SSLA, **27**, 53–78. Printed in the United States of America.
DOI: 10.1017/S0272263105050035

GAPS IN SECOND LANGUAGE SENTENCE PROCESSING

Theodore Marinis, Leah Roberts, Claudia Felser, and Harald Clahsen
University of Essex

Four groups of second language (L2) learners of English from different language backgrounds (Chinese, Japanese, German, and Greek) and a group of native speaker controls participated in an online reading time experiment with sentences involving long-distance *wh*-dependencies. Although the native speakers showed evidence of making use of intermediate syntactic gaps during processing, the L2 learners appeared to associate the fronted *wh*-phrase directly with its lexical subcategorizer, regardless of whether the subjacency constraint was operative in their native language. This finding is argued to support the hypothesis that nonnative comprehenders underuse syntactic information in L2 processing.

The real-time processing of sentences involving displaced constituents, or *filler-gap dependencies*, has been the focus of a considerable body of psycholinguistic research on monolingual sentence comprehension. A syntactically dislocated constituent—such as the fronted *wh*-phrase *which book* in *Which book did you read in only one hour?*—poses a challenge for the human sentence processing mechanism insofar as it cannot be fully integrated immediately into the emerging semantic or discourse representation. Instead, it must

Theodore Marinis is now working at the Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, and Leah Roberts is at the Max-Planck-Institute for Psycholinguistics, Nijmegen. The research reported here was supported by the Leverhulme Trust (grant no. F/00 213B to H. Clahsen, C. Felser, and R. Hawkins), which is gratefully acknowledged. We thank Bob Borsley, Roger Hawkins, Andrew Radford, the audiences at EUROSLA 12, the 24th Deutsche Gesellschaft für Sprachwissenschaft Meeting, the 27th annual Boston University Conference on Language Development, EUROSLA 13, three anonymous *SSLA* reviewers for helpful comments and discussion, and Ritta Husted and Michaela Wenzlaff for helping with the data collection. We also wish to thank Ted Gibson and Tessa Warren for making their prepublication manuscript available to us.

Address correspondence to: Claudia Felser, Department of Language and Linguistics, University of Essex, Colchester CO4 3SQ, United Kingdom; e-mail: felsec@essex.ac.uk.

© 2005 Cambridge University Press 0272-2631/05 $12.00

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

be retained in short-term memory until it can be linked to its subcategorizer or thematic role assigner. As the computational cost incurred by temporarily storing a filler in short-term memory increases with the distance between the filler and its associated gap (see, among others, Gibson, 1998; King & Just, 1991; King & Kutas, 1995; Kluender & Kutas, 1993), the human sentence processing mechanism will normally attempt to integrate a dislocated element at the earliest grammatically possible point during parsing. This well-documented preference for keeping filler-gap dependencies as short as possible is known as the Active Filler Hypothesis (Clifton & Frazier, 1989).

Linguistic theories differ with respect to the way filler-gap dependencies are analyzed. Within the generative-transformational tradition, a displaced constituent is assumed to form a syntactic dependency with an empty category at its base position and is thus only indirectly linked to its subcategorizer. According to the copy theory of movement (Chomsky, 1995), the empty category (e_i) involved in filler-gap dependencies is a silent but otherwise identical copy of the displaced constituent itself, as in (1).

(1) *Which book*$_i$ did you read e_i in only an hour?

On the other hand, some syntactic frameworks, including recent variants of Head-Driven Phrase Structure Grammar (HPSG), assume that a dislocated element is linked directly to its lexical subcategorizer (Pollard & Sag, 1994; Sag & Fodor, 1995). This linguistic controversy has given rise to different hypotheses as to how filler-gap dependencies are processed. The Trace Reactivation Hypothesis (TRH) assumes that the human parser postulates empty categories—*traces*—during the online comprehension of sentences containing such dependencies (e.g., Bever & McElree, 1988; Love & Swinney, 1996; Nicol & Swinney, 1989), whereas the Direct Association Hypothesis (DAH) claims that establishing a filler-gap dependency is a lexically driven process triggered by the automatic mental reconstruction of the subcategorizer's argument structure when it is encountered (Pickering & Barry, 1991; Sag & Fodor).

Crosslinguistic findings from studies on monolingual sentence comprehension suggest that two distinct mental processes may be involved in the processing of filler-gap dependencies: (a) a phrase-structure-based mechanism that triggers the retrieval of a filler from short-term memory at a specific structural position (the processing equivalent of inserting a copy of the filler into a particular syntactic slot, as predicted by the TRH); and (b) a lexically driven process of semantically integrating a displaced constituent with its thematic role assigner or other licenser, as predicted by the DAH. Although these two processes are usually difficult to dissociate empirically in head-initial languages like English (but see Lee, 2004; Nicol, 1993), evidence for the TRH can be gathered from studies on the processing of scrambling or *wh*-object movement in verb-final languages such as Japanese (Nakano, Felser, & Clahsen, 2002) or German (Clahsen & Featherston, 1999; Featherston, 2001; Fiebach, Schlesewsky, & Friederici, 2002), which found filler-reactivation effects before the subcategorizing verb had been encountered.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

Regardless of whether a filler is assumed to be linked to its lexical subcategorizer via empty categories located within the subcategorizer's extended projection, many contemporary syntactic theories agree that some kind of intermediate linguistic structure is present at intervening clause boundaries for dependencies spanning more than one clause. This intermediate structure mediates between the filler and its ultimate gap (or subcategorizer). Transformational theories (e.g., Chomsky, 1995) assume that long filler-gap dependencies are mediated by empty categories at intervening clause boundaries. More specifically, intermediate copies of the filler are assumed to be located in the specifier of intervening complementizer phrases (CPs). An example of what is commonly referred to as successive cyclic *wh*-movement is provided in (2).[1]

(2) *Who*$_i$ do you think e'_i (that) John says e'_i (that) Mary likes e_i?

In phrases containing a gap, nontransformational theories such as Generalized Phrase Structure Grammar (Gazdar, Klein, Pullum, & Sag, 1985) and many versions of HPSG (Levine & Hukari, 2004; Pollard & Sag, 1994) assume that the presence of a SLASH or GAP feature indicates an unresolved dependency. While remaining neutral as to the precise nature of the intermediate linguistic structure that has been claimed to be present at clause boundaries, we will, for expository reasons, refer to the positions marked e'_i in (2) as *intermediate gaps*.

Psycholinguistic evidence for successive cyclicity has been found, for example, in a study by Kluender and Kutas (1993) using event-related brain potentials (ERPs) and in a reading-time study by Gibson and Warren (2004). Relative to sentences in which subjacency was respected, Kluender and Kutas observed that the processing difficulty for sentences containing subjacency violations—such as (3)—increased both at the intervening *wh*-pronoun and at the filler's base position.[2]

(3) **Who*$_i$ couldn't you decide *who* should sing something for e_i at the family reunion?

Gibson and Warren (2004) investigated native English speakers' processing of grammatical sentences containing long-distance *wh*-dependencies. In line with Kluender and Kutas (1993), Gibson and Warren found that the availability of an intermediate *landing site* facilitated a filler's integration with its subcategorizer, thus providing indirect evidence for the psychological reality of intermediate gaps in first language (L1) sentence processing. Gibson and Warren's reading-time study provided the model for the present study and is discussed in more detail in the following section.

There is ample evidence that the mental representations constructed during L1 sentence processing are built up rapidly and in an incremental fashion (Fodor, 1995; Pickering, 1999) and that these representations also include abstract linguistic structure, such as empty categories or syntactic gaps. Nonetheless, surprisingly little is known to date about the way second language (L2) learners process the L2 input in real time. Our previous studies of L2

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

processing indicate that, although L2 learners, like native speakers (NSs), are guided by lexical information during parsing, they rely on phrase-structure information to a lesser extent than NSs, irrespective of their language background (Felser, Roberts, Marinis, & Gross, 2003; Papadopoulou & Clahsen, 2003; Roberts, 2003). If this is correct, then we might expect that, when processing *wh*-dependencies, L2 learners perform in accordance with the DAH but do not postulate any intermediate syntactic gaps.

PREVIOUS STUDIES OF L2 LEARNERS' PROCESSING OF *WH*-DEPENDENCIES

Only a few published studies have examined the real-time processing of *wh*-movement by L2 learners using online tasks.[3] Juffs and Harrington (1995) reported the results from two online grammaticality judgment experiments that measured Chinese-speaking learners' accuracy and reading times (RTs) for grammatical and ungrammatical sentences involving either subject or object extractions. Their results indicated that the learners had difficulties with grammatical sentences involving subject—but not object—extraction (cf. White & Juffs, 1998). The analysis of the reading-time data revealed that the Chinese-speaking participants slowed down significantly more at the region following the matrix verb in subject extractions from finite clauses (4a) than in object extractions (4b). No such effect was attested in the group of NS controls.

(4) a. *Who*$_i$ did Ann say e_i likes her friend? (subject extraction)
b. *Which man*$_i$ did Jane say her friends like e_i? (object extraction)

The authors argued that the learners' relatively poorer performance on subject extractions reflects processing rather than competence problems (cf. Juffs & Harrington, 1996). Observe that in sentences like (4a), the gap following the verb *say* may initially be analyzed as an object gap, a decision that must be revised as soon as the embedded verb *likes* is encountered. Even though this kind of reanalysis causes no or little processing difficulty for NSs, it does pose a problem for L2 learners, according to Juffs and Harrington.

Note, however, that given the nature of Juffs and Harrington's (1995) materials, their results do not provide any unequivocal evidence for the learners' use of empty categories during processing. As the purported trace position is adjacent to the subcategorizing verb, the slowdown observed in the postgap region may also be due to the learners' attempt to link the fronted *wh*-phrase directly to its subcategorizer in accordance with the DAH. The possibility that the learners may have a lexically or verb-driven processing strategy is strengthened by the fact that the learners (but not the NSs) also showed elevated RTs at the matrix verb *say*, a region prior to the locus of reanalysis. Juffs and Harrington (1996) speculated that the learners may be confused by the lack of semantic fit of the *wh*-pronoun *who* as the object of *say* at this point.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

Another reading time study, Williams, Möbius, and Kim (2001), investigated filled-gap effects in L2 processing. Specifically, this study examined whether L2 learners are sensitive to plausibility constraints during parsing. The experimental sentences involved adjunct extractions in two plausibility conditions, as shown in (5a) and (5b).

(5) a. *Which friend*$_i$ did the gangster hide the car for e_i late last night? (plausible-at-V)
b. *Which cave*$_i$ did the gangster hide the car in e_i late last night? (implausible-at-V)

In (5a), the fronted *wh*-phrase is a plausible object of the verb *hide*, whereas in (5b) it is not. Previous studies have shown that English NSs initially attempt to analyze the displaced *wh*-phrase as a direct object, a misanalysis that gives rise to increased processing difficulty when the real object *the car* is encountered (see, e.g., Stowe, 1986). In Williams et al.'s self-paced reading experiment, Chinese-, Korean-, and German-speaking learners of English read sentences presented on a computer screen in a word-by-word fashion and pressed a stop button to indicate the point at which they thought the sentence had become implausible. Assuming that online sentence comprehension is incremental in nature, the authors predicted that if the learners adopt a filler-driven or gap-as-first-resort strategy, then the *wh*-phrase in both conditions would initially be analyzed as the object of the verb. A filled-gap effect, reflected in longer RTs, would then be observed on the postverbal noun phrase (NP) due to the need for reanalysis at this point. If, on the other hand, a gap is posited only as a last-resort strategy to avoid ungrammaticality (see Fodor, 1978), then no such slowdown would be expected at the postverbal NP.

In the *stop-making-sense* task, the learners behaved similarly to the NSs. Participants made more stop decisions at or immediately after the verb in the implausible-at-V condition than in the corresponding plausible condition, which suggested that both the learners and the NSs were sensitive to plausibility information. The analysis of the reading-time data showed that, for all participant groups, the postverbal noun in the plausible-at-V condition elicited longer RTs than the postverbal noun in the implausible-at-V condition. This indicates that both the NSs and the learners analyzed the *wh*-filler as the direct object of the verb and that the plausibility of the *wh*-filler as a direct object affected the ease of reanalysis. The learners' L1 did not appear to have any effect on how they processed the experimental sentences. However, only the NSs showed an effect of plausibility at the determiner that introduced the postverbal NP. According to the authors, the earlier onset of the filled-gap effect observed in the NS group may indicate a greater sensitivity to the syntactic cue provided by the determiner, which signaled an incoming NP.

Williams et al.'s (2001) online experiment was complemented by an offline acceptability judgment task that investigated whether the learners were able to recover from misanalysis. The results showed that the learners—but not the NSs—judged the plausible-at-V sentences unacceptable significantly more

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

often than the implausible-at-V sentences. The authors concluded that the learners had more difficulty recovering from an initial misanalysis than NSs, particularly when this analysis is plausible, which suggested an overcommitment to a strongly plausible first analysis.

In summary, Williams et al.'s (2001) results suggested that L2 learners—like NSs—employ a filler-driven parsing strategy when processing *wh*-dependencies, irrespective of L1 background.[4] Observe, however, that like the results from Juffs and Harrington's (1995) study, Williams et al.'s results do not bear directly on the question of whether L2 learners postulate empty categories during processing. It is possible that the participants associated the *wh*-filler with the verb directly, a decision that they were forced to undo when the actual theme or patient argument became available. As the authors pointed out, the filled-gap effect observed on the postverbal noun in the nonnative participants may reflect purely thematic—rather than thematic and syntactic—reanalysis processes. The current study aims to dissociate potentially verb-driven integration effects from syntactic gap-filling by examining L2 learners' processing of successive-cyclic *wh*-movement structures.

THE PRESENT STUDY

Our study was modeled after Gibson and Warren's (2004) study on the processing of long-distance *wh*-dependencies by adult NSs of English. Using a self-paced reading task, Gibson and Warren investigated how NSs process sentences such as (6a) and (6b).

(6) a. The manager who_i the consultant claimed e'_i that the new proposal had pleased e_i will hire five workers tomorrow.
b. The manager who_i the consultant's claim about the new proposal had pleased e_i will hire five workers tomorrow.

The sentences in (6) differ in that (6a)—but not (6b)—provides an intermediate landing site for the fronted *wh*-pronoun. This is because in (6a), *wh*-movement has crossed a clause boundary that signals the beginning of a new cyclic domain, whereas (6b) involves extraction across a NP. Crucially, the linear distance between the filler and its ultimate gap (as measured in terms of the number of intervening words) was kept the same in both experimental conditions. To control for a possible confounding effect of subject-verb distance, Gibson and Warren's materials also included sentences that did not involve any *wh*-movement but differed in the relative distance between the verb *pleased* and the head of its subject (7a–b).

(7) a. The consultant claimed that the new *proposal* had pleased the manager who will hire five workers tomorrow.
b. The consultant's *claim* about the new proposal had pleased the manager who will hire five workers tomorrow.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

Gibson and Warren found an interaction between extraction and intervening phrase type at the region containing the *wh*-filler's subcategorizer *pleased*. RTs were shorter for sentences such as (6a) with an intermediate landing site than for sentences such as (6b). This effect was not present in the nonextraction conditions and thus cannot be attributed to any differences in subject-verb distance between the VP and NP conditions. Furthermore, the RTs elicited by the complementizer *that* in (6a) were found to be longer than in the corresponding nonextraction condition (7a). However, as the preposition *about* was also read more slowly in the extraction condition (6b) than in the nonextraction condition (7b), and the interaction between extraction and intervening phrase type proved only marginally significant, this observation does not by itself provide any conclusive evidence for filler reactivation at this position.

The intermediate gap effect observed by Gibson and Warren (2004) supports a strong version of the Active Filler Hypothesis, according to which a filler is reactivated cyclically to break up long dependencies into a series of shorter dependencies (see, e.g., Crocker, 1996; Frazier & Clifton, 1989). Note, however, that there was an asymmetry in Gibson and Warren's experimental materials between the extraction and nonextraction conditions; as compared to the nonextraction conditions, the extraction conditions contained more words preceding the embedded verb and one additional level of embedding before the critical segments. This asymmetry may have introduced a confounding factor such that lower RTs in the nonextraction conditions were partly due to the differences in length, structural complexity, or both, between the extraction and nonextraction conditions.

The present study has two major aims: (a) to replicate Gibson and Warren's (2004) finding with NSs of English using improved materials, and (b) to investigate whether L2 learners of English from different language backgrounds process long-distance *wh*-dependencies in the same way as—or differently from—NSs. To test whether the learners' L1 background has an effect on their processing of long-distance *wh*-dependencies in L2 English, we examined learners from both *wh*-movement (Greek and German) and *wh*-in-situ backgrounds (Chinese and Japanese).[5]

METHOD

Participants

Four groups of learners of L2 English participated in the current study: 34 Chinese-speaking learners (mean age = 25, range = 17–33), 26 Japanese-speaking learners (mean age = 27, range = 20–40), 24 German-speaking learners (mean age = 24, range = 19–46), and 30 Greek-speaking learners (mean age = 25, range = 20–37). Additionally, a group of 24 native English speakers served as a control group (mean age = 24, range = 19–34). The participants were recruited from the undergraduate and postgraduate student communities at the University of Essex and were paid a small fee for their participation. All participants

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

had normal or corrected-to-normal vision and were naïve with respect to the purpose of the experiment. The Chinese, Japanese, and German learners had first been exposed to English around the age of 11 in a classroom setting, and the Greek learners around the age of 8. Finally, the Chinese-speaking learners were all NSs of Mandarin Chinese.

To determine the learners' general proficiency in English at the time of testing, all completed a standardized proficiency test, the Oxford Placement Test (OPT; Allen, 1992). As our experimental materials involved structurally complex sentences, only learners at or above the upper intermediate level (i.e., 145/200 points or above) were included in our study. In addition to the OPT, the learners completed an offline questionnaire, whose purpose was to ensure that they were able to comprehend complex sentences of the kind that were later used in the online task. The questionnaire consisted of 20 sentences slightly adapted from those used in the self-paced reading task. There were five sentences corresponding to each of the four experimental conditions in the online experiment, as described in the Materials section. Each sentence was followed by a comprehension question and three choices, as illustrated by (8). For the full set of questionnaire materials, see Appendix A.

(8) The captain who the officer decided that the young soldier had displeased will write a formal report next week.
Who made a decision?
the captain *the officer* *the soldier*

The participants were instructed to read the sentences and indicate which of the three answers they considered the most appropriate.

Table 1 provides a summary of the learners' biographical data. Specifically, this table presents the learners' age at the time of testing, their age of first exposure to English, the amount of time they had spent in the United Kingdom at the time of testing, and the results from the OPT and offline questionnaire. All participants scored at 75% or above correct on the questionnaire, with mean scores exceeding 90% for all groups. This suggests that the learners could handle the types of sentences used in the self-paced reading experiment in an offline task.

Materials

The materials for the online task comprised a total of 88 sentences, including 8 practice items, 20 experimental sentences, and 60 filler sentences. The reason for including such a large number of fillers was to disguise the purpose of the experiment, thus preventing the participants from developing any response strategies. Each of the experimental sentences came in four versions in a 2×2 design with the extraction conditions (extraction vs. nonextraction) crossed by phrase type (VP vs. NP), as illustrated by (9a–d). The full set of experimental sentences is provided in Appendix B.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

Table 1. Summary of the learners' biographical data and pretest scores

Learner groups	Age[a]	Age of first exposure[a]	Time spent in U.K.[a]	OPT scores[b]	Questionnaire scores
Chinese					
M	25.06	11.94	0.85	156.35	92.5%
SD	3.92	2.17	1.38	7.16	6.75
Japanese					
M	26.54	11.77	2.15	169.15	92.31%
SD	4.21	0.91	1.58	11.52	6.96
German					
M	24.00	11.36	1.60	176.84	98.00%
SD	6.22	1.87	1.09	12.44	4.33
Greek					
M	24.80	8.67	2.48	172.50	96.17%
SD	3.11	2.19	2.20	9.90	4.29

[a]In years.
[b]The means represent scores out of a possible maximum of 200 points.

(9) a. Extraction across a VP (+ intermediate gap)
The nurse who_i the doctor argued e'_i that the rude patient had angered e_i is refusing to work late.
b. Extraction across an NP (– intermediate gap)
The nurse who_i the doctor's argument about the rude patient had angered e_i is refusing to work late.
c. Nonextraction, local subject-verb integration (VP)
The nurse thought the doctor argued that the rude patient had angered the staff at the hospital.
d. Nonextraction, nonlocal subject-verb integration (NP)
The nurse thought the doctor's argument about the rude patient had angered the staff at the hospital.

The sentences used in the two extraction conditions were structurally identical to those used by Gibson and Warren (2004). In the extraction conditions, an initial NP (*the nurse*) was followed by a relative clause that was introduced by a *wh*-pronoun (*who*) functioning as the object of the embedded verb (*angered*). The intermediate verb in the extraction-VP conditions (*argued*) was always a bridge verb (i.e., one that permits *wh*-extraction out of its complement clause). Although the [+human] relative pronoun *who* did not make a plausible direct object for the bridge verbs used in the extraction-VP condition, it is at least conceivable that the parser initially misanalyzes the filler as the object of the higher verb (i.e., *argued*) on purely structural, least effort grounds. To ensure as far as possible that the filler would not be mistaken for the object of the higher verb, we used only verbs that were strongly biased toward taking a sentential complement. Six verbs (*claim*, *argue*, *prove*, *suggest*, *conclude*, and *decide*) were selected from Garnsey, Pearlmutter, Myers, and Lotocky's (1997) list of sentential complement verbs, and a further six were

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

tested independently for their complement bias. To this end, 10 English NSs were given both free- and forced-choice sentence completion tasks, and out of the six verbs tested, we selected three (*dream*, *state*, and *think*) that showed a strong sentential complement bias of 73% or above.

The sentences in the two nonextraction conditions differed from those in Gibson and Warren (2004) in that they contained exactly the same number of words (up to the embedded verb) as the sentences in the corresponding extraction conditions. To avoid an asymmetry between the extraction and corresponding nonextraction conditions with respect to the degree of structural complexity, we added a further level of embedding to the sentences in the nonextraction conditions.

Four different experimental scenarios were created, each of which contained only one version of each experimental sentence. The conditions were distributed evenly across the four versions, so that each participant saw the same number of sentences per condition. The experimental sentences were pseudorandomized and mixed with the filler sentences. Each of the four experimental scenarios was preceded by the same practice items. All experimental sentences and half of the fillers were followed by a comprehension question, the purpose of which was to ensure that the subjects read the sentence properly and made an active effort to comprehend them.

Procedure

The pretests (OPT and questionnaire) and the online task were administered in two separate sessions, with approximately 1 week in between. Reading time and comprehension accuracy data were collected using the noncumulative moving-window procedure (Just, Carpenter, & Woolley, 1982). The presentation of the stimuli and the recording of reaction times and end-of-sentence responses was controlled by the MS-DOS version of the new experimental setup (NESU) software package (Baumann, Nagengast, & Klaas, 1993). The stimulus sentences were presented in a segment-by-segment fashion, in white letters (Arial 24 point) on a black background in the center of a 17-inch monitor. The experimental sentences were divided into six segments as indicated in (10).

(10) The nurse who / the doctor argued / that / the rude patient /
1 2 3 4
had angered / is refusing to work late.
5 6

As the purpose of our experiment was to test predictions for two specific sentence regions—namely, the beginning of the embedded clause and the subcategorizing verb—the complementizer *that* (or the preposition *about*) and the verb group *had angered* were presented as separate segments. The remainder of the sentences was divided into larger chunks so as to allow for a more natural reading than would be possible with word-by-word presentation.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

Participants were instructed to read each segment as quickly as possible for comprehension and then to press a pacing button to bring up the next segment. The end of each sentence was indicated by a full stop after the last word of the final segment. The last segment of each experimental sentence and of half the filler sentences was followed by a comprehension question (e.g., *Who angered the nurse?*). Following the presentation of the question, two answer options appeared on the screen, one on the left- and one on the right-hand side. For half of the questions, the correct answer appeared on the left-hand side of the screen, and for the other half, the answer on the right-hand side was correct. Participants were instructed to press either the left or right button of a dual push-button box, depending on which of the answers (left or right) they chose as correct. After the end of each trial, a message appeared on the screen instructing participants to press a predetermined key on the keyboard to trigger the start of the next trial. All participants completed the online task in approximately 30 minutes.

RESULTS

Accuracy

All participant groups scored highly in answering the comprehension questions that followed the experimental sentences. The NSs' mean accuracy score was 79.5%. The Chinese group scored 79%, the Japanese group 74.5%, the German group 84.75%, and the Greek group 79.75% correct. This demonstrates that the participants were paying attention to the task and that they were reading the sentences properly. The data from one Chinese-speaking participant who only scored 42% (= 2 *SD* below the group mean) correct were excluded from further analysis. A mixed three-way ANOVA with extraction (extraction vs. nonextraction) and phrase type (VP vs. NP) as within-subjects factors and language (English, Chinese, Japanese, German, Greek) as the between-subjects factor showed no significant main effects or interactions. This indicates that the learner groups did not differ from each other or from the NSs with respect to their ability to comprehend the experimental sentences. Moreover, neither the presence versus absence of a filler-gap dependency nor the type of the intervening phrase (VP vs. NP) appeared to have influenced the participants' accuracy scores.

Reading Times

Segments 3 and 5 are the segments relevant for determining whether intermediate syntactic gaps were postulated during processing. Segment 3 in the extraction-VP condition (9a) contains the complementizer *that*, which is hypothesized to trigger the reactivation of the filler *who* (= *the nurse*) at this position, whereas in the corresponding nonextraction condition (9c), the complementizer merely indicates the beginning of an embedded clause. Thus, if the participants postulate an intermediate gap at this point during process-

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

ing in sentences such as (9a), we expect segment 3 to elicit longer RTs in the extraction-VP condition than in the nonextraction-VP condition. Segment 5 contains the verb to which the filler ultimately needs to be linked in the extraction conditions. Recall that in sentences where an intermediate gap is possible (i.e., in the extraction-VP condition), the relevant distance between the filler and its subcategorizer is shorter than in sentences that do not permit the insertion of an intermediate gap, if the filler is mentally reactivated at the intermediate gap site. Following Gibson and Warren (2004), we therefore predict that, if filler integration is facilitated by the presence of an intermediate gap, RTs to segment 5 should be shorter in the extraction-VP condition than in the extraction-NP condition.

Furthermore, if the learners process long-distance *wh*-dependencies in the same way as NSs, we expect to find no statistical differences between the five participant groups. If, however, properties of the L1 have an impact on the way long-distance *wh*-dependencies are processed in L2 English, we might expect to find differences between the Chinese- and Japanese-speaking learners on the one hand, and the German- and Greek-speaking learners on the other. This is because, unlike German or Greek, Chinese and Japanese lack (overt) successive-cyclic *wh*-movement. Finally, if L2 sentence processing differs from L1 processing but is not influenced by properties of the L1 grammar, we should find differences between the NSs and the learners, but not among the individual learner groups.

Data Trimming. Following standard practice in this type of experiment, we included only RTs from correctly answered trials in the statistical analysis. Prior to the analysis of the data, responses were screened for trials whose total RT exceeded a timeout of 20,000 ms for the NSs and 25,000 ms for the learners.[6] This affected between 0.79% and 1.67% of the data across the language groups. Additionally, we screened the participants' RTs to each segment for outliers and eliminated individual data points beyond two standard deviations from the mean RTs for each condition per subject and item. This affected 3.86% of the data from the NSs, 3.66% of the Chinese learners' data, 0.73% of the Japanese learners' data, 3.08% of the German learners' data, and 2.04% of the Greek learners' data. Finally, the data sets from one Chinese- and two German-speaking participants were excluded because they were incomplete. The remaining data from 32 Chinese-, 26 Japanese-, 22 German-, and 30 Greek-speaking learners of English, and from 24 NS controls were included in the statistical analysis.

Between-Groups Analyses. Table 2 provides an overview of the five participant groups' mean RTs to each segment for all conditions. To determine whether there were any differences in processing the experimental sentences between the groups, we carried out a mixed three-way ANOVA with the factors extraction (extraction vs. nonextraction) and phrase type (VP vs. NP) as within-subjects factors and language (English, Chinese, Japanese, German, Greek) as a between-subjects factor. Recall that segments 3 and 5 are the crucial ones for the issue under investigation.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

Table 2. Mean RTs in milliseconds and standard deviations (in parentheses) per segment and condition

	Segment					
Condition	1	2	3	4	5	6
L1 English						
Extraction-VP	753	1069	825	1268	1075	1359
	(419)	(478)	(310)	(654)	(567)	(609)
Extraction-NP	717	1507	833	1366	1307	1343
	(353)	(755)	(340)	(619)	(750)	(558)
Nonextraction-VP	712	1195	729	1237	811	985
	(252)	(632)	(272)	(764)	(268)	(479)
Nonextraction-NP	718	1099	657	1066	820	1073
	(223)	(449)	(193)	(525)	(249)	(360)
L1 Chinese						
Extraction-VP	1585	2665	1062	2155	1630	3065
	(799)	(1101)	(716)	(1070)	(826)	(1329)
Extraction-NP	1434	3227	814	2492	1813	2822
	(474)	(1221)	(589)	(1290)	(925)	(1160)
Nonextraction-VP	1723	2691	836	1963	1349	1876
	(793)	(1059)	(473)	(880)	(611)	(719)
Nonextraction-NP	1783	2329	857	1676	1503	1996
	(808)	(864)	(365)	(744)	(826)	(765)
L1 Japanese						
Extraction-VP	1547	2652	956	2053	1560	2994
	(609)	(962)	(694)	(832)	(529)	(1438)
Extraction-NP	1507	3052	1126	1972	1910	2859
	(959)	(1105)	(587)	(861)	(765)	(905)
Nonextraction-VP	1470	2512	955	1675	1420	1929
	(478)	(1107)	(506)	(505)	(446)	(597)
Nonextraction-NP	1590	2282	918	1643	1523	1833
	(556)	(815)	(279)	(513)	(689)	(565)
L1 German						
Extraction-VP	844	1562	977	1628	1609	1599
	(278)	(803)	(353)	(563)	(888)	(424)
Extraction-NP	845	1801	935	1351	1374	1343
	(299)	(601)	(290)	(506)	(537)	(409)
Nonextraction-VP	958	1659	925	1196	959	1273
	(302)	(778)	(309)	(449)	(229)	(556)
Nonextraction-NP	968	1533	753	1265	925	1214
	(445)	(809)	(220)	(470)	(339)	(424)
L1 Greek						
Extraction-VP	851	1636	838	1757	1330	1786
	(464)	(906)	(485)	(1354)	(902)	(885)
Extraction-NP	909	2170	837	2022	1394	1682
	(387)	(1297)	(365)	(1547)	(1036)	(621)
Nonextraction-VP	1173	2015	875	1634	1086	1209
	(740)	(1160)	(417)	(1103)	(606)	(452)
Nonextraction-NP	1288	2052	664	1417	1008	1262
	(868)	(1216)	(254)	(764)	(475)	(407)

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

At segment 3, we found a main effect of extraction, $F_1(1, 129) = 8.412$, $p < .01$, $\eta^2 = .061$; $F_2(1, 94) = 6.321$, $p < .05$, $\eta^2 = .063$, a main effect of phrase type in the items analysis that was approaching significance in the subjects analysis, $F_1(1, 129) = 3.081$, $p = .085$, $\eta^2 = .023$; $F_2(1, 94) = 8.814$, $p < .01$, $\eta^2 = .086$, as well as a main effect of language, $F_1(1, 129) = 2.443$, $p = .05$, $\eta^2 = .070$; $F_2(4, 94) = 8.082$, $p < .001$, $\eta^2 = .256$. The analysis also revealed an interaction of phrase type and language that was significant in the items analysis, $F_1(4, 129) = 0.718$, $p = .581$; $F_2(4, 94) = 2.534$, $p < .05$, $\eta^2 = .097$. Additionally, a three-way interaction between extraction, phrase type, and language was found to approach significance, $F_1(4, 129) = 2.418$, $p = .05$, $\eta^2 = .070$; $F_2(4, 94) = 2.181$, $p = .077$, $\eta^2 = .085$.

The analysis of segment 5 revealed a main effect of extraction, $F_1(1, 129) = 49.970$, $p < .001$, $\eta^2 = .279$; $F_2(1, 94) = 56.882$, $p < .001$, $\eta^2 = .377$, a main effect of language, $F_1(4, 129) = 6.916$, $p < .001$, $\eta^2 = .177$; $F_2(4, 94) = 30.554$, $p < .001$, $\eta^2 = .565$, and an interaction of phrase type and language in the items analysis, $F_1(4, 129) = 1.887$, $p = .117$; $F_2(4, 94) = 3.088$, $p < .05$, $\eta^2 = .116$. As our results showed interactions with the factor language at both segments 3 and 5, we went on to analyze the data from each of the five participant groups separately.[7]

Native Speakers. For the NS data, we performed separate two-way ANOVAs for segments 3 and 5 with the factors extraction and phrase type. The analysis of the NSs' RTs to segment 3 revealed a main effect of extraction, $F_1(1, 23) = 4.578$, $p < .05$, $\eta^2 = .166$; $F_2(1, 19) = 9.672$, $p < .01$, $\eta^2 = .337$, which reflected the fact that RTs for the extraction conditions were significantly slower than those for the nonextraction conditions. Additionally, we found a main effect of phrase type in the items analysis, $F_1(1, 23) = 0.844$, $p = .368$; $F_2(1, 19) = 4.967$, $p < .05$, $\eta^2 = .207$. Like Gibson and Warren (2004), however, we found no significant interaction between extraction and phrase type.

At segment 5, the two extraction conditions again produced longer RTs than the nonextraction conditions, and sentences of the extraction-VP condition were read more quickly than those of the extraction-NP condition. The ANOVA for this segment showed a main effect of extraction, $F_1(1, 23) = 11.054$, $p < .01$, $\eta^2 = .325$; $F_2(1, 19) = 20.355$, $p < .001$, $\eta^2 = .517$, a main effect of phrase type, $F_1(1, 23) = 4.759$, $p < .05$, $\eta^2 = .171$; $F_2(1, 19) = 4.715$, $p < .05$, $\eta^2 = .199$, and an interaction between extraction and phrase type, $F_1(1, 23) = 4.994$, $p < .05$, $\eta^2 = .178$; $F_2(1, 19) = 4.364$, $p = .05$, $\eta^2 = .187$. Subsequent pairwise comparisons revealed significant differences between the extraction- and nonextraction-VP conditions, $t_1(23) = 2.560$, $p < .05$; $t_2(19) = 2.389$, $p < .05$, and between the extraction- and nonextraction-NP conditions, $t_1(23) = 3.551$, $p < .01$; $t_2(19) = 4.322$, $p < .001$, which reflected the additional processing cost associated with integrating a filler with its subcategorizer in the extraction conditions. Crucially, the difference between the extraction-VP and extraction-NP conditions also proved significant, $t_1(23) = 2.441$, $p < .05$; $t_2(19) = 2.220$, $p < .05$, which suggested that filler integration was comparatively less costly in the extraction-VP condition. The fact that the RTs for the two non-

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

extraction conditions (811 vs. 820 ms) did not differ significantly suggests that the relative distance between the critical verb and its subject did not affect processing time at the segment containing the verb.

In sum, the results from our NSs replicate Gibson and Warren's (2004) findings. Like Gibson and Warren, we observed elevated RTs at the intervening clause boundary in the extraction-VP condition compared to the corresponding nonextraction condition and shorter RTs to the segment containing the filler's subcategorizer for the extraction-VP than for the extraction-NP condition. These results support the hypothesis that NSs of English postulate intermediate gaps during the processing of long-distance *wh*-dependencies, which facilitates the filler's integration with its subcategorizer.

L2 Learners. As for the NSs, we carried out separate two-way ANOVAs with the factors extraction and phrase type for each learner group for the critical segments. At segment 3, the analyses of the data from the Chinese- and Japanese-speaking learners showed no main effects or interactions. For the German learners, we found a main effect of phrase type in the items analysis that was marginally significant in the subjects analysis, $F_1(1, 21) = 3.925$, $p = .061$, $\eta^2 = .157$; $F_2(1, 18) = 10.569$, $p < .01$, $\eta^2 = .370$, and a main effect of extraction in the subjects analysis only, $F_1(1, 21) = 4.388$, $p < .05$, $\eta^2 = .173$; $F_2(1, 18) = 2.186$, $p = .157$. The analysis of the Greek learners' data showed a main effect of phrase type in the items analysis that approached significance in the subjects analysis, $F_1(1, 29) = 3.593$, $p = .068$, $\eta^2 = .110$; $F_2(1, 19) = 6.315$, $p < .05$, $\eta^2 = .249$, an interaction between phrase type and extraction in the subjects analysis, $F_1(1, 29) = 7.095$; $p < .05$, $\eta^2 = .197$; $F_2(1, 19) = 1.493$; $p = .237$, but no effect of extraction. The interaction observed in the Greek group is due to their relatively short RTs in the nonextraction-NP condition, whereas their RTs to the complementizer *that* in the two VP conditions are actually longer in the nonextraction than in the extraction condition. Unlike the NSs, none of the learner groups showed a reliable main effect of extraction at segment 3, which indicates that the presence of a dislocated element earlier in the sentence did not significantly affect their processing of this segment.

At segment 5, all learner groups showed longer RTs in the extraction conditions than in the nonextraction conditions. The ANOVAs revealed main effects of extraction for all learner groups: Chinese: $F_1(1, 31) = 6.784$, $p < .05$, $\eta^2 = .180$; $F_2(1, 19) = 4.157$, $p = .056$, $\eta^2 = .180$; Japanese: $F_1(1, 25) = 8.162$, $p < .01$, $\eta^2 = .246$; $F_2(1, 19) = 5.216$, $p < .05$, $\eta^2 = .215$; German: $F_1(1, 21) = 15.175$, $p < .01$, $\eta^2 = .419$; $F_2(1, 19) = 27.241$, $p < .001$, $\eta^2 = .602$; Greek: $F_1(1, 29) = 9.149$, $p < .01$, $\eta^2 = .240$; $F_2(1, 19) = 12.415$, $p < .01$, $\eta^2 = .395$. For the Chinese group, the extraction effect was significant in the subjects analysis and very nearly approached significance in the items analysis. Additionally, we found a main effect of phrase type for the Japanese group, $F_1(1, 25) = 5.031$, $p < .05$, $\eta^2 = .168$; $F_2(1, 19) = 7.482$, $p < .05$, $\eta^2 = .283$, that approached significance in the items analysis for the German group, $F_1(1, 21) = 2.551$, $p = .125$; $F_2(1, 19) = 4.278$, $p = .053$, $\eta^2 = .192$. However, no interactions between extraction and phrase type were observed in any of the learner groups.[8]

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

These results suggest that the learners integrated the filler with its subcategorizer at segment 5 in both extraction conditions but that filler integration was not facilitated by the availability of an intermediate syntactic gap in the extraction-VP condition. Note that, for the German group, the extraction-VP condition actually elicited longer RTs at segment 5 than the extraction-NP condition. The observed effects of phrase type reflect the fact that some of the learner groups read segment 5 more quickly in either the VP or the NP conditions, independently of the presence of extraction.

Recall that the most crucial result from the NSs was the interaction between extraction and phrase type observed at segment 5. In the sentences involving extraction, the extraction-VP condition elicited shorter RTs than the extraction-NP condition, but no such difference was found between the corresponding nonextraction conditions. This pattern indicates that NSs of English associate the filler with an intermediate gap when processing sentences involving *wh*-extraction from an embedded clause, which facilitates filler integration later on. By contrast, none of the four learner groups showed any such interaction or intermediate gap effect.

DISCUSSION

The purpose of the current study was to investigate whether NSs and L2 learners of English make use of intermediate gaps during the processing of long-distance *wh*-dependencies. The main results of the self-paced reading experiment can be summarized as follows: (a) both the NSs and the learners proved equally good at comprehending sentences involving long-distance *wh*-dependencies; (b) all participants showed an effect of extraction at segment 5, which indicated that they attempted to integrate the displaced *wh*-phrase with its subcategorizing verb at this point during the parse (filler-integration effect); (c) for the NSs, filler integration was facilitated by the presence of an intermediate gap in sentences involving extraction across a clause boundary (intermediate gap effect); and (d) for the L2 learners there was no intermediate gap effect.

Intermediate Gaps in L1 Processing

Several studies of the processing of filler-gap dependencies have shown that NSs of English reactivate a displaced *wh*-constituent at the position of its associated syntactic gap. However, as the purported *wh*-gaps are located immediately after the verb or other lexical subcategorizer in English, results from these studies do not provide unambiguous evidence for phrase-structure- or trace-based reactivation; the results can also be explained by direct lexical association. To eliminate this potential confounding factor, we followed Gibson and Warren (2004) in using sentences that involved long-distance dependencies. According to both transformational (e.g., Chomsky, 1995) and nontransformational (Gazdar et al., 1985; Pollard & Sag, 1994) theories, sentences involving long-distance *wh*-movement also contain some kind of inter-

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

mediate syntactic structure (empty categories or GAP features) that links the filler to its subcategorizer.

Comparing sentences with and without extraction across either VPs or NPs, we found that English NSs postulate such intermediate gaps during real-time processing (thus replicating Gibson and Warren's 2004 findings). In the condition involving an intermediate gap, the NS group showed elevated RTs to the complementizer *that* relative to the same word in the corresponding nonextraction condition. This result may reflect an increase in the processing load while the parser consults a memory representation of the filler at this point and links it to the gap. As in Gibson and Warren's study, however, the preposition *about* in the NP conditions also elicited longer RTs in the extraction condition, so that the results from segment 3 must be interpreted with some caution. For example, it is conceivable that the main effect of extraction observed at this segment merely reflects the additional processing cost associated with storing the filler in working memory in the two extraction conditions. Alternatively, it is possible that, although extraction from a NP containing a specified subject is illicit, the parser attempts to analyze the *wh*-filler as the object of the preposition *about* at this point during processing (cf. Pickering, Barton, & Shillcock, 1994). If the parser considers the preposition *about* as a potential host for the *wh*-filler, then this hypothesis will not be sustained for long: Not only does such a dependency violate grammatical constraints but *about* is also immediately followed by its own object NP.

The crucial evidence that intermediate gaps do indeed form part of the mental representations constructed during parsing comes from our analysis of segment 5, the point at which the filler is integrated with its subcategorizer. The elevated RTs elicited by the two extraction conditions compared to the nonextraction conditions can be taken to reflect the cost of filler integration at this point. Assuming that the processing load associated with filler integration increases with the distance between the filler and its subcategorizer (Gibson, 1998, among others), the shorter RTs observed in the extraction-VP condition can be explained by the presence of intermediate linguistic structure that breaks up the long dependency into two shorter ones. Taken together, these results provide evidence that English NSs postulate intermediate gaps during the processing of long-distance *wh*-dependencies, in accordance with the subjacency constraint that forms part of their grammar. Our findings are consistent with a successive-cyclic version of the Active Filler Hypothesis, according to which a filler is retrieved from working memory at every grammatically possible gap position and not just upon encountering the lexical subcategorizer.

Native Speakers versus L2 Learners

The results from Williams et al.'s (2001) reading-time study suggested that L2 learners also employ a filler-driven strategy when processing *wh*-dependencies. Like NSs, their learners appeared to try to integrate a fronted *wh*-phrase with

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

a potential subcategorizer as soon as possible. Furthermore, both Williams et al.'s and Juffs and Harrington's (1995) results indicated that L2 learners may have more difficulty recovering from an initial misanalysis than NSs. Recall, however, that the elevated RTs observed by both Williams et al. and Juffs and Harrington at the reanalysis region are consistent both with an empty category-based account of gap-filling (TRH) and with direct lexical association (DAH). To separate these two possibilities, we used materials that syntactic theory claims also contain intermediate syntactic gaps. We found that the presence of such gaps facilitated a *wh*-filler's integration with its subcategorizer for the NSs but not for the learners. This suggests that the L2 learners did not postulate any intermediate syntactic gaps during processing but instead tried to link the filler directly to its lexical subcategorizer irrespective of the availability of an intermediate landing site. Gap-filling in L2 processing, then, appears to be driven by the lexicon rather than by requirements of the grammar, such as the subjacency constraint.

The extraction effect observed for the L2 learners at segment 5—the region containing the subcategorizing verb—is in accordance with the predictions made by the DAH. However, in contrast to the learners, the interaction between phrase type and extraction that we found for the NSs indicates that they postulated intermediate syntactic gaps during processing. The absence of any reliable extraction effects on segment 3 suggests that the prior encounter of a displaced element did not influence the learners' processing of the complementizer or preposition and that the *wh*-filler was not mentally reactivated at this point. In short, we found no evidence to suggest that L2 learners employ a filler-driven strategy when processing sentences containing long-distance *wh*-dependencies or that they postulate any syntactic gaps at all. Note that our results are not incompatible with Williams et al.'s (2001) findings, given that the effects they observed also occurred at or after a potential subcategorizer.

However, the learners' failure to make use of intermediate syntactic gaps did not seem to compromise their ability to understand the experimental sentences. Even though the NSs and the L2 learners were equally good at comprehending the experimental sentences, the two groups appeared to employ different processing mechanisms. English NSs made use of intermediate syntactic gaps, which facilitated filler integration further downstream. The L2 learners, on the other hand, attempted to establish a direct dependency between the fronted *wh*-phrase and its lexical subcategorizer despite the possibility of breaking down the *wh*-dependency into a series of smaller steps.

Our results are compatible with those of other studies that indicate that L2 learners underuse syntactic information during L2 processing as compared to NSs. Although there is evidence from various studies suggesting that L2 learners—like NSs—are guided by lexical-semantic and plausibility information during L2 sentence comprehension (Frenck-Mestre & Pynte, 1997; Juffs, 1998; Williams et al., 2001), studies by Felser et al. (2003), Papadopoulou and Clahsen (2003), and Roberts (2003) have shown that L2 learners do not seem to apply any locality principles based on phrase structure when processing

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

temporarily ambiguous sentences. The hypothesis that L2 processing differs from L1 processing with respect to the role of phrase-structure information in online processing is further supported by the results from event-related brain potential (ERP) studies. Hahne (2001), Hahne and Friederici (2001), and Isel (2002) found qualitative differences in first-pass parsing for sentences that were syntactically ill formed between NSs and L2 learners of German. Although sentences containing phrase-structure violations elicited two different ERP components associated with the processing of syntactic information in the NSs, an early anterior negativity and a P600, the L2 learners showed no early anterior negativity at all and, depending on their proficiency level, a delayed or no P600 component. Anterior negative-going waves occurring within 100–300 ms after stimulus onset have been argued to reflect early automatic structure-building processes, whereas the somewhat later P600 component has been associated with integration and reanalysis processes (Friederici, 2002). The observed L1/L2 differences indicate that L2 learners process syntactic category information differently from NSs and that the application of automatic, structure-based processing routines may be restricted during L2 sentence comprehension.

L1 Transfer in L2 Processing

The role of L1 transfer in nonnative language processing is controversial. Although some studies have found evidence of processing transfer in L2 sentence comprehension (e.g., Frenck-Mestre, 2002; Juffs, 1998), several other studies—including the present one—have found no differences in processing performance among learners from typologically different language backgrounds (Felser et al., 2003; Papadopoulou & Clahsen, 2003; Williams et al., 2001). Our study included four groups of L2 learners with similar levels of proficiency in L2 English, two from *wh*-movement backgrounds (German and Greek) and two from *wh*-in-situ backgrounds (Chinese and Japanese). If properties of the learners' L1 grammar influence the way they process sentences from the L2, then we might have expected both the German and Greek speakers to pattern with the NSs and differently from the Chinese and Japanese speakers. Our results, however, indicate that all L2 groups processed the experimental sentences in essentially the same way but differently from NSs.[9] Specifically, none of the learner groups appeared to postulate any intermediate gaps during real-time processing, irrespective of whether the subjacency constraint was operative in their L1. This shows that, even though the German- and Greek-speaking learners' L1 grammatical representations include intermediate syntactic gaps, they do not make use of such gaps when processing long-distance *wh*-dependencies in L2 English. The absence of such transfer effects may be at least partially explained by our hypothesis that L2 learners' sensitivity to syntactic information during real-time processing is more limited than that of NSs. Instead, L2 learners seem to rely more on lexical-semantic and other nonsyntactic cues to sentence interpretation.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

CONCLUSION

Our results show that NSs postulate intermediate syntactic gaps when processing long-distance *wh*-dependencies in English, whereas L2 learners do not. The learners' failure to make use of intermediate gaps during parsing proved independent of the presence of successive-cyclic *wh*-movement in their L1. This finding lends further support to the hypothesis that L2 learners' sensitivity to syntactic information during L2 processing is restricted relative to that of NSs. The observed dissociation between the learners' comprehension abilities and online sentence processing shows that, although learners are able to comprehend sentences containing long-distance *wh*-extractions, they do not use nativelike, phrase-structure-based processing mechanisms to achieve this goal during online comprehension.

(*Received 10 June 2004*)

NOTES

1. Chomsky (2000) proposed that additional copies of fronted elements are located at the left edge of (transitive) verb phrases. This proposal has not been tested in the present study, however.

2. The term *subjacency* refers to the requirement that nonlocal movements must take place in a series of small steps (Chomsky, 1973).

3. For reasons of space, and because the present paper focuses on sentence processing, we will not review the extensive literature on the L2 acquisition of *wh*-movement and subjacency here (for an overview of previous findings, see chapter 7 of Hawkins, 2001).

4. A potential problem with this study, however, is that there is no evidence that the learners interpreted the experimental items correctly. Recall that, in the offline task, the learners judged many of the experimental sentences as unacceptable even though they were both grammatical and fully plausible by the end of the sentence.

5. Some linguists have argued that (some analogue of) *wh*-movement also exists in Japanese (e.g., Watanabe, 2001). Although a critical evaluation of these proposals is beyond the aim and scope of this article, Japanese evidently lacks overt cyclic *wh*-movement of the kind that is found in *wh*-raising languages such as English, German, or Greek.

6. The reason for using different cutoff points for the learners and the NS controls was that the learners' RTs were slower overall than the NSs'.

7. Other differences in the RT data include the observation that for all groups, segment 2 elicited longer RTs in the extraction-NP condition than in all other conditions. A possible reason for this is that the extraction-NP condition is the only one in which no verb has been provided by the end of segment 2. As such, none of the three NPs presented thus far (*the nurse*, *who*, *the doctor's argument*) can be properly integrated into the emerging semantic or discourse representation at this point, which leads to a relatively higher local processing cost (see Gibson, 1998). As our predictions pertained to segments 3 and 5 only, statistics from the remaining segments are not reported here.

8. The Kolmogorov-Smirnov test was run on each of the groups' RT data for segments 3 and 5, and all but the data from the Greek group on segment 5 were found to be normally distributed. Therefore, an additional nonparametric (Friedman) test was run for the data from this group. The results confirmed those of the ANOVA reported, and there was no significant difference between the two extraction conditions on this segment ($p = .679$).

9. Although both German and Greek pattern with English in that they show successive-cyclic *wh*-movement, the direct translation equivalent of our extraction-VP condition is actually ungrammatical in German (but not in Greek). German also differs from Greek in that relative clauses are introduced by a relative pronoun rather than by a complementizer, as is possible in Greek. Despite the availability of a maximally close translation equivalent in Greek, however, the Greek-speaking learners did not appear to postulate any intermediate gaps.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

REFERENCES

Allen, D. (1992). *The Oxford Placement Test*. Oxford: Oxford University Press.

Baumann, H., Nagengast, J., & Klaas, G. (1993). *New Experimental Setup (NESU)*. Unpublished manuscript, Max-Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Bever, T. G., & McElree, B. (1988). Empty categories access their antecedents during comprehension. *Linguistic Inquiry*, *19*, 35–43.

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

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

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step* (pp. 89–155). Cambridge, MA: MIT Press.

Clahsen, H., & Featherston, S. (1999). Antecedent priming at trace positions: Evidence from German scrambling. *Journal of Psycholinguistic Research*, *28*, 415–437.

Clifton, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. M. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273–317). Dordrecht: Kluwer.

Crocker, M. (1996). *Computational psycholinguistics*. Dordrecht: Kluwer.

Featherston, S. (2001). *Empty categories in sentence processing*. Amsterdam: Benjamins.

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, *24*, 453–489.

Fiebach, C. J., Schlesewsky, M., & 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*, 250–272.

Fodor, J. D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry*, *9*, 427–473.

Fodor, J. D. (1995). Comprehending sentence structure. In L. R. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science: Vol. 1. Language* (2nd ed., pp. 209–246). Cambridge, MA: MIT Press.

Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, *4*, 93–126.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217–236). New York: Elsevier.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology*, *50*, 119–148.

Friederici, A. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84.

Garnsey, S., Pearlmutter, N., Myers, E., & Lotocky, M. (1997). The contribution of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93.

Gazdar, G., Klein, E., Pullum, G., & Sag, I. (1985). *Generalized phrase structure grammar*. Oxford: Blackwell.

Gibson, E. (1998). Syntactic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–75.

Gibson, E., & Warren, T. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax*, *7*, 55–78.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*, 251–266.

Hahne, A., & Friederici, A. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*, 123–141.

Hawkins, R. (2001). *Second language syntax*. Oxford: Blackwell.

Isel, F. (2002, March). Auditory processing of German sentences by French late bilinguals: Neurodynamics of syntactic and semantic processes. Paper presented at the 24th annual meeting of the German Linguistics Society (DGfS), Mannheim, Germany.

Juffs, A. (1998). Main verb versus reduced relative clause ambiguity resolution in second language sentence processing. *Language Learning*, *48*, 107–147.

Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in *wh*-extraction. *Studies in Second Language Acquisition*, *17*, 483–516.

Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language sentence processing. *Language Learning*, *46*, 283–326.

Just, M., Carpenter, P., & Woolley, J. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*, 228–238.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

King, J. W., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*, 580–602.
King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, *7*, 376–395.
Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*, 196–214.
Lee, M.-W. (2004). Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research*, *33*, 51–73.
Levine, R., & Hukari, T. (2004). *The unity of unbounded dependencies*. Stanford, CA: CSLI Publications.
Love, T., & Swinney, D. (1996). Coreference processing and levels of analysis in object-relative constructions: Demonstration of antecedent reactivation with the cross-modal priming paradigm. *Journal of Psycholinguistic Research*, *25*, 5–24.
Nakano, Y., Felser, C., & Clahsen, H. (2002). Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research*, *31*, 531–571.
Nicol, J. L. (1993). Reconsidering reactivation. In G. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The second Sperlonga meeting* (pp. 321–350). Mahwah, NJ: Erlbaum.
Nicol, J. L., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, *18*, 5–20.
Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, *25*, 501–528.
Pickering, M. (1999). Sentence comprehension. In S. Garrod & M. Pickering (Eds.), *Language processing* (pp. 123–153). Hove, East Sussex, UK: Psychology Press.
Pickering, M., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, *6*, 229–259.
Pickering, M., Barton, S., & Shillcock, R. (1994). Unbounded dependencies, island constraints, and processing complexity. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 199–224). Mahwah, NJ: Erlbaum.
Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
Roberts, L. (2003). *Syntactic processing in learners of English*. Unpublished doctoral dissertation, University of Essex, Colchester, United Kingdom.
Sag, I. A., & Fodor, J. D. (1995). Extraction without traces. In R. Aranovich, W. Byrne, S. Preuss, & M. Senturia (Eds.), *Proceedings of the 13th annual meeting of the West Coast Conference on Formal Linguistics* (pp. 365–384). Stanford, CA: CSLI Publications.
Stowe, L. (1986). Parsing *wh*-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, *1*, 227–245.
White, L., & Juffs, A. (1998). Constraints on *wh*-movement in two different contexts of non-native language acquisition: Competence and processing. In S. Flynn, G. Martohardjono, & W. O'Neil (Eds.), *The generative study of second language acquisition* (pp. 111–130). Mahwah, NJ: Erlbaum.
Watanabe, A. (2001). *Wh*-in-situ languages. In M. Baltin & C. Collins (Eds.), *The handbook of contemporary syntactic theory* (pp. 203–225). Oxford: Blackwell.
Williams, J., Möbius, P., & Kim, C. (2001). Native and non-native processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics*, *22*, 509–540.

APPENDIX A

SENTENCES USED IN THE QUESTIONNAIRE

1. The manager thought the secretary claimed that the new salesman had pleased the boss in the meeting.
2. The student who the headmaster's thoughts about the clever teacher had surprised does not usually do any homework.
3. The nurse who the doctor argued that the rude patient had angered is refusing to work late.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

4. The witness said the lawyer's proof about the evil criminal had confused the judge in court at the trial.
5. The actress who the journalist's suggestion about the talented writer had inspired will accept the role in the new play.
6. The customer thought the receptionist stated that the lazy cleaner had annoyed the manager in the hotel that morning.
7. The farmer said the builder's thoughts about the dedicated worker had amazed the boss last week at work.
8. The singer who the musician stated that the drunken guitarist had offended does not want to perform the concert this evening.
9. The schoolboy said the teacher's proof about the aggressive child had distressed the class at school last week.
10. The girl who the policeman concluded that the nasty boy had frightened has stopped going to school.
11. The coach who the manager's decision about the violent footballer had annoyed will cancel the match next week.
12. The politician thought the minister stated that the TV journalist had upset the president on the talk show.
13. The chef who the cook argued that the head waitress had bothered wants to find a new job.
14. The director said the agent's suggestion about the unpleasant dancer had disappointed the other members of the ballet.
15. The film star said the interviewer suggested that the horrible photographer had embarrassed the editor of the newspaper.
16. The man who the customer's thoughts about the shop assistant had amused was trying not to laugh.
17. The therapist said the patient dreamed that the strange woman had fascinated the members of the group.
18. The man who the detective's conclusion about the dangerous thief had distressed will buy a new alarm for his house.
19. The captain who the officer decided that the young soldier had displeased will write a formal report next week.
20. The tourist believed the guide's claim about the hotel manager had angered everybody on the holiday.

APPENDIX B

EXPERIMENTAL SENTENCES USED IN THE SELF-PACED READING EXPERIMENT

1a. The manager who the secretary claimed that the new salesman had pleased will raise company salaries.
b. The manager who the secretary's claim about the new salesman had pleased will raise company salaries.
c. The manager thought the secretary claimed that the new salesman had pleased the boss in the meeting.
d. The manager thought the secretary's claim about the new salesman had pleased the boss in the meeting.
2a. The student who the headmaster thought that the clever teacher had surprised does not like doing homework.
b. The student who the headmaster's thoughts about the clever teacher had surprised does not like doing homework.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

c. The student believed the headmaster thought that the clever teacher had surprised everybody at school last week.
d. The student believed the headmaster's thoughts about the clever teacher had surprised everybody at school last week.

3a. The nurse who the doctor argued that the rude patient had angered is refusing to work late.
b. The nurse who the doctor's argument about the rude patient had angered is refusing to work late.
c. The nurse thought the doctor argued that the rude patient had angered the staff at the hospital.
d. The nurse thought the doctor's argument about the rude patient had angered the staff at the hospital.

4a. The witness who the lawyer proved that the evil criminal had confused does not want to testify.
b. The witness who the lawyer's proof about the evil criminal had confused does not want to testify.
c. The witness said the lawyer proved that the evil criminal had confused the judge during the trial.
d. The witness said the lawyer's proof about the evil criminal had confused the judge during the trial.

5a. The actress who the journalist suggested that the talented writer had inspired will go on stage tonight.
b. The actress who the journalist's suggestion about the talented writer had inspired will go on stage tonight.
c. The actress thought the journalist suggested that the talented writer had inspired everybody with the new play.
d. The actress thought the journalist's suggestion about the talented writer had inspired everybody with the new play.

6a. The customer who the receptionist stated that the lazy cleaner had annoyed will not pay his bill.
b. The customer who the receptionist's statement about the lazy cleaner had annoyed will not pay his bill.
c. The customer thought the receptionist stated that the lazy cleaner had annoyed the manager of the hotel.
d. The customer thought the receptionist's statement about the lazy cleaner had annoyed the manager of the hotel.

7a. The farmer who the builder thought that the dedicated worker had amazed will give everybody extra money.
b. The farmer who the builder's thoughts about the dedicated worker had amazed will give everybody extra money.
c. The farmer said the builder thought that the dedicated worker had amazed the new boss last week.
d. The farmer said the builder's thoughts about the dedicated worker had amazed the new boss last week.

8a. The singer who the musician stated that the drunken guitarist had offended will not perform this evening.
b. The singer who the musician's statement about the drunken guitarist had offended will not perform this evening.
c. The singer thought the musician stated that the drunken guitarist had offended the drummer after the performance.
d. The singer thought the musician's statement about the drunken guitarist had offended the drummer after the performance.

9a. The schoolboy who the teacher proved that the aggressive child had distressed will complain at the meeting.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

b. The schoolboy who the teacher's proof about the aggressive child had distressed will complain at the meeting.
c. The schoolboy said the teacher proved that the aggressive child had distressed the class at school yesterday.
d. The schoolboy said the teacher's proof about the aggressive child had distressed the class at school yesterday.

10a. The girl who the policeman concluded that the nasty boy had frightened has stopped going to school.
b. The girl who the policeman's conclusion about the nasty boy had frightened has stopped going to school.
c. The girl said the policeman concluded that the nasty boy had frightened the children at the school.
d. The girl said the policeman's conclusion about the nasty boy had frightened the children at the school.

11a. The coach who the manager decided that the violent footballer had annoyed will cancel the match today.
b. The coach who the manager's decision about the violent footballer had annoyed will cancel the match today.
c. The coach said the manager decided that the violent footballer had annoyed his fans at the match.
d. The coach said the manager's decision about the violent footballer had annoyed his fans at the match.

12a. The politician who the minister stated that the TV journalist had upset will not give an interview.
b. The politician who the minister's statement about the TV journalist had upset will not give an interview.
c. The politician thought the minister stated that the TV journalist had upset the president on the program.
d. The politician thought the minister's statement about the TV journalist had upset the president on the program.

13a. The chef who the cook argued that the head waitress had bothered wants to find another job.
b. The chef who the cook's argument about the head waitress had bothered wants to find another job.
c. The chef said the cook argued that the head waitress had bothered the manager of the restaurant.
d. The chef said the cook's argument about the head waitress had bothered the manager of the restaurant.

14a. The director who the agent suggested that the unpleasant dancer had disappointed will cancel the performance tonight.
b. The director who the agent's suggestion about the unpleasant dancer had disappointed will cancel the performance tonight.
c. The director said the agent suggested that the unpleasant dancer had disappointed the others in the ballet.
d. The director said the agent's suggestion about the unpleasant dancer had disappointed the others in the ballet.

15a. The film star who the interviewer suggested that the horrible photographer had embarrassed will not answer any questions.
b. The film star who the interviewer's suggestion about the horrible photographer had embarrassed will not answer any questions.
c. The film star said the interviewer suggested that the horrible photographer had embarrassed the editor of the newspaper.
d. The film star said the interviewer's suggestion about the horrible photographer had embarrassed the editor of the newspaper.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press

16a. The man who the customer thought that the shop assistant had amused was trying not to laugh.
b. The man who the customer's thoughts about the shop assistant had amused was trying not to laugh.
c. The man believed the customer thought that the shop assistant had amused everybody in the store yesterday.
d. The man believed the customer's thoughts about the shop assistant had amused everybody in the store yesterday.

17a. The therapist who the patient dreamed that the strange woman had fascinated is writing a new book.
b. The therapist who the patient's dream about the strange woman had fascinated is writing a new book.
c. The therapist said the patient dreamed that the strange woman had fascinated the members of the group.
d. The therapist said the patient's dream about the strange woman had fascinated the members of the group.

18a. The man who the detective concluded that the dangerous thief had distressed will buy a new alarm.
b. The man who the detective's conclusion about the dangerous thief had distressed will buy a new alarm.
c. The man thought the detective concluded that the dangerous thief had distressed the people in the neighborhood.
d. The man thought the detective's conclusion about the dangerous thief had distressed the people in the neighborhood.

19a. The captain who the officer decided that the young soldier had displeased will write a formal report.
b. The captain who the officer's decision about the young soldier had displeased will write a formal report.
c. The captain said the officer decided that the young soldier had displeased the colonel at training today.
d. The captain said the officer's decision about the young soldier had displeased the colonel at training today.

20a. The tourist who the guide claimed that the hotel manager had angered wants to return home now.
b. The tourist who the guide's claim about the hotel manager had angered wants to return home now.
c. The tourist believed the guide claimed that the hotel manager had angered everybody in the holiday party.
d. The tourist believed the guide's claim about the hotel manager had angered everybody in the holiday party.

https://doi.org/10.1017/S0272263105050035 Published online by Cambridge University Press