Syntactic Experimental Evidence on Acceptability Variations of *Wh*-Island Effects in Chinese

Mengkai Wang

School of English and International Studies

Beijing Foreign Studies University

taylericy@bfsu.edu.cn

July 2024

Abstract

Wh-island effects, as an ungrammatical phenomenon that presents two *wh*-phrases at the fronted and the intervener positions in a sentence, have been reported to be ameliorated in terms of acceptability by self-repetition and morpho-syntactic features argued by an influential syntactic theory of *Featural Relativized Minimality*. However, languages low-susceptible to *wh*-island effects like Chinese are poorly verified. This study is aimed to test the previous findings and explore new recommendations for *Featural Relativized Minimality*, employing two acceptability experiments in Chinese to separately examine satiation effects and amelioration variations, both carefully involving morpho-syntactic and semantic features. Fifty participants were recruited to rate sentences with *wh*-island violations. The results show that no amelioration will appear only by self-repetition and that the *wh*-amelioration variations can be accounted for by an incorporation of morpho-syntactic features and semantic representations. The implications of the findings are that more languages and syntactic experimental conditions should be studied from multiple perspectives to enrich the understanding of *wh*-island amelioration.

Keywords: *Wh*-island effects; Acceptability variations; Amelioration; Featural Relativized Minimality; Similarity-based interference

Table of Contents

List of Figures

List of Tables

1 Introduction

Wh-island effects have been investigated by acceptability experiments in much syntactic work (e.g. Atkinson, 2016; Goodall, 2011; Hiramatsu, 2000; Hofmeister et al., 2007; Keshev & Meltzer-Asscher, 2018; Snyder, 2000; Sprouse, 2009). It has been argued that such ungrammatical phenomena can be ameliorated through syntactic satiation or morpho-syntactic devices (Atkinson, 2016; Snyder, 2000; Sprouse et al., 2012, 2013; Villata et al., 2016). They claim that native speakers may entirely accept those ungrammatical *wh*-sentences after reading them several times or increase their acceptability by adding morpho-syntactic features to the extracted *wh*-phrase and the intervener position, predicted by *Featural Relativized Minimality* (henceforth Featural RM) (Belletti et al., 2012; Friedmann et al., 2009; Rizzi, 2013). This theory deserves ample attention because firstly it predicts acceptability and amelioration variations across categories of *wh*-islands unlike the binary division proposed by other syntactic theories and secondly it acknowledges memory constraints in parsing, both of which require further exploration. Additionally, the above findings are most often reported in languages highly susceptible to *wh*-island effects, and some even cannot be successfully replicated. It is still under research how satiation effects and amelioration variation within Featual RM function in low-susceptible languages like Chinese and how they can be compared together through experiments.

Therefore, the present paper is of great significance to test Featural RM through satiation and amelioration variations in Chinese. It firstly reviews literature about the situations of *wh*-island effects and acceptability experiments as well as Featural RM, introducing research questions. Based on research questions, the experimental design is prepared, followed by the results and discussion. At last, the major findings along with any implications or limitations will be summarized, shedding light on *wh*-island syntactic phenomena and further research in generally exploring the integration between syntactic features and memory constraints and what else external can be included into Featural RM.

2 Literature Review

2.1 *Wh*-Island Effects

2.1.1 Definition of *Wh*-Island Effects

In generative syntax exists a phenomenon called *movement*. For instance, if a *wh*-interrogative sentence is generated, the *wh*-phrase is required to move to the beginning of this sentence. Chomsky (1964) found that movement phenomena have distance constraints that the movement is disallowed if a movement element crosses two nodes of the same type, concluded as A-over-A principle. In Figure 1, X can only cross the circled A node and will violate A-over-A principle when it crosses two A nodes to the topmost. A refers to any types of syntactic structures. In (1)b below, *London* performs topicalization (topic movement) required to cross two NP nodes, disallowed by A-over-A principle. If such a movement happens on *the journey to London* as in (1)c, it only crosses one NP node, accepted by the principle.

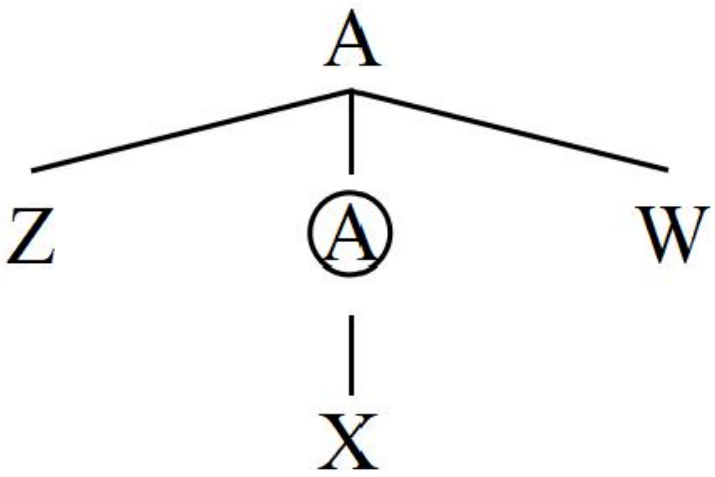
(1) a. I hate [the journey to London].

b. *London$_i$, I hate [the journey to $\_\__i$].

c. [The journey to London], I hate.

Figure 1

A-over-A principle.



Note. Adopted from "*Constraints on variables in syntax*" by Ross, J. R., 1967, Doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

A-over-A principle became the prototype of Island Effect and seemingly explained movement phenomena perfectly. However, it was found that A-over-A principle cannot involve all the situations where movement cannot be performed (Ross, 1967). Ross summarized various structures that cannot be crossed by movement and firstly described these structures as *island*. He thought that all the syntactic operations in a sentence are performed island by island, in which syntactic operations can only transform the interior of islands and the margins of the next island, being invisible of interiors of other islands. On this basis, he almost listed all types of islands, among which *wh*-island defines that subordinate clauses introduced by any *wh*-phrases are all islands. With generative syntactic theories entering *X*-bar era[1], Chomsky (1973) proposed the concept of *subjacency* to specifically explain *wh*-island constraint, yet arguing that all islands that Ross had listed are generated by a syntactic law called *subjacency condition*. This concept was improved and simplified by a number of researchers[2] (Chomsky, 1977, 1986, 1995, 2000, 2001, 2004; Huang, 1982a; Nunes, 1995; Uriagereka, 1999).

Although what subjacency is in generative syntax still remains complicated and uncertain, this concept encouraged to a great extent the development of island-effect studies (e.g.,, Pesetsky, 1987; Rizzi, 1990; Schutze, 1996; Tsai, 1994). The definition of island effect has gradually formed and it thereby refers to syntactic phenomena where a syntactic element is an island and cannot perform movement outside itself, in which *wh*-island effect requires the element to be subordinate clauses introduced by any *wh*-phrases (including *whether*), a generally acknowledged definition highly similar to what Ross proposed.

2.1.2 *Wh*-Island Effects in Chinese

It has been assumed that Chinese is not sensitive to island effects (Cheng, 2009; Huang, 1982a, 1982b; Tsai, 1994). However, Zhang (2002) proposed that island effects in Chinese were highly related to the types of verbs inside the islands, as in (2)[3]. Verbs like *beat* in (2)b more directly influence the objects, causing them to hardly move outside the islands, while

---

[1] *X*-bar era marks a totally different research period from what generative syntax did in the past, when theories and empirical studies on island effects greatly increased.

[2] Here *subjacency* is not required to explain for two reasons: firstly, it is a huge and complicated concept still under research; secondly, it is not concerned with this study’s procedures. See the Context citations for more reference if necessary.

[3] For the sake of readability, all the Chinese characters are romanized in this study.

psychological verbs like *admire* in (2)d exert no physical influence on the objects, hence facilitating their movement. According to Pan's analysis (2017), island effects occur in Chinese, as verbs like beat lacks the object after the movement. Psychological verbs are followed by a zero-pronoun *pro*, preventing the generation of island effects. These studies somehow demonstrate the sensitivity of island effects in Chinese.

(2) a. *[Laoban ouda nage qigai de] shier, chuanbian le zheli.*
The boss beat that beggar REL matter, spread ASP here.
'The matter that the boss beat that beggar was spread here.'

b. **Nage qigai$_i$, [laoban ouda $\_\__i$ de] shier chuanbian le zheli.*
That beggar, boss beat REL matter spread ASP here.
'That beggar, the matter that the boss beat was spread here.'

c. *[Dajia dou xinshang xinlaide laoshi de] chuanwen shi zhende.*
Everyone all admires new teacher REL rumor is true.
'The rumor that everyone all admires the new teacher is true.'

d. *Xinlaide laoshi$_i$, [dajia dou xinshang $\_\__i$ de] chuanwen shi zhende.*
New teacher, everyone all admires REL rumor is true.
'The new teacher, the rumor that everyone admires is true.'

Although there are a number of studies besides the above ones on existence or sensitivity of island effects in Chinese, including studies on *wh*-movement, the only certain thing is that *wh*-island effects exist in Chinese and that there is lack of studies specifically on this type (Lu et al., 2020; Sprouse et al., 2011, 2012; Sprouse & Norbert, 2013). One typical example of *wh*-island effects can be shown in (3) below. When the *wh*-phrase *shenme* replaces *dangao* in (3)b, this is syntactically acceptable because Chinese is not required to perform *wh*-movement to generate an interrogative sentence, in which no element crosses two nodes of the same type, following *wh*-island constraints. If the *wh*-phrase *shenme* is moved to the head of the sentence in (3)c, it becomes syntactically unacceptable, implicating that Chinese is actually dominated by *wh*-island effects.

(3) a. *Ni xiangzhidao shei [chi le dangao]?*
You wonder who eat ASP cake?
'Do you wonder who ate this cake?'

b. *Ni xiangzhidao shei [chi le shenme]?*
You wonder who eat ASP what?
'Do you wonder who ate what?'

c. *$Shenme_i$ ni xiangzhidao shei [chi le $\_\__i$]?*
What you wonder who eat ASP?
'What do you wonder who ate?'

2.2 Acceptability Variation in Syntax

Acceptability variation has been a research topic in generative syntax in recent decades, as it is believed that syntactic acceptability is uncertain but flowing across different structural environments. Schutze (1996) is the first one to systematically point out that linguistics is an experimental science in which syntax requires not only theoretical investigations but experimental evidence. Acceptability judgment paradigm is one of the experimental methods. Schutze (1996) analyzed almost all the linguistic studies adopting acceptability judgment paradigm and proposed one prototype of calculation model to explain what *acceptability* is. Acceptability in syntax thereby refers to judgment reactions towards various sentences in terms of whether they are syntactically correct, and further variation is how acceptability transforms.

However, later studies found that acceptability variation in syntax is greatly influenced by a number of intervening factors such as individual differences, experimental structures, and features of Test sentences, etc. (Chaves & Dery, 2019; Goodall, 2011; Sprouse, 2009; Synder, 2000). Therefore, why acceptability variation transforms and what it is controlled by remain huge issues in experimental syntactic field.

2.3 Theoretical Framework of Featural Relativized Minimality (Featural RM)

In order to better explain island effects, Rizzi (1990) came up with a theory of Relativized Minimality. Friedmann and his co-researchers (Belletti et al., 2012; Friedmann et

al., 2009; Rizzi et al., 2013) revised it into Featural Relativized Minimality (henceforth Featural RM) to focus on studying *wh*-island effects. There are two core components of Featural RM. Firstly, Featural RM argues that great acceptability variations happen across different types of *wh*-islands and particularly predicts how the acceptability of *wh*-island violations can ameliorate depending on the similarity of *wh*-phrases (Atkinson, et al., 2016). Secondly, Featural RM resembles memory constraints on sentence syntactic processing, where the similarity of competing words or phrases often predicts comprehension difficulties (Rizzi, 2013). Rizzi (2013) has summarized the definition of Featural RM constraint as in (4):

(4) Relativized Minimality (revised): in the configuration

… X … Z … Y …

The functioning condition in (4) suggests that X and Y cannot create a dependency if Z *c-commands* Y and has the same syntactic structural type as X. C-command is defined as in (5) (Rizzi, 1990):

(5) X c-commands Y iff neither X dominates Y nor vice versa, and the first projection dominating X dominates Y as well, as in Figure 2 (*Dominate* is that in a tree diagram G dominates X and Y if G is higher than X and Y).

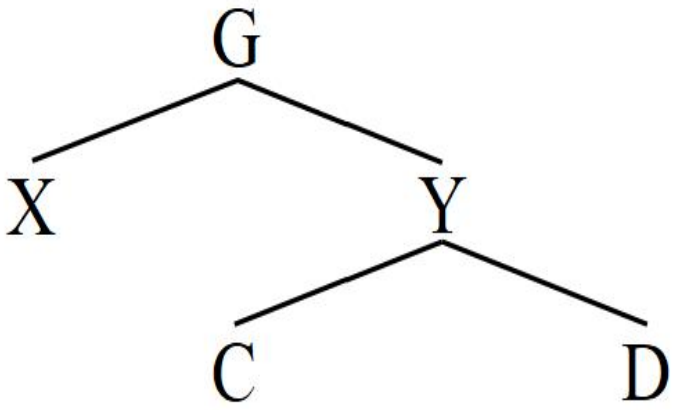
C-command is required to satisfy three rules: 1) X and Y is not the same as each other; 2) There is no dominating relation between X and Y; 3) X and Y are both dominated by the same node. Based on these three rules, in Figure 2, G dominates X and Y, indicating that X c-commands Y and vice versa. Once X c-commands Y, X c-commands all the subordinate constituents (X c-commands C and D in Figure 2). Therefore, as stated in (4), a *wh*-dependency cannot form when Z c-commands Y, where Z is more syntactically closer to Y than the fronted *wh*-phrase X. Featural RM recognizes such a violation coined by structural conflicts as a consequence by morpho-syntactic features of those constituents.

Before Featural RM, there was an empirical study that reported the amelioration effects of morpho-syntactic features on *wh*-island violations (Pesetsky, 1987). He found that

D(iscoursed)-linked *wh*-phrases are somehow related to the discourse and can be used to replace original *wh*-phrases to ameliorate the syntactic acceptability. As in (6)a, this is a sentence traditionally regarded by generative syntax as syntactically incorrect because of the movement of the single *wh*-phrase *what*. However, the *D*-linked *wh*-phrase *what kind of party* in (6)b is considered as marginally syntactically correct. This indicates that *wh*-island effect is ameliorated to some extent, even though its syntactic acceptability is still lower than the correct *wh*-movement in (6)c.

Figure 2

C-command and dominate.



(6) a. *What do you wonder who went to?

b. ?What kind of party do you wonder who went to?

c. What kind of party do you think that he went to?

Based on this observation, Rizzi and his colleagues declared that different degrees of overlapping in morpho-syntactic features account for amelioration in acceptability variation (Belletti et al., 2012; Friedmann et al., 2009; Rizzi et al., 2013). For instance, feature relations of *wh*-phrases can be categorized as identity in (6)a, inclusion in (6)b, and disjunction in (6)c (Friedmann et al., 2009). In (6)a, both the extracted *what* and the intervener *who* contain a [+Q(uestion)] feature, resulting in great degradation of acceptability. In (6)b, the extracted

D-linked *wh*-phrase *what kind of party* contains two features of [+N(oun)] and [+Q], while the intervener only contains one [+Q] feature. This configuration is defined as inclusion, as the *D*-linked *wh*-phrase is more specified in terms of features than the intervener, ameliorating the *wh*-island effects compared to that in (6)a. In (6)c, the subordinate clause is not introduced by a *wh*-phrase, containing no [+Q] feature. Therefore, there are no feature specifications between the *D*-linked *wh*-phrase and the potential intervener. This disjunction causes no violation of Featural RM. These configurations and their amelioration degrees can be concluded in Table 1.

Table 1

Taxonomy of feature set and well-formedness in Featural RM.

X **Fronted phrase**	**Z** **Intervener**	**Y** **Thematic position**	**Well-formedness**	**Type**
+A	+A	<+A>	Ungrammatical (*)	Identity
+A, +B	+A	<+A, +B>	Marginal (?)	Inclusion
+A	+B	<+A>	Grammatical (√)	Disjunction

Note. Adopted from "Relativized relatives: types of intervention in the acquisition of A-bar dependencies" by Friedmann, N., et al., 2009, *Lingua 119*, 67–88.

In summary, Featural RM is that it focuses on the morpho-syntactic features of the extracted and intervener constituents: the overlapping of features can create different degrees of acceptability, in which amelioration is generated when the extracted *wh*-phrase contains more or distinct morpho-syntactic features than the intervener. It should be noted that Featural RM somehow allows memory constraints to function when a sentence is parsed, which will be discussed in the section of discussion below.

2.4 Empirical Studies in *Wh*-Island Effects by Acceptability Experiments

With island effects heatedly discussed and studied since 1970s, experiments have been applied to explain such syntactic phenomena more scientifically (Kluender & Kutas, 1993). Astonishingly, the study reported that *wh*-island effects might not be a syntactic constraint but a memory/cognitive constraint. Though the result manifested overhasty due to experimental deviations, it initiated the debate whether island effects are a syntactic phenomenon and the tendency where acceptability judgment experiments are performed in island-effect studies.

One highly-appreciated study about satiation in island effects (Snyder, 2000) found that only types of *wh*-island, subject island, and complex NP island could achieve acceptability satiation[4] while others could not, challenging the uniformity of island effects. However, the findings were not stably repeated in later studies. Whether satiation exists still requires more experimental research. Other scholars (Hofmeister et al., 2013; Hofmeister & Sag, 2010; Hofmeister & Vasishth, 2014; Sprouse, 2009; Sprouse et al., 2012, 2013; Sprouse & Norbert, 2011) further explored island effects from the perspective of psychological linguistics and super-additive interaction. All these studies are based on acceptability experiments focusing on its variation of island effects including *wh*-islands and as pioneering research inspired the recent studies, almost all adopted experiments to study acceptability of *wh*-islands (e.g., Atkinson et al., 2016; Christensen, Kizach, & Nyvad, 2012; Keshev & Meltzer-Asscher, 2018; Shlonsky, Villata, & Franck, 2020; Stepanov, Mušič, & Stateva, 2018).

2.5 Summary

In general, current studies in the field of island effects focus on the general situations, single languages, and morpho-syntactic amelioration devices. However, the amelioration in *wh*-island effects cannot be generalized due to the lack of evidence from more languages. Additionally, whether *wh*-island effects can achieve satiation and how amelioration of acceptability is observed only in terms of morpho-syntactic means are under-research. In light of the huge research gap and remaining issues in generative syntax, the present study attempts to contribute to this field, with addressing the following questions:

---

[4] Acceptability satiation is the final stage of acceptability variation, in which one sentence violating island effects can be totally acceptable in terms of syntax.

1) Do *wh*-island effects in Chinese achieve satiation (without any amelioration devices)? If *yes*, can this be evidence for *wh*-island satiation across all languages? If *no*, what new findings can be drawn?

2) How can acceptability variation of *wh*-island effects in Chinese be ameliorated by morpho-syntactic devices and how well will the amelioration be?

3) What can the amelioration in Chinese *wh*-island effects predict about or add to Featural RM?

3 Methodology

3.1 Rationale

The study adopts two acceptability experiments to explore whether satiation exists in Chinese *wh*-island effects and how morpho-syntactic features ameliorate *wh*-island effects in Chinese.

Experiment 1 does not follow the typical design of syntactic satiation, in which the task is to judge whether a sentence is grammatically correct with *yes/no* answers (Snyder, 2000). This experiment revises this design and requires all the participants to rate sentences following a 7-point scale because in this way the satiation process can be presented more explicitly and the results can be explained with more details. Additionally, this experiment does not employ the magnitude estimation design[5] with participants freely rating the sentences, which hardly ensures that the understanding of rating across participants can be on the same level, thus yielding tremendous deviations. Because of the refined design and the fact that Chinese is not sensitive to island effects (Cheng, 2009; Huang, 1982a, 1982b; Tsai, 1994), it is predicted that the possibility of satiation existing in Chinese will be lower than that of other languages with more sensitivity to the island effects.

Experiment 2 employs a 2×2×2 factorial design, manipulating three factors of matrix *wh*-phrases (simple vs. *D*-linked), feature relations (identity vs. inclusion), and embedded structures of island (*wh*-island vs. non-island). This experiment aims to explain subtle and

[5] Magnitude estimation was once used in Sprouse's study of satiation effects (2009). This design renders great freedom to participants. For example, if one participant regards one sentence half grammatical, then he can give a rating of 50, and if he regards one sentence doubly grammatical, then he can give a rating of 200.

complicated acceptability intuition differences by using a quantitative method of a 7-point scale as Experiment 1, which has been considered credible for syntactic variation studies (e.g., Hofmeister & Sag, 2010; Lu et al., 2020; McDaniel & Cowart, 1999; Sprouse et al., 2012; Sprouse & Hornstein, 2013). Before presenting the details of this experiment, it should be clarified that this experiment does not serve to explore how real-time parsing processes in terms of wh-island effects, but to estimate whether predictions by Featural RM of acceptability variation can be verified by data from Chinese sentences.

3.2 Experimental Design

3.2.1 Participants

A total of 50 participants were recruited via social media, with each experiment involving 25 participants. They reported that they had acquired Chinese as their first language and that they were in good healthy condition. In Experiment 1, the average age of participants is 20.32 (range: 18–26), while 7 are male and 18 are female. In Experiment 2, the average age is 19.72 (range: 18–23), while 12 are male and 13 are female. They participated in the experiments via their own electronic devices (e.g., laptops).

All the participants had successfully completed the pre-tests in the two experiments (see more details in Procedures section), and provided informed consent. They were all paid 20 RMB as remuneration, supported by *the Fundamental Research Funds for the Central Universities*. The number of participants was antecedently analyzed out by GPower, a platform for predicting how many participants a quantitative study requires to achieve statistical significance.

3.2.2 Materials and Stimuli

Both experiments were designed on PC Ibex Farm, a web-platform for linguistic research (Drummond, 2020). However, with regard to participants' accessibility and practicality, all the test information were transferred to and presented on a survey platform named *Wenjuanxing*, where all the participants successfully accomplished the experiments.

In Experiment 1, 70 items were used in total (see sample set in Table 2), and each of them contained a context sentence providing the background information. Based on it, a test

sentence developed in the *wh*-interrogative form. All the items were averagely distributed into 10 blocks, with each involving 5 totally grammatical items as fillers and 5 *wh*-island ungrammatical ones as ready-to-rate sentences. These ungrammatical item types are listed in (7). Compared to the previous research (e.g., Snyder, 2000; Sprouse, 2009), this design enlarged the test types and estimated *wh*-island satiation more specifically such that the results could yield stronger explanations for this phenomenon.

Table 2

A sample set of stimuli for Experiment 1.

Context	*Ni renwei women diuqi le jian ai zheben shu.* You think we desert ASP *Jane Eyre* this book. 'You think that we deserted the book *Jane Eyre*.'
Test sentence	*Shenme ni renwei shei diuqi le?* What you think who desert ASP? 'What do you think who deserted?'

(7) a. *What*-island with a *what* fronted phrase

Shenme ni shengcheng shenme yijing biancheng le?

What you claim what have become ASP?

'What did you claim what had become?'

b. *What*-island with a *who* fronted phrase

Shenme ni rending shei yao yanchi?

What you assume who will delay?

'What do you assume who will delay?'

c. *Who*-island with a *who* fronted phrase

Shei ni kandao shei jiehun le?

Who you see who marry ASP?

'Who did you see who married?'

d. *Who*-island with a *what* fronted phrase

Shei ni xiangxin shenme zai zhemo?

Who you believe what be torture?

'Who do you believe what is torturing?'

e. *Who*-island with a *which* fronted phrase

Nayige ni caixiang shei xihuan?

Which you suppose who like?

'Which do you suppose who likes?'

f. *Whether*-island with a *what* fronted phrase

Shenme ni xiangzhidao shifou ta you?

What you wonder whether she have?

'What do you wonder whether she has?'

g. *Whether*-island with a *who* fronted phrase

Shenme ni xiangzhidao shifou women xiangnian?

What you wonder whether we miss?

'What did you wonder whether we missed?'

The stimuli of Experiment 2 were comprised of 24 sets/items of bi-clausal *wh*-interrogative sentences, with each set containing 8 sentences (see sample set in Table 3). All of them were counterbalanced across eight lists such that every participant simply saw one condition of each set, following a Latin square design. According to Featural RM, all the island sentences/conditions are less acceptable than those non-island ones, and the identity type is considered as the least acceptable compared to all other conditions, in which the inclusion type create an amelioration effect and the disjunction type ameliorates the sentences into grammatically acceptable ones. The further understanding of the principles should be that in the inclusion type non-island sentences should be more acceptable than those island counterparts, which actually are still more acceptable than non-island sentences in the identity type. It seems that predictions of Featural RM are not made explicit in these aspects. Therefore, based on these observations, three factors of the stimuli were designed to test amelioration effects in *wh*-island effects predicted by Featural RM. Besides, the stimuli only included fronted phrase types of *who* and *which* and verbs with discernible interrogative

features (e.g., *wonder*) to be the exemplars for three reasons. Firstly, they could guarantee the explicitness of *wh*-island phenomena in order that participants could make decisions without much comprehension difficulty. Secondly, this allowed only morpho-syntactic amelioration to function without other intervening factors like semantic features.[6] Thirdly, single test sentence types would keep the experiment consistent all through the process.[7] Forty-eight filler items were randomly interspersed with all the test items. This study manipulated the varying syntactic acceptability of these filler items into three categories: fillers expected to receive higher ratings (apparently acceptable), fillers expected to receive lower ratings (apparently unacceptable), fillers expected to receive ratings between the higher and lower. The filler items involved declarative and interrogative sentences. However, this study amended the previous studies (Atkinson et al., 2016) by balancing the number of three categories and two types of sentences such that participants would not be specifically influenced by any certain kind of sentences. There are basically two reasons for employing filler items. Firstly, this can foster participants to rate all the sentences from a largely variable range, which can reveal more details on amelioration. Secondly, all the filler items with distinct acceptability features can be regarded as a baseline to rate sentences. All the eight lists were pseudorandomized such that sentences of the same condition would not appear adjacently. All the stimuli and filler items in Experiment 1 and 2 are provided in Appendices.

Table 3

A sample set of stimuli for Experiment 2.

	Island	*Shei ni xiangxin shei xihuan ?* Who you believe who like? 'Who do you believe who likes?'
Identity		*Shei xiangxin shei xihuan zhege*

[6] If *wh*-phrases like *what* and *who* appear in the same sentence separately as the fronted and the intervener, they will carry not just morpho-syntactic amelioration effects but also possible semantic amelioration effects of whether it is animate, which could definitely interfere the accuracy of data. However, this does not mean that no semantic features exist, but instead morpho-syntactic features themselves contain semantic features.

[7] In this way, participants would not be burdened with recognizing sentences with changes of wh-phrases, providing more valid and analyzable data.

		Non-island	Who believe who like this *xuesheng?* student? 'Who believes who likes this student?'
Simple *wh*-phrase	Inclusion	Island	*Shei ni xiangxin nage jiaoshou* Who you believe which professor *xihuan?* like? 'Who do you believe which professor likes?'
		Non-island	*Shei xiangxin nage jiaoshou* Who believe which professor *xihuan zhege xuesheng?* like this student? 'Who believes which professor likes this student?'
	Identity	Island	*Nage xuesheng ni xiangxin nage* Which student you believe which *jiaoshou xihuan?* professor like? 'Which student do you believe which professor likes?'
		Non-island	*Nage tongxue xiangxin nage* Which classmate believe which *jiaoshou xihuan zhege xuesheng?* professor like this student? 'Which classmate believes which professor likes this student?'
D-linked *wh*-phrase		Island	*Nage xuesheng ni xiangxin shei* Which student you believe who *xihuan?* like?

		'Which student do you believe who likes?'
Inclusion	Non-island	*Nage tongxue xiangxin shei xihuan zhege xuesheng?* Which classmate believe who like this student? 'Which classmate believes who likes this student?'

3.2.3 Procedures

In both experiments, the participants were required to rate sentences from 1 (unacceptable) to 7 (acceptable) in terms of syntactic acceptability. Before the experiments started, participants had been instructed to determine acceptability based on whether a native speaker or such an environment could generate the given sentence and whether they could accept the sentence's grammatical structure when they received it. They were required to entirely read every sentence without carefully analyzing its grammatical structure. They were informed that only the syntactic acceptability should be considered whatever the semantic or pragmatic situation was and however the potential textual or social background was interpreted. Each time participants saw only one test sentence. They did not refer back to any previous choices whenever they were making judgments.

Additionally, each participant in both the two experiments conducted a pre-test with six items. These trials included two highly acceptable sentences, two marginal sentences, and two highly unacceptable ones. This practice encouraged the participants to avoid only using the two polarized ends to rate sentences such that the participants would not treat the experiments as binary choices. The data from participants who confined their ratings to the extreme ends (i.e., 1 and 7) would be excluded from experimental analyses because they would greatly distort acceptability contrasts.[8] Actually, none of these data was identified as invalid for participants' appropriate judgment.

[8] This step would not change the analysis pattern as it only included representative data with variation such that the accuracy of results could be guaranteed. The data of these six items were regarded as practice ones and were not analyzed.

3.3 Data Analysis

In Experiment 1, the descriptive results were first attained with a Wilcoxon sign test done. Unlike what previous studies did, all the mean ratings were converted to *z*-score calculated respectively in Block 1–3 and Block 5–7, among which Block 4 served as a transition layer. Through the *z*-score mean ratings, whether there was any satiation effect was easily determined. This study provided statistical analyses. To begin with, An independent-samples *t*-test was employed to compare the mean ratings between Block 1–3 and Block 5–7, including each participant's data (Baayen, 2008). Moreover, a one-way repeated-measures analysis of variance (ANOVA) was performed between Block 1–3 and Block 5–7 by subtracting the mean rating of Block 1–3 from that of Block 5–7 to draw a direct comparison across different sentence types of *wh*-islands. If both the *t*-test shows a correct interpretation and the ANOVA test presents an increase of ratings for *wh*-islands, then it can be predicted that a significant effect exists. However, these procedures might not directly obtain results of whether satiation effects happen because binary choices or exceedingly loose magnitude estimation were not given. If no satiation effects are observed but acceptance ratings still increase in this way, it is safe to state that these traditional estimation ways cannot exactly predict satiation effects.

In Experiment 2, the raw data of test sentences and filler sentences were processed into *z*-scores within participants (Schütze & Sprouse, 2013). This could amend participants' potential to use a different scale of ratings (e.g., only using a subset of given ratings), as it standardized all the participants' ratings. Additionally, all the raw ratings were analyzed as well, and still the same results were drawn. For the preciseness and accuracy, only the *z*-score results were reported in this study. The Linear Mixed-effect Models were then employed, which allowed that random participant and random effect variables were simultaneously included (Baayen et al., 2008). These models were run using the lme4 package in the R environment (Bates et al., 2005; R Core Development Team, 2015). The Sattherwaite approximation in the lmerTest package was used to estimate *p*-values for the random effects (Kuznetsova et al., 2015). If the results presented a significant effect, pairwise comparisons would be conducted to test significance between single conditions, adopting mixed-effect models with maximal random effects structure.

Since the two experiments employed the same 7-point scale, the results of them could be analyzed together to lead to discussion in comparative perspective.

4 Results and Discussion

4.1 Syntactic Satiation

The revised design could to some extent suspect all the previous observed findings adopting the magnitude estimation design or the traditional *yes/no* design, because it might cause unreasonable rating comprehension and it might appear that participants did not totally accept one sentence's grammatical structure, but amelioration (did not) indeed happen(ed), causing participants to choose an answer tending towards "totally acceptable" or "totally unacceptable", which were simply "yes" or "no" answers. That could as well yield Data collected in these ways somehow affected the result analyses and discussions. Furthermore, this experiment focused on *wh*-island type, and greatly increased the stimuli from one sentence per block in previous studies to five sentences of different *wh*-island subcategories per block (Goodall, 2005; Hiramatsu, 2000; Snyder, 2000; Sprouse, 2009).

Table 4 shows the basic rating situations of all participants, which provides a straight-forward descriptive version of data. The *p*-value of the sign test ($p = 0.20$) result reveals no significance between the expected amelioration ratings and the difference ratings by Block 1-3 (T) and Block 5-7 (T). Figure 3 clearly describes the relationship among the descriptive elements and directly supports what Table 4 tells about the relationship among participants, differences of *z*-scores, and amelioration effects. Figure 4 presents mean *z*-score ratings of target and filler sentences by rating types of Block 1-3 and Block 5-7. It can be clearly observed that filler sentences receive much higher ratings than target ones, suggesting that participants have rated the sentences understandingly and as expected (Block 1-3 (F) *z*-score mean = 0.76, Block 5-7 (F) *z*-score mean = 0.96). Specifically, there is no significant increase between Block 1-3 (T) and Block 5-7 (T) (Block 1-3 (T) *z*-score mean = −0.92, Block 5-7 (T) *z*-score mean = −0.80), indicating non-existence of satiation effects. Besides, the descriptive results can reveal that only extremely marginal amelioration has happened to *wh*-island effects in Chinese. A possible concern is that analyzing descriptive statistics

somehow do not contain powerful explanations. To address it, the independent samples *t*-test and one-way ANOVA results provide further analyses, as in Table 5. The results both support the explanation from descriptive data that no significant increase is attained between two groups ($t(28) = -1.14$, $F(1, 28) = 1.29$, $p = 0.27$).

Table 4

The distribution of descriptive results in Experiment 1.

	Descriptive statistics		**Sign test**	
Results	Increase rating number	Mean increase rating	Z	p
Data	21(25)	0.24	−1.28	0.20

Table 5

T-test and one-way ANOVA results in Experiment 1.

Rating Types	***T*-test**					**ANOVA**			
	SE	**SD**	**df**	***t***	**Sig.**	***F***	**df**	**N**	**Sig.**
Block 1-3 (T)	0.59	0.15	28	−1.14	0.27	1.29	1	15	0.27
Block 5-7 (T)	0.65	0.17						15	

Contrary to previous findings that satiation effects (do not) exist in English syntax (Snyder, 2000; Sprouse, 2009), this study found that there is even no experimental amelioration. One possible explanation for this can be that languages follow distinct sentence processing procedures and constraints of syntax. For instance, in Chinese sensibility to *wh*-movement is much lower such that acceptability hardly varies when one sentence is judged as ungrammatical. It can be argued that satiation effects correspond to cross-linguistic variations mostly not because of native speakers' individual grammar competence but depending on how human sentence processing functions across languages, and this is evidence for the long proposed theory by Berwick and Weinberg (1986). Synder (2000) has argued that *satiation* and *non-satiation* therefore can be take as a dimension along which the perception of ungrammaticality varies. Based on this, this study could reform it somewhat.

Another dimension of *amelioration* and *non-amelioration* can be considered, and the pattern is abstracted in (8), in which the marks *???*, *??*, ***, and ** represent *satiation* and *non-satiation*, *?* and * represent *non-amelioration* and *amelioration*, and *?** serves as a transition layer. This is a correspondence between the superficial ratings and the underlying acceptability variation.

(8) ??? / ?? / ? / ?* / * / ** / ***

Figure 3

The alluvial plot among participants, difference z-score ratings, and amelioration effects in Experiment 1.

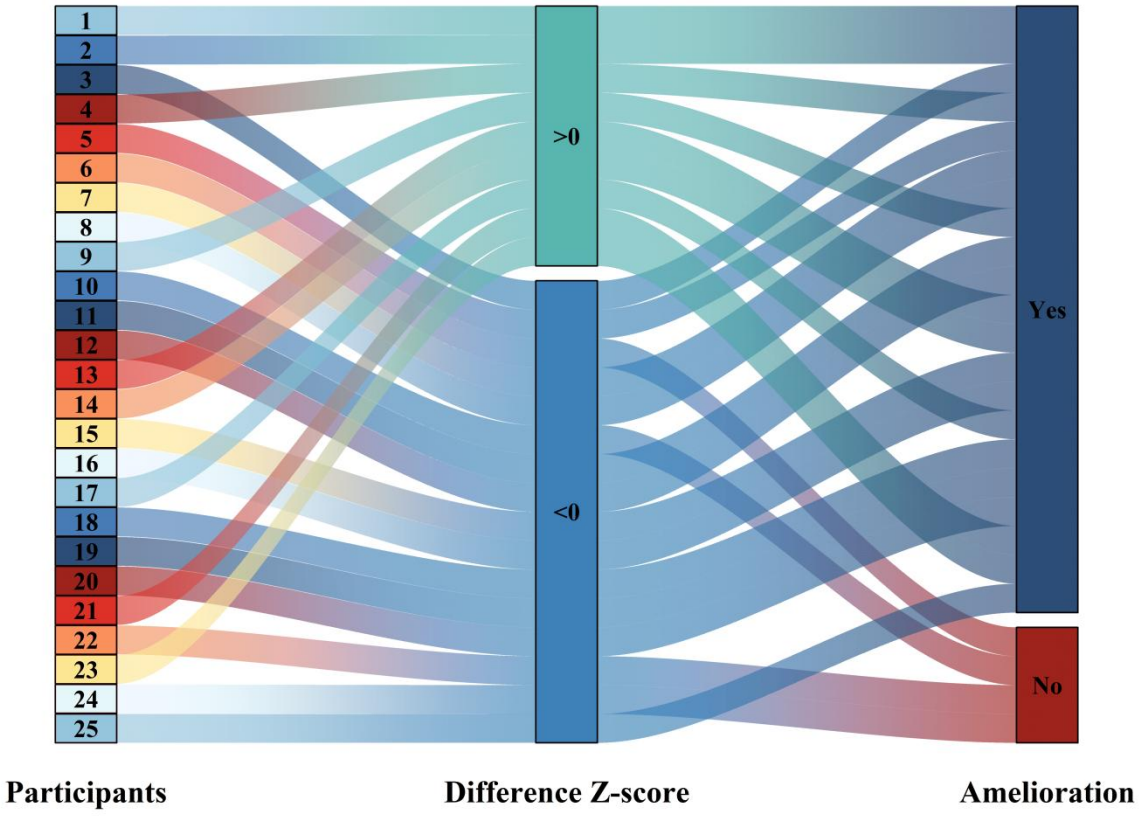
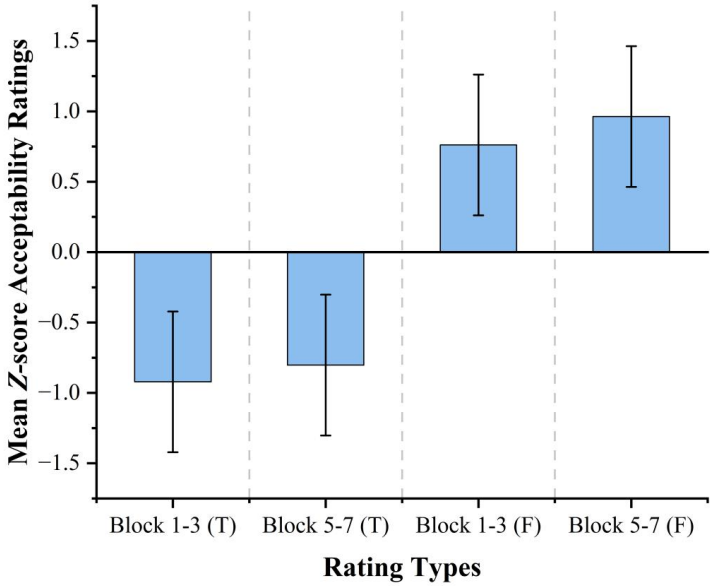


Figure 4

Mean z-score acceptability ratings of target and filler sentences by rating types of Block 1–3 and Block 5–7 in Experiment 1.



Note. Within the *X*-axis, *T* refers to target sentences and *F* refers to filler sentences, which applies to all the same situations in this study. Error bars represent ±1 SE.

4.2 Morpho-syntactic Amelioration in Acceptability Variation

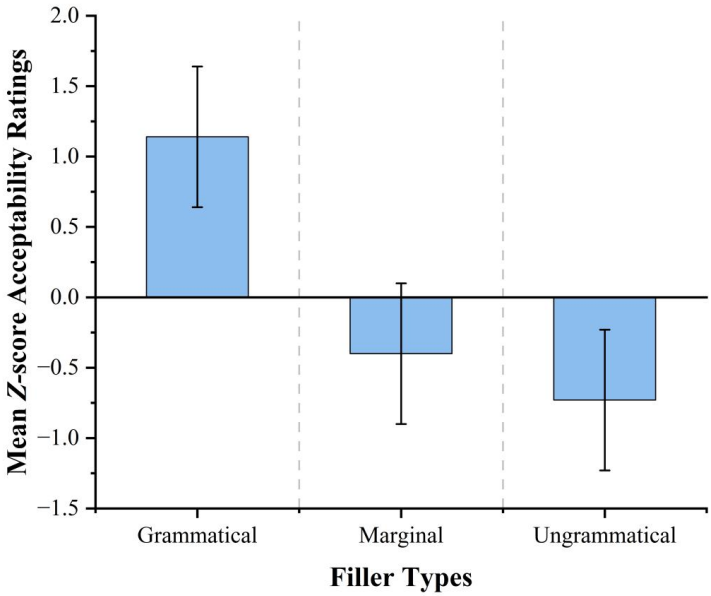
Figure 5 presents mean *z*-score acceptability ratings for filler sentences in Experiment 2. Grammatical filler sentences were rated as the most acceptable (mean *z*-score = 1.14), and ungrammatical ones were rated as the least acceptable (mean *z*-score = −0.73). Marginal filler sentences obtained ratings slightly near a zero (mean *z*-score = −0.40). These different ratings indicate that participants basically understood the experiment and have done it successfully.

Figure 6 shows the target sentences' mean *z*-score ratings under different matrix *wh*-phrases. All the island types were rated as less acceptable than their non-island counterparts (mean *z*-score of island types = −0.73, mean *z*-score of non-island types = 0.82). Among non-island types, the highest rating type is the non-island *D*-linked *wh*-phrase identity type (mean *z*-score = 1.26). The other non-island types received various ratings (mean *z*-scores = 0.77, 1.17, and 0.06). Among island types, the highest rating type is the island *D*-linked *wh*-phrase identity condition (mean *z*-score = −0.36), except which the other island types obtained similar ratings (mean *z*-scores = −1.34, −1.17, and −0.39).

The Linear Mixed-effect Model results in Table 6 present estimated coefficients, standard errors, and *p*-values for fixed effects with different conditions as fixed effects and by-participant and by-item random intercepts. The *p*-values in Table 6 and the ratings in Figure 6 are thermodynamically visualized in Figure 7 to clearly show the differences of these data.

Figure 5

Mean z-score acceptability ratings of filler sentences by filler types in Experiment 2.



Note. Error bars represent ±1 SE.

Obviously, there exist significant main effects of islandhood and matrix *wh*-phrases, but there exists no main effect of feature relations. Additionally, there is a significant main effect of the interaction between islandhood and matrix *wh*-phrases, indicating that matrix *wh*-phrases somehow ameliorate the acceptability of islandhood. Planned pairwise interaction comparisons were conducted among island conditions, shown in Table 7, suggesting a significant main effect between the simple *wh*-phrase condition and the *D*-linked *wh*-phrase condition (F = 49.57, SE = 0.07, $p < 0.001$). This indicates that *D*-linking amelioration is observed. On the one hand, there is no significant effects between the simple *wh*-phrase

identity condition and the simple *wh*-phrase inclusion condition (F = 1.72, SE = 0.08, p = 0.19), and between the *D*-linked *wh*-phrase identity condition and the *D*-linked *wh*-phrase inclusion condition (F = 0.10, SE = 0.08, p = 0.76). However, on the other hand, the *D*-linked *wh*-phrase identity condition is significantly more acceptable than simple *wh*-phrase identity condition (F = 41.94, SE = 0.08, p < 0.001). Similarly, the simple *wh*-phrase inclusion condition is significantly more acceptable than the *D*-linked *wh*-phrase inclusion condition (F = 24.74, SE = 0.08, p < 0.001). These two matrix-fixed models suggest that the *D*-linked *wh*-phrase condition reliably ameliorates island effects whatever the feature relation is, as plotted in Figure 8.

Table 6

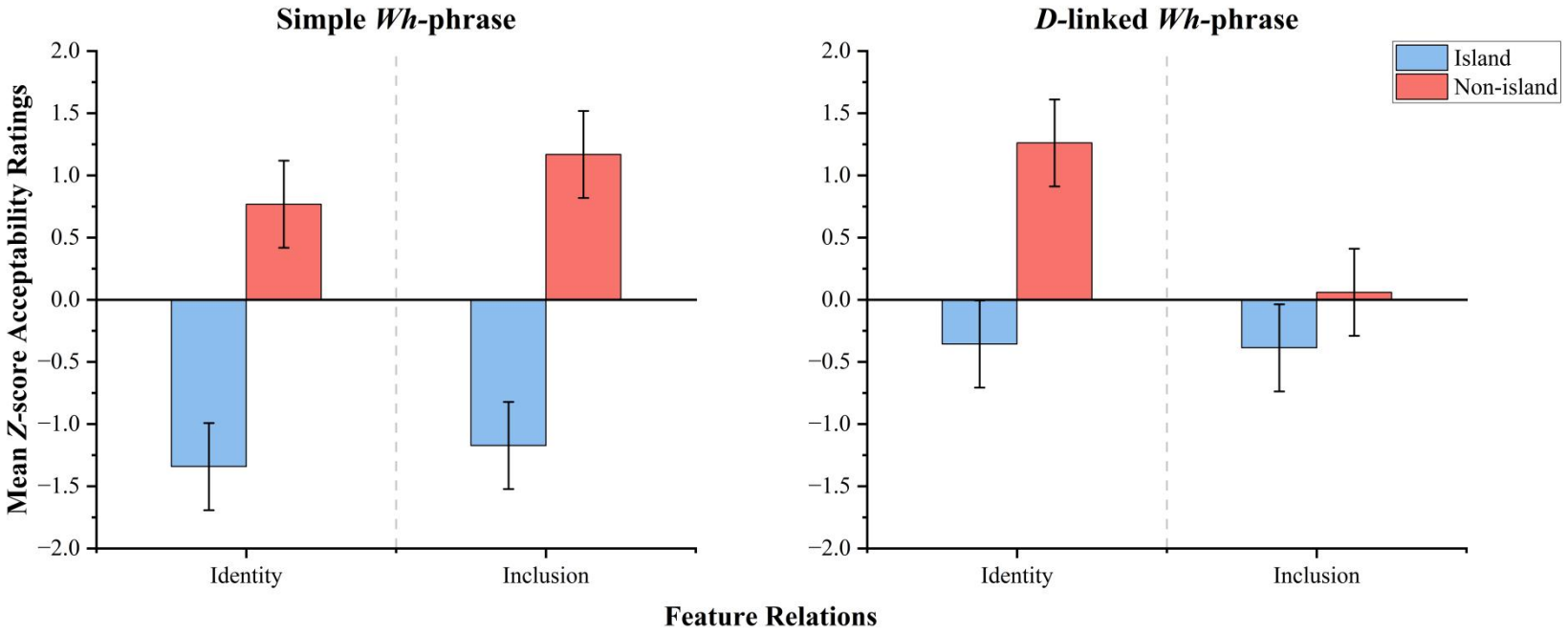
Fixed effects of the Linear Mixed-effect Models with by-participant and by-item intercepts for different conditions in Experiment 2.

Fixed Effects	Estimate	SE	*t*
Intercept	−0.78	0.58	−1.33
Islandhood (Embedded Structures of Islands)	2.28***	0.38	6.07
Feature Relations	0.28	0.33	0.87
Matrix *Wh*-phrases	1.26***	0.33	3.85
Islandhood × Feature Relations	−0.05	0.21	−0.24
Islandhood × Matrix *Wh*-phrases	−0.61**	0.21	−2.93
Feature Relations × Matrix *Wh*-phrases	−0.13	0.21	−0.63
Islandhood × Feature Relations × Matrix *Wh*-phrases	>−0.01	0.13	>−0.01

Note. This model does not converge and provides random slopes for the conditions. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

Figure 6

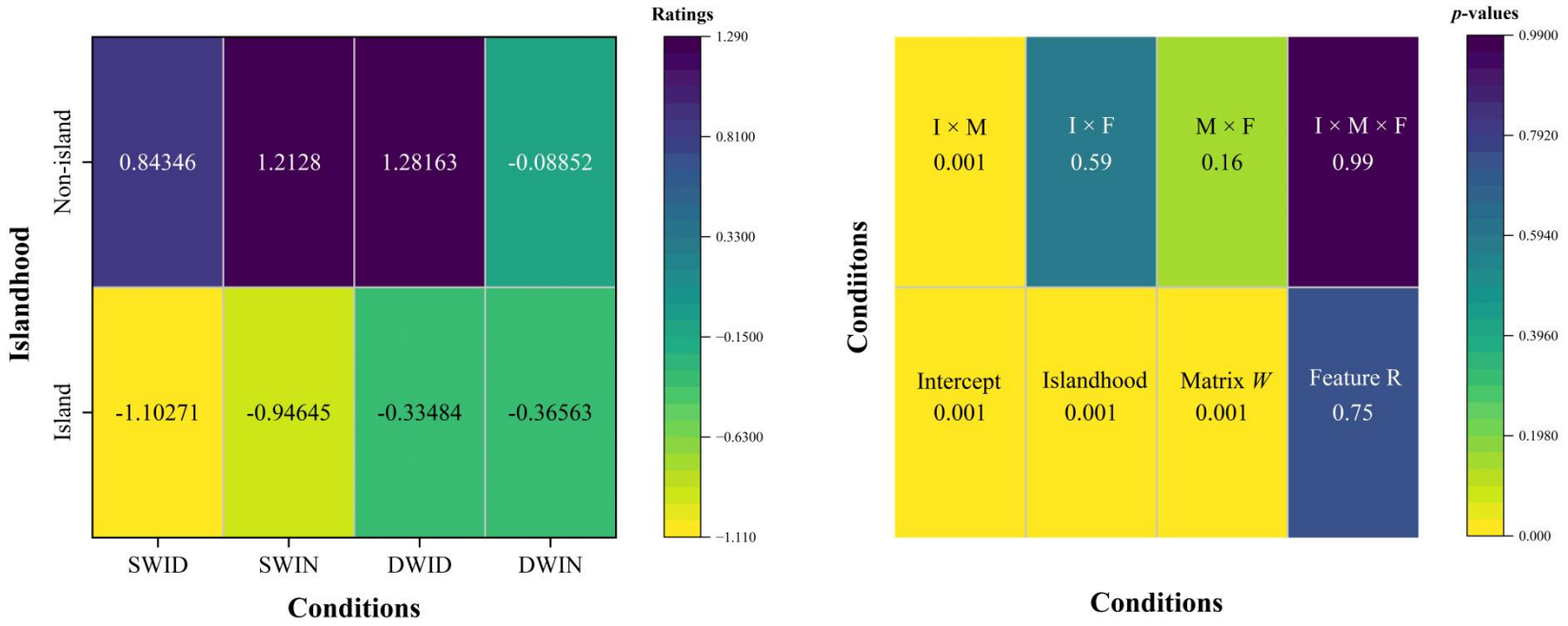
Mean z-score acceptability ratings of matrix wh-phrases by feature relations in Experiment 2.



Note. Error bars represent ±1 SE.

Figure 7

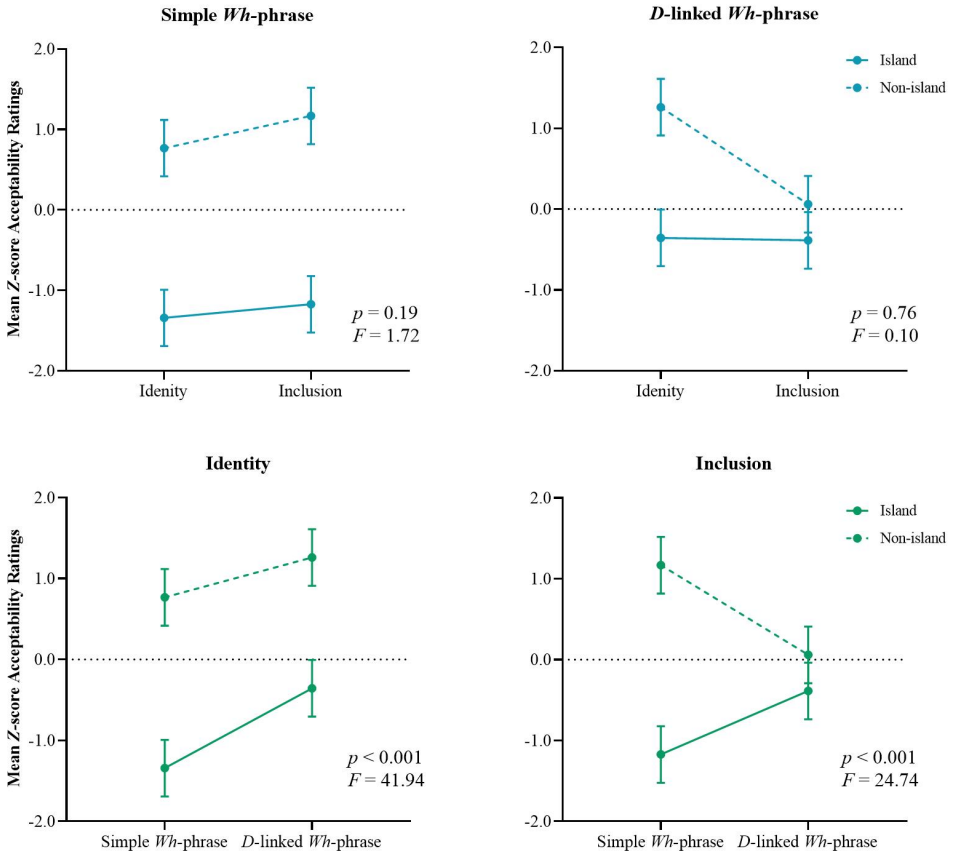
The thermodynamic plot for mean z-score acceptability ratings across conditions in Experiment 2.



Note. SW refers to *Simple Wh-phrase*, *DW* refers to *D-linked Wh-phrase*, *ID* refers to *Identity*, and *IN* refers to *Inclusion*. *I* refers to *Islandhood*, *M* refers to *Matrix Wh-phrase*, *F* refers to *Feature Relation*, *W* refers to *Wh-phrase*, and *R* refers to *Relation*.

Figure 8

The interaction plot across conditions in Experiment 2.



Note. Error bars represent ±1 SE.

Table 7

Summary of pairwise interaction comparisons among island conditions.

Interaction Comparison Models	*F*	SE	*p*
Simple *Wh*-phrase × *D*-linked *Wh*-phrase	49.57	0.07	***
Simple *Wh*-phrase × *D*-linked *Wh*-phrase (Identity)	41.94	0.08	***
Simple *Wh*-phrase × *D*-linked *Wh*-phrase (Inclusion)	24.74	0.08	***
Identity × Inclusion (Simple *Wh*-phrase)	1.72	0.08	0.19
Identity × Inclusion (*D*-linked *Wh*-phrase)	0.10	0.08	0.76

Note. *Note. F* values are used here instead of *β*, both of which have the same explanatory power. *p<0.05, **p<0.01, ***p<0.001.

The results reveal that *wh*-island effects generally lead to degradation of acceptability in Chinese. This acceptability degradation is modulated by the matrix *wh*-phrases between two feature relation conditions: the *D*-linked *wh*-phrase identity and inclusion conditions separately receive greater ratings than the simple *wh*-phrase identity and inclusion conditions. Besides, the *D*-linked *wh*-phrase identity is rated as the most acceptable among all the island conditions. This is consistent with previous literature in English (Atkinson et al., 2016), proving that the *D*-linked *wh*-phrase identity functions robust amelioration effects across experiments and languages. However, in the same matrix *wh*-phrase conditions, there is no evidence of amelioration in *wh*-island effects by feature relations. Surprisingly, the absence of amelioration in feature relations is against the amelioration effects in the overlapping feature configurations predicted by Featural RM, considering that the amelioration by feature relations have been widely reported in previous studies (Alexopoulou & Keller, 2007; Atkinson et al., 2016; Goodall, 2015; Pesetsky, 1987). The findings about the ranking of amelioration effects among different *wh*-island violations can be concluded in (9). The study will discuss below the theoretical recommendations for Featural RM.

(9) *D*-linked *wh*-phrase identity ≥ *D*-linked *wh*-phrase inclusion ≥ Simple *wh*-phrase inclusion ≥ Simple *wh*-phrase identity

4.3 Acceptability Variation in Comparative Perspective

Experiment 1 and 2 both employed a 7-point scale with a basic task of rating sentences for different testing purposes. However, the underlying core of them was to test any amelioration effects in *wh*-island violations, with one by self-amelioration and the other by morpho-syntactic devices. Their results can, to some extent, be discussed comparatively. A simplified version of findings are concluded in Table 8.

Figure 9 presents the relationship between mean ratings and mean *z*-score ratings across participants in two experiments, with bubbles and a centroid. It indicates that the two experiments basically follow the same distribution pattern with the majority centralized and the minority scattered. Figure 10 represents the dispersion degrees and distribution situations of two experiments. Obviously, Experiment 1 is similarly dispersed to Experiment 2, indicating a similar range of ratings across participants. Slightly the scattering spots are symmetrical based on the median across the two experiments. Experiment 1 and 2 contain few outliers and occupy a small number compared to their own whole groups, conforming to normal distribution. Additionally, the basic statistics of the two experiments are internally coordinating.[9] In light of the observations in two figures, Experiment 1 and 2 are both within the normal fluctuation range of data, thus having a comparative foundation.

Figure 11 presents the scattering and regression analysis of the two experiments. Experiment 2 has bigger absolute values of its regression coefficient and R^2 value in the equation than that of Experiment 1. Besides, the regression coefficients and R^2 in Experiment 1 are extremely near a zero and similar to each other. Therefore, in Experiment 1, there was no amelioration after participants repeated the same type of sentences several times, while in Experiment 2, amelioration appeared in certain morpho-syntactic conditions. It seems that morpho-syntactic devices show a greater amelioration effect than self-amelioration. From self-amelioration to morpho-syntactic features, behind the acceptability variation is the trajectory of amelioration variation. This can elicit two possible findings. Firstly, not all conditions can ameliorate *wh*-island effects and why the *D*-linked *wh*-phrase identity condition presents the strongest amelioration effects still requires exploring. At the same time, it cannot be roughly concluded that no sentence processing has existed in those

[9] The ameliorated results of the two experiments show the analogous violin shape.

non-functioning conditions due to their same amelioration effects as self-amelioration. It is only safe to say that the *D*-linked *wh*-phrase identity condition is triggered as the most effective amelioration device compared to the least effective of self-amelioration in Chinese.

Secondly, Chinese is restrictively susceptible to *wh*-island effects. In the previous literature, how much Chinese is (not) susceptible to this syntactic phenomenon is not answered (Cheng, 2009; Huang, 1982a, 1982b; Lu et al., 2020; Tsai, 1994). However, it now can be inferred that *wh*-island effects can be observed by native speakers and ameliorated through certain devices, where morpho-syntactic features can be a baseline. To be more specific, the simple *wh*-phrase inclusion condition is the baseline, across which Chinese is becoming susceptible to be ameliorated. Self-amelioration and the *D*-linked *wh*-phrase identity condition are relatively two extreme ends of susceptibility.[10]

Table 8

Comparisons between previous and present findings.

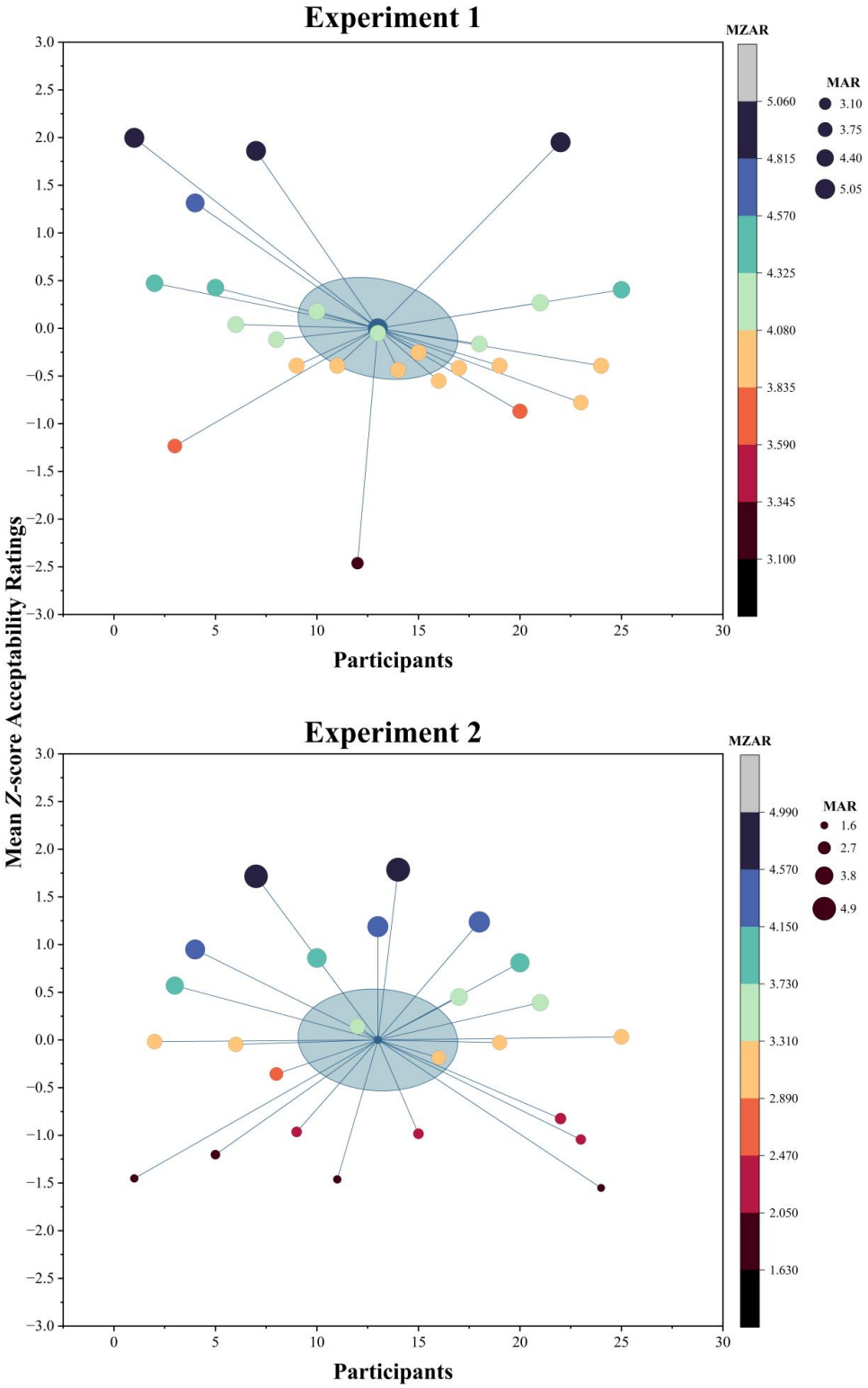
Findings	Previous	Present
Wh-island effects	√	√ +
Satiation	√ / ×	× +
Amelioration	√	√ +
Theoretical predictions	√	√
New recommendations	√	√

Note. The mark + means something different and new from the previous literature.

[10] This is explained in the following aspects. Firstly, the only estimated amelioration device is morho-sentactic features. Therefore, among all the possible amelioration devices, that is considered as "relatively". Secondly, within the tested situations, self-amelioration does not refer to no amelioration at all, and similarly, the *D*-linked *wh*-phrase inclusion does not mean the totally corrected effects. Instead, they simply represent the weakest and strongest amelioration effects.

Figure 9

The bubble plot for two experiments across participant.



Note. MZAR refers to *Mean Z-score Acceptability Ratings*, *MAR* refers to *Mean Acceptability Ratings*.

Figure 10

The violin plot with a box and scattering points across two experiments.

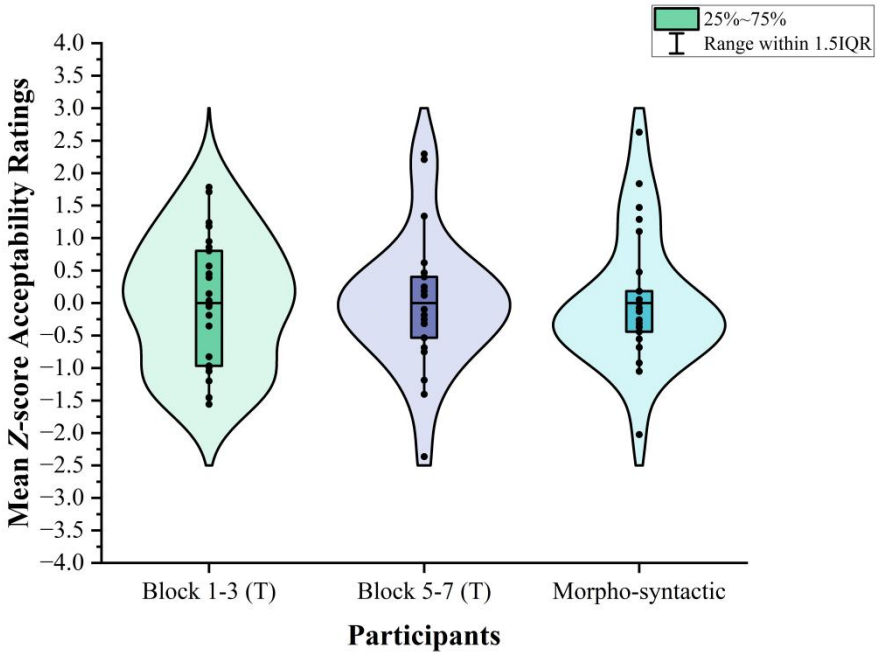
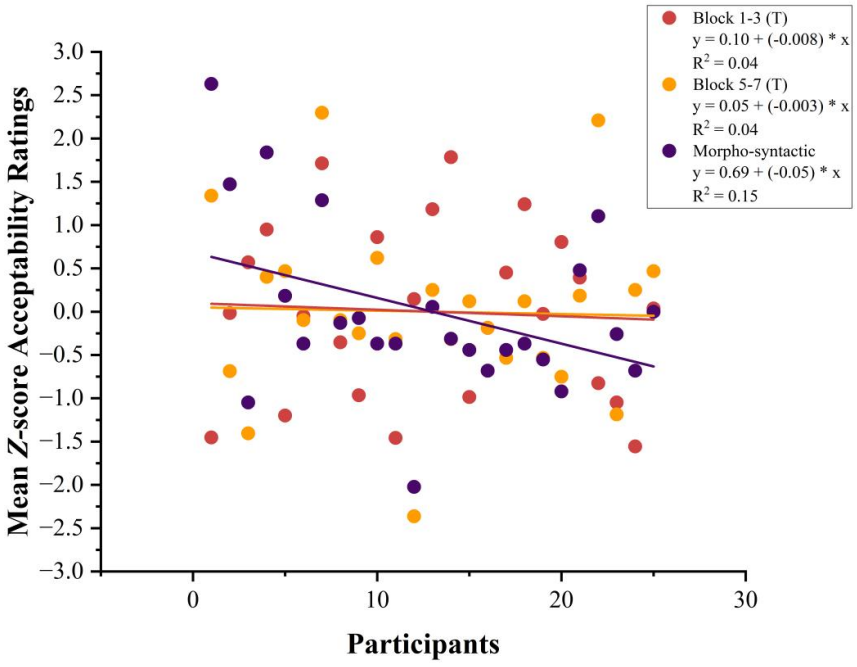


Figure 11

The scatter plot of regression models across two experiments



4.4 Discussion

4.4.1 Recommendations for Featural RM

The discussion in this study indicates that Featural RM cannot completely account for the *wh*-island amelioration effects by morpho-syntactic features, except that the *D*-linked *wh*-phrase identity gives rise to a robust amelioration effect. However, the discussion does not hold this statement against Featural RM in nature, but at least it requires amendment and addition. One potential recommendation is that the morpho-syntactic amelioration effects proposed by Rizzi and his colleagues should be cross-linguistically explained.

It is suggested that languages of low susceptibility to *wh*-island effects be ameliorated not mainly by feature relations but by matrix *wh*-phrases, while those of high susceptibility remain the same[11], as summarized in Table 9. In this way, Featural RM should as well predict that morpho-syntatic features indeed play a significant role in ameliorating *wh*-violations but this theoretical argument cannot reliably explain why different configurations achieve different levels of amelioration across languages and whether a general certain pattern of amelioration exists in all languages.

In the previous literature, recommendations for Featural RM include adding an animacy feature (Atkinson et al., 2016). However, in that study, the *D*-linked *wh*-phrase inclusion condition did not present a robust amelioration effect, and this feature could well offset that condition to somehow achieve amelioration. This study has proved a similar strong amelioration effect by the *D*-linked *wh*-phrase inclusion condition such that this recommendation is not accepted. It is not obvious whether this addition only functions in high-susceptibility languages or this is caused by experimental errors.

In short, it is not evident whether the recommendations proposed pertinently according to the experimental results can account for the amelioration effects by different featural configurations because all the manipulations are still fully internal to the principles of Featural RM in syntax. If the amelioration effects cannot be explained by the newly additional adjustments, then at least something beyond morpho-syntactic features should be considered into Featural RM to account for *wh*-island effect amelioration patterns.

---

[11] Here it refers to the version revised by Atkinson and the co-authors (2016).

Table 9

The summary of revised predictions by Featural RM.

Conditions	Simple *Wh*-phrase^{1*}	Simple *Wh*-phrase^{2*}	*D*-linked *Wh*-phrase^{1*}	*D*-linked *Wh*-phrase^{2*}
Identity^{1*}	×	Blank	√	Blank
Identity^{2*}	Blank	×	Blank	√*
Inclusion^{1*}	×	Blank	√	Blank
Inclusion^{2*}	Blank	×	Blank	×

Note. The superscript *1** refers to low-susceptible languages and *2** refers to high-susceptible languages. The marks × and √ represent (not) achieving amelioration effects compared to the counterpart in the same feature relation condition, the mark √* represents the amelioration compared to the counterpart in the same matrix *wh*-phrase condition, and *Blank* represents no interaction.

4.4.2 Memory Constraints and Similarity-based Interference

For a long time, *similarity-based interference* has been studied to explain that sentence processing is a joint consequence by syntactic and semantic features (Lewis & Vasishth, 2005; Van Dyke & McElree, 2006). There are two realization mechanisms. Firstly, the most researched type is named *retrieval interference* (Gordon et al., 2001, 2004). When a parser understands *wh*-interrogative relative clauses, he is required to retrieve the extracted *wh*-phrase back to its original thematic position. Restricted by memory, the retrieval mechanism will activate all the noun phrases based on cue search, leading to comprehension difficulties of phrases with similar features. Another type is called *encoding interference* (Gordon et al., 2002; McElree, 2006; Nairne, 2002; Van Dyke, 2007), observed when a parser processes words or phrases with similar features. Encoding and storing them as independent items in memory will be interrupted, resulting in less exact or robust representations and requiring retrieving in the sentences. This could yield questions about whether self-amelioration and morpho-syntactic amelioration are influenced by similarity-based interference.

According to this prediction, the genuine situation should be: the *D*-linked *wh*-phrase identity condition experiences the least similarity-based interference, while self-amelioration and those simple *wh*-phrase conditions attain the most. It is presupposed that all the *wh*-phrases carry semantic distinct features, summarized with morpho-syntactic features in Table 10[12]. According to the additions of semantic conditions, a well-ameliorated version for Chinese and English *wh*-islands is presented in (10), in which (10)a represents the least acceptable form and (10)b shows the relatively best amelioration version. The syntactic tree clearly shows the process of movement and the same syntactic structure of the raw sentences and the ameliorated version, as in Figure 12. The *D*-linked *wh*-phrase identity and inclusion conditions are supposed to receive more semantic distinctness to ease the retrieval as they carry more complex fronted and intervener *wh*-phrases. This hypothesis has been partially testified in a number of studies because the *D*-linked *wh*-phrase inclusion condition was not observed with an amelioration (Atkinson, 2016; Hofmeister, 2007, 2013; Villata, 2016). The previous studies (e.g. Atkinson, 2016) still wonder whether an amelioration effect happens in the *D*-linked *wh*-phrase inclusion condition after matching it with semantic distinctness. This is justified in this study. However, no conclusion that such a condition presents significant amelioration effects can be drawn because the single experiment is limited to account for all the same situations. It calls for attention that morpho-syntactic and semantic features somehow explain the phenomena together and that the combination of them requires more research to test amelioration in *wh*-violations.

(10) a. *Shei ni jieshou shei pei zhe?*
Who you accept who accompany ASP?
[+ Q, + Everybody] [+ Q, + Everybody]
'Who do you accept who accompanies?'

b. *Nage pengyou ni jieshou nage laoshi pei zhe?*
Which friend you accept which teacher accompany ASP?
[+ Q, + N, + Somebody] [+ Q, + N, + Somebody]

[12] This is a brief summary for reference in order to indicate that different levels of overlapping result in different retrieval processes.

'Which friend do you accept which teacher accompanies?'

Table 10

The distribution of morpho-syntactic and semantic feature distinctness.

***Wh*-phrases**	**Morpho-syntactic Features**	**Semantic Features**	**Semantic Distinctness**
Who	+Q	+ Everybody	×
What	+Q	+ Everything	×
Which person	+Q, +N	+ Somebody	√
Which thing	+Q, +N	+ Something	√

Note. What has an overlap because *everything* includes *everybody* and *something*, and *who* has an overlap because *everybody* includes *somebody.*

Additionally, when self-amelioration appears, the parser is required to repeat the retrieval process several times. In this study, it seems that these trials could not evoke any changes of comprehension for native speakers simply because no *wh*-phrases in any positions of a sentence varied in the certain place of each block. However, it is noted that each block contained different types of *wh*-phrases, and when they were pseudorandomized every participant rated them across seven blocks. It should have appeared that a cross influence happened and due to similarity-based interference some sentences received satiation or amelioration effects at last. The opposite fact has seldom referred to and two explanations might account for this. Firstly, low-susceptible languages requires more sentences with higher frequency to prepare for the final amelioration. Experiment 1 had fewer sentences than Experiment 2, achieving different results. Secondly, the retrieval process is eased by a certain degree of distinctness. When a sentence is not typical enough in terms of its morpho-syntactic and semantic features, it cannot be identified or ameliorated by the parser.

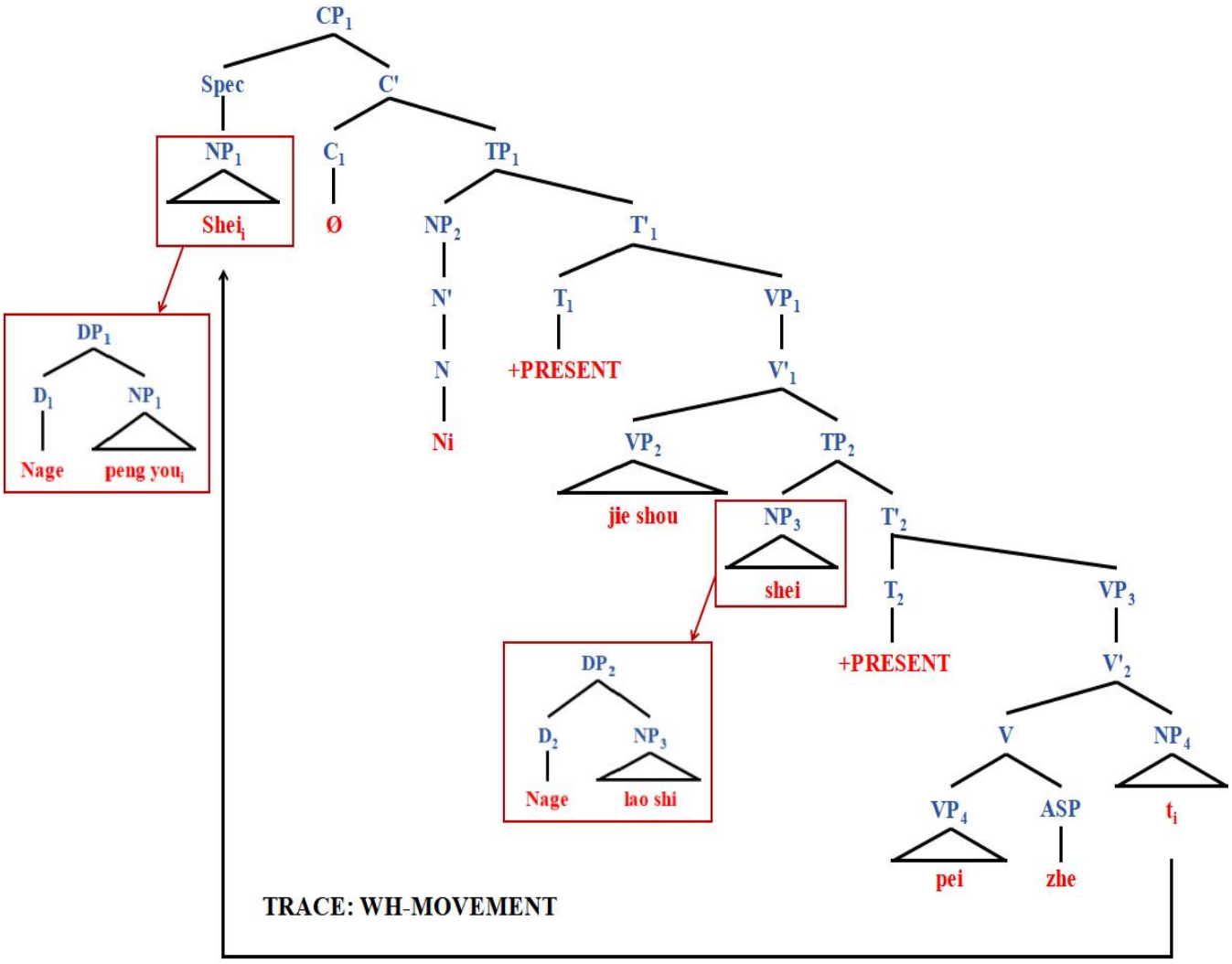
In summary, amelioration in *wh*-violations is, to some extent, a combination of psycholinguistic/semantic[13] and syntactic functions. However, based on these observations,

[13] Similarity-based interference is researched in psycholinguistics.

it is highly possible that more features should come into the featural system of *wh*-amelioration.

Figure 12

The syntactic tree adopting X-bar theory for the wh-island and ameliorated sentences as the sample finding.



Note. The square-framed parts are the supplementary structures of the ameliorated sentences for the *wh*-island ones.

5 Conclusion

5.1 Major findings

The present study investigated the satiation and amelioration effects with focusing on morpho-syntactic and semantic features through two acceptability experiments. There are two main findings. Firstly, it is discovered that no satiation effects exist in Chinese by speakers' self repetition, even nor does amelioration. Surprisingly, the expected appearing semantic melioration did not show up. Secondly, morpho-syntactic features fails to fully account for *wh*-island amelioration which presents different amelioration effects cross-linguistically: in high-susceptible languages the *D*-linked *wh*-phrase inclusion condition loses its amelioration, while in low-susceptible ones it shows a highly similar amelioration effect to the *D*-linked *wh*-phrase identity condition observed in almost all languages. At the same time, similarity-based interference with semantic features is supposed to function to cause memory constraints, eased by semantic distinctness. It arouses an argument that *wh*-island phenomena are not simply syntactic but multi-factorial. The comparison between self-repetition and morpho-syntactic devices together superficially reveals simply the amelioration effects and deeply triggers sentence processing for languages at different susceptible levels.

5.2 Limitations and Implications

The present study is limited for two aspects. Firstly, the number of stimuli in Experiment 1 is insufficient, leading to less discussion on examining memory constraints. Secondly, the test language itself has a limited explanatory power, which otherwise can directly solve some uncertain questions remaining in the study.

It is suggested that future research should focus on more languages low-susceptible to *wh*-island effects to further test acceptability variations, which might, in the end, draw a general conclusion to explain *wh*-amelioration in syntax. Moreover, it is predicted that a more specific and entire explanation of *wh*-island acceptability variations should involve semantic features not only stimulated by memory constraints of similarity-based interference but demonstrated by any other possible theories. This appeals to more investigations to examine amelioration effects by acceptability variations in multiple syntactic experimental environments considering more features including semantic constraints on parsing mechanisms.

References

Alexopoulou, T., & Keller, F. (2007). Locality, cyclicity, and resumption: at the interface between the grammar and the human sentence processor. *Language 83*, 110–160. https://10.1353/lan.2007.0001

Atkinson, E., Apple, A., Rawlins, K., & Omaki, A. (2016). Similarity of *wh*-Phrases and Acceptability Variation in *wh*-Islands. *Frontiers in Psychology, 6*, 2048. https://doi.org/10.3389/fpsyg.2015.02048

Baayen, R. H. (2008). Analyzing linguistic data: A practical introduction to statistics using R. *Journal of Child Language, 37*(2), 465–470. https://doi.org/10.1017/cbo9780511801686

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language 59*, 390–412. https://10.1016/j.jml.2007.12.005

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear MixedEffects Models using Eigen and S4. R Package Version* 1.1-9. Available at: https://CRAN.R-project.org/package=lme4

Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua 122*, 1053–1069. https://10.1016/j.lingua.2012.02.007

Berwick, R. C., & Weinberg, A. (1986). *The Grammatical Basis of Linguistic Performance: Language Use and Acquisition*. Cambridge, Mass.: MIT Press.

Chaves, R. P. & Dery, J. E. (2019). Frequency effects in subject islands. *Journal of Linguistics, 55*(3), 475–521. https://doi.org/10.1017/S0022226718000294

Cheng, L. L. (2009). *Wh*-in-situ, from the 1980s to now. *Language and Linguistics Compass 3*:767–791. https://doi.org/10.1111/j.1749-818X.2009.00133.x

Chomsky, N. (1964). *Current issues in linguistic theory.* The Hague: Mouton.

Chomsky, N. (1973). Conditions on transformations. In *A festschrift for Morris Halle*, ed. S. Anderson & P. Kiparsky, 232–286. Rinehart: Holt.

Chomsky, N. (1977). *On wh-movement.* New York: Academic Press.

Chomsky, N. (1986). *Barriers.* Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: The framework. In *Step by step: Essays in minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels & J. Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, N. (2001). Derivation by phase. In *Ken Hale: A life in language*, ed. M. Kenstowicz, 1–50. Cambridge, MA: MIT Press.

Chomsky, N. (2004). Beyond explanatory adequacy. In *The cartography of syntactic structures: Vol. 3. Structures and beyond*, ed. A. Belletti, 104–131. Oxford: Oxford University Press.

Christensen, K. R., Kizach, J., & Nyvad, A. M. (2012). Escape from the Island: Grammaticality and (Reduced) Acceptability of *wh*-island Violations in Danish. *Journal of Psycholinguistic Research, 42*, 51 – 70. https://doi.org/10.1007/s10936-012-9210-x

Drummond, A. (2020). *Ibex farm. 2013.* Available at: http:spellout.net/ibexfarm

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua 119*, 67–88. https://10.1016/j.lingua.2008.09.002

Goodall, G. (2011). Syntactic satiation and the inversion effect in English and Spanish *Wh*-questions. *Syntax, 14*, 29-47. https://doi.org/10.1111/j.1467-9612.2010.00148.x

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. Journal of experimental psychology. *Learning, memory, and cognition, 27*(6), 1411–1423. https://doi.org/10.1037//0278-7393.27.6.1411

Gordon, P. C., Hendrick, R., & Johnson, M. R. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language, 51*, 97–114. https://doi.org/10.1016/J.JML.2004.02.003

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-Load Interference in Syntactic Processing. *Psychological Science, 13*(5), 425–430. http://www.jstor.org/stable/40063874

Hiramatsu, K. (2000). *Accessing linguistic competence: Evidence from children's and adults' acceptability judgments*. [Doctoral Dissertation]. Mansfield: University of Connecticut, Storrs.

Hofmeister, P., Jaeger, T., Sag, I., Arnon, I. & Snider, N. (2007). Locality and accessibility in wh-questions. In *Roots: Linguistics in Search of its Evidential Base*, ed. S. Featherston & W. Sternefeld, 185-206. Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110198621.185

Hofmeister, P., Jaeger, T. F., Arnon, I., Sag, I. A., & Snider, N. (2013). The source ambiguity problem: Distinguishing the effects of grammar and processing on acce ptability judgments. *Language and Cognitive Processes, 28*(1-2), 48–87. https://doi.org/10.1080/01690965.2011.572401

Hofmeister, P., & Sag, I. A. (2010). Cognitive Constraints and Island Effects. *Langua ge, 86*(2), 366–415. https://doi.org/10.1353/lan.0.0223

Hofmeister, P., & Vasishth, S. (2014). Distinctiveness and encoding effects in online sentence comprehension. *Frontiers in Psychology, 5*: 1237. http://dx.doi.org/10.3389/fpsyg.2014.01237

Huang, C. T. (1982a). *Logical relations in Chinese and the theory of grammar*. [Doctoral dissertation]. Cambridge, MA: MIT dissertation.

Huang, C. T. (1982b). Move *wh* in a language without *wh*-movement. *The Linguistic Review 1*:369–416. https://doi.org/10.1515/tlir.1982.1.4.369

Keshev, M., & Meltzer-Asscher, A. (2018). A processing-based account of subliminal wh-island effects. *Natural Language & Linguistic Theory, 37*, 621–657. https://doi.org/10.1007/S11049-018-9416-1

Kluender, R. E., & Kutas, M. (1993). Subjacency as a processing phenomenon.*Langua ge and Cognitive Processes, 8*, 573–633. https://doi.org/10.1080/01690969308407588

Kuznetsova, A., Brockhoff, B., & Christensen, H. B. (2015). *lmerTest: Tests in Linear Mixed Effects Models. R Package Version* 2.0–29. Available at: http://CRAN.R-project.org/package=lmerTest

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive science, 29*(3), 375–419. https://doi.org/10.1207/s15516709cog0000_25

Lu, J., Thompson, C. K., & Yoshida, M. (2020). Chinese *Wh*-in-Situ and Islands: A

Formal Judgment Study. *Linguistic Inquiry, 51*, 611–623. https://doi.org/10.1162/ling_a_00343

McDaniel, D., & Cowart, W. (1999). Experimental evidence for a minimalist account of English resumptive pronouns. *Cognition 70*, B15–B24. https://10.1016/S0010-0277(99)00006-2

McElree, B. (2006). Accessing recent events. In *The psychology of learning and motivation: Advances in research and theory*, ed. B. H. Ross, 155 ‑ 200. Elsevier Academic Press. https://doi.org/10.1016/S0079-7421(06)46005-9

Nairne J. S. (2002). The myth of the encoding-retrieval match. *Memory (Hove, England), 10*(5-6), 389–395. https://doi.org/10.1080/09658210244000216

Nunes, J. M. (1995). *The copy theory of movement and linearization of chains in the minimalist program*. [Doctoral dissertation]. College Park: University of Maryland.

Pan, V. J. (2017). Phasal Agree as Precondition on Multiple Transfers: Evidence From A-bar Dependency in Chinese. In *2017 The Syntax-Morphology Interface in Generative Grammar: Proceedings of The 19th Seoul International Conference on Generative Grammar*, ed. Changguk Yim, 261–276. The Korean Generative Grammar Circle.

Pesetsky, D. (1987). Wh-in-situ: movement and unselective binding. In *The Representation of in Definiteness*, ed. E. J. Reuland and A. G. B. ter Meulen, 98–129. Cambridge, MA: MIT Press

R Core Development Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rizzi, L. (1990). *Relativized minimality.* Cambridge, MA: MIT Press.

Rizzi, L. (2013). Locality. *Lingua. 130*, 169–186. https://10.1016/j.lingua.2012.12.002

Ross, J. R. (1967). *Constraints on variables in syntax*. [Doctoral dissertation]. Cambridge, MA: MIT dissertation.

Schütze, C. T. (1996). *The Empirical Base of Linguistics: Grammaticality Judgment and Linguistic Methodology*. Chicago: The University of Chicago Press.

Schütze, C. T., & Sprouse, J. (2013). Judgment data. In *Research Methods in Linguistics*, ed. R. J. Podesva & D. Sharma, 27–50. New York, NY: Cambridge University Press.

Shlonsky, U., Villata, S., & Franck, J. (2020). Covert Movement in Multiple-*Wh* Questions: Experimental and Theoretical Investigations. *Syntax, 23*, 185–202. https://doi.org/10.1111/synt.12192

Snyder, W. (2000). An experimental investigation of syntactic satiation effects. *Linguis tic Inquiry 31*:575–582. https://doi.org/10.1162/002438900554479

Sprouse, J. (2009). Revisiting Satiation: Evidence for an Equalization ResponseStrategy. *Linguistic Inquiry, 40*(2), 329–341. https://doi.org/10.1162/ling.2009.40.2.329

Sprouse, J., Matt W., & Colin P. (2012). A test of the relation between working memory capacity and syntactic island effects. *Language 88*:82–123. https://doi.org/10.1353/LAN.2012.0004

Sprouse, J., & Norbert H. (2013). *Experimental syntax and island effects.* Cambridge: Cambridge University Press.

Sprouse, J., Shin F., Hajime O., & Robert K. (2011). Reverse island effects and the backward search for a licensor in multiple wh-questions. *Syntax 14*:179–203. https://doi.org/10.1111/j.1467-9612.2011.00153.x

Stepanov, A., Mušič, M. & Stateva, P. (2018). Two (non-) islands in Slovenian: A study in Experimental Syntax. *Linguistics, 56*(3), 435–476. https://doi.org/10.1515/ling-2018-0002

Tsai, W. D. (1994). On nominal islands and LF extraction in Chinese. *Natural Language & Linguistic Theory 12*: 121–175. https://doi.org/10.1007/BF00992747

Uriagereka, J. (1999). Multiple spell-out. In *Working minimalism*, ed. S. D. Epstein & N. Hornstein, 251–282. Cambridge, MA: MIT Press.

Van Dyke J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of experimental psychology. Learning, memory, and cognition, 33*(2), 407–430. https://doi.org/10.1037/0278-7393.33.2.407

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of memory and language, 55*(2), 157–166. https://doi.org/10.1016/j.jml.2006.03.007

Villata, S., Rizzi, L., & Franck, J. (2016). Intervention effects and Relativized Minimality: New experimental evidence from graded judgments. *Lingua, 179*, 76–96. https://doi.org/10.1016/J.LINGUA.2016.03.004

Acknowledgements

For insightful advice and checking the development process of this study, I would express special thanks to Lian Zhang.

Funding

The experiment expenditure was paid by *the Fundamental Research Funds for the Central Universities* (Grant Number: 2024JX078).

Appendices

Appendix A

The Stimuli for Experiment 1 (Translated Version)

Experiment 1	Block	Item	Condition	Sentence
1-5 Non-island	A	1	Context	You guess that he likes bananas and apples.
6-10 *Wh*-Island			Test sentence	What kinds of fruits do you guess that he likes?
		2	Context	You claimed that we were afraid of that scary dog.
			Test sentence	What animal did you claim that we were afraid of?
		3	Context	You are whispering that Tom has stolen his mum's wallet because he simply lacks money.
			Test sentence	Why are you whispering that Tom has stolen his mum's wallet?
		4	Context	You believe that this team will win the least rivalrous competition.

			Test sentence	What competition do you believe that this team will win?
		5	Context	Professor Green announced that all the students would participate in the meeting.
			Test sentence	Which professor announced that all the students would participate in the meeting?
		6	Context	You are sure that the armed forces attacked that puny fan because he broke the on-site order .
			Test sentence	Why are you sure what attacked that puny fan?
		7	Context	Peter is complaining that the boss has made an offense.
			Test sentence	What is Peter complaining who has made?
		8	Context	The radical party

				has promised that the conservative party would never come back to politics.
			Test sentence	What has the radical party promised who would never come back to?
		9	Context	I doubt whether Jimmy gossiped about the poorest victim.
			Test sentence	Which do I doubt whether Jimmy gossiped about?
		10	Context	You wonder whether we excluded Sarah from our group.
			Test sentence	Who do you wonder whether we excluded from our group?
	B	1	Context	You suspect whether we hated *Gone with the Wind*.
			Test sentence	Did you suspect

				whether we hated *Gone with the Wind*?
		2	Context	You asked whether we moved away from Beijing.
			Test sentence	Did you ask whether we moved away from Beijing?
		3	Context	My youngest sister complained that we twins ate the cake.
			Test sentence	Who complained that we twins ate the cake?
		4	Context	Karl complained Mr. White assigned so much work.
			Test sentence	Which student complained that Mr. Green assigned so much work?
		5	Context	Little John complained that we felt ashamed of Joe.
			Test sentence	Which child complained that we felt ashamed of Joe?
		6	Context	The childish boy

				complained that we put extra stress on all students for more relaxation.
			Test sentence	Why did the childish boy complain who we put extra stress on?
		7	Context	You think that we deserted the book *Jane Eyre*.
			Test sentence	What do you think who deserted?
		8	Context	You think that this miraculous man could live without any person.
			Test sentence	Who do you think who could live without?
		9	Context	You imagine that he will replace the position of cochariman.
			Test sentence	Which do you imagine who will replace?
		10	Context	You wonder whether we care

				about that baby cat.
			Test sentence	What do you wonder whether we care about?
	C	1	Context	You assume that a seed can grow with comfortable temperature, fertile soil, and enough water.
			Test sentence	What conditions do you assume that a seed can grow with?
		2	Context	You assume that a panda is good at rolling around and climbing trees.
			Test sentence	What skills do you assume that a panda is good at?
		3	Context	That police in black doubted whether you caused the accident.
			Test sentence	Which police doubted whether you caused the accident?

		4	Context	Mary guessed that I passed all the exams.
			Test sentence	Who guessed that I passed all the exams?
		5	Context	Mrs. Maria wondered whether he cheated in his homework.
			Test sentence	Who wondered whether he cheated in his homework?
		6	Context	You imagine that your sister can cook seafood for you because she will be back.
			Test sentence	Why do you imagine who can cook seafood for you?
		7	Context	You believed that she had committed suicide.
			Test sentence	What did you believe who had committed?
		8	Context	You observed that

				he had married that elegant lady.
			Test sentence	Who did you observe who had married?
		9	Context	You knew that I had been writing to the linguists in Oxford.
			Test sentence	Which did you know who had been writing to?
		10	Context	You wonder whether the car hit your dad.
			Test sentence	Who do you wonder whether the car hit?
	D	1	Context	He declared that the person had become his nightmare.
			Test sentence	What did he declare that the person had become?
		2	Context	He declared that no one could finish that work because of its difficulty.
			Test sentence	Why did he declare that no one could

				finish that work?
		3	Context	He declared that you could never achieve your dream.
			Test sentence	What did he declare that you could never achieve?
		4	Context	He declared that he would never forgive his previous best friend.
			Test sentence	Who did he declare that he would never forgive?
		5	Context	He declared that we would receive severe punishment.
			Test sentence	What did he declare that we would receive?
		6	Context	He declared that nobody would be the hero to save the world.
			Test sentence	What did he declare who would be the hero to save?
		7	Context	He declared that she had divorced with

				her husband.
			Test sentence	Who did he declare who had divorced with?
		8	Context	He declared that you would reach your destination.
			Test sentence	What did he declare who would reach?
		9	Context	He declared that she had cruelly killed her dog.
			Test sentence	What did he declare who had cruelly killed?
		10	Context	He declared that he had seen the whole process of the car accident.
			Test sentence	What did he declare who had seen?
	E	1	Context	You argue that he does not love you at all because he has not contacted you for three months.
			Test sentence	Why do you argue that he does not love

				you at all?
		2	Context	You hope that we pick the smallest type.
			Test sentence	What category do you hope that we pick?
		3	Context	You said that we should contemplate this course.
			Test sentence	What did you say that we should contemplate?
		4	Context	You are wondering whether I waited for my friends.
			Test sentence	Are you wondering whether I waited for my friends?
		5	Context	You advise that we can use blue.
			Test sentence	Which colour do you advise that we can use?
		6	Context	He said that all of us could succeed because we were equipped with all the skills.

			Test sentence	Why did he say who could succeed?
		7	Context	You assume that one of your classmates is going to delay graduation.
			Test sentence	What do you assume who is going to delay?
		8	Context	You believe that this illness is torturing your mum.
			Test sentence	Who do you believe what is torturing?
		9	Context	You think that your sister should choose that taller guy as boyfriend.
			Test sentence	Which do you think who should choose as boyfriend?
		10	Context	He wonders whether she has a boyfriend.
			Test sentence	Who does he wonder whether she has?
	F	1	Context	The qualified employee complains

				that we have been docking wages.
			Test sentence	Who complains that we have been docking wages?
		2	Context	The opponent guesses that you have got five offers from top universities.
			Test sentence	Who guesses that you have got five offers from top universities?
		3	Context	You felt that he held great anger towards you.
			Test sentence	What kind of feelings did you feel that he held towards you?
		4	Context	You suppose that your birthday gift will a teddy bear.
			Test sentence	What do you suppose that your birthday gift will be?
		5	Context	You said that

				nobody would like to play with the shyest kid in your class.
			Test sentence	Which kid did you say that nobody would like to play with?
		6	Context	You believe that the professor can publish two books for his creativity and diligence.
			Test sentence	Why do you believe who can publish two books?
		7	Context	You claimed that your dream had become something around the corner.
			Test sentence	What did you claim what had become?
		8	Context	You are hoping that you can watch the concert with your lover.
			Test sentence	Who are you hoping who can watch the concert with?

		9	Context	You suppose that Miss Zhang likes winter best.
			Test sentence	Which do you suppose who likes?
		10	Context	You wonder whether we missed our hometown.
			Test sentence	What did you wonder whether we missed?
	G	1	Context	You have insisted that he is innocent for not being confessed by him.
			Test sentence	Why have you insisted that he is innocent?
		2	Context	He claimed that we were constructing a novel model.
			Test sentence	What did he claim that we were constructing?
		3	Context	She believed that we had hidden the first group of people from observation.

			Test sentence	Who did she believe that we had hidden from observation?
		4	Context	He was complaining that we were not fair.
			Test sentence	Who was complaining that we were not fair?
		5	Context	She wonders whether she is a candidate.
			Test sentence	Does she wonder whether she is a candidate?
		6	Context	You are proposing that the United Nations should protect the remote area for more peace.
			Test sentence	Why are you proposing what should protect the remote area?
		7	Context	You predict that your talented brother will conquer the conundrum.
			Test sentence	What do you predict

				who will conquer?
		8	Context	We two should point that the new project is being discussed.
			Test sentence	Who should point what is being discussed?
		9	Context	You decide that we will visit the biggest city in this country.
			Test sentence	Which do you decide who will visit?
		10	Context	He wondered whether his parents adopted a new baby.
			Test sentence	Who did he wonder whether his parents adopted?

Appendix B

The Stimuli for Experiment 2 (Translated Version)

Experiment 2	Item	Condition	Sentence
a. Simple *wh*-phrase, identity, island	1	a	Who do you believe who likes?
b. Simple *wh*-phrase, identity, non-island		b	Who believes who likes this student?
c. Simple *wh*-phrase, inclusion, island		c	Who do you believe which professor likes?
d. Simple *wh*-phrase, inclusion, non-island		d	Who believes which professor likes this student?
e. *D*-linked *wh*-phrase, identity, island		e	Which student do you believe which professor likes?
f. *D*-linked *wh*-phrase, identity, non-island		f	Which classmate believes which professor likes this student?
g. *D*-linked *wh*-phrase, inclusion, island		g	Which student do you believe who likes?
h. *D*-linked *wh*-phrase, inclusion, non-island		h	Which classmate believes who likes this student?

	2	a	Who did you assume who waited for?
		b	Who assumed who waited for this stranger?
		c	Who did you assume which tour guide waited for?
		d	Who assumed which tour guide waited for this stranger?
		e	Which stranger did you assume which tour guide waited for?
		f	Which traveller assumed which tour guide waited for this stranger?
		g	Which stranger did you assume who waited for?
		h	Which traveller assumed who waited for this stranger?
	3	a	Who does he guess who has married?
		b	Who guesses who has married the marketing manager?
		c	Who does he guess which employee has married?
		d	Who guesses which

			employee has married the marketing manager?
		e	Which manager does he guess which employee has married?
		f	Which colleague guesses which employee has married the marketing manager?
		g	Which manager does he guess who has married?
		h	Which colleague guesses who has married the marketing manager?
	4	a	Who did you think who would focus on?
		b	Who thought who would focus on the new-born baby?
		c	Who did you think which nurse would focus on?
		d	Who thought which nurse would focus on the new-born baby?
		e	Which infant did you think which nurse would focus on?
		f	Which doctor thought which nurse would focus on the new-born baby?
		g	Which infant did you think

			who would focus on?
		h	Which doctor thought who would focus on the new-born baby?
	5	a	Who do you wonder who is complaining about?
		b	Who wonders who is complaining about the staff members?
		c	Who do you wonder which passenger is complaining about?
		d	Who wonders which passenger is complaining about the staff members?
		e	Which group of people do you wonder which passenger is complaining about?
		f	Which driver wonders which passenger is complaining about the staff members?
		g	Which group of people do you wonder who is complaining about?
		h	Which driver wonders who is complaining about the staff members?

	6	a	Who did you claim who gossiped about?
		b	Who claimed who gossiped about the little girl?
		c	Who did you claim which bully gossiped about?
		d	Who claimed which bully gossiped about the little girl?
		e	Which victim did you claim which bully gossiped about?
		f	Which classmate claimed which bully gossiped about the little girl?
		g	Which victim did you claim who gossiped about?
		h	Which classmate claimed who gossiped about the little girl?
	7	a	Who do you predict who will replace?
		b	Who predicts who will replace the deputy manager?
		c	Who do you predict which section chief will replace?
		d	Who predicts which section chief will replace the deputy manager?
		e	Which position do you

			predict which section chief will replace?
		f	Which clerk predicts which section chief will replace the deputy manager?
		g	Which position do you predict who will replace?
		h	Which clerk predicts who will replace the deputy manager?
	8	a	Who have you found who betrays?
		b	Who has found who betrays the team leader?
		c	Who have you found which underling betrays?
		d	Who has found which underling betrays the team leader?
		e	Which superior have you found which underling betrays?
		f	Which enemy has found which underling betrays the team leader?
		g	Which superior have you found who betrays?
		h	Which enemy has found who

			betrays the team leader?
	9	a	Who did you doubt who punished?
		b	Who doubted who punished the disobedient student?
		c	Who did you doubt which teacher punished?
		d	Who doubted which teacher punished the disobedient student?
		e	Which student did you doubt which teacher punished?
		f	Which parents doubted which teacher punished the disobedient student?
		g	Which student did you doubt who punished?
		h	Which parents doubted who punished the disobedient?
	10	a	Who does she doubt who has surpassed?
		b	Who doubts who has surpassed the best competitor?
		c	Who does she doubt which hero has surpassed?
		d	Who doubts which hero has

			surpassed the best competitor?
		e	Which competitor does she doubt which hero has surpassed?
		f	Which judge doubts which hero has surpassed the best competitor?
		g	Which competitor does she doubt who has surpassed?
		h	Which judge doubts who has surpassed the best competitor?
	11	a	Who do you imagine who will see?
		b	Who imagines who will see the superstar that has the most awards?
		c	Who do you imagine which fan will see?
		d	Who imagines which fan will see the superstar that has the most awards?
		e	Which superstar do you imagine which fan will see?
		f	Which critic imagines which fan will see the superstar that has the most awards?

		g	Which superstar do you imagine who will see?
		h	Which critic imagines who will see the superstar that has the most awards?
	12	a	Who did he imagine who would select as president?
		b	Who imagined who would select this mayor as president?
		c	Who did he imagine which voter would select as president?
		d	Who imagined which voter would select this mayor as president?
		e	Which candidate did he imagine which voter would select as president?
		f	Which representative imagined which voter would select this mayor as president?
		g	Which candidate did he imagine who would select as president?
		h	Which representative imagined who would select

			this mayor as president?
	13	a	Who do you warn who should stay away from?
		b	Who warns who should stay away from all the murderers?
		c	Who do you warn which kid should stay away from?
		d	Who warns which kid should stay away from all the murderers?
		e	Which kind of people do you warn which kid should stay away from?
		f	Which expert warns which kid should stay away from all the murderers?
		g	Which kind of people do you warn who should stay away from?
		h	Which expert warns who should stay away from all the murderers?
	14	a	Who did he propose who should fight against?
		b	Who proposed who should fight against invaders?
		c	Who did he propose which

			army should fight against?
		d	Who proposed which army should fight against invaders?
		e	Which kind of people did he propose which army should fight against?
		f	Which leader proposed which army should fight against invaders?
		g	Which kind of people did he propose who should fight against?
		h	Which leader proposed who should fight against invaders?
	15	a	Who do you propose who should protect?
		b	Who proposes who should protect the refugees?
		c	Who do you propose which party should protect?
		d	Who proposes which party should protect the refugees?
		e	Which colony do you propose which party should protect?
		f	Which citizen proposes

			which party should protect the refugees?
		g	Which colony do you propose who should protect?
		h	Which citizen proposes who should protect the refugees?
	16	a	Who did you understand who cursed?
		b	Who understood who cursed the initiator?
		c	Who did you understand which victim cursed?
		d	Who understood which victim cursed the initiator?
		e	Which bully did you understand which victim cursed?
		f	Which audience understood which victim cursed the initiator?
		g	Which bully did you understand who cursed?
		h	Which audience understood who cursed the initiator?
	17	a	Who does he assume who will become?
		b	Who assumes who will

			become a successful businessman?
		c	Who does he assume which trainee will become?
		d	Who assumes which trainee will become a successful businessman?
		e	Which kind of people does he assume which trainee will become?
		f	Which trainer assumes which trainee will become a successful businessman?
		g	Which kind of people does he assume who will become?
		h	Which trainer assumes who will become a successful businessman?
	18	a	Who did they believe who chatted with?
		b	Who believed who chatted with a group of mathematicians?
		c	Who did they believe which talent chatted with?
		d	Who believed which talent chatted with a group of mathematicians?

		e	Which group did they believe which talent chatted with?
		f	Which professor believed which talent chatted with a group of mathematicians?
		g	Which group did they believe who chatted with?
		h	Which professor believed who chatted with a group of mathematicians?
	19	a	Who do you guess who is pursuing?
		b	Who guesses who is pursuing this lovely girl?
		c	Who do you guess which boy is pursuing?
		d	Who guesses which boy is pursuing this lovely girl?
		e	Which girl do you guess which boy is pursuing?
		f	Which teacher guesses which boy is pursuing this lovely girl?
		g	Which girl do you guess who is pursuing?
		h	Which teacher guesses who is pursuing this lovely girl?

	20	a	Who did she predict who hated?
		b	Who predicted who hated the doctors?
		c	Who did she predict which patient hated?
		d	Who predicted which patient hated the doctors?
		e	Which profession did she predict which patient hated?
		f	Which journalist predicted which patient hated the doctors?
		g	Which profession did she predict who hated?
		h	Which journalist predicted who hated the doctors?
	21	a	Who do you wonder who taunted?
		b	Who wonders who taunted the minority?
		c	Who do you wonder which student taunted?
		d	Who wonders which student taunted the minority?
		e	Which community do you wonder which student

			taunted?
		f	Which participant wonders which student taunted the minority?
		g	Which community do you wonder who taunted?
		h	Which participant wonders who taunted the minority?
	22	a	Who did he argue who carried?
		b	Who argued who carried the little baby?
		c	Who did he argue which family carried?
		d	Who argued which family carried the little baby?
		e	Which child did he argue which family carried?
		f	Which partner argued which family carried the little baby?
		g	Which child did he argue who carried?
		h	Which partner argued who carried the little baby?
	23	a	Who did she wonder who deserted?
		b	Who wondered who deserted

			one of the family members?
		c	Who did she wonder which indifferent people deserted?
		d	Who wondered which indifferent people deserted one of the family members?
		e	Which necessary individual did she wonder which indifferent people deserted?
		f	Which listener wondered which indifferent people deserted one of the family members?
		g	Which necessary individual did she wonder who deserted?
		h	Which listener wondered who deserted one of the family members?
	24	a	Who does he suspect who questions?
		b	Who suspects who questions the speaker?
		c	Who does he suspect which hearer questions?
		d	Who suspects which hearer questions the speaker?
		e	Which side of people does he

			suspect which hearer questions?
		f	Which onlooker suspects which hearer questions the speaker?
		g	Which side of people does he suspect who questions?
		h	Which onlooker suspects who questions the speaker?

Appendix C

The Filler Items for Experiment 2 (Translated Version)

Filler Sentences	Question Type	Item	Sentences
	Grammatical	1	Which season do you suppose that Miss Wang likes best?
		2	Which part do you suggest that I should finish?
		3	Who did you say that this activity would invite?
		4	Who are you writing that you miss so much?
		5	What do you argue that nobody cares about?
		6	What have you promised to her that you would never do again?
		7	What kind of beliefs do you swear that no member will offend against?
		8	What kind of comfort did your mum who had just beaten you up offer?
	Middle	1	Which child do you make the decision that

			we select?
		2	Which building did the government appeal to protection of?
		3	Who are you worried that would never succeed?
		4	Who did the country construct the university that aims at?
		5	What do you complement the statement that lacks?
		6	What did he perfect this proposal that required?
		7	What category did the demand to classify still fade?
		8	What sort of party have you made the party list that was topped with?
	Ungrammatical	1	Which part did you assert that he had finished the introduction?
		2	Which class do you ask that they belong to the upper?

		3	Who took responsibility for the matter about that he did?
		4	Who do you assume that he will forget her?
		5	What are you hoping for that your dreams will come true?
		6	What do you hate most that someone lies to you?
		7	What parts of a person attract you most about that you have a deep impression?
		8	What city do you prefer to live that you can stay forever?
	Statement Type		
	Grammatical	1	The cat with black spots that I was feeding went away quickly.
		2	I don’t know the truth that the teacher who has retired announces.
		3	Our friend who claimed never to marry is now a dad of two daughters.

		4	The girl who is standing far away from us has offended the new classmate.
		5	Making friends with various people is something that one must learn.
		6	Complaining about the environment is always said to be a loser's behaviour.
		7	The professor has agreed that we could leave in advance.
		8	You should explain the fact that you are absent again.
	Middle	1	I loved that you always wore perfume.
		2	She appreciated that we gathered around after quite a long time.
		3	I hate the person that I am not sure whether you like.
		4	I knew about the discipline that they

			wondered whether I had studied.
		5	I make the decision that I will quit the job that I have tolerated it.
		6	It is on this house that the huge fire has influenced.
		7	It was this person that he who likes departed without hesitation.
		8	Nothing could stop me from going further where there is better scenery.
	Ungrammatical	1	It is our dress that we should try on new clothes.
		2	The teacher all of us would not say somebody could like.
		3	You are holding a cup that I hope I have some water.
		4	Jim will pay the bill that we are eating too much.
		5	The concert my best friend has frequently

			recommended me to watch.
		6	I have run the company for many years that I sell electronic products.
		7	There is no architecture consists of cotton.
		8	A big mistake we have urged that the students ought to correct.