British Journal of Education, Society & Behavioural Science
4(2): 176-190, 2014

SCIENCEDOMAIN *international*
www.sciencedomain.org

A Study of Filler-gap Dependencies and Extractions of English WH-Questions by Chinese English Learners

Wang Huili[1*], Wu Miao[1*] and Chen Hongjun[1*]

[1]School of Foreign Languages, Dalian University of Technology, Dalian, China.

Authors' contributions

This work was carried out in collaboration between all authors. Author WM designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Authors WH and CH managed the analyses of the study. All authors read and approved the final manuscript.

Case Study

Received 21st July 2013
Accepted 26th September 2013
Published 22nd October 2013

ABSTRACT

The study aims to investigate and discuss behavioral responses of Chinese English learners to dependencies and WH-extractions in processing of English WH-questions. Special attention is paid to subject extraction and object extraction. Meanwhile, all the target sentences were characterized into those with one clause and those with double clauses due to different dependencies. Using E-Prime, the present study examines whether second language speakers of English with advanced levels turn to working memory in the on-line comprehension of WH-extractions, and to what degree working memory gets involved. Performance was recorded and analyzed related to reaction time (RT) and accuracy to probe into other strategies when participants read. RT, as well as accuracy, is taken as dependent variable. They reflect the difficulty of various aspects in language processing.

The findings showed that long-dependency WH-questions with object-extraction are the most memory-costing and time-consuming for Chinese English-majors to process. Male participants, as well as those who have done a better job in test for English major (TEM8), got a higher accuracy. Therefore, the results suggest that not only dependency and type of extraction in WH-questions, but also gender and level of language ability have an influence

**Corresponding author: Email: heisey117@163.com;*

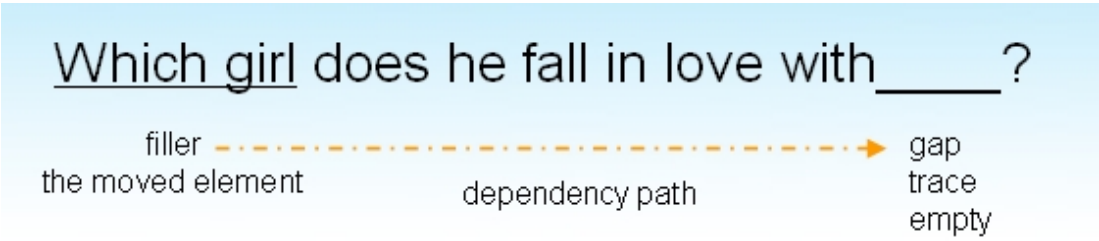
on the processing.

Keywords: WH-questions; dependency; processing preference; gender; linguistic level.

1. INTRODUCTION: THE SOUTH AFRICAN SITUATION

Sentence, as one of the most basic forms of communication, is a very significant part of discourse. Therefore, the comprehension or the processing of sentences is reckoned as one of the most essential cognitive activities for human beings. Composed of several types, sentence can be subdivided into some descriptions, and interrogative sentence, one of its subdivisions, is always a salient one in syntax, second language acquisition (SLA) and cognitive linguistics.

What is known is that most interrogative pronouns in English start with ***wh*** (**eg.** what, where, who, etc.) with the exception of ***how***. Hence, WH-question is much in evidence refers to the question which begins with a fronted interrogative phrase, and the process can be named as WH-movement. As illustrated in Pic. 1, with the movement of "which girl" to the initial position, an empty category is generated. The moved element is called filler, and the position freed up is called a gap or a trace.



Pic. 1. An illustration of WH-movement question

Sentence comprehension critically depends upon the reconstruction of grammatical relations between arguments and the predicate of a sentence by assigning a hierarchical structure to the input of words [1]. Working memory not only maintains the finite amount of information in an active status, but helps people deal with that information as well, making itself necessary and critical for the processing of language.

1.1 WH-Questions

WH-movement has become very productive. Chomsky propounded it in 1977 and revised the approach in 1981. WH-movement is not only favored in generative linguistics but also in the cognitive field. Here, attention is will be paid to cognitive linguistics specifically. Researchers in this field came up with a series of theories, and an example is Active Filler Strategy (AFS). AFS was brought up for describing the subject-object ambiguity of the WH-filler [2]. Some other models have also been introduced by scholars in the matter of WH-movement, such as Lexical Expectation Model (transitivity preference) of Fodor in 1978, and All Resorts Model (plausibility) of Stowe in 1984 [3]. Stowe managed to testify Lexical Expectation Model of Fodor's. Whereas, the results were contrary to the predictions of

Stowe's All Resorts Model, the Last Resort Model as well as the simple version of the First Resort Model.

With the development of more and more mature technologies, the methods used by researchers varied from behavioral study to neuroscience. One of the examples of latter is event-related potentials (ERPs). Researchers have come a long way to take advantage of ERPs to examine the processing of dependencies of WH-movement [3,4], and constraints or violations on movement in processing [5,6]. Other languages like German and Japanese of WH-movement have also been studied by ERPs in terms of memory costs in the syntactic processing [7,8,9].

Languages differ, so do the orders within them. Variations in word order make sentences more difficult to process. As for English, WH-movement can cause a noncanonical word order. Compared with English, Chinese is very different since it is a wh-in-situ language, distinct from English in the formation of this sort of movement.

a. Tom ai shang le shui"?
b. Tom is falling in love with whom?

What can be seen from above is that the interrogative phrase *shui* (whom) is not moved to the very beginning of the sentence, instead, it is located after the verb phrase. The interrogative movement in Chinese is covert, unlike the overt WH-movement in English. In other words, Chinese is a WH-in-situ language. Many researchers focus on the acquisition of WH-questions in SLA of Chinese or other language learners. Kim found the island constraints could be observed in all learner groups, but for native controls, that-trace effect occurs by means of experimental syntax [10]. Universal Grammar is another constraint for adult EFL learners' competence during the processing of WH-movement [11]. Except for constrains in SLA of WH-questions, the asymmetry of extractions in WH-questions for learners of different nationalities is also to be discussed in the present study.

WH-extractions can be varied due to their parts of speech. However, they can be fronted over an indefinitely large amount of words or phrases. The distance between WH-filler and trace is called unbounded dependency or dependency for short. For this reason, the type of extraction and dependency between gap and filler are essential factors affecting the processing of WH-questions.

Previous researches have shown that there are different preferences for subject-extraction or object-extraction WH-movement. On one hand, there is a tendency for the preference of subject-extraction WH-sentences. Participants were found to offer a faster reply to the WH-subject extraction sentences than the WH-object extraction ones when they were provided with Italian WH-questions [12]. Those questions manipulated the WH-extractions site in order to figure out that the parser is sensitive to the type of dependencies following the Minimal Chain Principle (MCP). The result showed that there had been a processing advantage for the WH-subject extractions. WH-questions, which also include relative clauses, show a preference for those with subject extractions over the ones with object extractions [13]. Evidence comes from a variety of measures, such as acceptability judgments, reading times as well as other neural scientific methods like ERPs, PET, or fMRI. Researchers attested to the preference not only owing to longer reading time with a lower acceptability ratings during processing object WH-dependencies, but also for the detection of slow anterior negative brain potentials. Therefore, the processing load was increased during processing.

On the other hand, disagreement has been put forward for all time. An accumulation of literature has demonstrated that readers found it more effort-consuming to maintain a long dependency between filler and its gap when they managed to process WH-movement sentences. For instance, no matter native speakers or L2 learners, neither of them managed to judge grammatical subject extractions, while they both have no difficulty in judging sentences involving object extractions [14,15,16]. Schachter and Yip found native and nonnative English speakers to finish their designed grammaticality tasks with object and subject structures. The finding demonstrated that native speakers exhibited a processing preference, as did nonnatives whose mother tongue did not bias them towards another certain structure. Juffs selected three groups of participants speaking different native languages. Either of the groups has done an online work to judge the grammaticality of WH-questions with varying extractions. It tuned out that Subject extraction from a finite clause is particularly difficult for all learners, and object extraction from a finite clause was the easiest. Among those three groups, the Japanese had the greatest difficulty with subject extractions, then, the Chinese and Spanish respectively. All of them showed a rough sledding in the sentences with the subject extractions than with the object extractions. The outcomes suggested that two finite verbs verbs next to one another may be an important factor causing parsing breakdown. As for the influence of L2, if there is no WH-movement in the L1, word order in the L1 has an additional negative effect on processing. Some type of subject extractions, as the ones from nonfinite clauses and that-trace, are more difficult to process and judge [17]. The acquisition of those types of extractions in WH-questions has also been studied with the result that object questions are acquired at the same age or earlier than subject ones [18], which seems to be contradictory with the result from Amanda and her colleagues, though the preference of subject-object asymmetry was not much mentioned [19]. They found that there were an appropriate response to subject-question from infants by 15 months of age, and infants responded properly to both subject- and object-questions by 20 months.

1.2 The Present Study

Four objectives remain to be proved in the present study. First, which kind of WH-extraction would Chinese English learners generate, subject extraction or object extraction? Second, the study intends to find out to which degree the extraction affects the processing and comprehension of the English WH-questions. Third, evidence can be found regarding whether the length of dependency can be counted as an element influencing the processing of English WH-questions for Chinese English learners with high proficiency. Finally, it is expected to seek out the reasons for Chinese English learners' difficulties when they process specific type of English WH-questions. When the demonstrations come out, it is promising that the processing of language could be integrated with cognition in a better way. Meanwhile, the author anticipates drawing inspirations on English WH-questions teaching.

Three presumptions will be made in this study. Firstly, since the clause inserted will prolong the dependency, it is presumed that the more syntactic information is maintained during the processing, the more working memory of longer duration would be required.

Secondly, it is assumed that WH-questions with object extraction of long dependency would require longer reaction time (RT), and that Chinese English learners would show a higher accuracy in object-extraction WH-questions than in subject-extraction ones.

The last presumption is that gender or levels of language might also have an impact on the results.

2. METHODOLOGY

2.1 Materials

In order to testify the hypotheses proposed above, a series of stimuli were set and targeted. Half of the stimuli came from the original design of Juffs and Harrington [20]. The rest of them were written based on the original edition and partly modified regarding dependency and WH-extraction. As this study mainly focuses on the syntactic factors contributing in the processing of comprehension, the semantic impact should be minimized. 64 plausible stimuli from Juffs and Harrington were later expanded from 8 or 9 words to 14 or 15 words when another clause was embedded. For this reason, 64 pairs of experimental sentences were applied here. Half of the pairs (64 sentences) were composed of subjective fillers for the main clause verbs, and the rest of the pairs contained the WH-extraction as objective fillers for the verbs in main clauses. Except the extraction site, dependency was also taken into consideration. The stimuli with long dependency got two embedded clauses, and the later added clause was a relative one. None of the verbs checked appeared more than twice. Consequently, the judgments cannot be affected by the frequency or verb bias. Each type of experimental sentence is exemplified below.

(2-1) Who did the client mention shot the banker? (Short-dependency subject-extraction WH-question)
(2-2) Who did the client mention the banker shot? (Short-dependency object-extraction WH-question)
(2-3) Which one did the client who had been in great trouble believe shot the tricky banker? (Long-dependency subject-extraction WH-question)
(2-4) Which one did the client mention the tricky banker who had been in great trouble shot? (Long-dependency object-extraction WH-question)

Apart from the stimulus sentences, 96 filler sentences were added, including 16 grammatical interrogative sentences (examples from 2-1 to 2-4) and 80 ungrammatical ones (examples as 2-6 or 2-7). All the ungrammatical WH-questions were restricted in certain types of island constraints, and they were also classified into two parts concerning length. The rest of the filler sentences were grammatical. Additionally, they began with other types of extraction instead of subject or object ones in the target sentences. All target sentences are direct questions.

(2-5) When did the CEO who had been appointed recently claim new divisions to be set up? (Grammatical filler)
(2-6) *Who do you meet Tom after you met? (Adjunct island constraint)
(2-7) *Who did you buy a picture of Sue and? (The coordinate structure constraint)

2.2 Participants

Before the formal experiments, every participant was asked to fill in a questionnaire concerning personal basic information, including their language learning history. During the pretest, 20 postgraduates were invited to do an off-line test, in which they were asked to do the grammatical judgment of all the stimuli. They should point out the unfamiliar words or phrases in order to get rid of the lexical impact while reading. The participants of the pretest have either passed the Test for English Majors-Band 8 (TEM 8), or achieved the same level with English-major students. Another 76 participants from Dalian University of Technology

(DUT) and Liaoning Normal University (LNNU), including 62 female and 14 male with the certificate of TEM 8, finally participated in the formal experiments. They are all right-handed participants between 22 to 26 years old. The English periods of studying as a second language varied from 10 to 15 years. Every participant received a payment in reward for this study. The scores of their test results range from 60 to 80.

Table 1. Grades of Test for English Majors-Band 8 (TEM 8) on two levels

Language level	Grades of TEM8		
	N	Mean	SD
Good	26	75.54	2.79
Pass	40	64.88	2.93

In order to verify the influence of language ability, participants were divided into two groups. The good group (M=75.54, SD=2.79) scored 70 or above, and the pass group's (M=64.88, SD=2.93) ranged from 60 to 69. With the help of SPSS, the grades between those groups were tested through Independent T-test, and a significant difference ($p < 0.01$) was found.

2.3 General Procedure

Participants accomplished the experiments in two phases, and they finished each part individually in a closed and quiet room. Each section of the experiments comprised four parts, including practice, instructions, experimental part, and concluding words. Facing a monitor, participants were seated in front of a personal computer. Before the experiments, they were all required to do a practice so as to get familiar with the experimental environment. The design of experiments was completed by E-Prime 2.0 edition, and their reaction time (RT) and accuracy were recorded. The data collected afterwards were calculated by SPSS 16.0 edition.

All the participants were given a number of target sentences. The whole experiments were divided into two parts with a two-week interval, not only because of the participants' capacity of workload, but also due to some repeated filler sentences.

In the process of data collection, every stimulus sentence was read word by word in a moving window. The speed of reading was set as follows:

- Instructions of the experiment were displayed on the screen at the participants' disposal. As soon as they finished reading, they could start the experiment by pressing any button.
- A fixation mark appeared on the screen at the very beginning of each sentence for 500ms.
- Each word was presented for 700ms.
- The duration of the blank screen shown between every two words was 400ms.

Words could be seen in black case letter in size 18 on a white background. A question for judging the grammaticality of each stimulus was raised after the last word of each sentence. Participants should react by pressing the key "F" or "J" on the key board.

3. RESULTS

3.1 Results Based on Different Proficiencies

3.1.1 Results of reaction time (RT)

The reaction time (RT) for both genders with different levels of language is shown below.

Table 2. Results of reaction time based on different proficiencies

Gender	good		pass		Sig.
	M	SD	M	SD	
Male	2.48	1.59	2.41	1.43	0.095
Female	3.12	1.76	2.04	1.12	0.101
Sig.	0.572		0.587		

The t-test results demonstrate that there exists no significant difference ($p > 0.05$) between male and female in terms of RT or between good and pass group in terms of RT.

3.1.2 Results of accuracy

The accuracy for both genders with different language levels is shown below.

Table 3. Results of accuracy based on different proficiencies

Gender	good		pass		Sig.
	M	SD	M	SD	
Male	83.40	10.62	59.15	3.21	0.013
Female	67.50	16.63	55.57	15.57	0.060
Sig.	0.163		0.002		

The t-test results indicate that within the pass group there is a significant difference ($p < 0.05$) between male and female in terms of accuracy. Male do a much better job than female because the mean of male (M=59.15) is higher than that of female (M=55.57). Meanwhile, there also exits a significant difference ($p < 0.05$) between good and pass group for male. Looking at female alone, a marginal significant difference ($0.05 < p < 0.10$) exists between good and pass group. The mean of good group (M=83.40) is higher than that of pass group (M=59.15) with respect to male. Therefore, male with good grades judged more accurately than those who just passed the language test. Female with good marks were more accurate than those with lower marks despite the marginal significant difference, because the mean of female in good group (M=67.50) is higher than that of pass group (M=55.57).

To sum up, language level has no effect on RT but influences accuracy of judgment. The good group obviously has a more efficient and better judgment of the tasks. Gender is also a factor that has an effect on accuracy, and male did better judgments.

3.2 Results for Participants with the Same English Proficiency on Different Types of Sentences

The results of RT and accuracy for the participants on the same level of language while reading different types of sentences are shown in the following tables.

3.2.1 Results of reaction time (RT)

Table 4. Results of reaction time based on different types of sentences for participants with the same English level

Structure		Good			Pass		Sig.
	Sig.	M	SD	Sig.	M	SD	
LS-ex.	0.059	3.30	2.39	0.004	2.75	2.15	0.021
SS-ex.		2.66	1.54		1.89	1.05	0.055
LO-ex.	0.015	3.65	2.28	0.032	2.12	1.23	0.307
SO-ex.		2.27	1.18		1.67	0.98	0.239

The statistics showed in the table above were calculated by one-way ANOVA. The participants in good group spent longer time in judging long-dependency WH-questions with subject extractions (LS-ex., M=3.30) than short-dependency subject-extraction ones (SS-ex., M=2.66), and the difference is marginal ($0.05 < p < 0.10$). Concerning the RT- related comparison between LO-ex. (M=3.65) and SO-ex. (M=2.27) WH-questions, the difference is significant ($p < 0.05$) within the same group. So is the situation in the pass group. The difference is significant either in the RT of LS-ex. WH-questions (M=2.75) and SS-ex. ones (M=1.89), or in that of LO-ex. WH-questions (M=2.12) and So-ex. ones (M=1.67).

With respect to the rest of the data, obtained from paired-samples t-test, it can be found that despite a marginal significant difference ($0.05 < p < 0.10$), the participants with higher marks took longer time to judge SS-ex. WH-questions (M=2.66) than the pass group (M=1.89). The pass group (M=2.75) also had faster responses on the LS-ex. than the good group (M=3.30), and the difference is significant ($p < 0.05$). However, the comparison of WH-questions with object extractions was not that significant between this pair of groups.

3.2.2 Results of accuracy

Table 5. Results of accuracy based on different types of sentences for participants with the same English level

Structure		Good			Pass		Sig.
	Sig.	M	SD	Sig.	M	SD	
LS-ex.	0.000	57.69	17.34	0.000	55.62	20.35	0.566
SS-ex.		71.63	16.31		63.88	22.88	0.516
LO-ex.	0.125	63.22	22.57	0.003	54.50	15.44	0.300
SO-ex.		69.71	15.28		67.25	20.97	0.414

Regarding the accuracy for participants with different levels of language, there is no significant difference in one-way ANOVA between good and pass groups. Turning to the good group, participants were more accurate when judging SS-ex., as the mean of it (M=71.63) is much higher than that of LS-ex. (M=57.69), and the difference is significant (p

<0.05). Speaking of the other participants, the difference is also significant ($p<0.05$) between LS-ex. and SS-ex. that is more accurate to judge, since the mean of the latter type (M=63.88) is higher than that (M=55.62) of the former. The situation is similar between LO-ex. and SO-ex. within pass group. SO-ex. seems easier to judge with a higher mean (M=67.25) than LO-ex. (M=54.50), and their difference is significant ($p<0.05$).

3.3 Results of Participants of the Same Gender on Different Types of Sentences

When it comes to different types of sentences, the results of RT and accuracy for male and female are shown below.

3.3.1 Results of reaction time (RT)

Table 6. Results of reaction time based on different types of sentences for participants of the same gender

Structure		Male			Female		Sig.
	Sig.	M	SD	sig.	M	SD	
LS-ex.	0.175	2.76	1.79	0.002	2.54	1.86	0.958
SS-ex.		2.56	1.51		2.01	1.32	0.918
LO-ex.	0.741	2.46	1.64	0.001	3.13	2.31	0.154
SO-ex.		1.98	1.11		1.90	0.97	0.694

From the figures above, it can be found that the difference between male and female by one-way ANOVA is not significant, regardless of type of structure. Nevertheless, within the female group, a significant difference ($p<0.05$) did appear between WH-questions with LS-ex. and those with SS-ex. respectively. Similarly, concerning the understanding of LO-ex. WH-questions and SO-ex. ones, the former type (M=3.13) required longer time than the latter (M=1.90), and their difference is significant ($p<0.05$). However, as for male, no significant difference can be found either between the LS-ex. WH-questions and SS-ex. ones, or between the LO-ex. WH-questions and SO-ex. ones.

3.3.2 Results of accuracy

Table 7. Results of accuracy based on different types of sentences for participants of the same gender

Structure		Male			Female		Sig.
	Sig.	M	SD	Sig.	M	SD	
LS-ex.	0.012	68.31	23.76	0.000	53.93	15.56	0.166
SS-ex.		79.02	19.32		65.82	18.99	0.060
LO-ex.	0.043	62.94	16.78	0.007	55.95	20.24	0.181
SO-ex.		81.70	16.58		62.91	19.94	0.049

Turning to the situations of accuracy, it can be detected that male (M=79.02) did a better job than female (M=65.82) in terms of the processing of SS-ex. WH-questions, though the significant difference is marginal ($0.05<p<0.10$). A similar situation was found in the processing of SO-ex. ones. The significant difference ($p<0.05$) between the two genders

indicates that male (M=81.70) made more precise judgments than female (M=62.91). Looking at male or female alone, the paired-samples t-test demonstrates that there exist significant differences ($p<0.05$) within the LS-ex. and SS-ex. pair, and the LO-ex. and SO-ex. one. Both genders showed a higher accuracy when judging SS-ex. and SO-ex. WH-questions.

3.4 Results of Different Genders or Proficiencies on Judgment

The results of RT and accuracy for participants of different genders or with different levels of language are shown as follows.

Table 8. Results of different genders on judgment

Gender	Male		Female		Sig.
	M	SD	M	SD	
RT	2.44	1.37	2.39	1.42	0.935 0.007
ACC	73.01	15.10	59.40	16.64	0.055 3.931

It can be seen that the difference in RT between participants of different genders is not significant, but there is a marginal difference ($0.05<p<0.10$) in accuracy. Male (M=73.01) judged more accurately than female (M=59.40).

Table 9. Results of different proficiencies on judgment

Level of Language	good		pass		Sig.
	M	SD	M	SD	
RT	2.97	1.68	2.10	1.15	0.112 3.542
ACC	65.57	14.92	60.01	18.02	9.346 0.912

As for the participants with different levels of language, there is no significant difference in either RT or accuracy.

3.5 Correlations of Reaction Time (RT) and Accuracy with English Proficiency

In order to have a better overview of the relations between proficiency of English majors and the results of different types of sentences, the correlations were calculated and shown below. In terms of RT below, it can be seen that all the participants have a significant correlation ($p<0.05$) between RT and their English proficiency in the processing of LS-ex. WH-questions. As for SS-ex. ones, the correlation is found to be marginal significant ($0.05<p<0.10$).

Table 10. Correlations between RT and level of language

Level of language	RT			
	LS-ex.	SS-ex.	LO-ex.	SO-ex.
Sig.	0.030	0.066	0.108	0.229

Concerning the correlations between accuracy and language level as shown below, it is clearly seen that except for SS-ex. WH-questions, the other types have revealed a significant ($p<0.05$), or a marginal correlation ($0.05<p<0.10$).

Table 11. Correlations between accuracy and level of language

Level of language	ACC.			
	LS-ex.	SS-ex.	LO-ex.	SO-ex.
Sig.	0.058	0.112	0.011	0.067

4. DISCUSSION

4.1 The Influence of Dependency on the Processing of Different Types of WH-Questions

The comprehension of language is made up of a pair of procedures, which basically include processing and storage. The achievement of syntax is a part of processing, and storage provides necessary input materials for various levels of processing. Working memory system, a kind of temporarily processing and storing system for linguistic information, can be considered as the resource of cognition and provides linguistic materials. Thus, the processing and storage that occur at the same time may have an interactional influence on each other. From the analysis above, it is not hard to conclude that the longer the dependency of is, the more working memory a WH-question requires.

The results from the analysis of reaction time (RT) and accuracy of various types of WH-questions indicated that, within each group with different English proficiencies, there did exist a significant difference between SS-ex. and SL-ex. WH-questions. The participants in the pass group showed another significant difference between SO-ex. and LO-ex. WH-questions. participants were more accurate on SS-ex. and SO-ex. WH-questions. Therefore, dependency does influence accuracy. Since the longer dependency a WH-question has, the more working memory it requires during the processing, WH-questions with short dependency are definitely easier to process, and they do not require that much working memory in the storage system.

Regarding RT of different language proficiencies, each group spent more time in judging long-dependency WH-questions than short-dependency ones regardless of the kind of extraction the questions have. Thus, it is not hard to see that all the participants do well in judging short-dependency WH-questions, which is consistent with the result mentioned above. Therefore, dependency is an influential factor during the processing. As the result stated, accuracy for SS-ex. WH-questions surpasses that for SO-ex. WH-questions. Since the dependency of object-extraction WH-questions is longer than that of subject-extraction ones, the result makes sense proving that long dependency will prolong RT and lower accuracy in processing.

As the long-dependency WH-questions in this paper refer to the ones with another clause which will interrupt the understanding of the main clause, the inserted clause will undoubtedly prolong the filler and gap no matter which type of extraction it belongs to. Since the more clauses inserted there are, the more they will interrupt the processing, the information presented before the clause has to be retained in working memory, or, can only be activated afterwards. In addition, different roles of fillers given by types of extraction in WH-questions make working memory different. For example, a NP in a subject-extraction WH-questions is taken as subject either in the main clause or in the subordinate clause, whereas it will be both a subject and an object in the object-extraction WH-questions, let alone double clauses inserted. Hence, the load of working memory is bigger in processing object-extraction WH-questions. All the participants had certain working memory capacity; it is no wonder that as the length of dependency increases for the participants on the same level, the accuracy will decrease.

4.2 The Influence of Type of Extraction on the Processing of Different Types of WH-Questions

The results indicate that there is a marginal difference for all participants with respect of SO-ex. and SS-ex. WH-questions. The former type takes longer time to process, so when the dependency remains unchanged, extraction affects the results of RT, and subject-extraction WH-questions are easier to process. As a result, type of extraction influences the processing. According to the results of correlations in 3.5, it is clear to get the verdict that there is a positive correlation between accuracy and language ability. For this part, the more proficient the participants have, the more accurately they perform when judging LO-ex. as well as short-dependency WH-questions. Thus, the type of extraction does have an effect on the processing of WH-questions. Due to the positive correlations found between English proficiency and accuracy, and the fact that there is a longer dependency between the WH-filler and its gap, it can be deduced that WH-questions with object-extraction do not easily make sense and are hard to process. In terms of the correlation between RT and language ability, it can be analyzed that no matter how long the dependency is, subject-extraction WH-questions decides the judging RT. As for the type of extraction, object-extraction requires more resources of working memory because of the longer dependency between filler and gap.

4.3 The Roles of Gender and Proficiency in Processing WH-Questions

The experiments manifest that gender and English proficiency are meaningful for Chinese English learners in the course of processing and understanding.

On the one side, gender has an impact on the processing of various types of WH-questions. Concerning accuracy, significant differences were found within each group or between different groups in 3.1. As for participants with the same language level (pass group), male showed a higher accuracy than female, and the difference was significant, which indicates that gender is a factor that influences accuracy, and male obviously have an advantage over female. This has also been proved in the results in 3.3, as male judged more accurately than female in terms of object-extraction WH-questions with short and long dependencies respectively, with a significant difference between the two groups. It has been confirmed again that gender differed in the processing of WH-questions and male outranged female. Moreover, it has been found in 3.4 that no matter which language level the participants have, male were more accurate in the judging tasks than female, in spite of a marginal significant

difference. Knowledge experience or information structure of different genders may explain those differences. It has been investigated in some domestic researches that male have a stronger logical capability, a broader spatial vision, and more extra-curriculum experiences, resulting in the fact that they performed better than female in reading and processing.

On the other side, participants with a better performance in TEM8, regardless of gender, performed more accurately when judging. From the data and analysis in 3.1, it is not hard to see that level of language for Chinese English-majors, as well as gender, can be counted as an element influencing accuracy. Furthermore, the result in 3.2 corresponds with that mentioned above. During the processing of LS-ex. WH-questions SS-ex. ones, a significant difference and a marginal significant difference were found in good group. Therefore, the participants were asked to process the same type of WH-question, level of language becomes a reason for those differences. No significant difference was found in some respects between participants with different language levels due to some possible reasons. Firstly, there may be some nonintellectual factors which have an impact on the RT, such as the adaption to the environment, mood, or physical condition, etc, Secondly, RT can measure the reaction time activated by representation instead of its maintenance, so the weight balance of RT and accuracy need to be adjusted . Although there is no significant difference between participants with different proficiencies, the function of working memory on the understanding and processing of types of WH-questions still owes to the significant difference of accuracy.

5. CONCLUSION

It cannot be neglected that differences do exist between Chinese and English WH-questions (Liang, 2012). On one hand, unlike Chinese, English is inclined to obey the principle of subjacency. On the other hand, these two languages select different values for the uninterpretable feature [wh-] that is overt in English and thus triggering WH-movement while covert in Chinese with no WH-movement as mentioned. For these reasons, it can be concluded as follows:

1) Gender can be counted as a factor which influences the processing of WH-questions. Generally, male showed more accuracy than female in the experiments.
2) Level of language for the Chinese English-majors also plays a significant role in processing. The more proficient their English is, the more accurately they judge.
3) Dependencies and types of extraction in WH-questions are essential in the processing, too. Long dependency makes Chinese English-majors more time-consuming and memory-costing to process grammatical WH-questions. Participants are inclined to judge grammatical WH-questions with short dependency more quickly and accurately. They preferably processed short WH-questions with subject extraction. Therefore, short-dependency subject-extraction WH-questions are the most preferred to process for Chinese English-majors.

Although the results are of importance in academic research, there still exist some flaws in the experiments. For example, some participants reported in the feedback questionnaire that they had made mistakes by confusing the two judging keys on the keyboard once or twice during the experiments. Furthermore, the sample of participants is not large enough. Moreover, the ratio between two genders is unbalanced. Such defects may have an effect on the accuracy of the experiments to some extent, which can be smoothed over by a further study in this field.

ACKNOWLEDGEMENT

The research was supported by National Social Science Foundation (13BYY072).

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Christian JF, Matthias S, Angela DF. Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions. Journal of Memory and Language. 2002;47:250-272.
2. Laurie AS, Michael KT, Greg N C. Filling Gaps On-line: Use of Lexical and Semantic Information in Sentence Processing. Language and Speech. 1991;34(4):319-340.
3. Robert K, Marta K. Bridging the Gap: Evidence from ERPs on the Processing of Unbounded Dependencies. Journal of Cognitive Neuroscience. 1993;5(2):196-214.
4. Henk JH, Katherine AC, Daniel SR. Short-term semantic retention during on-line sentence comprehension. Brain potential evidence from filler-gap constructions. Cognitive Brain Research. 2003;15(2):178-190.
5. Richard M, Lee O. Constraints on Movement Phenomena in Sentence Processing: Evidence from Event-related Brain Potentials. Language and Cognitive Processing. 1996;11(5):495-523.
6. H W Cowles, Robert K, Marta K, Maria P. Violations of information structure: An electrophysiological study of answers to wh-quesitons. Brain and Language. 2007;102:228-242.
7. Christian JF, Matthias S, Angela DF. Syntactic Working Memory and the Establishment of Filler-Gap Dependencies: Insight from ERPs and fMRI. Journal of Psycholinguistic Research. 2001;30(3):321-338.
8. Claudia F, Harald C, Thomas M. Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and wh-movement in German. Brain and Language. 2003;87(3):345-354.
9. Mieko U, Robert K. On the processing of Japanese wh-questions: An ERP study. Brain Research. 2009;1290:63-90.
10. Kim B, Goodall G. Age-Related Effects on Constraints on Wh-movement. // Julia Herschensohn and Darren Tanner. Proceedings of the 11th Generative Approaches to Second Language Acquisition Conference (GASLA 2011). Somerviller, MA: Cascadilla Proceedings Project. 2001:54-62.
11. Ghasem T. The Availability of Universal Grammar to Second Language Learners: A Case of WH-movement. International Journal of English. Linguistics. 2012;2(3):34-43.
12. Marica DV. Syntactic analysis in sentence comprehension: Effects of dependency types and grammatical constrains. Journal of Psycholinguistic Research. 1996;25(1):117-133.
13. Cynthia KT, Mary ET, Kirrie JB, Stephen CF. Agrammatic Aphasic Subjects' Comprehension of Subject and Object Extracted WH-Questions. Brain and Language. 1999;67(3):169-187.
14. Jacquelyn S, Virgina Y. Grammaticality judgments: why does anyone object to subject extraction. Studies in Second Language Acquisition. 1990;12:379-392.
15. Alan J. The influence of first language on the processing of wh-movement in English as a second language. Second Language Research. 2005;21(2):121-151.

16. Carrie N J, Susan C B. The processing and comprehension of wh-questions among second language speakers of German. Applied Psycholinguistics. 2009;30(4):603-636.
17. Alyona B, Lydia W. Evidence for the fundamental difference hypothesis or not? Island constraints revisited. Studies in Second Language Research. 2009;31(2):199-223.
18. Karin S. The Acquisition of Subject and Object WH-questions. Language Acquisition. 1995;4(1-2):5-48.
19. Amanda S, George H, Peter WJ. Early Understanding of Subject and Object Wh-Questions. Infancy. 2003;4(3):423-436.
20. Juffs A, Harrington M. Parsing Effects in Second Language Sentence Processing. Studies in Second Language Acquisition. 1995;17(4):483-516.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=302&id=21&aid=2329