



HAL
open science

The influence of first language on the processing of wh-movement in English as a second language

Alan Juffs

► **To cite this version:**

Alan Juffs. The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, 2005, 21 (2), pp.121-151. 10.1191/0267658305sr255oa . hal-00572088

HAL Id: hal-00572088

<https://hal.science/hal-00572088>

Submitted on 1 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The influence of first language on the processing of *wh*-movement in English as a second language

Alan Juffs *University of Pittsburgh*

Adult learners of English as a second language who speak Chinese ($n = 30$), Japanese ($n = 28$) or Spanish ($n = 46$) as a first language (L1), and a comparison group of native speakers ($n = 22$) read sentences that contain: (a) ungrammatical *wh*-extractions that violate island constraints; and (b) grammatical long-distance Subject and Object extractions from finite and nonfinite clauses. Word-by-word reading times for each sentence were collected using the self-paced reading technique. Results suggest that the presence or absence of *wh*-movement in the L1 and the headedness of the verb phrase in the L1 are unable to explain all of the variation between the nonnative speaker groups. Severe garden path effects were observed in Subject extractions from finite clauses, but not in extractions from nonfinite clauses, suggesting that two finite verbs next to one another may be an important factor in causing parsing break-down. Individual variation in reading time was not predictable from measures of reading span or word span in either the first or second language.

I Introduction

The distinction between the steady-state grammar (competence) and the way the grammar is put to use (performance) is central to the generative approach to linguistics. Comprehension constitutes performance in that the hearer or reader must assign a structure to words/morphemes that is compatible with the grammar, or competence. This process is termed parsing. A complex relationship exists between the grammar and the

parser (e.g., Crain and Fodor, 1985; Gibson, 1998; Pritchett, 1992a), and many questions about this relationship remain unresolved (e.g., Frazier and Clifton, 1996; Gibson and Schütze, 1999; Grodner *et al.*, 2003). However, it is now well established that native speakers of a language parse incoming strings incrementally; that is, they do not wait for several words and then create a structure as was previously believed (e.g., Fodor, 1978), but instead they immediately assign a place in a structure being built to each word (Tanenhaus *et al.*, 1989). In experimental processing research which is based on the principles and parameters theory of competence (Chomsky, 1981; 1986), it has been claimed that empty categories have the same status as words in that they must also be assigned a position in the structure being built (Stowe, 1986; Bever and McElree, 1988; for recent data supporting the psychological reality of traces, see Nakano *et al.*, 2002).

The field of second language acquisition has focused on comprehension performance because comprehension plays a central role in making data available to the learner about the structural properties of the target language (e.g., Krashen, 1987). Second language researchers have long understood that the primary linguistic data (PLD) that learners receive do not necessarily become 'intake' leading to the development of second language (L2) competence (Corder, 1982). With this problem in mind, VanPatten has developed a technique of instruction based on 'input processing' which fosters a specific kind of comprehension, which, he claims, is necessary and sufficient for the development of L2 competence (VanPatten, 1996; 2002; see also DeKeyser *et al.*, 2002). Given this central theoretical and pedagogical role of comprehension in classroom L2 development, it is surprising that relatively little research exists into exactly how L2 speakers comprehend sentences in their L2 as they actually listen or read word by word (Sharwood-Smith, 1993). This gap in understanding is now being filled (Fender, 2001; Juffs, 2001).

While the relationship between formalisms that describe competence, on the one hand, and principles of parsing, on the other, is controversial, this article assumes that there is in fact a clear relationship. Perhaps even more controversially, I will assume that the null hypothesis is that linguistic principles developed to describe competence are valuable in investigating incremental parsing decisions (Weinberg, 1999; Juffs

2004). This position does not exclude an important role for discourse context, real-world knowledge, lexical information and other factors, but it does permit the use of formalisms developed to describe linguistic competence to be directly applied to the investigation of linguistic performance. One major advantage of formalisms is that they permit precise predictions concerning the time course of processing, whereas other approaches as yet lack such precision or ‘granularity’ (Hsiao and Gibson, 2003: 15).

In this context, Juffs and Harrington (1995; 1996) used the moving window technique (Just *et al.*, 1982) to develop a better understanding of a phenomenon that Schachter and Yip (1990) and White and Juffs (1998) had noticed in their investigations of L2 knowledge of constraints on *wh*-movement.¹ Specifically, White and Juffs (1998) found that Chinese learners were more accurate at judging grammatical long-distance Object extraction, as in (1a), than grammatical long-distance Subject extraction, as in (1b).

- 1) a. What does Mary believe John teaches ___? (Object extraction)
 b. Who does Mary believe ___ teaches linguistics? (Subject extraction)

In a follow-up study, Juffs and Harrington (1995) showed that a word-by-word reading profile of sentences such as (1b), which had been used in earlier research, showed a dramatic increase in the time taken to read the verb ‘teaches’, the location of the Subject gap. They concluded that the results supported the hypothesis that a parsing deficit (performance), rather than a competence deficit, was the source of lower accuracy on Subject extractions compared to Object extractions. However, they also pointed out that it was necessary to validate the results by replication, and that data from speakers of other first languages was necessary.

The aims of this article are modest given theoretical developments in theory (Chomsky, 1995; Weinberg, 1999), and in comparison with first language (L1) psycholinguistic research (e.g., Nakano *et al.*, 2002; Hsiao and Gibson, 2003). The first goal is a replication of Juffs and

¹The research reported in White and Juffs (1998) was completed before that reported in Juffs and Harrington (1995). Moreover, it has been proposed that constraints on *wh*-movement can be considered a parsing problem, not a grammatical constraint (e.g., Pritchett, 1992b; Kluender and Kutas, 1993). However, data from Akan suggest that a parsing account cannot be correct (Saah and Goodluck, 1995).

Harrington (1995) with another group of Chinese-speaking learners, as well as groups of Japanese-speaking and Spanish-speaking learners. In addition, measures of working memory are included to address the issue of individual variation within the groups. The article is organized as follows. First, a more detailed review of the parsing framework assumed in the article is provided. An experiment that investigates the effect of L1 on processing *wh*-movement in English as a second language is reported, and the results are discussed with reference to other recent research.

II The grammar and the parser

This article reports on the processing performance of nonnative speakers with grammatical and ungrammatical *wh*-movement within the general framework of Pritchett (1992a). This framework remains compatible with Minimalist approaches to linguistic competence (Weinberg, 1999; Juffs, 2004).² Pritchett's claim, in (2), is that as a sentence is constructed each principle of the grammar is satisfied as early as possible.

- 2) (Syntactic) parsing is the local application of global grammatical principles. (Pritchett, 1992a). This can be paraphrased as follows: As each word comes through the parser, try to interpret fully the local string, i.e., take it as a complete sentence, and satisfy Theta attachment, Case, Binding, etc. as early as possible.

In order to make this statement more concrete, consider the parsing decisions required to assign structure to the sentence in (3a), which contains a *wh*-Subject extracted out of a finite clause. (Less important details are omitted for clarity.)

- 3) a. Who does Jane believe ___ likes her friend?
 b. Who_i did Jane believe t_i? (matrix Object trace)
 c. Who_i did Jane believe [_{IP} t_i likes] (first reanalysis)
 d. Who_i did Jane believe [_{IP} t_i likes her (friend)] (end parse)

In (3b) the *wh*-phrase 'who' is encountered first. The *wh*-phrase permits the parser to construct a CP because the grammar contains the information

²The insights that Pritchett brought to parsing issues remain influential (Inoue and Fodor, 1995: 47; Weinberg, 1999). Weinberg's (1999: 289) summary of when Spell-out in Minimalism can and cannot occur recalls Pritchett's (1992a) theta-reanalysis constraint, that is, 'if reanalysis occurs within a domain where spell out has not applied, the material can be accessed and the preferred reading can be transformed into a dispreferred reading.' In fact, it might be argued that Minimalism uses aspects of Pritchett's approach in that 'chunks' are matched for agreement of features.

that *wh*-phrases can occur in the clause initial Spec of CP position. (The CP is the locus of the [+Strong] *wh*-feature in English.) As soon as the *wh*-phrase is identified, the parser seeks a possible gap, which is the location of the trace of the moved *wh*-phrase. In (3b) the Subject position of the matrix clause is quickly unavailable, hence the parser posits a gap in the matrix Object position, so that the *wh*-phrase is licensed by receiving Case and a theta role through an A-bar chain from the trace that is governed by the verb 'believe'. In this way, all principles of the grammar are satisfied locally as soon as is possible: the theta criterion is satisfied and the *wh*-phrase is licensed for visibility by the Case assignment from the matrix verb. However, on encountering the verb 'likes', the parse of the gap as the matrix Object is no longer viable because 'likes' is a finite verb with two theta roles to be discharged. In order to satisfy the theta requirements of both 'believe' and 'likes', the alternative analysis is to posit a CP complement of 'believe', and establish the gap in the Subject position of the embedded clause, illustrated in (3c). When the Object NP 'her (friend)' is encountered, the parse may terminate. Hence, parsing a sentence containing a *wh*-Subject extracted from an embedded clause requires a rapid reanalysis from a matrix Object trace to an embedded Subject trace.

Parsing of a Subject *wh*-phrase extracted from a nonfinite clause is actually more complex, as the steps in (4) show.

- 4) a. Who does Jane expect ___ to fire Bill?
- b. Who_i does Jane expect t_i?
 (one clause; matrix Object trace)
- c. Who_i does Jane expect [_{IP} t_i to]
 (first reanalysis: Object → Subject)
- d. Who_i does Jane_k expect [CP [_{IP} PRO_k to fire t_i]]
 (second reanalysis: Subject trace → PRO and Object trace)
- e. Who_i does Jane expect [_{IP} t_i to fire Bill]
 (third reanalysis: PRO → Subject trace; deletion of Object trace; end parse)

The analysis proceeds as in (3) until the parser reaches the embedded verb. In (4c), the parser can satisfy the theta and Case requirements of the *wh*-phrase by assuming that 'expect', an Exceptional Case Marking verb, can assign Case to a trace in Spec IP of the infinitival clause. However, in (4d) this analysis must (very) rapidly be revised because if the parser assumes that the trace is in Object position and that there is an empty Subject PRO, all of the theta requirements of both 'expect' and 'fire' can be satisfied (4c). When the NP 'Bill' is encountered,

however, the parser must return to the analysis of the gap as a Subject of 'to fire'.

In the L1 processing literature, considerable evidence exists that the parser posits *wh*-gaps as early as possible (e.g., Stowe, 1986; Gibson and Hickok, 1993; Fiebach *et al.*, 2002). Juffs and Harrington (1995; 1996) found increased reading time at Subject gaps with advanced Chinese-speaking learners of English as a second language. They interpreted this finding as evidence for a processing deficit in parsing sentences such as (3), and argued that the inferior judgement ability of the Chinese speakers on these sentences was because A-bar chains for *wh*-phrases are not necessary in Chinese. (A-bar chains are required for topicalization, and *wh*-chains at LF may be required (Huang, 1982; see also Xu, 1990; Aoun and Li, 1993). However, several issues remained unresolved by Juffs and Harrington's early work. First, it was not clear that the absence of *wh*-movement in Chinese was really the cause of the problems Chinese learners were having with sentences such as (3). Comparison groups of learners whose L1 does have *wh*-movement are also necessary because such learners should not have as great a problem processing *wh*-movement. If they do, then the influence of the L1 will be less convincing as an explanation.

A reviewer correctly notes that when considering L1 transfer in studies of this kind, both L1 grammar and L1 processing preferences should factor into research questions. My assumption in this article is that Spanish-speaking learners build L2 *wh*-chains in the same way that English native speakers do, but that they might be affected by L1 processing preferences because of other properties, e.g., pro-drop and the requirement that C⁰ contain lexical material in embedded clauses. For Japanese and Chinese, both *wh*-in-situ languages, the assumption is that no *wh*-phrase filler-gap relationships are created because they are unnecessary in the L1 (Aoun and Li, 1993). Even if Chinese speakers have to process gaps in other structures in the L1, evidence exists that gap processing in Chinese need not be similar to English. Hsiao and Gibson's (2003) study of relative clause processing in Chinese concluded that Subject relatives are harder than Object relatives because of the head-final nature of Chinese relative clauses and the storage load that head-final relatives impose. In other words, Chinese speakers process according to the demands of the L1, not some generalized parsing algorithm that makes object relatives difficult in

all languages. Inoue and Fodor (1995) point out that many questions remain about how head-final language speakers parse their L1. Some evidence exists that, despite the verb being clause final, parsing is also incremental in Japanese. Data that support this assumption come from processing studies of structural ambiguity in relative clauses in Japanese. For example, many relative clause structures in Japanese are not distinguishable from simple sentences until the head NP appears (Yamashita *et al.*, 1993: 253). This is because Japanese allows null subjects and null objects, and no syntactic or morphological cues are present that disambiguate such structures. Three assumptions follow from the type of research just discussed. First, both Chinese and Japanese have to deal with ambiguity and reanalysis in their respective L1s. Secondly, the parsing of the L1 is tuned to the L1 grammar, and not some set of universal parsing procedures. Thirdly, this research supports the notion that the L1 grammar and L1 processing preferences are closely linked. It therefore makes sense to hypothesize that the L1 grammar transfers as well as L1 processing preferences, and that both can affect L2 processing.

The second problem area for the Juffs and Harrington (1995) study was that the stimuli provided insufficient evidence that Object gaps were not a problem because the Object gaps in the stimuli coincided with the last word in the sentence. In the moving window technique, the last word in the sentence has inflated reading times because participants are beginning to make grammaticality decisions as well as reading that word. This second problem was caused by use of stimuli from the previous study (White and Juffs, 1998), but use of these stimuli had been necessary to replicate those earlier findings. The present study corrects this problem.

Finally, in studies of L2 sentence processing, within-group variation in reaction times is quite large (Juffs and Harrington, 1995; Juffs, 1998). One possible reason for this variation is that individual differences exist with regard to working memory; this has been the case in studies of L1 sentence processing (for a review, see Just and Carpenter, 1992). Although the construct of working memory, and the methods to measure it, remain extremely controversial (Baddeley, 2000; Just and Varma, 2002; MacDonald and Christiansen, 2002), one goal of the present research was to investigate the potential of working memory differences to explain within-group variation in L2 processing.

Researchers have approached Working Memory (WM) from two main points of view. The first is phonological short-term memory (STM), usually measured, by the ability to repeat in exact order a sequence of digits, words or nonsense syllables (e.g., Baddeley *et al.*, 1998; for a discussion of this type of WM in second language acquisition research, see Ellis, 2001). The second component is a storage and processing capacity, sometimes referred to as part of the Central Executive (CE). One common measure of this capacity is the classic reading span task (RST) of Daneman and Carpenter (1980), in which participants must read sentences aloud and then recall words that occurred in those sentences. It remains unclear at present that working memory as a construct is fully understood outside L1 college-age and intellectual-level participants (Baddeley, 2000: 86–87; Roberts and Gibson, 2003). It is possible that the first type of working memory is a component of the second, but researchers report that it is only the RST that has predictive power where individual differences in on-line language processing are concerned (Daneman and Carpenter, 1980: 451; King and Just, 1991: 582).

WM as measured by the reading span task has not been included in any studies of parsing in L2 in a generative framework. Harrington and Sawyer (1992) and Berquist (1997) have used this measure to look at individual differences in proficiency, and Kroll *et al.* used a variation of this measure to investigate vocabulary. Myles *et al.* (1998; 1999), Mackey *et al.* (2002), and Robinson (2002) have looked at working memory in production data, learning an unfamiliar language, and at different stages of learning French syntax. However, many issues remain unresolved in the use and interpretation of these measures, and space does not permit a complete review of these problems (see Juffs, 2003). It is therefore important to see this study as only preliminary in using WM measures in on-line processing.

III Research questions

The research questions in this study appear in (5–8):

- 5) Can the Juffs and Harrington (1995) study be replicated? That is, are Chinese-speaking learners as accurate as native speakers in rejecting ungrammatical extraction, but less accurate than native speakers on grammatical extraction? Do they have a reliably longer reading time in critical regions of the sentence containing Subject extraction?

- 6) Is there an effect for L1? Specifically:
- a. whether there is *wh*-movement in the L1?
 - b. the basic word order of the L1, in particular the headedness of VP?

One expects learners with L1s that do not have *wh*-fronting to have problems similar to those of the Chinese speakers, but that learners whose L1 does have *wh*-movement will not. On the other hand, there might be an effect for word order: for instance, Spanish speakers may pattern with Chinese speakers (mainly SVO), whereas Japanese speakers (SOV) might pattern differently. This latter prediction is based on Fender (2003), who reports that Japanese-speaking learners have particular problems with integration when reading simple sentences when compared to Arabic-speaking learners. There are potentially two areas where differences might be detectable: first in terms of accuracy in judging the sentences, and second in terms of word-by-word parsing decisions. The third question, therefore, relates to where effects discussed in (6) are found in the data:

- 7) If effects are found, are the effects in (6) clear in:
 - a. accuracy;
 - b. word-by-word reading times.
- 8) To what extent can intra-group variation be explained by individual differences in working memory?

IV Method

1 Stimuli

In order to answer the research questions in (5)–(8) a series of stimuli were developed based on the original White and Juffs (1998) and Juffs and Harrington (1995) stimuli. Twenty-seven structures (160 items) were included in a yes–no forced-choice grammaticality judgement task. Only 7 structures, listed in (9), are reported on in this article:

- 9) a. * Who did Tom believe the claim that Ann saw ___ at school?
(noun complement)
- b. * Who did Tom hear the woman who saw ___ on television?
(relative clause)
- c. * Who did Ann meet the teacher after she saw ___ last week? (adjunct)
- d. Who does the nurse know ___ saw the patient at the hospital?
(finite, Subject)
- e. Who does the nurse know the doctor saw ___ in his office?
(finite, Object)
- f. Who does the boss expect ___ to meet the customers next Monday?
(nonfinite, embedded Subject)
- g. Who does the boss expect to meet ___ next Monday? (nonfinite, Object)

All the grammatical sentences could plausibly have a matrix clause Object gap as well as an extraction from an embedded clause. In the earlier studies, the *wh*-extractions of some sentences could only have come from an embedded clause, e.g., ‘Who did Mary insist — buy the tickets?’ This is because ‘insist’ allows only a CP complement.

2 Participants

Participants who did not achieve an overall accuracy of 60% or more in the acceptability judgement tests on declarative clauses with complex NPs and relative clauses were excluded from the analysis. It is necessary to require participants to correctly accept complex sentences with embedding. This is because if they cannot accurately judge such declarative sentences, they cannot be expected to judge *wh*-extractions out of those sentences (Schachter, 1989). Twelve Chinese, 8 Japanese and 11 Spanish speakers were excluded based on this screening. The remaining participants were as follows:

- 30 Chinese-speaking learners (*wh* in-situ, predominantly SVO, with the exception of the ‘ba’ construction);
- 28 Japanese-speaking learners (*wh* in-situ, SOV); and
- 46 speakers of Spanish-speaking learners (*wh*-movement, predominantly SVO with the exception of clitic pronouns).

There was also a comparison group of monolingual native speakers of English ($n = 22$). All participants received payment for their participation in the study.

The grammar and vocabulary sections of the Michigan English (Corrigan *et al.*, 1978) was administered to ensure that the learners were equal in general proficiency. As can be seen from Table 1, some differences among the groups was apparent as measured by a GLM ANOVA. One way to make the groups equal in proficiency would have been to eliminate many participants of higher proficiency from the Chinese and Spanish groups. However, elimination because of these scores would mean reducing the participant pool somewhat artificially on the basis of this test that used discrete points. I therefore decided to include as many participants as possible, and remain cautious in interpreting any between-group differences.

Table 1 Michigan test results: raw scores

Michigan	Chinese		Japanese		Spanish		F	df	p
	M*	SD	M*	SD	M*	SD			
Vocabulary	28.33 ^a	7.67	20.39 ^a	6.21	26.65	7.58	9.6	2,102	.0002
Grammar	29.8 ^b	6.0	25.07 ^b	5.28	26.89	7.26	4.042	2,102	.0205
Total	58.03 ^c	12.59	45.46 ^{c,d}	10.32	53.45 ^d	13.96	7.29	2,102	.0011

Note: *Means that are co-superscripted are reliably different

3 Moving window procedure

Before the test, a list of all the words in the sentences was given to the participants. The list was in alphabetical order, and participants checked any of the words they did not know. On arrival at the test, the participants knowledge of the vocabulary items was checked. The method used to collect the data was the moving window procedure (Just *et al.*, 1982; Papadopoulou, this issue). In this procedure, a participant reads the stimulus sentence one word at a time; pressing a button each time he or she is ready for the next word. In this way, the participant determines the reading speed, and not the experimenter. As each new word appears, the previous word disappears, thereby preventing the participant from ever seeing the entire sentence on the screen at once. The time each participant takes to read every word is recorded. Once the last word has disappeared, the program prompts the participant to record a decision of 'possible' or 'not possible'. The stimuli not reported on in this article included 6 tokens of a *wh*-extraction from a matrix clause followed by an adjunct clause. In addition, there were 3 tokens of *wh*-questions in one clause, and 3 tokens of monoclausal declaratives. These were included to prevent an expectation in the participants that there would only be successive cyclic *wh*-movement, or only (long) biclausal sentences.

4 Working memory

The following measures of working memory were collected:

- A Daneman and Carpenter (1980) test of working memory in English based on Harrington and Sawyer (1992).

- A reading-span test in the L1. The test, developed by Osaka and Osaka (1992), was used for the Japanese-speaking participants, whereas in-house tests were developed for those participants who were Chinese-speaking or Spanish-speaking. (Both of these tests, like Osaka and Osaka (1992), took their sentences from High School textbooks.) An important modification to standard reading-span tests was that the target words for recall were not the final word in the sentence for any of the tests. This is because in Japanese the final word will always be a verb. Since the words in the Japanese test are sentence-internal, a decision was taken to make all the RST tests have sentence-internal words as the target of recall.
- Two word-span tests, in L1 and L2. A real word-span test was used rather than a nonword test of the phonological loop (Baddeley *et al.*, 1998).

V Results

I first present results for accuracy in judgements and then for word-by-word reading profile. Alpha was set at .05 for all statistical tests. If a significant omnibus *F* statistic was found, the Tukey procedure was used to identify any reliable differences among the means of the groups.

1 Accuracy on judgement of ungrammatical and grammatical wh-extraction

Accuracy on judging ungrammatical extractions appear in Table 2. The means in Tables 2 and 3 are reported as decimals because participants received a score of 1 for each correct response and a score of 0 for each incorrect response: the numbers in the table are means out of 6 calculated from each participant's scores in each group. All groups correctly reject ungrammatical *wh*-movement at a level above chance. For example, a score of 0.66 indicates 4 out of 6 correct, a score of 0.75 means 4.5 correct and so on.

A General Linear Model (GLM) Type III sums of squares is a type of ANOVA used when cell sizes are unequal. This procedure was used to explore the data with Language and Structure as independent variables. The L1 effect ($F_{3, 122} = 2.62, p = .0538$) comes very close to

Table 2 Mean accuracy scores on ungrammatical sentences

Structure	Chinese		Japanese		Spanish		English	
	M	SD	M	SD	M	SD	M	SD
Complex NP islands ($n = 6$) (9a)	0.72	0.44	0.66	0.47	0.73	0.44	0.79	0.41
Relative clause islands ($n = 6$) (9b)	0.67	0.47	0.65	0.48	0.72	0.45	0.72	0.45
Adjunct islands ($n = 6$) (9c)	0.66	0.47	0.70	0.45	0.72	0.45	0.73	0.44

reliability. This is because the mean for the native speakers, about 4.5 out of 6 correct, is not that much higher than the nonnatives. The Tukey procedure did show all comparisons to be reliable except that between the Japanese and Chinese groups. The accuracy scores here are low. However, in the self-paced reading paradigm, even accuracy by native speakers in their own language can be low. For example, Hsiao and Gibson (2003) report similar percentage correct for native speakers of Chinese reading complex relative clauses in Chinese. Clearly, the processing load of reading, retaining and integrating on line makes comprehension and judgements more difficult on line than in paper tests where the full sentence can be reviewed at leisure.

The accuracy results for grammatical long-distance extractions from embedded clauses appear in Table 3. All groups correctly accept grammatical *wh*-movement at above chance level, but there is some variation by language and structure type. Subject extraction from a finite clause is particularly difficult for all learners, including the Spanish-speaking learners.

Table 3 Mean accuracy scores on grammatical sentences

Structure	Chinese		Japanese		Spanish		English	
	M	SD	M	SD	M	SD	M	SD
Subject extraction finite ($n = 6$) (9d)	0.59	0.49	0.63	0.48	0.64	0.47	0.73	0.44
Object extraction finite ($n = 6$) (9e)	0.70	0.45	0.68	0.46	0.75	0.44	0.79	0.41
Subject extraction nonfinite ($n = 6$) (9f)	0.69	0.46	0.62	0.49	0.74	0.44	0.76	0.43
Object extraction nonfinite ($n = 6$) (9g)	0.67	0.47	0.70	0.45	0.70	0.46	0.79	0.41

An analysis of the differences among the participants was carried out using the GLM procedure, with Language (4) and Structure (4) as independent variables. There was a main effect for Language ($F_{3, 122} = 4.51, p = 0.0049$). Once again, all groups were different from each other with the exception of the Japanese and the Chinese learners, who were not reliably different. There was also a main effect for Structure ($F_{3, 366} = 4.02, p = 0.0078$). All structures were reliably different from one another, with the exception of the extractions from nonfinite clauses. This means that Subject extraction from a finite clause was reliably more difficult than other extraction types, and that Object extraction from a finite clause was the easiest. Extractions from nonfinite clauses fell between these two. There was no Language by Structure interaction.

2 Reading profiles of correct responses to sentences with grammatical wh-extraction

Unless otherwise stated, all reading times are from responses where the participant correctly judged the item possible in the case of grammatical long-distance extraction. Reading times for error data were replaced with missing values for all participants. Each statistical analysis is also based on data that were screened, by Language group for each word in each sentence type, for outliers of reading times that exceeded 3 standard deviations from the mean for that language group. Reading times that were removed based on this procedure were replaced with missing values. Elimination of data creates unequal cell sizes for the statistical analysis; moreover, reaction time data tend to be positively skewed (Cowan and Hatasa, 1994). To compensate for these characteristics in the data, all analyses of reading times used log transformations of the data using the General Linear Model (GLM), Type III Sums of Squares, with the Tukey procedure used for all *post hoc* analyses. Two points are important when considering the reading profiles presented here. One point is that the initial few words and last words of each sentence are read at approximately the same speed by all participants, including the native speakers. This result indicates that mid-sentence differences in reading times reflect true differences in processing decisions at those points, and are not due to differences in script from the L1 (e.g.,

Chinese characters or kana) and the alphabetic script of the L2, English. Secondly, the native speakers were not included in the statistical analysis because their accuracy is higher, and the variation much lower; the inclusion of the native speakers would thus always lead to an overall reliable *F* statistic. The profile of the native speakers should be used for comparative descriptive purposes only (see also Fender, 2003: 298).

Figure 1 provides an illustration of Subject extraction; Object extraction is shown in Figure 2. The spike in reading time on the embedded verb that can be observed in Figure 1 is similar to that found in Juffs and Harrington (1995); however, the Japanese also increase in reading time on the head noun of the Object in Figure 1, whereas all other learner groups decrease in reading time. Moreover, encountering a Subject NP in the lower clause of an Object extraction in Figure 2 does not 'surprise' the learners. In order to explore the differences between the two reading profiles, a three factor GLM procedure (Language (3), Structure (2), Trial (2)) was conducted on the main verb, 'suggest' in the example figure, and the following word, the embedded verb in Figure 1 and the first element of the embedded Subject NP in Figure 2. The relevant means and standard deviations are in Table 4.

The analysis shows a reliable effect for Language ($F_{2, 102} = 10.23$, $p = 0.0001$), with all groups reliably different from one another. In addition, there were reliable effects for Structure ($F_{1, 101} = 51.80$, $p = 0.0001$), a reliable effect for Trial ($F_{1, 101} = 50.29$, $p = 0.0001$), a reliable Structure by Trial Interaction ($F_{1, 102} = 40.79$, $p = 0.0001$) and a reliable Language by Structure by Trial interaction ($F_{2, 102} = 3.54$, $p = 0.0326$). Essentially, these results show that the all language groups are different from one another: the order of processing difficulty is greatest for the Japanese, then the Chinese and Spanish respectively. Clearly, the Japanese have the greatest difficulty with Subject extraction. Moreover, the reliable Language by Structure by Trial interaction shows that the Japanese, Chinese and Spanish all had relatively more difficulty at this point in the sentence with the Subject extractions than with the Object extractions. Hence, the processing data reflect the accuracy data here, with the exception that the Chinese participants are reliably different from the Japanese.

The parsing profiles of nonfinite clauses appear in Figures 3 and 4. In the case of nonfinite clauses, the issue is whether encountering the

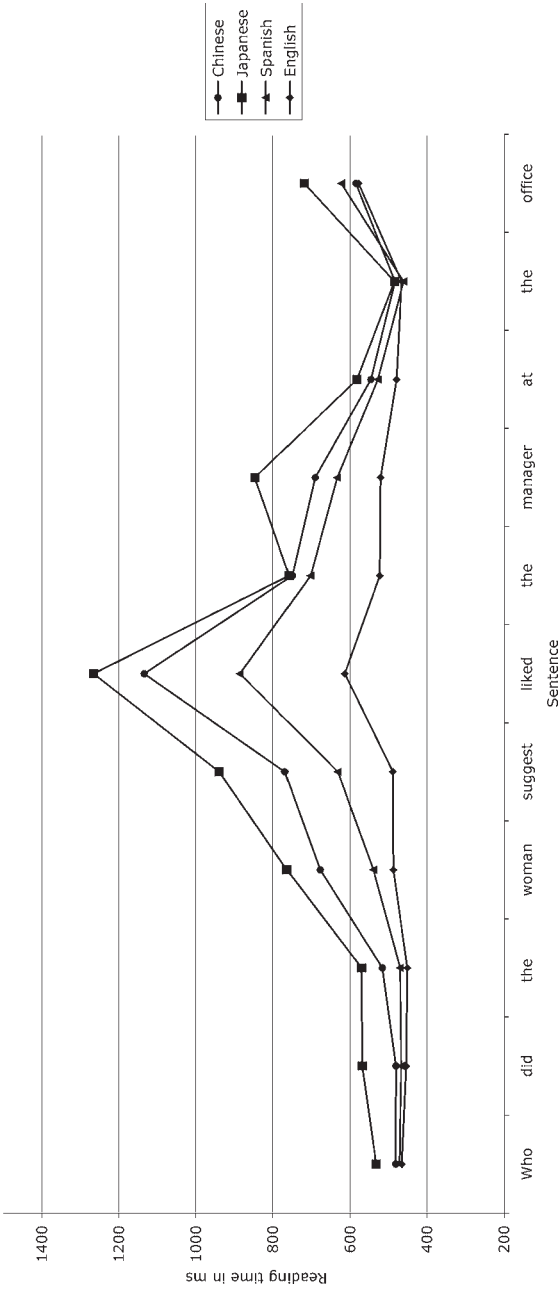


Figure 1 Word-by-word reading times of extraction of a subject from a finite clause

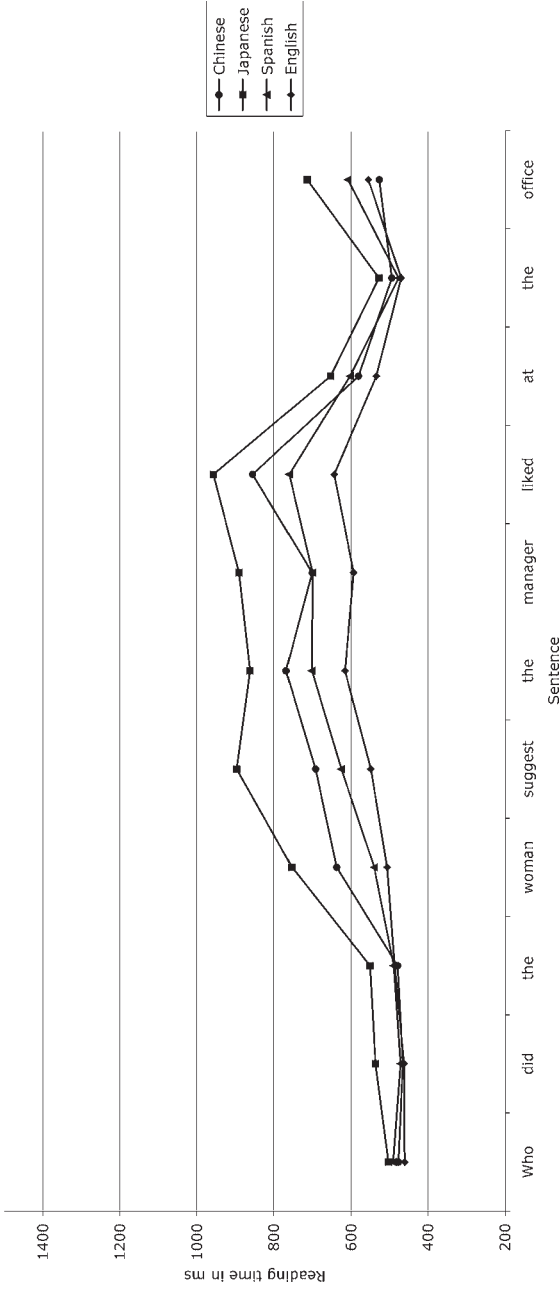


Figure 2 Word-by-word reading times of the extraction of an object from a finite clause

Table 4 Means of critical area from Figures 1 and 2

Structure	Trial	Mean	SD
<i>Chinese</i>			
Subject	Word 5	776.35	480.30
Subject	Word 6	1148.47	728.13
Object	Word 5	692.66	441.34
Object	Word 6	768.74	442.90
<i>Japanese</i>			
Subject	Word 5	939.30	559.79
Subject	Word 6	1265.70	693.80
Object	Word 5	896.62	515.94
Object	Word 6	862.07	512.46
<i>Spanish</i>			
Subject	Word 5	635.94	391.46
Subject	Word 6	892.04	580.57
Object	Word 5	625.09	366.04
Object	Word 6	701.85	481.61

Object gap in the embedded clause leads to processing break-down. Clearly, these problems are not the same as for extraction from finite clauses. Recall that the processing load will occur once the Object gap is encountered because the previous analysis of a Subject extraction must be revised at this point. Therefore, a three factor GLM procedure (Language (3), Structure (2), Trial (2)) was conducted on the embedded verb, e.g., ‘hire’ in the example in Figure 4, and the following word, the preposition in Figure 4 and the first element of the embedded Object NP in Figure 3. The result of this analysis is that there was a main effect for Language ($F_{2, 102} = 10.19, p = .0001$). In this case, the Spanish and the Chinese are not reliably different from one another, but they are both reliably different from the Japanese. Since the Japanese were not matched with the other two groups in proficiency, no clear conclusion can be made about the effect of the head position of the VP. However, the data do suggest that the Japanese speakers have special problems in processing verbs. There were no other reliable effects or interactions, except a near reliable interaction of Language and Trial ($F_{2, 101} = 2.66, p = 0.0745$). These results suggest the profile of the Japanese tends to drop in Object extraction more sharply than the other language groups. These results reflect those reported in Juffs (2004) for the same participants processing garden path sentences.

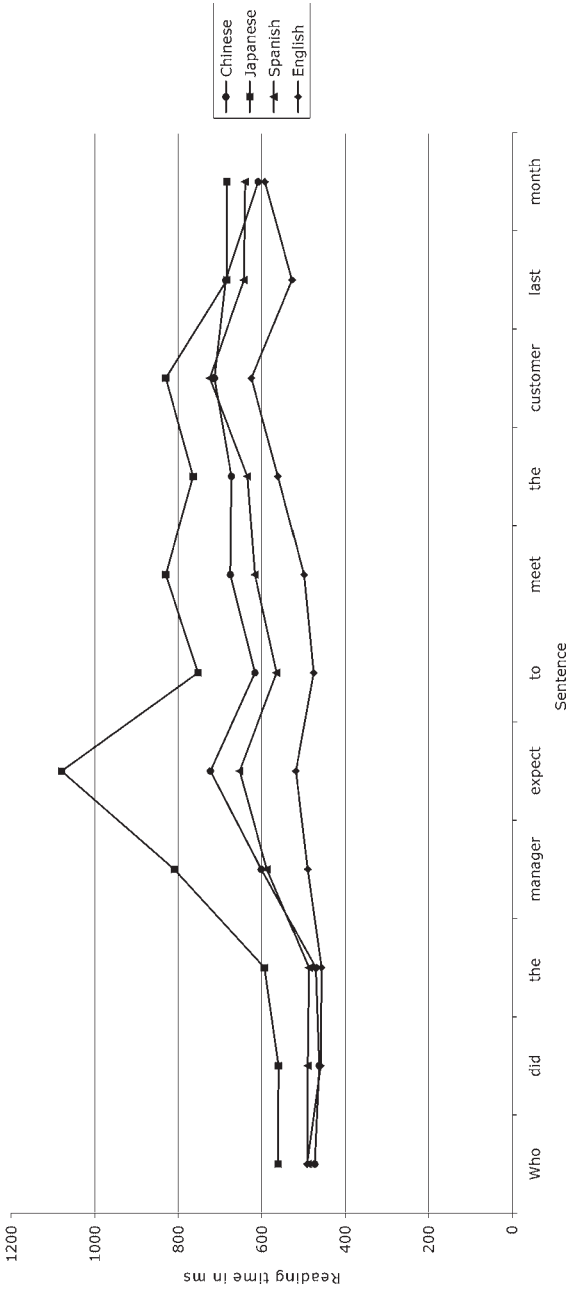


Figure 3 Word-by-word reading time for a subject extraction from a nonfinite clause

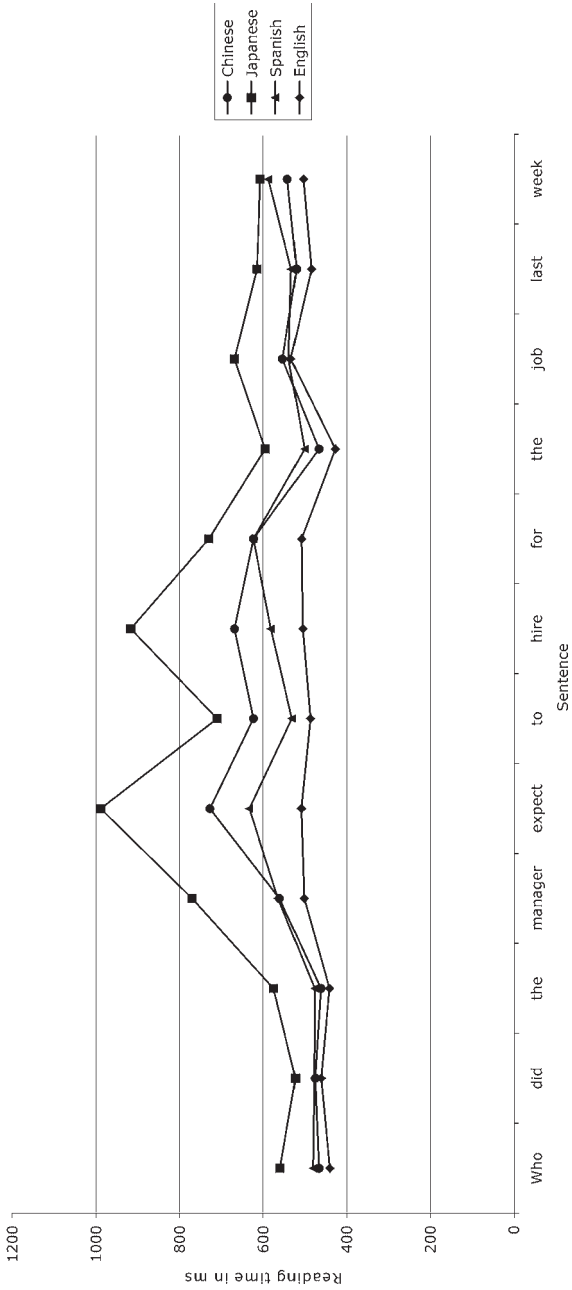


Figure 4 Word-by-word reading times for an object extraction from a nonfinite clause

Table 5 English-speaking participants ($n = 22$, accurate *wh*-data only)

	English word span	English reading span	<i>wh</i> -Subject extraction second verb 1
English word span	1		
English reading span	0.62 0.0018	1	
<i>wh</i> -Subject extraction second verb 1	-0.32	-0.20	1

3 Working memory

The relationship between the measures of working memory and reading times are reported in Tables 5–8. Results are reported as correlations, where in Tables 6–8 ‘Total’ refers to the total number of points in the Michigan vocabulary and grammar sections. Correlations are reported with the reading times on the embedded verb from the Subject extractions from finite clauses that are illustrated in Figure 1. This verb was taken as the point where processing load was highest, and where individual differences in working memory would most likely show an effect (Just *et al.*, 1996: 778–79). The figures in the table, in descending order, are the r statistic and the p value if it was reliable.

The data in Tables in 5–8 show reliable relationships among the measures of working memory but not between these measures and the overall proficiency or the reading times. The finding that the working memory measures for the Reading Span and the Word Span correlate,

Table 6 Chinese-speaking learners of ESL ($n = 30$, accurate *wh*-data only)

	Total Michigan score	Native word span	Native reading span	English word span	English reading span	<i>Wh</i> -subject extraction second verb
Total	1					
Native word span	0.18	1				
Native reading span	-0.01	0.09	1			
English word span	0.34 0.0593	0.33 0.07	0.16	1		
English reading span	0.11	-0.13	0.68 0.0001	0.13	1	
<i>Wh</i> -subject extraction second verb	0.05	-0.03	-0.13	-0.11	-0.15	1

Table 7 Japanese-speaking learners of ESL (*n* = 28, accurate *wh*-data only)

	Total	Native word span	Native reading span	English word span	English reading span	<i>Wh</i> -subject extraction second verb
Total	1					
Native word span	0.12	1				
Native reading span	0.29	0.59 0.0009	1			
English word span	0.18	0.44 0.019	0.73 0.0001	1		
English reading span	0.12	0.37 0.053	0.57 0.0015	0.62 0.0004	1	
<i>Wh</i> -subject extraction second verb 1	-0.34 0.08	-0.001	0.08	-0.045	-0.12	1

even for the native speakers, is remarkable in itself and is discussed in greater detail in Juffs (2003).

VI Discussion

The results reported in the previous section permit the following answers to the research questions for this study. First, the Juffs and Harrington (1995) results are largely replicated. Chinese-speaking learners of English show sensitivity to ungrammatical vs. grammatical *wh*-movement. Moreover, learners show a reliable asymmetry between

Table 8 Spanish-speaking learners of ESL (*n* = 46, accurate *wh*-data only)

	Total	Native word span	Native reading span	English word span	English reading span	<i>Wh</i> -subject extraction second verb
Total	1					
Native word span	0.24	1				
Native reading span	0.27	0.45 0.0013	1			
English word span	0.23	0.47 0.0008	0.23	1		
English reading span	0.21	0.28	0.58 0.0001	0.11	1	
<i>Wh</i> -subject extraction second verb	0.08	-0.13	0.08	-0.14	0.07	1

accuracy on grammatical Subject *wh*-movement from an embedded finite clause with Object extraction from a finite clause. The accuracy results, even for the on-line measures, are lower than those reported by Juffs and Harrington (1995). This result may be due to the greater length of the sentences that were used in this study. The stimuli were 11–12 words long compared to 7–8 words long in previous research.

The L1 clearly affected accuracy for the grammatical *wh*-extractions, but marginally so for the ungrammatical extractions. Although this result must be treated with caution because the Japanese group is not proficiency matched to the others, it suggests that in accuracy, the lack of *wh*-movement in the L1 is a disadvantage for overall for these structures even when other proficiency measures show no difference. The availability of *wh*-fronting in the L1 provides a clear advantage for judging *wh*-movement in English in this on-line task, even though Spanish requires a complementizer in the head of embedded lower clause and does not require an overt Subject. However, the statistical analysis of the results in Table 3 shows that the Spanish-speaking participants, although they are more accurate overall, show the same asymmetry on the accuracy measures for grammatical Subject and Object extraction. This result indicates that they have similar difficulties with parsing these structures.

The reading profiles confirm that the locus of difficulty is indeed the embedded verb in Subject extractions from finite clauses. Clearly, the Japanese-speaking participants encounter the greatest difficulty, even though there were not reliably different from the Chinese on accuracy. The Japanese-speaking participants take the longest time to read both the finite and the nonfinite embedded clauses. The reading profiles suggest that the Japanese have problems not only with Subject extraction, but also with integrating objects into the parse. This result confirms the integration problems found by Fender (2003) and Juffs (1998) for speakers of SOV languages. Once again, the Spanish-speaking participants are not immune from the effect of encountering a second verb in the embedded clause. Although they are able to parse these structures more quickly, they show similar effects. This result, taken with the accuracy data, casts doubt on Juffs and Harrington's (1995) suggestion that the inability to process changes in the features of a *wh*-chain is the source of the Chinese learners' problems with Subject extraction.

The word-by-word processing results from the sentences containing *wh*-extractions from nonfinite clauses do not provide strong support for the *wh*-processing account either. The parsing profile is much flatter than that of Subject extractions from finite clauses, even though the predicted parsing difficulties are similar in terms of gap re-analysis. The obvious difference is the juxtaposition of two tensed verbs in embedded finite clauses. Therefore, it is essential to examine further the contribution of finiteness, because tense plays a role in Case assignment to a *wh*-chain and is marked by overt morphology in the verb in English. The syntactic context of two tensed verbs next to one another, as in Object relative clauses (e.g., ‘the tiger that the lion chased climbed the tree’) in contrast to Subject relative clauses (e.g., ‘the tiger that chased the lion climbed the tree’) was examined by Jelliffe and Juffs (2004). These authors found a similar parsing profile for the two verbs in an Object relative clause as was found in these data for Subject *wh*-extraction from a finite clause. An Occam’s Razor account could therefore be that the learners experience severe garden path effects with two finite verbs next to one another, an effect also experienced by native speakers, albeit far less severe. Williams *et al.* (2001) made a similar suggestion. They found that processing of *wh*-structures by nonnative speakers was similar to native speakers, but that when nonnative speakers were garden-pathed, they had much greater difficulty in recovering from a misparse. It may therefore be possible that all of the learners are being garden-pathed by unfamiliar structures rather than failing to build *wh*-chains.

Marinis *et al.* (2003) point out that it is not possible with the type of data presented in Juffs and Harrington (1995), and *ceteris paribus* in this article, to distinguish between theoretical approaches that analyse *wh*-gaps as independent from the verb that subcategorizes for the displaced filler and theories that posit independent *wh*-traces. This is because in English the verb is always next to the gap, unlike in German where the finite verb and gap-site need not be adjacent. Remarkably, Marinis *et al.* (2003) found no L1 effects in processing *wh*-gaps in learners of English as a second language. This finding also suggests that L2 learners are well able to form and interpret *wh*-chains.

The results reported in this article might be construed as supporting an argument in Felser *et al.* (2003: 478) that L2 learners do not parse incrementally, but rather delay integration in a way earlier proposed for

native speakers in Fodor (1978). Although that claim may be correct for the adjunct attachment that these authors investigate (relative clause attachment in ambiguous sentences such as ‘the reporter phoned the boss of the secretary who was reading a book’, where ‘who was reading the book’ could conceivably modify either ‘the boss’ or ‘the secretary’), it does not seem to be correct in general for L2 processing. If this were the case, one would not expect learners to experience garden path effects (Juffs, 2004; Juffs and Harrington, 1996) or to exhibit the kind of accuracy and parsing profile asymmetries reported here and in Juffs (1998). The less radical position is that L2 learners used verb-argument structure information, but not a filler-gap strategy.

The working memory results failed to show any relationship between parsing the Reading Span measure of working memory, or the Word Span measure of working memory. This was true of the measures taken in the first and second languages of the participants. It is possible to divide learners into groups of high working memory and low working memory. However, Juffs (2004) showed no reliable results with the processing of garden path sentences using this approach. Since no reliable correlations were found in those data, further analysis was not conducted due to space limitations. Juffs (2003) analyses these and other data in greater detail.

VII Conclusions

The results of this study replicate Juffs and Harrington’s (1995; 1996) results to some extent. They had shown that advanced nonnative speakers show sensitivity to grammatical and ungrammatical *wh*-extraction, even in an online task. In this respect, their L2 performance resembles that of native speakers. However, the experiment in this article also shows that success in on-line processing of *wh*-movement in the L2 depends only to some degree on whether the L1 has overt *wh*-movement because all learners experienced Subject–Object asymmetries in processing grammatical *wh*-extraction. If there is no *wh*-movement in the L1, word order in the L1 has an additional negative effect on processing. The word-by-word reading time data suggest that a garden path caused by two finite verbs appearing next to each other may be a better explanation for these data than the one originally proposed by Juffs and

Harrington (1995). If this is true, the implication is that not only do L2 learners have sensitivity to grammatical vs. ungrammatical *wh*-movement, but that they also have an ability to reassign features to *wh*-chains during processing performance.

Several important limitations of the study require attention. First, because the processing of nonfinite clauses appears to be easier than a Subject *wh*-extraction from a finite clause, it is clear that parsing effects that the theoretical framework had predicted were not sufficiently precise. In future work, the role of tense and/or tense morphology in processing and the presence of two adjacent finite verbs call for careful study with the same group of learners. Secondly, the Japanese group was not proficiency matched to the Chinese and Spanish groups, making direct comparisons difficult. Further research is needed to investigate the role of the L1 and the role of particular configurations of verbs within complex sentences (e.g., relative clauses) to see whether L2 learners are using processes that are either based on their L1, or some other algorithm that is not compatible with formalisms that have been developed to describe competence. Finally, the role of working memory needs to be carefully investigated with more sophisticated statistical analysis.

Acknowledgements

The research reported in this article was supported in part by the National Science Foundation (grant SBR-9709152 to Alan Juffs), for which I am grateful. Preliminary results were presented at the Annual Meeting of the American Association for Applied Linguistics in Vancouver in 2000.

VIII References

- Aoun, J.** and **Li, A. Y.-H.** 1993: *Wh*-elements in situ: syntax or LF? *Linguistic Inquiry* 24, 199–238.
- Baddeley, A.** 2000: Short-term and working memory. In Tulving, E. and Craik, F., editors, *The Oxford handbook of memory*, New York: Oxford University Press, 77–92.
- Baddeley, A., Gathercole, S.** and **Papagno, C.** 1998: The phonological loop as a language learning device. *The Psychological Review* 105, 158–73.

- Berquist, B.** 1997: Individual differences in working memory span and L2 proficiency: capacity or processing capacity? Paper presented at the Proceedings of the GALA '97 Conference on Language Acquisition, Edinburgh, UK.
- Bever, T.G. and McElree, B.** 1988: Empty categories access antecedents during comprehension. *Linguistic Inquiry* 19, 35–43.
- Chomsky, N.** 1981: *Lectures on Government and Binding*. Dordrecht: Foris.
 — 1986: *Knowledge of language*. New York: Praeger.
 — 1995: *The Minimalist Program*. Cambridge, MA: MIT Press.
- Corder, S.P.** 1982: *Interlanguage and error analysis*. Oxford: Oxford University Press.
- Corrigan, A., Dobson, B., Kellman, E., Palmer, A., Peterson, J., Spaan, M., Strowe, L., Tyma, S. and Upshur, J.** 1978: *English placement test*. Ann Arbor, MI: University of Michigan: University of Michigan English Language Institute.
- Cowan, R. and Hatasa, Y.A.** 1994: Investigating the validity and reliability of native speaker and second language. In Tarone, E., Gass, S. Cohen, A., editors, *Research methodology in second-language acquisition: second language acquisition research: theoretical and methodological issues*. Hillsdale, NJ: Erlbaum, 287–302.
- Crain, S. and Fodor, J.D.** 1985: How can grammars help parsers? In Dowty, D., Karttunen, L. and Zwicky, A., editors, *Natural language parsing: psychological, computational and theoretical perspectives*, Cambridge: Cambridge University Press, 94–128.
- Daneman, M. and Carpenter, P.** 1980: Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19, 450–66.
- DeKeyser, R.M., Salaberry, R., Robinson, P. and Harrington, M.W.** 2002: What gets processed in processing instruction? A commentary on Bill VanPatten's 'Processing instruction: an update'. *Language Learning* 52, 805–23.
- Ellis, N.** 2001: Memory for language. In Robinson, P., editor, *Cognition and second language instruction*, New York: Cambridge University Press, 33–68.
- Felser, C., Roberts, L. and Marinis, T.** 2003: The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics* 24, 453–89.
- Fender, M.** 2001: A review of L1 and L2 ESL word integration skills and the nature of L2/ESL word integration development of lower level text processing. *Language Learning* 51, 319–96.
 — 2003: English word recognition and word integration skills of native Arabic-speaking and Japanese-speaking learners of English as a Second Language. *Applied Psycholinguistics* 24, 289–315.

- Fiebach, C.J., Schlesewsky, M. and Friederici, A.D.** 2002: Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German *wh*-questions. *Journal of Memory and Language* 47, 250–72.
- Fodor, J.D.** 1978: Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9, 427–74.
- Frazier, L. and Clifton, C.** 1996: *Construal*. Cambridge, MA: MIT Press.
- Gibson, E.** 1998: Linguistic complexity: locality of syntactic dependencies. *Cognition* 68, 1–76.
- Gibson, E. and Hickok, G.** 1993: Sentence processing with empty categories. *Language and Cognitive Processes* 8, 147–61.
- Gibson, E. and Schütze, C.T.** 1999: Disambiguation preferences in noun phrase conjunction do not mirror corpus frequency. *Journal of Memory and Language* 40, 263–79.
- Grodner, D., Gibson, E., Argaman, V. and Babyonyshev, M.** 2003: Against repair-based reanalysis in sentence in comprehension. *Journal of Psycholinguistic Research* 32, 141–66.
- Harrington, M. and Sawyer, M.** 1992: L2 working memory capacity and L2 reading skills. *Studies in Second Language Acquisition* 14, 25–38.
- Hsiao, F. and Gibson, E.** 2003: Processing relative clauses in Chinese. *Cognition* 90, 3–27.
- Huang, C.-T.J.** 1982: Move *wh* in a language without *wh* movement. *Linguistic Review* 1, 369–416.
- Inoue, A. and Fodor, J.D.** 1995: Information-paced parsing of Japanese. In Mazuka, R. and Nagai, N., editors, *Japanese sentence processing*. Hillsdale, NJ: Erlbaum, 9–64.
- Jelliffe, J.D. and Juffs, A.** 2004. Processing of subject and object relative clauses by Chinese-speaking and Spanish-speaking learners of English. Poster presented at GASLA VII, Bloomington, Indiana, April 2004.
- Juffs, A.** 1998: Main verb vs. reduced relative clause ambiguity resolution in second language sentence processing. *Language Learning* 48, 107–47.
- 2000. Working memory and L1 influences in ambiguity resolution in L2 English sentence processing. American Association of Applied Linguistics Annual Meeting, Vancouver, BC, March 13, 2000.
- 2001: Psycholinguistically-oriented second language research. *Annual Review of Applied Linguistics* 21, 207–23.
- 2003: Working memory: understanding issues in mainstream psychology. Paper presented at the Second Language Research Forum, Tucson, Arizona. October 18, 2003.
- 2004: Representation, processing and working memory in a second language. *Transactions of the Philological Society*, 102, 199–226.
- Juffs, A. and Harrington, M.** 1995: Parsing effects in L2 sentence processing: subject and object asymmetries in *wh*-extraction. *Studies in Second Language Acquisition* 17, 483–516.

- 1996: Garden path sentences and error data in second language sentence processing research. *Language Learning* 46, 286–324.
- Just, M.A. and Carpenter, P.** 1992. A capacity theory of comprehension: individual differences in working memory. *Psychological Review* 99, 122–49.
- Just, M.A., Carpenter, P. and Keller, T.** 1996. The capacity theory of comprehension: new frontiers of evidence and arguments. *The Psychological Review* 103, 773–80.
- Just, M.A., Carpenter, P. and Woolley, J.D.** 1982. Paradigms and processes and in reading comprehension. *Journal of Experimental Psychology: General* 3, 228–238.
- Just, M.A. and Varma, S.** 2002: A hybrid architecture for working memory: reply to MacDonald and Christianson 2002. *Psychological Review* 109, 55–65.
- King, J. and Just, M.A.** 1991: Individual differences in syntactic processing: the role of working memory. *Journal of Memory and Language* 30, 580–602.
- Kluender, R. and Kutas, M.** 1993: Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8, 573–633.
- Krashen, S.D.** 1987: *Principles and practice in second language acquisition*. Englewood Cliffs, NJ: Prentice-Hall.
- Kroll, J., Michael, E., Tokowicz, N. and Dufour, R.** 2002: The development of lexical fluency in a second language. *Second Language Research* 18, 137–71.
- MacDonald, M.C. and Christiansen, M.H.** 2002: Reassessing working memory: comment on Just and Carpenter 1992 and Waters and Caplan 1996. *Psychological Review* 109, 35–54.
- Mackey, A., Philp, J., Egi, T., Fujii, A. and Tatsumi, T.** 2002: Individual differences in working memory, noticing of interactional feedback and L2 development. In Robinson, P., editor, *Individual differences and instructed language learning*, Philadelphia: John Benjamins, 181–209.
- Marinis, T., Roberts, L., Felser, C. and Clahsen, H.** 2003: Gaps in second language sentence processing. *Essex Research Reports in Linguistics* 45, 43–79.
- Myles, F., Hooper, J. and Mitchell, R.** 1998: Rote or rule? Exploring the role of formulaic language in the foreign language classroom. *Language Learning* 48, 323–364.
- Myles, F., Mitchell, R. and Hooper, J.** 1999: Interrogative chunks in French L2: a basis for creative construction? *Studies in Second Language Acquisition* 21, 49–80.
- Nakano, Y., Felser, C. and Clahsen, H.** 2002: Antecedent Priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research* 31, 531–71.

- Osaka, M. and Osaka, N.** 1992: Language independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society* 30, 287–89.
- Pritchett, B.L.** 1992a: *Grammatical competence and parsing performance*. Chicago: Chicago University Press.
- 1992b: Parsing with grammar: Islands, heads, and garden paths. In Goodluck, H. and Rochemont, M., editors, *Island constraints: theory, acquisition, and processing*, Dordrecht: Kluwer, 321–50.
- Roberts, R. and Gibson, E.** 2003: Individual differences in sentence memory. *Journal of Psycholinguistic Research* 31, 573–598.
- Robinson, P.** 2002: Effects of individual differences in intelligence, aptitude and working memory on incidental SLA. In Robinson, P., editor, *Individual differences and instructed language learning*. Philadelphia, PA: John Benjamins, 211–51.
- Saah, K.K. and Goodluck, H.** 1995: Island effects in parsing and grammar: evidence from Akan. *The Linguistic Review* 12, 381–409.
- Schachter, J.** 1989: Testing a proposed universal. In Gass, S. and Schachter, J., editors, *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press, 73–88.
- Schachter, J. and Yip, V.** 1990: Grammaticality judgments: why does anyone object to subject extraction? *Studies in Second Language Acquisition* 12, 379–92.
- Sharwood-Smith, M.** 1993: Input enhancement in instructed SLA: theoretical bases. *Studies in Second Language Acquisition* 15, 165–179.
- Stowe, L.** 1986: Parsing *wh*-constructions: evidence for on-line gap location. *Language and Cognitive Processes* 2, 227–46.
- Tanenhaus, M.K., Carlson, G. and Trueswell, J.C.** 1989. The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes* 4, 211–34.
- VanPatten, B.** 1996: *Input processing and grammar instruction in second language acquisition*. Norwood, NJ: Ablex.
- 2002: Processing instruction: An update. *Language Learning* 52, 755–803.
- Weinberg, A.** 1999: A minimalist theory of human sentence processing. In Epstein, S. and Hornstein, N., editors, *Working minimalism*, Cambridge, MA: MIT Press, 287–315.
- White, L. and Juffs, A.** 1998: Constraints on *wh*-movement in two different contexts of non-native language acquisition: Competence and processing. In Flynn, S., Martohardjono, G. and O’Neill, W., editors, *The generative study of second language acquisition*, Hillsdale, NJ: Erlbaum, 111–30.
- Williams, J.N., Möbius, P. and Kim, C.** 2001. Native and non-native processing of English *wh*-questions: parsing strategies and plausibility constraints. *Applied Psycholinguistics* 22, 509–40.

- Xu, L.** 1990: Remarks on LF movement in Chinese questions. *Linguistics* 28, 355–82.
- Yamashita, H., Stowe, L. and Nakayama, M.** 1993. Processing of Japanese relative clause constructions. In Clancy, P., editor, *Japanese/Korean Linguistics*, Volume 2. Stanford, CA: Center for the Study of Language and Information, 248–63.