter surfaced only in Modern Chinese; such later usage can be also considered as a “second split” in the grammaticalisation process of the lexical verb *dé* ‘to obtain’ (Ziegeler 2003, 251).

⁵ In this study, 要 *yào* also occurs with volitional or futurity readings, especially when retrieved with the token *should have* [tab. 5].

⁶ The glosses follow the general guidelines of the Leipzig Glossing Rules. Additional glosses include: BA = ‘preposition introducing the object in the *ba*-construction’; DE = ‘structural particle *de*’; INC = ‘inchoative’; SFP = ‘sentence-final particle’.

In a cartographic perspective, adjusting our terminology and taxonomy into Tsai's (2015) proposal, the anankastic 必须/要 *bìxū/yào* are hosted in the inflectional layer, between the outer and the inner subject, while the deontic 应该 *yīnggāi* is hosted in the complementiser layer, as its epistemic counterpart. Finally, Sparvoli (2012) identified a set of symmetrical traits of the deontic/anankastic contrast. In this context, the more relevant is related to the different behaviour in perfective contexts, where anankastic modals trigger actuality entailment while the deontic get a counterfactual reading. This corpus-based study is therefore aimed at testing this stipulation, but before presenting the method and the results, we need to present the issue related to the factual reading of modalised expression and introduce the notion of 'actuality entailment'.

2.2 Modals and Factuality

Since Kiefer (1987) and Chung and Timberlake (1985) and, even before, with Lü Shuxiang ([1942] 1944, 187), modality has been related to the notion of 'non-factuality', implying that when an eventuality is *possible* or *necessary*, it is by default *non-factual*. However, the implicative feature of the semi-modal *get* and the lexical verb *manage to* has been identified already by Karttunen (1971), who observed that, in a past environment, sentences like (3a) imply (3b) and express that a given event was actualised; therefore, they are not compatible with a continuation which negates the actualisation of the state of affairs.

3. a. John **managed/got/happened** to solve the problem,
#but he didn't solve it.
b. = John solved the problem.
(Karttunen 1971, 342, 346 slightly modified)

From a typological approach to modality, Van der Auwera and Plungian (1998, 103-4) underscored that most markers, such as *manage*, in the perfective form mark the completion of the process.⁷ From the possible world semantics, Bhatt (1999) describes this phenomenon as "actuality entailment" (hereafter AE), referred to a modalised proposition whose event holds in the actual world. Hacquard (2006) provided a unified account where AE is inferred contextually through the combination of two ingredients: the scopal properties

⁷ Van der Auwera and Plungian (1998, 103-4) classified *manage* as a demodalised marker expressing participant-internal "actuality" and underscored that most markers of participant-internal actuality, in the perfective form, when paralleled to their imperfective counterparts, mark the completion of the process

of the modal and the identity of the event. If the modal scopes below aspect and the event is anchored in a bound interval, then we have AE. In this way, the actuality implication is analysed not only with reference to Ability – that is, considering perfective ability modals as underlyingly implicative, *à la* Bhatt – but it is also accounted for other root modalities:

modal interpretations that did yield actuality entailments were those with a *circumstantial* modal base (abilities, goal-oriented and pure circumstantials); the ones that didn't were those with an *epistemic* or a (truly) *deontic* interpretation.⁸ (Hacquard 2006, 113)

The circumstantial feature seems to play a crucial role in the actuality reading of modalised expressions in past environment.⁹ Moreover, in languages with perfective-imperfective morphology, a deontic modal occurring with an anankastic interpretation, as *devoir* in (4a), in the perfective form yields AE. In the imperfective form instead (4b), depending on the context and the continuation, it can have a counterfactual, progressive/habitual or generic interpretation (Hacquard 2006, 103).

4. a. *Pour aller au zoo, Jane a dû prendre le train... [#but did not]*
to go to.the zoo Jane **must.PST.PFV** take the train
- b. *Pour aller au zoo, Jane devait prendre le train... [but did not]*
to go to.the zoo Jane **must.PST.PFV** take the train (Hacquard 2006, 14)

In Hacquard's framework, the implicative reading arises from the perfective aspect outscoping the modal. More specifically, aspect starts as an argument of the verb and moves out yielding two nodes of type *t*: TP and VP. This allows a root modal to appear either right above TP or right above VP, with aspect moving right above the modal (Hacquard 2017, 52).¹⁰ When low, the modal is bound by the aspect of the VP event; when high, it is bound by the speech event or, in embedded contexts, by attitude events. This, in turn, implies that

⁸ Hacquard (2006, 41) uses the label 'real' deontic with reference to someone granting permission or imposing an obligation on someone else.

⁹ The circumstantial reading is also underscored by Van der Auwera and Plungian (1998, 103-4) with reference to participant-internal actuality.

¹⁰ For a cartographic account on the scopal property with respect to aspect of Chinese modals, see Tsai 2015.

in each configuration, the modal has different relational time: it is anchored, respectively, to the event time, the utterance time and the attitude time. As a result, AE effect is not expected when the modal occurs in embedded sentences. In § 5.2, we will take into account this feature while discussing our results concerning the tokens in embedded position (shown in chart 2).

2.3 Counterfactuality and Temporal Orientation

For both Bhatt (1999) and Hacquard (2017), the lack of AEs in imperfective modals, as in (4b), is due to an additional layer of modality associated with the latter. In Reischenbachian terms, the difference between perfective and imperfective aspect is accounted for with reference to the specular relation between reference and event time whereby the perfective locates the event within the *reference time*, whereas the imperfective locates the reference time within the *time of the event*, hence its typical features of ongoingness, repetition, and regularity. We do not need to discuss here in more detail the perfective/imperfective contrast, but we should recall that the imperfective morphology can give rise to a number of different readings, such as the *progressive* and *non-progressive continuous* interpretations, the *habitual* (including generic/dispositional meanings), and also the *circumstantial habitual*. The latter encompasses “a type of discourse in which a type of setting is first introduced, and then sequences of events that typically occur within that setting are enumerated” (Carlson 2012, 838).

Moreover, in modalised expressions, the imperfective can trigger a past counterfactual interpretation. Generated by the opposite inference of AE, the counterfactual reading conveys that “the speaker believes a certain proposition not to hold” (Iatridou 2000, 231); a counterfactual interpretation implies that the situation at stake has already been ‘settled’, and that such an (unactualised) state of affairs cannot be reversed. In other words, past counterfactual modals tell us how the world *should* or *could have* turned out to be, if a state of affairs had obtained (Condoravdi 2002), as in (5):

5. At that point he **should/might** (still) have won the game but he didn’t in the end. (Condoravdi 2002, 62 slightly modified)

As emphasised by Condoravdi, (5) conveys that “we are now located in a world whose past included the (unactualised) possibility of his winning the game” (2002, 60); in general terms, *should have* expresses that it is “necessary at the present moment that a certain state of affairs obtained in the past” (60) and is thus compatible with both the epistemic and counterfactual interpretation. The latter reading

stems from a future temporal orientation of the modal combined with a past perspective, that is, its reference time is an interval “starting at some past time and extending to the end of time” (75). These elements point to a future-in-the-past orientation of the counterfactual construal. Now that we have set the main coordinates of the theoretical framework, we can turn our attention to the language-specific issues related to counterfactual and AE in Chinese, which will be addressed, respectively, in §§ 2.4 and 3.

2.4 Counterfactuality in Chinese

Since Bloom (1981), the investigation on the encoding of counterfactuality in Chinese (Nevins 2002; Jiang 2000, 2019a; Yong 2016; Jing-Schmidt 2017; Liu 2019, among others) has been primarily focused on counterfactual conditionals, as in (6). Using the terminology adopted in § 2.2, we could say that in these constructions, the antecedent conveys an hypothesis (as ‘it had rained yesterday’) which is opposite to what happens (or happened) in reality; the consequent instead states what would or would have turned out to be, if that state of affairs had obtained (that is, ‘I would have gone’ in (6)).

6. 要是昨天下雨了,我(就)回去。(Liu 2019, 41)
yàoshi zuótiān xià yǔ le wǒ jiù huí qù
if yesterday fall rain SFP I (then) return go
‘If it had rained yesterday, I would have gone’.
NOT: *‘If it rained yesterday, I will go’.

While in Indo-European languages the reality status of each proposition is typically signalled through tense morphology, the Chinese encoding of counterfactuality can hardly be captured by a clear-cut syntactic account. The relevant literature has in fact shed light on the role of the combination of hypothetical conjunctions like 要不是 *yàobushi* ‘were it not for’ with other markers, such as the aspectual and the sentence final particle 了 *le*, the temporal marker 早 *zǎo* ‘early’, negative operators or discourse markers such as 真的 *zhēnde* ‘really’.¹¹ Due to the diverse elements at stake, the investigations on Chinese counterfactual conditionals are characterised by a constructionist approach and typically aim at producing a pragmatic or semantic account, without relying on a specific syntactic derivation. This composite scenario is described as a “cluster of unnoticeable weak features or lexical items that contribute, sometimes jointly, to reaching of counterfactual meaning” (Jiang 2019b, 283). For instance, in (7),

¹¹ For a detailed account of this topic, see Jiang 2019b, 284 ff.

we have the combination of a conditional conjunction 要是 *yàoshi*, a past time-reference and the distal 那个 *nàge* ‘that’ which contributes to locating the event in a hypothetical past event. As observed by Jiang, by replacing it with the proximal 这个 *zhège*, the sentence could be interpreted as “if this free-kick is in, the match will go into overtime” (285). The subtle, though essential, contribution of the distal 那个 *nàge* is thus a good example of what is meant by ‘weak feature’, that is, a feature which is neither sufficient nor essential but yet contributes to the ‘construction’ of the counterfactual interpretation.

7. 要是那个任意球罚进了,就会踢加时赛了。(Jiang 2019b, 285)
yàoshi nà-ge rènyì-qiú fá-jìn le jiù huì tī
if that-clf free-kick shoot-in SFP hence will kick
jiā-shí-sài le
extra-time-match SFP
‘If that free-kick had been in, the match would have gone into overtime’.
NOT: *‘if this free-kick is in, the match will go into overtime’.

Despite this ‘weak feature’, unified accounts are being formulated, especially with reference to past counterfactuals, which, starting from Ziegeler, are considered as the only environment in which the “counterfactual construal can be obtained reliably” (2000, 104), as in (6). Similarly, Liu (2019) stressed the role of the combination of the past time reference and the conditional setting, while Jiang (2019a) highlighted the “tense mismatch” which locates the event in a hypothetical past, obtained either by pointing to a relative tense (as in 7) or by the use of time adverbs as 早 *zǎo* ‘early’.¹² It must be emphasised that the proposals above are consistent with Condoravdi’s emphasis on the combination between a past perspective and a future temporal orientation of the modal, as the aspectual 了 *le* in the antecedent, and 会 *huì* in the consequent, in (6) and (7).

In a corpus-based approach, Yong (2016) shed light on the correlation with past-oriented temporality, negation, emphatic modal adverbs, optative mood, first person pronouns, and demonstratives. Focusing on the pragmatic dimension, Jing-Schmidt (2017) paired a set of discourse functions with five bi-clausal hypothetical constructions and provided an analysis of the co-occurring modality markers, including modal verbs, adverbs, and modal particles. Based on 3,698 tokens of 要不是 *yàobushi*, she singled out 35 modal items (Jing-Schmidt 2017, 37) wherein the two highest ranked expressions are the futuri-

¹² Jiang (2019a) also mentioned a second type of encoding of counterfactual conditional, having impossible or absurd antecedents, where the counterfactual meaning is only triggered by a ‘pure inference’, but those instances are not relevant in the context of this paper.

ty markers 不会 *búhuì* ‘won’t’ and 会 *huì* ‘will’.¹³ Further discussion is in order on the contribution of 会 *huì*, which can be classified either as a futurity marker or, following Jing-Schmidt, as a speaker stance marker signalling ‘epistemic certainty’. In the discussion of current data, we will address this topic in § 5.1. Here we need to recall that Jing-Schmidt observed that those 35 modal combinations uniformly signal speaker stance; thus, she emphasised the evaluative nature of this construal, describing it as the result of the idiosyncratic combination of different counterfactual ingredients.

To conclude, in the study on Chinese counterfactual, the issue of the contribution offered by necessity modals is addressed only peripherally. Importantly, Feng and Yi (2006), following Wu (1994), included 原来应该 *yuánlái yīnggāi*, glossed as ‘should have been’, among the markers used to elicit a counterfactual reading by the participants in their study; for two out of three respondents, the deontic modal preceded by 原来 *yuánlái* proved to be the most productive marker, triggering counterfactual reading in 92% of the 200 statements. This result directly leads us to the working hypothesis of present studies.

3 Hypothesis and Prediction

3.1 Anankastic Strength and Actuality Entailment

We propose that in Chinese, in past contexts, deontic and anankastic modals can be a likely index of the (counter)factual reading (Sparvoli 2012).¹⁴ For outlining our proposal, we will start by focusing on the factuality reading of necessity modals in past contexts.

In a formal semantic perspective, Chen (2012) observed a lack of AE of 应该 *yīnggāi* and 必须 *bìxū* due to a covert prospective aspect of Mandarin deontic and anankastic (in her terminology, “goal-oriented”) modals. From a typological framework and based on the semantic contents of the notional ideas underlying modalities, our working hypothesis is that AE effects are correlated to the modal prominence of the necessity marker: it is high with anankastic markers and it is

¹³ Jing-Schmidt labels them as “modals that express high epistemic certainty” (2017, 36). In the framework entertained here, futurity is a post-modal marker (Van der Awera, Plungian, 1998, 194 ff.), developed from epistemic necessity (Li 2004, 256).

¹⁴ As an anticipation of this claim, cf. Alleton 1984 and Myhill, Smith 1995, 266, who underscored the counterfactual value played by 该 *gāi*. For a diachronic account, cf. Meisterernst 2017. Liu (2019) also suggested the need for more investigation on the role of modality in the making of counterfactual reading.

null with pure deontic ones (Sparvoli 2012, 2015).¹⁵ Our framework suggests that full-fledged AE is typically found with negative forms or forms combined with the exclusive focus marker 只 *zhǐ* ‘only, just’ (Sparvoli 2019). Regarding the latter, it must be stressed that:

表示可能的词，加一“只”字，如“只能”、“只好”、“只得”、“只会”，把他的可能性缩小，就成为表示必要或必然。

By adding the character 只 *zhǐ* before words expressing possibility, as in 只能 *zhǐnéng*, 只好 *zhǐhǎo*, 只得 *zhǐdé*, 只会 *zhǐ huì*, their possibility feature is reduced, and they are turned into expressions of necessity or certainty. (Lü Shuxiang [1942] 1944, 256)¹⁶

As emphasised by Li Renzhi, in these cases we do not have a real semantic shift into the necessity domain, but rather the extension of a possibility expression “to its extreme” (2004, 190). The underlying principle is that there is a continuum from possibility to necessity. Along the same lines, we propose a cline from deontic to strong anankastic modals, based on their anankastic strength.

Table 1 Anankastic strength of necessity modals (Sparvoli 2012, 217; 293)

Anankastic strength	Necessity modal	Modality	Logic implication in a conditional period
++++	不得不 <i>bùdébù</i> , 非得 <i>fēiděi</i> , 只好 <i>zhǐhǎo</i>	Anankastic necessity	Only possibility
+++-	必须 <i>bìxū</i>		Mere necessity
++--	得 <i>děi</i>		(necessary condition)
+---	要 <i>yào</i>		Sufficiency condition*
----	应该 <i>yīnggāi</i>	Deontic necessity	Simple implication, alternatives are available

* Typically, bouletic meaning in the antecedent of a conditional period. In the consequent it typically occurs combined with the focus marker 只 *zhǐ* expressing sufficiency condition. For a more detailed account of the different modal distribution in conditional construction, in combination with 才 *cái* and 就 *jiù*, see Sparvoli 2012, 273 ff.

¹⁵ Sparvoli (2019) suggests that the occurrence of AE in the negative form points to an aspectual coercion, arguably the neutralisation of the modal prospectivity feature, triggered by the negation.

¹⁶ Unless otherwise indicated all translations are by the Author.

3.2 The Working Hypothesis

We have seen that, with a circumstantial reading, the perfective forces the complement to hold in the actual world (Hacquard 2006, 14), and that an imperfective modalised form is typically compatible with a counterfactual, habitual/circumstantial, progressive, and generic reading. In Chinese, morphological tense marking is not available, while anankastic and deontic modalities are lexicalised in two sets of items displaying opposite scopal properties with reference to negation (Lü [1942] 1944; Sparvoli 2012) and aspect (Tsai 2015). The working hypothesis of this paper is that, in such heavily isolating language, the strategy for denoting (counter)factuality could be offered by the shift to a different necessity modal. Practically speaking, a contrast like (4a) and (4b) above would be expressed shifting from a deontic marker, as 应该 *yīnggāi*, 该 *gāi*, 应当 *yīngdāng*, to an anankastic marker, as 不得不 *bùdébù*, 只好 *zhǐhǎo*, 必须 *bìxū*, 得 *děi*. This paper attempts to verify such an hypothesis through a corpus-based study. If confirmed, this proposal would make it possible to outline a tripartite typological classification of (counter)factuality marking:

- a. in languages perfective/imperfective morphology (French, Italian, Catalan, Bulgarian, Greek, Hindi): *mood* and *tense* shift (Hacquard 2006);
- b. in languages lacking perfective/imperfective morphology but having morphological tense-marking (English): both *mood*, *tense* and *modal shift*;
- c. in heavily isolating languages like Chinese: *modal shift* combined with *temporal markers*.

Now we can turn again to the prototypical examples by Hacquard (2006), mentioned in (3-4) and propose their Chinese equivalents as visible in (8), (9) and (10) below.

8. Factual [AE effect, 'Jane did take the train']
 - a. *Pour aller au zoo, Jane a dû prendre le train.*
[Indicative, past perfective, deontic *devoir*]
 - b. To go to the zoo, Jane **had to** take the train.
[Indicative, past, anankastic *have to*]
 - c. (那时候)去动物园珍妮不得不坐火车。
(*nà shíhòu*) qù dòngwùyuán Zhēnnī *bùdébù* zuò huǒchē
that time go zoo Jane cannot.but sit train
[Temporal marker + strongest anankastic marker 不得不 *budébu* 'cannot but']

9. Non-factual [maybe Jane took the train, or maybe not]
- Pour aller au zoo, Jane **devait** prendre le train.*
[Indicative, past imperfective, deontic, *devoir*]
 - To go to the zoo, Jane **would have had** to take the train.
[conditional, past, anankastic *have to*]
 - (那时候)去动物园珍妮得坐火车。
(*nà shíhòu*) qù dòngwùyuán Zhēnnī **děi** zuò huǒchē
that time go zoo Jane need.to sit train
[Temporal marker + anankastic 得 *děi* ‘need to’]
10. Counterfactual [Jane did not take the train]
- Pour aller au zoo, Jane **aurait dû** prendre le train.*
[Conditional, past, deontic, *devoir*]
 - To go to the zoo, Jane **should have taken** the train.
[Conditional, past, deontic *should*]
 - (那时候)去动物园珍妮[本来]应该坐火车。
(*nà shíhòu*) qù dòngwùyuán Zhēnnī [*běnlái*] yīnggāi
that time go zoo Jane originally should
zuò huǒchē
sit train
[Temp. marker + (counterfactual adverbial) + deontic 应该 *yīnggāi* ‘should’]

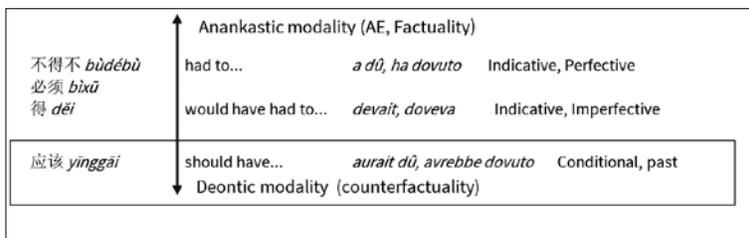


Figure 1 From Counterfactuality to Factuality (Sparvoli 2015)

3.3 The Prediction

Along these lines, the predictions are that: (i) the Chinese equivalents of the counterfactual occurrences of *should have* are marked by pure deontic markers such as (应)当/该 (*yīng*)*dāng*/*gāi* ‘should’, alone or in combination with the counterfactual marker 本(来) *běn*(*lái*); (ii) stronger anankastic markers, such as 不得不 *bùdébù* ‘cannot but’ or

只好 *zhǐhǎo* ‘can only’, are banned in counterfactual environments; (iii) 必须 *bìxū* ‘have to’ preferably gets a factual interpretation; (iv) weaker anankastic modals, such as 得/要 *děi/yào* ‘must’, have a distribution similar to imperfective modals in French or Italian, thus they are compatible with both counterfactual and factual environments, without yielding AE.

Table 2 Prediction: the distribution of Chinese necessity modal in (counter)factual statements

Modal	Counterfactual	Habitual	Non-factual	Factual*
不得不 <i>bùdébù</i> , 非得 <i>fēiděi</i> , 只好 <i>zhǐhǎo</i> , 只得 <i>zhǐdé</i> , 只能 <i>zhǐnéng</i>	x	x	x	√
必须 <i>bìxū</i>	x	√	√	√
得 <i>děi</i>	√	√	√	√
要 <i>yào</i>	√	√	√	√
应该 <i>yīngāi</i>	√	x	x	x

* By factual we intend a proposition that can only be understood as actualised, which would typically happen when we have a modal yielding AE effect.

4 The Method

To test our predictions, we browsed two subsets of the *E-C English-Chinese Parallel Concordancer*. More specifically, we consulted the datasets named *E-C English Novels* (0.807 million words) and the *E-C Chinese Novels* (0.181 million words), wherein each pair of source and target text is aligned at the sentence level. To facilitate the identification of Chinese modals in past contexts, we selected the most prominent English (counter)factual necessity markers (*should have* and *had to*), to then identify their Chinese equivalents in the bilingual tokens thus retrieved. In total, we processed 795 bilingual tokens; after filtering the invalid tokens, the remaining 527 valid ones were tagged against five types of eventualities. Table 3 shows the token distributions and the list of Chinese equivalents encountered for each type of eventuality.¹⁷

¹⁷ The specific distribution of Chinese markers per each eventuality is visible in Chart 2, which provides a comprehensive overview of the results. The distribution obtained for each keyword, separately, is shown in table 5 (*should have*) and table 7 (*had to*).

Table 3 Tokens and types of eventualities

Keywords	E-C subsets	Filtered tokens	Types					Total	
			Counterfactual	Factual	Habitual	Non-factual (matrix)	Non-factual (embedded)		
should have (385)	English novels	151	173 应该/当 <i>yīnggāi/dāng</i> ; 最好 <i>zuihǎo</i> ; 要 <i>yào</i> ; 一定 <i>yíding</i> ; 准 <i>zhǔn</i> ; 不见得 <i>bújiàndé</i> ; 可以 <i>kěyǐ</i> ; 能 <i>néng</i> ; 会 <i>huì</i> ; 本来 <i>běnlái</i> ; 早就 <i>zǎojiù</i> ; other	1 竟然 <i>jìngrán</i>					325
	Chinese novels	34	26 应该/当 <i>yīnggāi/dāng</i> ; 本来 <i>běnlái</i> ; 可以 <i>kěyǐ</i> ; 想要 <i>xiǎngyào</i> ; 须 <i>xū</i> ; other						60
had to (410)	English novels	42		112 不得/能不 <i>bùdé/néng bù</i> ; 只好 <i>zhǐhǎo</i> ; 非得 <i>fēiděi</i> ; 必须 <i>bìxū</i> ; 得 <i>děi</i> ; 需要 <i>xūyào</i> ; 要 <i>yào</i> ; other	18 要 <i>yào</i> ; 得 <i>děi</i> ; 必须 <i>bìxū</i> ; 非得 <i>fēiděi</i> ; other	15 必须 <i>bìxū</i> ; 得 <i>děi</i> ; 要 <i>yào</i> ; 早就 <i>zǎojiù</i> ; 非 <i>fēi</i> ; other	48 必须 <i>bìxū</i> ; 不得不 <i>bùdé</i> ; 应该 <i>yīnggāi</i> ; 需要 <i>xūyào</i> ; 能 <i>néng</i> ; 须 <i>xū</i> ; 得 <i>děi</i> ; other		235
	Chinese novels	41	3 应该 <i>yīnggāi</i> ; 不必 <i>búbì</i> ; other	97 不得/能不 <i>bùdé/néng bù</i> ; 只好/能 <i>zhǐhǎo/néng</i> ; 非不可 <i>fēi bùkě</i> ; 必须 <i>bìxū</i> ; 得 <i>děi</i> ; 要 <i>yào</i> ; 需要 <i>xūyào</i> ; other	7 要 <i>yào</i> ; 会 <i>huì</i>	20 要 <i>yào</i> ; 必须 <i>bìxū</i> ; 得 <i>děi</i> ; 该 <i>gāi</i> ; other	7 应该 <i>yīnggāi</i> ; other	175	
		268	202	207	25	35	58		795

The high rate of invalid tokens (34%, no. 268) is due to the characteristics of the major datasets used in this study. The *E-C English Novels Large Corpus* includes 13 classics from 19th-century English literature and their Chinese translation (typically conducted just before the turn of this century, see Appendix). In that variety of English, the usage of our first token, *should have*, encompassed a heterogeneous range of meanings, thus requiring an attentive process of selection for isolating the relevant tokens (as we will clarify below). Moreover, in that repertoire, even when occurring with a counterfactual meaning, *should have* is often used as an equivalent of *would have*, as in (11), thus providing data related to conditional counterfactuals rather than modalised counterfactual. However, since conditional counterfactuals attract a conspicuous number of deontic modals (Jing-Schmidt 2017), we also included this type of token in the scope of our analysis.

11. “and the effort which the formation and the perusal of this letter must occasion, *should have* been spared, *had not* my character required it to be written and read”. (Jane Austen, *Pride and Prejudice*)
[counterfactual conditional, *should have*=*would have*]

On the other hand, while the sampling size is limited, this repertoire offers the advantage of being easily accessible in full narrative context and in a variety of languages. Focusing on widely translated, easily accessible and relatively familiar classics facilitated the process of disambiguation of the factuality reading. In fact, when necessary, we also double-checked the results of our disambiguation analysing the perfective-imperfective morphology found in the Italian translation of the relevant passage. In this way, we could disambiguate each token in the light of the context of narration, independently from the morphology and the modal classes of the keyword. For instance, (12) was retrieved from the *E-C Chinese Novels* by selecting *had to*; in light of the continuation in full narrative context, the token including 该 *gāi* ‘should’ was tagged in the counterfactual type.

12. The Kianghsi bus did not cross over, so they *had to* transfer to the Hunan bus, which departed at noon.

江西公路车不开过去了, 他们该换坐中午开的湖南公路车。

Jiāngxī gōnglùchē bù kāi guo qu le tāmen gāi
Jiangxi bus NEG drive cross go SFP they should
huàn zuò zhōngwǔ kāi de Húnán gōnglùchē
transfer sit noon depart DE Hunan bus

Continuation: The next morning they arrived at Chiehualung, on the border between the provinces of Kinaghshi and Hunan. The Kianghsi bus did not cross over, so they **had to transfer** to the Hunan bus, which departed at noon. Of all the buses they had taken on the way, none had arrived at a station so promptly as this one; so rather than quarrel about

the short distance they felt that they'd come out a good half-day ahead and **decided to take a night's rest instead of catching the bus that day**. (Qian Zhongshu, *Wei cheng*. Engl. transl. *Fortress Besieged*, 2017, 255)

The token visible in (13), instead, has been retrieved with the key-word *should have* but tagged as factual, given the reading of *should have*, rendered in Chinese with the evaluative modal 竟然 *jìngrán*.

13. "It is astonishing [...] that my heart *should have* been so insensible!"
(Jane Austen, *Sense and Sensibility*)

简直令人吃惊, 我的心竟然那么麻木不仁!

jiǎnzhí lǐngrénchījīng wǒde xīn jìngrán
simply shocking my heart unexpectedly
name mámùbùrén
like.that insensitive
= I was insensitive

The first step in the disambiguation process was filtering all the invalid segments wherein the Chinese target does not correspond to the English source text or vice versa. When possible, we tried to retrieve the correct target segments. A case in point is (11), repeated in (14), which was already mentioned in the previous section. Such a segment has been classified as counterfactual and tagged as a conditional, namely, a case where *should have* is rendered in Chinese with the possibility modal 可以 *kěyǐ* 'can, may' preceded by a hypothetical conjunction.

14. "and the effort which the formation and the perusal of this letter must occasion, *should have* been spared, had not my character required it to be written and read". (Jane Austen, *Pride and Prejudice*)

“我曾经衷心地希望我们双方会幸福, 可是我不想在这封信里再提到这些, 免得使你痛苦, 使我自己受委屈。” Correct match: 我所以要写这封信, 写了又要劳你的神去读, 这无非是拗不过自己的性格, 否则便可以双方省事, 免得我写你读。

Entries wherein *should have* occurs as the conditional of the lexical verb 'to have', as (15), have also been filtered:

15. "As to the future," said the Doctor, recovering firmness, "I *should have* great hope". (Charles Dickens, *A Tale of Two Cities*)

The second step in the disambiguation process was filtering the segments whose reading is not counterfactual. As a point of fact, *should have* does not necessarily force the counterfactual meaning. It can also have an epistemic reading, as in (16a), and, in embedded claus-

es, a deontic meaning (16b). Considering the variety of English offered by the corpus, it also occurs in future-in-the-past interpretations, as in (16c).

16. a. “This wine-shop keeper was a bull-necked, martial-looking man of thirty, and he *should have* been of a hot temperament, for, although it was a bitter day, he wore no coat, but carried one slung over his shoulder”. (Charles Dickens, *A Tale of Two Cities*)
- b. “My mother”, said Monks, in a louder tone, “did what a woman *should have* done”. (Charles Dickens, *Oliver Twist*)
- c. “She had asked him not to leave London on any account, until he *should have* seen her again”. (Charles Dickens, *David Copperfield*)

Moreover, in a substantial group of filtered segments, *should have* has a purely illocutionary function. In these cases, the Chinese rendering relies on discourse markers, such as 我相信 *wǒ xiāngxìn* ‘I think’, as in (17).

17. [“Oh me, oh me!” exclaimed the wretched Emily,]¹⁸ in a tone that might have touched the hardest heart, I *should have* thought. (Dickens, *David Copperfield*)
- [...] 那声音我相信就连最铁石的硬心肠人听了也会被感动的
- | | | | | | | | |
|----------------|----------------------|------------|-----------------|------------|-------------|------------|----------------|
| <i>nà</i> | <i>shēngyīn</i> | <i>wǒ</i> | <i>xiāngxìn</i> | <i>jiù</i> | <i>lián</i> | <i>zuì</i> | <i>tiě-shí</i> |
| that | sound | I | believe | then | even | most | iron-stone |
| <i>de</i> | <i>yìng-xīncháng</i> | <i>rén</i> | <i>tīng-le</i> | <i>yě</i> | <i>huì</i> | <i>bèi</i> | |
| DE | hard-heart | person | hear-PFV | also | FUT | PASS | |
| <i>gǎndòng</i> | <i>de</i> | | | | | | |
| move | DE | | | | | | |

Table 4 Filtered tags (*should have*: All English novels)

Misalignments	Lexical verb 'to have'	Illocutionary	Future-in- the-past	Epistemic	Deontic	Total
28	43	17	45	16	2	151
19%	28%	11%	30%	11%	1%	100%

The segment with future-in-the-past reading covers 30% of the filtered items [tab. 4], and 14% of the entire 325 tokens retrieved from the *E-C English Novels* via *should have*.

¹⁸ In order to provide the contextual information needed for the factuality judgement, we included the relevant source text between square brackets.

5 The Study

5.1 Keyword 1. *Should Have*

In this section, we will first present the data retrieved from the *E-C English Novels*, that is, the English Chinese language combination. The first observation is that the tokens with counterfactual interpretation are embedded in the same environment described in the literature on Chinese counterfactual conditionals (see § 2.4), as 应该 *yīnggāi* in the consequent of a conditional construction, in (18).

18. “Well, sir, I think I *should have* known you, if I had taken the liberty of looking more closely at you”. (Charles Dickens, *David Copperfield*)
 “哦，先生我相信，如果我刚才能看你更仔细些，我应该认出你。”
 ó xiānshēng wǒ xiāngxìn rúguǒ wǒ gāngcái néng
 oh sir I believe if I just could
 kàn nǐ gèng zǐxì xiē wǒ yīnggāi rènchū nǐ
 look you more closely a.bit I should recognise you
 = I did NOT recognise you.

The results of the interrogation show that among the counterfactual tokens retrieved through the keyword *should have*, the most frequent non-epistemic necessity modal is the deontic (应)该/当 (*yīng*) *gāi/dāng*, followed by 要 *yào* and 最好 *zuihǎo*. In the taxonomy, 最好 *zuihǎo* is classified as deontic (Sparvoli 2012, 263), and it can safely be said that among the equivalents of *should have* with counterfactual meaning, anankastic modals are not found.

It also appears that the counterfactual reading is contributed by a number of other markers (see table 5, ‘Non-modals’) that typically occur in counterfactual conditionals, such as conditional conjunctions, focus markers, and temporal deictics that locate the sentence in a past context (Jiang 2000; Jing-Schmidt 2017; Liu 2019, among others).

Table 5 Modal distribution, counterfactual tokens (*should have*, *E-C English Novels*)¹⁹

Domain	Marker	No.	%
Deontic necessity (non-anankastic)	(应)该/当 (<i>yīng</i>) <i>gāi/dāng</i> 'should'	29	17% 18%
	最好 <i>zuihǎo</i> 'had better'	1	1%
Volition + Futurity*	要 <i>yào</i> 'want', 'is going to'	6	3%

¹⁹ Each modal can occur in combination with other counterfactual ingredients, such as a conditional constructions or other markers typically found in Chinese counterfactuals.

Epistemic necessity	一定 <i>yídìng</i> ‘certainly’	6	3%	6%
	准 <i>zhǔn</i> ‘certainly’	3	2%	
	不见得 <i>bújiànde</i> ‘not necessarily’	1	1%	
Non-epistemic possibility	可以 <i>kěyǐ</i> ‘may’	6	3%	11%
	能 <i>néng</i> ‘can’	14	8%	
Futurity	会 <i>huì</i> ‘futurity’	68		39%
Non-modals	本(来) <i>běn(lái)</i> ‘originally’	5	3%	18%
	早就 <i>zǎo jiù</i> ‘earlier, before’	8	5%	
	Conditionals	11	6%	
	Others	7	4%	
Underspecified	ND	8		5%
		173	100%	

* This dataset consists of bilingual segments translated from English into Chinese, obtained with the keyword *should have*; in this type of repertoire, 要 *yào* occurs in sentences with a first-person subject, as a ‘subjective necessity marker’, with volitional or futurity meaning, thus having the meaning of 想要 *xiǎngyào* ‘would like to’, 快要 *kuàiyào* ‘to be going to’, or 将要 *jiāngyào* ‘will’. For a comprehensive account of all 要 *yào* tokens, see chart 3.

The study also confirmed the crucial role of counterfactual chunks (Jiang 2019) like 早就 *zǎo jiù* in (19).

19. “I should have cried out, if I could”. (Charles Dickens, *Great Expectations*)
 如果我能够叫出声, 我早就大叫了起来。
rúguǒ wǒ nénggòu jiào-chu shēng wǒ zǎo jiù
 if I be.capable yell-exit voice I earlier then
dà jiào le qǐlai
 greatly yell PFV start
 = I did NOT yell

The constructionist feature of Chinese counterfactual is well represented by (20), which, paraphrasing Wang and Jiang (2011), displays virtually all the “ingredients of counterfactuality”, in addition to the deontic 该 *gāi*:

20. “I should have said this sooner, but for my long mistake”.
 (Charles Dickens, *Great Expectations*)
 “要不是我一向对这件事情的误解, 我本该早就说了。”
yàobúshì wǒ yíxiàng duì zhè shìqíng de wùjiě wǒ
 if.not.be I always towards this matter DE misread I
běn gāi zǎo jiù shuō le
 originally should early then tell SFP

There are also entries in which the counterfactual meaning is underspecified in Chinese (here signalled with ‘nd’), thus confirming a phenomenon already observed by Yong (2016).²⁰ An example from the present study is (21).

21. “[mimicking his poverty, his boots, his coat, his mother,] everything belonging to him that they **should have had** consideration for”. (Charles Dickens, *David Copperfield*, 242)
- [...] 一切他们注意到的属于他的, 都被他们取笑。
 yīqiè tāmen zhùyì-dào de shǔyú tā de
 all they notice-RES DE belong.to he DE

Importantly, as highlighted by Jing-Schmidt (2017), the futurity marker 会 *hui* is the most common equivalent (39%) of the counterfactual *should have* [tab. 5]. The typical scenario of the occurrence of 会 *hui* is in the consequent of a conditional period. In such an environment, the counterfactual reading is derived by implicature and signalled by a number of *weak features* described in § 2.3, such as a past temporal orientation combining with a negative or adversative presupposition, typically provided contextually or in the continuation of the narration (as in (22)) and, thus, difficult to capture syntactically.

22. “If I could have seen my mother alone, I should have gone down on my knees to her and besought her forgiveness”. (Charles Dickens, *David Copperfield*)
- 如果我可以单独看到母亲, 我会向她跪下, 请求她原谅
 rúguǒ wǒ kěyǐ dāndú kàn-dào mǔqīn wǒ huì xiàng
 if I can alone see-RES mother I FUT towards
 tā guìxia qǐngqiú tā yuánliàng
 she kneel.down plea she forgive
 Further contextual information: “but I saw no one [...] during the whole time” / “可是在那段日子里 [...]我看不到任何人” *kěshì zài nà duàn rìzi li*
 [...] *wǒ kànbudào rènhé rén*.

Jing-Schmidt relates Chinese counterfactuals to the prominence of the epistemic stance of the viewer. While agreeing in the epistemic nuance of futurity as conveyed by 会 *hui*, and in the modal component of the semantic of future in general (Giannakidou, Mari 2016), we prefer to single out the futurity reading from the epistemic certainty. This choice is based on two main reasons. Firstly, 10% of 会 *hui* occurrence

²⁰ In a corpus-based study, Yong (2016) used 13 different hypothetical conjunctions as keywords and, after collecting 3,000 conditionals, disambiguated 245 counterfactuals. Yong’s investigation also includes data from a parallel corpus, observing a tendency towards “counterfactual cancellation” occurring after being translated into Mandarin (Yong 2016, 909, 912).

es are in combination with necessity epistemic markers such as 一定 *yíding* and 准 *zhǔn*, which would confirm classic modal stacking *epistemic necessity* > *futurity* (23). Secondly, even though there are contexts in which 会 *huì* could be interpreted epistemically or even dynamically, as in (23), it could also be argued that without 会 *huì* the event would be anchored to the time of utterance (“I now know what you meant”) rather than to the event time (“at that time, I would have known what you meant”). Paraphrasing Condoravdi (2002), it could be said that 会 *huì* sets the reference time in an interval “starting at some past time and extending to the end of time”. Therefore, in the composite mechanism of Chinese counterfactuality, 会 *huì* expresses how the world *would have* turned out to be if a state of affairs had obtained.

23. a. “If I had never seen Charles, my father, I should have been quite happy with you”. (Charles Dickens, *A Tale of Two Cities*)
“若是我没遇到查尔斯, 爸爸, 我跟你也一定会很幸福的。”
ruòshì wǒ méi yùdào Chárlesì bàba wǒ gēn nǐ
if I not meet Charles dad I with you
yě *yíding huì* hěn xìngfú de
also certainly FUT quite happy DE
- b. “If you [had sent the message, ‘Recalled to Life,’ again,” muttered Jerry, as he turned,] “I *should have* known what you meant, this time”. (Charles Dickens, *A Tale of Two Cities*)
“即使你 [...] 我也会懂得你的意思的。”
jìshǐ nǐ [...] wǒ yě *huì* dǒngdé nǐde
even.though you I also FUT understand your
yìsi de
meaning DE

Moreover, the data also include examples wherein 会 *huì* cannot be spelled out with any other meaning than futurity. A case in point is (24), which refers to the topic of love commitment. The addressee is telling a third person that, even though Estella’s personality had been ruined, had she married him, he would have loved Estella anyway. Our understanding of the sentence in its narrative context is that the speaker’s heart here is crying out “I will always love her”, without the slightest *epistemic weakening* (Giannakidou, Mari 2017).

24. “I *should have* loved her under any circumstances—Is she married?”
(Charles Dickens, *Great Expectations*)
我在任何情况下都会爱她。[她现在结婚了吗?]
wǒ zài rènhé qíngkuàng xià dōu huì ài tā
I in whatever situation under even FUT love she
tā xiànzài jiéhūn le ma?
she now marry PFV Q

In summary, the results suggest that, in past conditionals, 会 *huì* can be considered as the equivalent of *would* in future-in-the-past expressions and that the combination with weak features as the past temporal orientation, the negative presupposition and the first person subject (Ziegeler 2000; Yong 2016) trigger a counterfactual inference.

5.1.1 Past Counterfactual of Wish

The data collected selecting the keyword *should have* in the English-Chinese combination seem to confirm Ziegeler's (2000, 104) claim that: "it is only in past temporal conditionals that a counterfactual construal may be reliably obtained in Chinese". But we also encountered examples where (应)该/当 (*yīng*)*gāi*/*dāng* does not occur in conditional contexts, as in (25). Such examples are labelled as counterfactual wishes, "whereby the subject expresses a desire for things to be different from what they are or were" (Iatridou 2000, 231).

25. "I might have been too reserved, and *should have* patronised her more".
(Charles Dickens, *Great Expectations*)

我是太谨小慎微了。我^{应该}多关怀她，更加地真诚友好
wǒ shì tài jǐnxiǎoshènweī le wǒ yīnggāi duō
 I be too cautious PFV I should more
guānhuái tā gèngjiā-de zhēnchéng yǒuhǎo
 take.care she even.more-ly be.sincere be.friendly

Even though it is clear that no linguistic category is independently responsible for the counterfactual interpretation (just as for any other construction, it could be said), the data also show that by adding an appropriate temporal marker such as 那时候 *nàshíhòu* 'at that time', the shift from counterfactual to factual reading can be obtained by replacing 应该 *yīnggāi* with 只好 *zhǐhǎo*; with the latter an AE effect is triggered and the sentence gets a factual reading (26).

26. a. 我是太谨小慎微了。[那时候]我^{应该}多关怀她
wǒ shì tài jǐnxiǎoshènweī le [nàshíhòu] wǒ yīnggāi
 I be too cautious SFP that.time I should
duō guānhuái tā
 more take.care she
 'I had been too reserved. At that time, I **should have** taken more care of her'.

- b. 我是太谨小慎微了。[那时候]我只好多关怀她,更加地真诚友好
wǒ shì tài jǐnxiǎoshènwēi le nàshíhòu wǒ zhǐhǎo
I be too cautious SFP that.time I can.only
duō guānhuái tā
more take.care she
'I had been too reserved. At that time, I **had to** take more care of her'.

The data from the *E-C English Novels* thus suggest that (i) unlike in anankastic modals, the cluster (应)该/当 (*yīng*)*gāi/dāng* is attracted by conditional counterfactual [tab. 5] and that (ii) (应)该/当 (*yīng*)*gāi/dāng* plays a crucial role in conveying a counterfactual meaning of the 'past wishes' type, as in (26a) and (26b).

5.1.2 Past Counterfactual of Reprimand

More evidence about the contribution of deontic modal in counterfactual environment is found by selecting the keyword *should have* in the *E-C Chinese Novels* (0.181 million words). In this way, we collected 60 tokens from texts originally written in Chinese, and then rendered in English via *should have*. Of the total 60, only 26 have counterfactual interpretation; moreover, in addition to these 26, we also found 5 tokens in which the counterfactual interpretation is present only in the English rendering. Importantly, while processing texts originally written in Chinese and subsequently rendered with the English *should have*, we found that out of 19 tokens including (应)该/当 (*yīng*)*gāi/dāng* only 2 are in conditional constructions. Moreover, in this repertoire, the prevailing *nuance* of the deontic tokens is the expression of reproach or reprimand (16 out 20 tokens) that performs the discourse function described by Myhill and Smith, in which "the speaker expresses dissatisfaction with the listener's failure to do something" (1995, 266). In a past context, this discourse function obtained a counterfactual reading, as in (27a). Though mostly addressed to second-person subjects, the reprimand can also be referred to a third party, as in (27b).

27. a. 方先生,你应该知道出典,你不比我们呀!(Qian Zhongshu, *Wei cheng*)
Fāng xiānshēng nǐ yīnggāi zhīdào-chu diǎn nǐ
Fang mr. you **should** know-RES classics you
bùbǐ wǒmen yā
be.unlike us SFP
'Mr Fang, you *should have* recognised the allusion. You're not like us!'
Continuation: 为什么也一窍不通?你罚两杯,来! Wèishéme yě
yīqiàobùtōng? Nǐ fá liǎng bēi, lái! 'How come you didn't have the
faintest idea about it either? You're fined two glasses. Come on'.

- b. [...] 说鸿渐父亲当初该要求至少两间里有一间大房。
(Qian Zhongshu, *Wei cheng*)
shuō Hóngjiàn fùqīn dāngchū **gāi** yāoqiú zhìshǎo
tell Hongjian father originally **should** request at.least
liǎng jiān li yǒu yī jiān dà fang
two CLF in have one CLF big room
‘[...] commenting that Hung-chien’s father *should have* insisted that
at least one of the two rooms be a large one’.

Table 6 Modality distribution, counterfactual tokens (should have, *E-C Chinese Novels*)

Deontic necessity	Volition + FUT		Possibility	Anankastic	No modal devise		Total
	(应)该/当 (yīng)gāi/ dāng	想要 xiǎng yào	可以 kěyǐ	须 xū	本(来) běn(lái)	other	
16 <i>reprimand</i> counterfactuals	20	1	1	1	1	2	26
4 <i>past wish</i> counterfactuals							
	73%	4%	4%	4%	4%	12%	100%

The distribution of modal markers in the tokens from the *E-C Chinese Novels* attests to the prominence of (应)该/当 (*yīng)gāi/dāng*, present in 20 out of 26 counterfactual tokens (73%). However, contrary to expectations, there is also one anankastic modal, 须 *xū* in (28), occurring in first-person direct speech, in a prose poem by Lu Xun (*死火 Sǐ huǒ*, *Dead Fire*, 1925).

28. 倘使你 不给我 温热, 使我 重行 烧起, 我 不久 就 须 灭亡。(Lu Xun, *Sǐ huǒ*)
tǎng shǐ nǐ bù gěi wǒ wēnrè shǐ wǒ chóng
if cause you NEG to me warm cause me again
xíng shāo qǐ wǒ bùjiǔ jiù **xū** mièwáng
do burn INC I not.long then **must** perish
‘If you had not warmed me and made me burn again, before long I *should have* perished’.

Other unexpected results found in first-person direct speech will be discussed in § 5.2.2.

5.2 Keyword 2. *Had to*

Selecting *had to*, 410 tokens were retrieved from the two datasets. Once filtered the invalid and irrelevant entries (83 in total), we obtained 327 segments in which *had to* occurs with a modal meaning.

The perfective morphology of *had to* does not necessarily force perfective aspect, being also compatible with habitual, generic, and progressive readings. Moreover, as emphasised by Hacquard (2017), AE is typically neutralised when the modalised proposition is an embedded clause (§ 2.2). Along these lines, each entry was manually tagged as *factual*, *habitual/generic/circumstantial*, *non-factual*, or *non-factual (embedded)*, as in table 7.

Table 7 Token distribution for the keyword *had to*

E-C English Novels														
	不得不 <i>bùdébù</i>	不能不 <i>bùnéngbù</i>	只好 <i>zhǐhǎo</i> 只得 <i>zhǐdé</i>	非可 <i>fēikě</i>	必须 <i>bìxū</i>	须 <i>xū</i>	得 <i>děi</i>	需要 <i>xūyào</i>	要 <i>yào</i>	(应)该 <i>(yīng)gāi</i>	不必 <i>búbì</i>	可以 <i>kěyǐ</i>	Other	Tot
Factual	21	3	16	3	21	0	20	1	2	0	0	0	25	112
Habitual	0	0	0	1	2	0	4	0	7	0	0	0	4	18
Non-factual	0	0	0	1	6	0	3	0	1	0	0	0	3	15
Non-factual embedded	6	0	0	0	7	1	1	2	13	3	1	1	14	48
Subtotal	27	3	16	5	36	1	28	3	23	3	1	1	46	193

E-C Chinese Novels														
	不得不 <i>bùdébù</i>	不能不 <i>bùnéngbù</i>	只好 <i>zhǐhǎo</i> 只得 <i>zhǐdé</i>	非可 <i>fēikě</i>	必须 <i>bìxū</i>	须 <i>xū</i>	得 <i>děi</i>	需要 <i>xūyào</i>	要 <i>yào</i>	(应)该 <i>(yīng)gāi</i>	不必 <i>búbì</i>	可以 <i>kěyǐ</i>	Other	Tot
Factual	3	3	25	2	2	0	6	1	11	0	0	0	44	97
Habitual	0	0	0	0	0	0	0	0	6	0	0	0	1	18
Non-factual	0	0	0	0	2	0	2	0	10	1	0	1	4	15
Non-factual embedded	0	0	0	0	0	0	0	0	3	2	0	0	2	48
Counterfactual	0	0	0	0	0	0	0	0	0	2	1	0	0	3
Subtotal	3	3	25	2	4	0	8	1	30	5	1	1	51	134
Total	30	6	41	7	40	1	36	4	53	8	2	2	97	327

We identified 112 tokens having *factual reading*. Excluding 3 tokens with dynamic prominent modals (需要 *xūyào* ‘need’, and 要 *yào* ‘must’), all the other modalised tokens (84 in total) include strong anankastic modals, such as 只好 *zhǐhǎo* in (29).

29. “he *had to* keep swallowing, he was so like to choke”. (Mark Twain, *Tom Sawyer*)

为了嗓子不哽塞住,只好把泪水往肚子里咽。

wèile sāngzi bù gěng sè-zhù zhǐhǎo
for throat NEG choke stop-RES can.only
bǎ lèishuǐ wǎng dùzi lǐ yàn
BA tears to stomach in swallow

Habitual entries also encompass *circumstantial habituals* (see § 2.3), that is, a sequence of events is enumerated within a setting previously created, as ‘cleaning and scraping’, introduced by 要 *yào* with a dynamic necessity meaning, as in (30):

30. “[...] The spoons *had to be cleaned* and the frying-pan *scraped*, and the mugs and pudding-basin **swilled** in the lake”. (Arthur Ransome, *Swallows and Amazons*)

[...] 有汤匙要清洗, 煎锅要刮洗, 还有杯子及布丁盘浸泡在湖里。

yǒu tāngchí yào qīngxǐ jiān-guō yào guā-xǐ
exist spoon need clean frying-pan need scrape-clean
hái yǒu bēizi jí bǔdīng pán jìnpào zài hú lǐ
also exist mugs and pudding basin soak to.be.at lake in

We have included in the habitual class also entries like (31), where an episode is depicted as something happening with a certain regularity (有时 *yǒushí* ‘now and then’) in a given setting. In languages with rich tense morphology, habitual eventualities are typically rendered with the imperfective; therefore, for double checking the reading, when available, we consulted their Italian translation, and found the indicative imperfective of *dovere* ‘must’, which is typically used for expressing a habitual ongoing event in the past, such as *doveva* in (31).

31. “You’d see [a muddy sow and a litter of pigs come lazing along the street and whollop herself right down in the way,] where folks *had to* walk around her”. (Mark Twain, *The Adventures of Huckleberry Finn*)

有时你会看见 [...] 人们走过时必须绕过它走。

yǒushí nǐ huì kànjiàn [...] rénmen zǒu guo shí
sometime you might see people walk pass time
bìxū rào guo tā zǒu
must go.round PASS it walk

*Ecco una scrofa coperta di fango che se ne andava a spasso per la via trotterellando con tutta la figliata dei maialini appresso, e la gente ci **doveva** must.IND.IPFV girare attorno.* (It. transl., 221)

The following is an example of *generic habitual*, expressing a generalisation which obtained some time in the past, as that for the duty of “a common servant” in (32).

32. “[But next minute I whirled in on a kind of an explanation how a valley was different from a common servant and] *had to go* to church [...] *on account of its being the law*”. (Mark Twain, *The Adventures of Huckleberry Finn*)

[...] 他非得上教堂去 [...] 因为这是法律上有了规定的。

tā fēiděi shàng jiàotáng qù [...] yīnwèi zhè shì
he must go church go because this be

fǎlǜ shàng yǒu-le guīdìng de
law on exist-PFV rule DE
Ma un attimo dopo mi sono lanciato in una spiegazione di come un vallet-
to è diverso da un servo qualsiasi, ed **era costretto**^{be.forced to.IND.IPFV} ad anda-
re in chiesa volente o nolente, e a sedersi con la sua famiglia, perché **così**
voleva^{want.IND.IPFV} **la legge**. (It. transl., 266)

(33) is an example of *non-factual reading*. Notwithstanding the perfective morphology in English, the full context reveals that the subject hasn't left the island yet (Ransome [1930] 2012, 486); therefore, the entry is tagged as *non-factual*.

33. “Besides, she *had to* say good-bye to the island”. (Arthur Ransome, *Swallows and Amazons*)

而且, 她也必须和小岛说再见。
érqiě tā yě **bìxū** hé xiǎo dǎo shuō zàijiàn
beside she also must with small island say goodbye

A considerable number of entries (tagged as ‘others’) are not modalised and convey factuality through other means, such as resultative constructions, perfective 了 *le* and the focus marker 才 *cái* ‘only then, not until’, as in (34).

34. 这是远绕了三里路才找到的。(Lu Xun, *Bēn yuè*)

zhè shì yuǎn rào le sānshí lǐ lù **cái**
this be far go.round PFV thirty li road only.then
zhǎodào de
find DE
'I *had to* go an extra thirty *li* to find it'.

5.2.1 Temporal Feature Bleach in Embedded Position

The eventuality types observed for deontic and anankastic modals in embedded position are in line with predictions (i) and (ii): as an equivalent of *had to*, 应该/当 *yīnggāi/dāng* is found only in this environment in which the AE effect is not triggered (cf. Hacquard 2017, 52; see § 2.2). In these cases, modals retain their non-factual orientation and their specific flavour, as for (35), where 该 *gāi* has a fully-fledged deontic reading without shifting to counterfactual reading.

35. “[Nor, did I look towards Wemmick] until I had finished all I *had to* tell”. (Charles Dickens, *Great Expectations*)

一直等我吃完了我该说的话
yízhí děng wǒ shuō-wán le wǒ **gāi**
straight.to wait I say-RES PFV I should

shuō de huà
say DE word

Similarly, in the same environment, the strongest anankastic modals occur without triggering AE, as in (36), having a futurity temporal orientation, as confirmed by the past conditional in the Italian translation.

36. I walked the last mile, **thinking** as I went along **of** what I *had to* do.
(Charles Dickens, *David Copperfield*)

我边走边考虑我不得不去做的事
wǒ biān zǒu biān kǎolù wǒ bùdébù qù
I while walk while think I cannot.but go
zuò de shì
handle DE matter

*Percorsi a piedi l'ultimo miglio pensando, lungo il cammino, a quello che
avrei fatto^{do.PST.COND} (It. transl., 749)*

Another interesting phenomenon is related to the counterfactual reading of 不必 *búbì* in past contexts, as an equivalent of ‘would not have had to’. Just as all the modals triggering AE are possibility markers combined with the negation or with the focus marker 只 *zhǐ*, in a similar and symmetric way, the anankastic negation 不必 *búbì* ‘there is no need to’ seems to yield a counterfactual reading. This is another element pointing to the role of focus-sensitive operators in the expression of factuality and counterfactuality (Sparvoli 2019), a topic that will need to be discussed separately.

5.2.2 Unexpected Data. Backshift in First-Person Narrative

Although the modal distribution in the factual domain meets the prediction, we did find one token in which 要 *yào* marks the anankastic modality and obtains a factual reading – recall that in our prediction the weak anankastic 要 *yào* should convey a non-factual meaning, open to both a factual and counterfactual reading, or a habitual reading. The case in point is (37), in which the event, described in a direct speech first-person narrative context, is only compatible with factual interpretation, as it can be inferred by the continuation (‘it produced various effects’) and confirmed by the perfective indicative (*passato remoto*) of the Italian *dovere* ‘must’ (*dovemmo*). Similarly, to the unexpected counterfactual reading of 须 *xū*, (28), it appears that, in first-person direct speech, the reading of necessity modals is elusive.

37. “said Traddles: ‘[...], [after Sarah was restored], we still *had to* break it to the other eight; [and it produced various effects upon them of a most pathetic nature]’”. (Charles Dickens, *David Copperfield*)

特拉德尔说道, “[...] 我们还要告诉其余那八个”

Tèlādélǎ'ěr shuōdào [...] wǒmen hái yào gàosù qíyú
Traddler say we still must inform the.others
nà bā ge
that eight CLF

Protestò Traddles: “[...] *Quando Sarah si fu ripresa, dovemmo*^{MUST.IND.IPFV} *affrontare le altre otto*”.

(Charles Dickens, *David Copperfield*, It transl., 563)

Another unexpected behaviour, again found in a first-person narrative context, is shown in (38) where 非得 *fēiděi* ‘must’ gets a non-factual interpretation.

38. “We’d GOT to find that boat now – *had to* have it for ourselves”. (Mark Twain, *The Adventures of Huckleberry Finn*)

我们得把那条小船找到, 马上找到——非得找来给我们自己用。

wǒmen děi bǎ nà tiáo xiǎochuán zhǎodào
we have.to BA that CLF boat find
mǎshàng zhǎodào fēiděi zhǎo lái gěi
immediately find must find come to
wǒmen zìjǐ yòng
we REFL use

Ora davvero *dovevamo*^{MUST.IND.IPFV} *trovare quella barca – per noi stessi*.
(It. transl., 114)

These phenomena, observed in first-person narrative contexts, could be interpreted as a temporal backshift of the speaker viewpoint. More precisely, in a modalised context, the evaluation of necessity is set back at a past time, that is, in (38), before finding the boat. Along these lines, the AE effect stemming from the strong anankastic is neutralised and the event is described as an ongoing state – as also suggested by the imperfective (*imperfetto*) of the Italian *dove-re* ‘must’ (*dovevamo*).²¹

²¹ Two types of backshifts, in the scenarios of *justification for a past action* and in the *narration context*, have been described by Hacquard (2017, 59) with reference to the epistemic modals.

5.3 Distribution of the 要 yào Tokens

Before presenting our concluding data, we need to focus on the modal distribution of the 要 yào tokens, which surface with five different meanings (see § 2.1). As shown in chart 1, the 要 yào tokens display a set of related behaviours which are consistent with our predictions for the anankastic and with the account by Bhatt (1999), Hacquard (2006) and Tsai (2015) for the dynamic domain. Firstly, the reality status of the segments including 要 yào is evenly distributed in all the types of eventualities, with the most frequent occurrences in habitual reading (34% in matrix position and 24% including embedded tokens). Secondly, the factual reading is mainly visible in the dynamic domain (8 out of 9, 89%); in the anankastic contexts, we only have one token, shown in (37). Thirdly, given the past contexts of all the tokens, 要 yào is compatible with the deontic meaning only in embedded position (see § 2.2); finally, 要 yào gets counterfactual reading only when occurring with a volitional or futurity reading, thus confirming the non-factual feature of this weak anankastic modal.

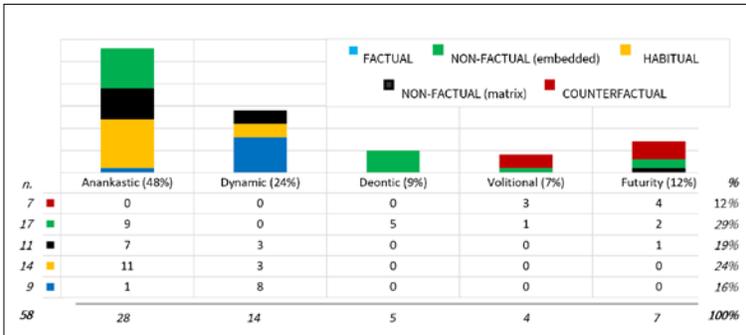


Chart 1 Distribution of 要 yào: Eventuality types per modal reading (58 tokens)

Finally, by aggregating all the data retrieved with the two keywords *should have* and *had to*, we obtained a tentative picture of the factuality reading of 386 tokens including Chinese modals, shown in chart 2.²² By including also modals in embedded position, we could observe that, consistent with what was anticipated in § 2.2, in such an environment strong anankastic modals do not have implicative reading, as in (36), while deontic modals retain their meaning without shift-

²² It should be noted that the data displayed in Chart 2 are the result of a filtering process: from the total of 795 tokens, we excluded 268 non-relevant tokens and, from the remaining 567, we also filtered 141 tokens whose Chinese segment does not include a modal marker, thus obtaining 386 tokens including modals in matrix and embedded position.

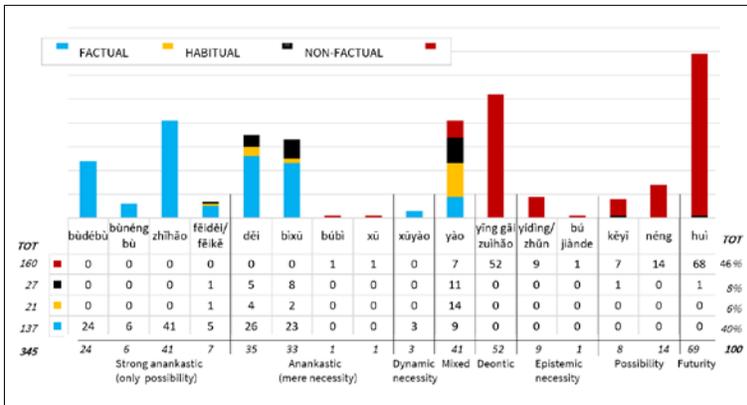
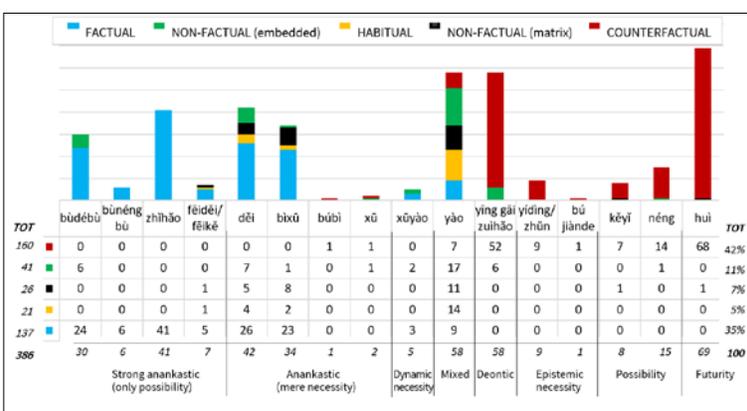


Chart 2 Eventuality types per Chinese modal (386 tokens of *had to* and *should have*)

Chart 3 Eventuality types of Chinese modals in matrix position (345 tokens)

ing to counterfactual reading, as in (35). Finally, to get a clearer picture of the modal distribution per eventuality type, we excluded the tokens in embedded position (41,11%), as seen in chart 3.

6 Conclusion

The results of the aggregated data for modals in matrix position [chart 3] show a gradient cline in which the two extreme poles obtain a unique reading: past counterfactual for pure deontic and factual for strong anankastic modals. In terms of factuality, the modal categories here observed are not discrete. Each class presents one mark-

er that partially overlaps with the adjacent modality. For instance, the distribution of the habitual reading ranges from the dynamic 要 yào (3.14%) to the anankastic 要 yào (11.52%), and can also be seen, albeit less frequently, with other anankastic markers such as 得 děi (4.19%) and 必须 bīxū (2.10%), and even the strong anankastic 非得 fēiděi (1.5%), as seen in (32). Since each modality contains a marker that shares (to a lesser extent) one reading with the adjacent class, the factuality value decreases across a cline from anankastic to deontic modals.

The results confirm our prediction (i): namely, pure deontic markers such as (应)该/当 (yīng)dāng/gāi, alone or in combination with the counterfactual marker 本(来) běn(lái) are the equivalents of counterfactual *should have*. As shown in chart 2, we can see that, out of all 160 tokens with counterfactual meaning, the deontic is the most prominent full-fledged modality, and it allows for counterfactual reading also when occurring without 本(来) běn(lái). However, the counterfactual distribution is twofold. On the one hand, deontic markers prevail in the Wish and Reprimand Counterfactuals retrieved by browsing the texts originally written in Chinese [tab. 6]. On the other hand, the data retrieved from material originally written in English and then translated into Chinese mainly returned counterfactual conditionals wherein the prominent role is played by the futurity marker 会 huì [tab. 5]. This latter result supports the constructionist view of Chinese counterfactual conditionals and points to the prominent role of futurity markers (Ziegeler 2000; Jiang 2000; Jing-Schmidt 2017; Liu 2019, among others). It also attests to a future-in-the-past orientation of the counterfactual construal, thus confirming Condoravdi's (2002) account. In this sense, we could say that, in the typical makeup of Chinese counterfactual conditional, the choice between a possibility modal (能 néng, 可以 kěyǐ), a deontic necessity modal (应)该/当 (yīng)gāi/dāng or a futurity marker (会 huì) tells us, respectively, how the world *could*, *should* or *would* have turned out to be if only the given state of affairs had obtained.

Prediction (ii) stipulated that stronger anankastic markers, such as 不得不 bùdébù 'cannot but' or 只好 zhǐhǎo 'can only', are banned from counterfactual environments. The data confirm this hypothesis, but we must also mention the occurrence of 非得 fēiděi with a non-factual reading. The relevant entry occurs in a first-person narrative context, thus it could be interpreted as a backshift, but we also found one token with generic habitual reading; therefore, it appears that, contrary to the predictions, 非得 fēiděi patterns more with the mere necessity markers than with the only-possibility ones.

We obtained a problematic result for prediction (iii), positing that 必须 bīxū 'have to' preferably gets a factual interpretation. We found one token with a counterfactual 须 xū (first-person direct speech), and the data point to a weaker anankastic strength of 必须 bīxū com-

pared with 得 *děi*. Prediction (iv), on the other hand, is confirmed. In general, mere necessity modals have a distribution similar to imperfective markers in Italian since they are compatible and commonly found in habitual and non-factual sentences. In sum, the data show a slightly different order in anankastic strength, namely, 只好 *zhǐhǎo* > 不得不/不能不 *bùdébù/bùnéngbù* > 非得 *fēiděi* > 得 *děi* > 必须 *bìxū* > 要 *yào*, whereas more data need to be collected for analysing the factuality of 须 *xū*.

Notwithstanding some minor discrepancies with the prediction, the data confirm the crucial role played by the deontic vs anankastic contrast in the marking of factuality in Chinese. Lastly, some pedagogical implications may be emphasised with reference to the equivalents of the tensed forms of the Italian *dovere* ‘must’. Namely, the two poles getting unique factual (只好 *zhǐhǎo*, 不得不 *bùdébù*) and counterfactual ((应)该 *(yīng)gāi* cluster) readings can be mapped onto, respectively, the past indicative and the past conditional of *dovere*; a good candidate as an equivalent of the imperfective of *dovere* can be found in 要 *yào* (especially for direct speech) or 得 *děi*. Finally, the data point to the equivalence between the role of the English *would* and 会 *huì* in past contexts.

Bibliography

E-C Concord (2008). *English Chinese Parallel Concordancer*. <https://corpus.eduhk.hk/paraconcord/search>. The Hong Kong Institute of Education. Project leader: Dr. Wang Lixun. Program designers: Chris Greaves, Wang Lixun.

Primary sources

- Dickens, C. (1859). *A Tale of Two Cities*. London: Chapman & Hall. Transl. by S. Spaventa Filippi as *Le due città*. Roma: Newton Compton editori, 2012.
- Dickens, C. (1872). *The Personal History and Experience of David Copperfield the Younger*. London: Chapman & Hall. Transl. by U. Dettore as *David Copperfield*. Milano: Garzanti, 2012.
- Dickens, C. (1881). *Great Expectations*. Boston: Estes & Laurait. Transl. by M.F. Melchiorri as *Grandi speranze*. Roma: Newton & Compton editori, 2010.
- Lu, X. 鲁迅 (2005). *Lu Xun quan ji. Di er juan* 鲁迅全集 第二卷 (The Complete Works of Lu Xun Vol. Two). Beijing: Renmin wenxue chubanshe. Transl. by X. Yang and G. Yang as *Selected Works. Vol. One*. Beijing: Foreign Language Press, 1957. <https://archive.org/details/in.ernet.dli.2015.184581/page/n1/mode/2up>.
- Lu, X. 鲁迅 [1925] (2005). “Ben yue” 奔月 (The Flight to the Moon). Lu 2005, 370-80. Transl. by Yang, X., Yang, G, 1957, 283-95.
- Lu, X. 鲁迅 [1926] (2005). “Si huo” 死火 (Dead Fire). Lu 2005, 200-1. Transl. by Yang, X., Yang, G, 1957, 342-4.
- Qian Z. 钱钟书 [1947] (2017). *Wei cheng* 围城. Beijing: Renmin wenxue chubanshe. Transl. by J. Kelly and N.K. Mao as *Fortress Besieged*. Bloomington; London: Indiana University Press, 1979.

- Ransome, A. [1930] (2012). *Swallows and Amazons*. London: Random House.
- Twain, M. (1876). *The Adventures of Tom Sawyer*. Toronto: Belford Brothers.
Transl. by L. Bigiaretti as *Le avventure di Tom Sawyer*. Milano: Giunti Editore, 2007.
- Twain, M. (1885). *Adventures of Huckleberry Finn (Tom Sawyer's Comrade)*. New York: C.L. Webster. Transl. by R. Reim as *Le avventure di Huckleberry Finn*. Roma: Newton Compton editori, 2007.

Secondary sources

- Alleton, V. (1984). *Les auxiliaires de mode en chinois contemporain*. Paris: Editions de la Maison des sciences de l'homme.
- Bhatt, R. (1999). *Covert Modality in Non-Finite Contexts* [PhD Dissertation]. Philadelphia: University of Pennsylvania.
- Blaszack, J. et al. (eds) (2016). *Tense, Mood, and Modality. New Perspectives on Old Questions*. Chicago: University of Chicago Press.
- Bloom, A.H. (1981). *The Linguistic Shaping of Thought. A Study in the Impact of Language on Thinking in China and the West*. Hillsdale (NJ): Lawrence Erlbaum Associates.
- Carlson, G. (2012). "Habitual and Generic Aspect". Binnick, R. (ed.), *The Oxford Handbook of Tense and Aspect*. Oxford: Oxford University Press, 828-52.
- Chen, S. (2012). "The Interaction of Modals and Temporal Marking in Mandarin Chinese". Goto, N.; Otaki, K.; Sato, A.; Takita K. (eds), *Proceedings of GLOW-in-Asia IX 2012*. Mie (Japan): Mie University, article 4.
- Chung, S.; Timberlake, A. (1985). "Tense, Aspect and Mood". Shopen, T. (ed.), *Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press, 241-58.
- Condoravdi, C. (2002). "Temporal Interpretations of Modals. Modals for the Present and for the Past". Beaver, D. et al. (eds), *Stanford Papers on Semantics*. Stanford: CSLI Publications, 59-87.
- Feng, G.; Yi, L. (2006). "What if Chinese Had Linguistic Markers for Counterfactual Conditionals? Language and Thought Revisited". Sun, R. (ed.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society (CogSci 2006)*. Mahwah, NJ: Lawrence Erlbaum Associates, 1281-6.
- Giannakidou, A.; Mari, A. (2016). "Epistemic Future and Epistemic MUST. Nonveridicality, Evidence and Partial Knowledge". Blaszack, J. et al. 2016, 75-117.
- Hacquard, V. (2006). *Aspects of Modality* [PhD dissertation]. Boston: Massachusetts Institute of Technology.
- Hacquard, V. (2017). "Modals. Meaning Categories?". Blaszack, J. et al. 2016, 45-74.
- Iatridou, S. (2000). "The Grammatical Ingredients of Counterfactuality". *Linguistic Inquiry*, 31(2), 231-70. <https://doi.org/10.1162/002438900554352>.
- Jiang Y. 蒋严 (2000). "Hanyu tiaojian ju de weishi jieshi" 汉语条件句的违实解释 (On the Counterfactual Reading of Chinese Conditionals). Zhang B. (ed.), *Yufa yanjiu he tansuo, 10 语法研究和探索(十) (Studies and Investigations on Chinese Grammar)*. Beijing: The Commercial Press, 257-79.
- Jiang, Y. (2019a). "Ways of Expressing Counterfactual Conditionals in Mandarin Chinese". *Linguistics Vanguard*, 5(3). <https://doi.org/10.1515/lingvan-2019-0009>.

- Jiang, Y. (2019b). "Chinese and Counterfactual Reasoning". Huang C.-R.; Jing-Schmidt, Z.; Meisterernst, B. (eds), *The Routledge Handbook of Chinese Applied Linguistics*. Abingdon; Oxford; New York: Routledge, 276-93.
- Jing-Schmidt, Z. (2017). "What Are They Good for? A Constructionist Account of Counterfactuals in Ordinary Chinese". *Journal of Pragmatics*, 113, 30-52. <https://doi.org/10.1016/j.pragma.2017.03.004>.
- Karttunen, L. (1971). "Implicative Verbs". *Language*, 47(2), 340-58. <https://doi.org/10.2307/412084>.
- Kiefer, F. (1987). "On Defining Modality". *Folia Linguistica*, 21(1), 67-94. <https://doi.org/10.1515/flin.1987.21.1.67>.
- Li, R. (2004). *Modality in English and Chinese. A Typological Perspective*. Boca Raton: Dissertation.com.
- Liu, H. (2019). "Encoding Counterfactuality in Chinese, Syntactically". *International Journal of Chinese Linguistics*, 6(1), 27-45. <https://doi.org/10.1075/ijchl.18002.liu>.
- Lü S. 吕叔湘 [1942] (1944). *Zhongguo wenfa yaolüe* 中国文法要略 (Outline of Chinese Grammar). Shanghai: Shangwu yinshuguan.
- Lü, S. 吕叔湘 (1984). *Xiandai Hanyǔ babai ci* 现代汉语八百词 (Modern Chinese 800 Words). Hong Kong: Shangwu yinshuguan.
- Meisterernst, B. (2017). "Modality and Aspect and the Thematic Role of the Subject in Late Archaic and Han Period Chinese. Obligation and Necessity". *Lingua Sinica*, 3, art. 10. <https://doi.org/10.1186/s40655-017-0027-2>.
- Myhill, J.; Smith, L.A. (1995). "The Discourse and Interactive Functions of Obligation Expressions". Bybee, J.; Fleischman, S. (eds), *Modality in Grammar and Discourse*. Amsterdam: John Benjamins, 239-91.
- Nevins, A. (2002). "Counterfactuality without Past Tense". Hirotani, M. (ed.), *The Proceedings of NELS 32*. Amherst (MA): GLSA University of Massachusetts, 441-51.
- Palmer, F.R. (1990). *Modality and the English Modals*. 2nd ed. London; New York: Longman.
- Sparvoli, C. (2012). *Deontico e anankastico. Proposta di ampliamento della tassonomia modale basata sull'analisi dei tratti distintivi dei modali cinesi inerenti dovere e necessità* [PhD Dissertation]. Venice: Ca' Foscari University of Venice. <http://hdl.handle.net/10579/1228>.
- Sparvoli, C. (2015). "Actuality Entailment in the Necessity Domain, a Case for Chinese". *9th Conference of the European Association of Chinese Linguistics* (Stuttgart, 25 September 2015). Stuttgart University.
- Sparvoli, C. (2019). "The Temporal Orientation of Chinese Necessity Modals. Some Observations on the Negated Form". *XXXIIèmes Journées de Linguistique d'Asie Orientale* (Paris 29 June 2019). Institut National des Langues et Civilisations Orientales (INALCO).
- Tsai, W.-T.D. (2015). "On the Topography of Chinese Modals". Shlonsky, U. (ed.), *Beyond Functional Sequence. The Cartography of Syntactic Structures*, vol. 10. Oxford: Oxford University Press, 275-94.
- Wang Y. 王宇婴, Jiang Y. 蒋严 (2011). "Hanyǔ weishi yi de goucheng yinsu" 汉语违反语义的构成因素 (The Ingredients of Counterfactuality in Chinese). Jiang, Y. 蒋严 (ed.), *Zoujin xingshi yuyongxue* 走近形式语用学 (Approaching Formal Pragmatics). Shanghai: Shanghai Educational Publishing House, 366-412.
- Wang, Y. (2013). *The Ingredients of Counterfactuality in Mandarin Chinese*. Beijing: China Social Science Press.

- Wu, C.H.-F. (1994). "If Triangle Were Circles,..." – A Study of Counterfactuals in Chinese and English. Taipei: The Crane Publishing.
- Yong, Q. (2016). "A Corpus-Based Study of Counterfactuals in Mandarin". *Language and Linguistics*, 17(6), 891-915. <https://doi.org/10.1177/1606822x16660505>.
- van der Auwera, J.; Plungian, A.V. (1998). "Modality's Semantic Map". *Linguistic Typology*, 2, 79-124.
- von Fintel, K., Iatridou, S. (2007). "Anatomy of a Modal Construction". *Linguistic Inquiry*, 38(3), 445-83. <https://doi.org/10.1515/lity.1998.2.1.79>.
- von Wright, G.H. (1963). *The Varieties of Goodness 1958-1960*. London: Routledge & Kegan Paul. <https://www.jstor.org/stable/40071397>.
- Ziegeler, D.P. (2000). *Hypothetical Modality. Grammaticalisation in an L2 Dialect*. Studies in Language Companion Series 51. Amsterdam; Philadelphia: John Benjamins.
- Ziegeler, D.P. (2003). "Redefining Unidirectionality. Insights from Demodalisation". *Folia Linguistica*, 37(Issue Historica-vol-24-1-2), 225-66. <https://doi.org/10.1515/flih.2003.24.1-2.225>.

Appendices

E-C English Novels

Files included in the consulted corpus.

Wordcount, title, author and translator retrieved from <https://corpus.eduhk.hk/paraconc/info>.

Title	Author and translator	English words	Chinese characters
<i>Alice in Wonderland</i> 爱丽丝梦游仙境	Lewis Carroll. Translator: not mentioned.	7,393	12,924
<i>A Tale of Two Cities</i> 双城记	Charles Dickens. Translator: not mentioned.	136,382	231,837
<i>David Copperfield</i> 大卫·科波菲尔	Charles Dickens. Translators: Shi Dingle 石定乐, Shi Dingrou 石定柔 (1999). 湖南文艺出版社. Changsha: Hunan wenshu chubanshe.	357,668	579,714
<i>Oliver Twist</i> 雾都孤儿	Charles Dickens. Translator: He Wenan 何文安 (1999). 译林出版社. Nanjing: Yilin chubanshe.	159,045	289,200
<i>Great Expectations</i> 远大前程	Charles Dickens. Translator: Luo Zhiye 罗志野 (2001). 译林出版社. Nanjing: Yilin chubanshe.	186,424	364,840
<i>Pride and Prejudice</i> 傲慢与偏见	Jane Austen. Translator: not mentioned.	126,950	225,307
<i>Sense and Sensitivity</i> 理智与情感	Jane Austen. Translator: not mentioned.	120,333	207,710

<i>Swallows and Amazon</i> 小水手探险记	Arthur Ransome. Translators: Lee Hsingchin 李幸瑾; Wang Lixun 王立勋 (2004). 台湾商务. Taipei: Taiwan shangwu.	99,291	177,225
<i>The Adventures of Tom Sawyer</i> 汤姆·索亚历险记	Mark Twain. Translators: Cao Xiaohong 曹晓红; Yu Xiaoguang 于晓光 (1999). 大众文艺出版社. Beijing: Dazhong wenyi chubanshe.	70,622	126,651
<i>The Adventures of Huckleberry Finn</i> 赫克尔贝里·芬历险记	Mark Twain. Translator: Xu Ruzhi 许汝襁 (1998). 译林出版社. Nanjing: Yilin chubanshe.	111,209	199,525
<i>Dog Tale</i> 一只小狗的故事	Mark Twain. Translator: Zhang Yousong 张友松 (2010). 光明日报出版社. Beijing: Guangming ribao chushe.	4,355	8,161
<i>The \$30,000 Bequest</i> 三万元遗产	Mark Twain . Translator: not mentioned.	10,977	19,818
Total		1,390,649	2,442,912

E-C Chinese Novels

Title	Author and translator	English words	Chinese characters
绿化树 <i>Mimosa</i>	Zhang Xianliang 张贤亮. Translator: not mentioned.	46,625	102,820
芙蓉镇 <i>A Small Town Called Hibiscus</i>	Gu Hua 古华. Translator: Gladys Yang (2015). Beijing: Foreign Languages Press.	66,346	136,974
鲁迅小说 <i>Lu Xun's Novels</i>	Lu Xun (鲁迅) Translator: not mentioned.	66,027	94,464
围城 <i>Fortress Besieged</i>	Qian Zhongshu 钱锺书. Translators: Jeanne Kelly, Nathan K. Mao (1979). Bloomington; London: Indiana University Press.	220,000	219,996
Total		398,998	554,254