

Gaps in Second Language Sentence Processing

**Theodore Marinis, Leah Roberts, Claudia Felser
and Harald Clahsen¹**

In: Essex Research Reports in Linguistics 45 (2003) 43-79.

Corresponding Author:

Claudia Felser
Department of Language and Linguistics
University of Essex
Colchester CO4 3SQ
United Kingdom

Phone: +44-1206-872113
Fax: +44-1206-873598
E-mail: *felsec@essex.ac.uk*

Abstract

Four groups of second language (L2) learners of English from different language backgrounds (Chinese, Japanese, German & Greek) and a group of native speaker controls participated in an on-line reading-time experiment with sentences involving long-distance *wh*-dependencies. While 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 subcategoriser, regardless of whether or not the subjacency constraint was operative in their native language. This finding is argued to support the hypothesis that L2 learners under-use syntactic information in L2 processing, which prevents them from processing the L2 input in a native-like fashion.

1. Introduction

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 but instead must be retained in short-time memory until it can be linked to its subcategoriser, 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 analysed. 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 subcategoriser. According to the copy theory of movement (Chomsky 1995, and later), the empty category (= e_i in example [1] below) involved in filler-gap dependencies is a silent but otherwise identical copy of the displaced constituent itself.

(1) *Which book* _{i} did you read e_i in only an hour?

Some lexically-based syntactic frameworks including variants of Head-Driven Phrase Structure Grammar, on the other hand, assume that a dislocated element is linked *directly* to its lexical subcategoriser (Pollard & Sag, 1994). This linguistic controversy has given rise to different hypotheses as to how filler-gap dependencies are processed, the *Trace Reactivation Hypothesis* (TRH), according to which the human parser postulates empty categories ('traces') during the on-line comprehension of sentences

containing such dependencies (Bever & McElree 1988; Love & Swinney, 1996; Nicol & Swinney, 1989; Swinney, Ford, Frauenfelder & Bresnan, 1988, among others), and the *Direct Association Hypothesis* (DAH), which maintains that establishing a filler-gap dependency is a lexically-driven process triggered by the automatic mental reconstruction of the subcategoriser's argument structure when this is encountered (Pickering & Barry, 1991; Sag & Fodor, 1994).

Results from a number of studies on monolingual sentence comprehension suggest that two distinct mental processes may in fact be involved in the processing of filler-gap dependencies: (i) a phrase structure-based mechanism that triggers a filler's retrieval 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 (ii) a lexically-driven process of semantically integrating a displaced constituent with its thematic role assigner or other licenser, as predicted by the DAH. Whereas these two processes are usually difficult to dissociate empirically in head-initial languages like English (but see Nicol, 1993), evidence for the TRH can be gathered from studies on the processing of filler-gap dependencies 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 subcategorising verb had been encountered.

Regardless of whether or not a filler is assumed to be linked to its lexical subcategoriser via empty categories located within the subcategoriser's extended projection, though, most contemporary syntactic theories agree that for dependencies spanning more than one clause, some kind of intermediate linguistic structure is present at intervening clause boundaries which mediates between the filler and its ultimate gap (or subcategoriser). An example of what is commonly referred to as 'successive-cyclic *wh*-movement' is provided in (2) below.

(2) Who_i do you think e_i (that) John says e_i (that) Mary likes e_i ?

Traditional evidence for the successive-cyclic nature of *wh*-movement includes various types of 'island' effect (Ross, 1967), *wh*-complementiser agreement in languages like Irish (McCloskey, 2001), children's use of medial *wh* in questions such as *Who do you*

think who's in the box? (Thornton, 1990), and *wh*-copying found in a number of languages including German, Frisian, Afrikaans, and Romani (see Felser, in press, and references cited there). Psycholinguistic evidence for successive-cyclicity has been found, for example, in a study by Kluender & Kutas (1993) using event-related brain potentials (ERPs), and in a reading-time study by Gibson & Warren (1999). Kluender & Kutas observed that the processing difficulty for sentences containing subjacency violations such as (3) below increased (relative to sentences in which subjacency was respected) 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 & Warren (1999) investigated native English speakers' processing of grammatical sentences containing long-distance *wh*-dependencies like that in sentence (4) below.

- (4) The manager *who_i* the consultant claimed *e_i* that the new proposal had pleased *e_i* will hire five workers tomorrow.

Similar to Kluender & Kutas (1993), the authors found that the availability of an intermediate 'landing site' facilitated a filler's integration with its subcategoriser, thus providing indirect evidence for the psychological reality of intermediate gaps in L1 sentence processing. Gibson & Warren's reading-time study provided the model for the present study, and will be discussed in more detail in section 3.1 below.

While there is ample evidence that the mental representations constructed during L1 sentence processing are built up rapidly and in an incremental fashion, and also include abstract linguistic structure such as empty categories, or syntactic gaps, surprisingly little is known to date about the way second language learners process the L2 input in real time. Instead, L2 research has traditionally focused on the acquisition of grammatical knowledge using off-line methodologies such as grammaticality judgement, elicitation techniques, or comprehension tasks. Previous studies of L1 sentence processing in a range of different languages have shown, however, that between-language variation is not restricted to differences in grammar, but that some

processing strategies may also be subject to cross-linguistic variation (Cuetos, Mitchell & Corley, 1996; Frazier & Rayner, 1988; Gibson, Pearlmutter, Canseco-Gonzalez & Hickok, 1996; Mazuka & Lust, 1990, among others). Hence, besides being faced with the task of acquiring the L2 grammar, L2 learners may also need to acquire any language-specific processing strategies that are used in the target language. The observation that sentence processing is not necessarily uniform across languages also raises the possibility of L1 processing transfer in L2 acquisition, an issue that has featured prominently in much research within the framework of the Competition Model of language acquisition and processing (Harrington, 1987; MacWhinney, 1997, 2002). It is conceivable, for example, that L2 learners from *wh*-in-situ backgrounds fail to process *wh*-dependencies in L2 English in a native-like way, whereas L2 learners whose L1 also shows overt *wh*-movement are indistinguishable from native speakers in this domain.

Our previous studies of L2 processing indicate that although L2 learners, like native speakers, are guided by lexical information during parsing, they rely on phrase-structure information to a lesser extent than native speakers do - irrespectively of their language background (Felser, Roberts, Gross & Marinis, 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.

2. Previous studies of L2 learners' processing of *wh*-dependencies

The vast majority of existing L2 studies on the acquisition of *wh*-movement and subjacency have used off-line tasks such as grammaticality judgements, and their results are not fully conclusive.³ Only a few published studies are available to date that have examined the real-time processing of *wh*-movement by L2 learners using on-line tasks. A reading-time study carried out by Juffs & Harrington (1995) addressed the issue of whether it is processing difficulties or a competence deficit that causes problems with certain types of filler-gap dependencies for learners of English whose native language does not show successive-cyclic *wh*-movement and thus arguably lacks the subjacency constraint. Juffs & Harrington report the results from two on-line grammaticality judgement experiments that measured Chinese-speaking learners' accuracy and reading times for grammatical and ungrammatical sentences involving either subject and object

extractions. The results from the full-sentence presentation version of the experiment showed that the Chinese-speaking learners' response accuracy was comparable to the native speakers' for ungrammatical subject and object extractions, indicating that they had acquired the subjacency constraint. They performed significantly worse than the native speakers, however, on grammatical sentences involving subject - but not object - extraction (compare also White & Juffs, 1998). The learners' difficulties with subject extractions were also reflected in their on-line reading times. In the experiment using word-by-word presentation, the two participant groups showed distinct patterns of processing the grammatical sentences. Specifically, the learners were found to slow down significantly more at the region following the matrix verb in subject extractions from finite clauses such as (5a) below than in object extractions, as in (5b). No such slowdown was attested in the group of native speaker controls.

- (5) 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 argue that the learners' relatively poorer performance on subject extractions reflects processing rather than competence problems (cf. Juffs & Harrington, 1996). Observe that in sentences like (5a) above, the gap following the verb *say* may initially be analysed as an object gap, a decision that must be revised as soon as the embedded verb *likes* is encountered. While this kind of reanalysis causes no or little processing difficulty for native speakers, it does, according to Juffs & Harrington, pose a problem for L2 learners.

Note, however, that given the nature of Juffs & Harrington's 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 subcategorising verb, the slowdown observed in the post-gap region may also be due to the learners' trying to link the fronted *wh*-phrase directly to its subcategoriser, 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 native speakers) also showed elevated reading times at the matrix verb *say*, a region prior to the locus of reanalysis. Juffs & Harrington (1996, p.300) speculate 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.

Another reading-time study by Williams, Möbius & Kim (2001) investigated so-called 'filled-gap' effects in L2 processing, and the question of whether or not L2 learners are sensitive to plausibility constraints during parsing. Their experimental sentences involved adjunct extractions in two plausibility conditions, as shown in (6a) and (6b) below.

- (6) 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 example (6a), the fronted *wh*-phrase is a plausible object of the verb *hide*, whereas in example (6b) it is not. Previous studies have shown that native speakers of English initially attempt to analyse 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 (compare e.g. Stowe, 1986). In Williams et al.'s self-paced reading experiment, Chinese, Korean, and German-speaking learners of English were asked to read sentences presented on a computer screen in a word-by-word fashion, and to indicate the point at which they thought the sentence had become implausible by pressing a 'stop' button. Assuming that on-line 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 analysed as the object of the verb when this is encountered. A filled-gap effect would then be observed on the post-verbal NP, reflected in longer reading times, due to the need for reanalysis at this point. If, on the other hand, a gap is posited only as a last-resort strategy (that is, to avoid ungrammaticality; compare Fodor, 1978), then no such slowdown would be expected at the post-verbal NP.

In the 'stop-making-sense' task, the learners behaved similarly to the native speakers. All but the Chinese-speaking participants made more 'stop' decisions at and immediately after the verb in the 'Implausible-at-V' condition than in the corresponding plausible condition, suggesting that both the learners and the native speakers were sensitive to plausibility information. The analysis of the reading time data showed that for all participant groups, the post-verbal noun in the 'Plausible-at-V' condition elicited longer reading times compared to the post-verbal noun in the 'Implausible-at-V'

condition. This indicates that both the native speakers and the learners analysed 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 background did not appear to have any effect on how they processed the experimental sentences. Only the native speakers showed an effect of plausibility at the determiner introducing the post-verbal NP, however. According to the authors, the earlier onset of the filled-gap effect observed in the native group may indicate a greater sensitivity to the syntactic cue provided by the determiner, which signalled an incoming NP.

Williams et al.'s on-line experiment was complemented by an off-line acceptability judgement task to investigate the different learner groups' ability to recover from misanalysis. The results showed that the learners but not the native speakers judged the 'Plausible-at-V' sentences unacceptable significantly more often than the 'Implausible-at-V' ones. Similarly to Juffs & Harrington (1995, 1996), the authors conclude that the learners have more difficulty than native speakers recovering from an initial misanalysis, particularly when this analysis is plausible, suggesting an over-commitment to a strongly plausible first analysis.

Summarising, Williams et al.'s results suggest that L2 learners, like native speakers, employ a filler-driven parsing strategy when processing *wh*-dependencies, irrespective of their language background. 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 off-line 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. Observe further that like the results from Juffs & Harrington's (1995) study, Williams et al.'s results do not bear directly on the question of whether or not 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 point out themselves, the filled-gap effect observed on the post-verbal noun in the non-native 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.

3. The present study

3.1 Background

Our study was modelled after Gibson & Warren's (1999) study on the processing of long *wh*-dependencies by adult native speakers of English. Using a self-paced reading task, Gibson & Warren (hereafter, G&W) investigated how native speakers process sentences such as (7a) and (7b) below.

- (7) 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 (7) above differ in that (7a) but not (7b) provides an intermediate landing site for the fronted *wh*-pronoun. This is because in (7a), *wh*-movement has crossed a clause boundary that signals the beginning of a new cyclic domain, whereas (7b) involves extraction across a noun phrase. 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. In order to control for a possible confounding effect of subject-verb distance, G&W's materials also included sentences of the following types, which did not involve any *wh*-movement but which differed in the relative distance between the verb *pleased* and the head of its subject (*viz.* *proposal* in [8a], and *claim* in [8b]).

- (8) 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.

The authors found an interaction between extraction and intervening phrase type at the region containing the *wh*-filler's subcategoriser *pleased*. Reading times were shorter for sentences such as (8a) that provided an intermediate landing site than for sentences such as (8b), an effect that was not present in the non-extraction conditions and thus cannot be attributed to any differences in subject-verb distance between the VP and NP

conditions. Furthermore, the reading times elicited by the complementiser *that* in (7a) were found to be longer than in the corresponding non-extraction condition (8a), although the interaction between extraction and intervening phrase type did not reach significance here.

The 'intermediate gap' effect observed by G&W supports a strong version of the Active Filler Hypothesis according to which a filler is reactivated cyclically so as to break up long dependencies into a series of shorter ones (compare Crocker, 1996; Frazier & Clifton, 1989). Note, however, that there was an asymmetry in G&W's experimental materials between the extraction and non-extraction conditions in that the extraction conditions contained more words than the non-extraction conditions, and one additional level of embedding before the critical segments. This asymmetry may have introduced a confound such that lower reading times in the non-extraction conditions might have been partly due to the differences in length and/or structural complexity between the extraction and non-extraction conditions.

The present study has two major aims: (i) to replicate G&W's finding with native speakers of English using improved materials, and (ii) to investigate whether L2 learners of English from different language backgrounds process long *wh*-dependencies in the same way, or differently from, native speakers. To test whether the learners' L1 background has an effect on their processing of long *wh*-dependencies in L2 English, we examined learners from both *wh*-movement (Greek, German) and *wh*-in-situ backgrounds (Chinese, Japanese).

3.2 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), as well as a group of 24 native English-speaking controls (mean age = 24, range = 19-34). The participants were recruited from among the undergraduate and postgraduate student communities at the University of Essex and were paid a small fee for their participation. All participants had normal or corrected-to-normal vision, and were naïve with respect to the purpose of the experiment.

The Chinese-speaking learners were all native speakers of Mandarin Chinese. All learners had first been exposed to English around the age of 11 in a classroom setting, and none of them considered themselves bilingual. Table 1 provides an overview of the learners' age at the time of testing, their age of first exposure to English, and the time the participants had spent in the UK at the time of testing.

Table 1: Summary of the learners' bio-data

Learners' Bio-data		Age	Age of First Exposure to English	Time Spent in UK
Chinese	Mean (yrs)	25.06	11.94	0.85
	SD	3.92	2.17	1.38
	Range	17-33	6-15	0.1-6
Japanese	Mean (yrs)	26.54	11.77	2.15
	SD	4.21	0.91	1.58
	Range	20-40	10-13	0.5-6.6
German	Mean (yrs)	24	11.36	1.6
	SD	6.22	1.87	1.09
	Range	19-46	8-16	0.3-5
Greek	Mean (yrs)	24.80	8.67	2.48
	SD	3.11	2.19	2.20
	Range	20-37	5-13	0.1-8

To determine the learners' general proficiency in English at the time of testing, all of them underwent a standardised 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., learners scoring 145/200 points or above) were included in our study. In addition to the OPT, the learners also completed an off-line questionnaire, the purpose of which was to ensure that they were able to comprehend complex sentences of the kind that were later used in the on-line

task. The questionnaire consisted of 20 sentences that were similar but not identical to the sentences used in the self-paced reading task. There were five sentences corresponding to each of the four experimental conditions in the on-line experiment, as described below in the *Materials* section. Each sentence was followed by a comprehension question and three choices, as illustrated by (9) below (for the full set of questionnaire materials, see Appendix A).

- (9) 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 2 presents a summary of the participant's scores in the OPT and in the off-line questionnaire.

Table 2: Learners' Oxford Placement Test and Questionnaire scores

<i>Pre-test Scores</i>		<i>Oxford Placement Test</i>	<i>Questionnaire</i>
<i>Chinese</i>	<i>Mean</i>	156.35	92.5%
	<i>SD</i>	7.16	6.75
<i>Japanese</i>	<i>Mean</i>	169.15	92.31%
	<i>SD</i>	11.52	6.96
<i>German</i>	<i>Mean</i>	176.84	98%
	<i>SD</i>	12.44	4.33
<i>Greek</i>	<i>Mean</i>	172.50	96.17%
	<i>SD</i>	9.90	4.29

All participants scored at 75% or above correct in the Questionnaire, suggesting that they could handle the types of sentences used in the self-paced reading experiment in an off-line task.

Materials

The materials for the on-line task comprised a total of 88 sentences, including 8 practice items, 20 experimental sentences, and 60 filler sentences. The reason of including such a large number of fillers was to prevent the participants from being able to guess the purpose of the experiment, and to keep them from developing any response strategies. Each of the experimental sentences came in four versions in a 2x2 design with the conditions (+/-) Extraction crossed by (VP/NP) Phrase Type, as illustrated by (10a-d) below (the full set of experimental sentences is provided in Appendix B).

(10)a. **EXTRACTION, VP**

The nurse *who_i* the doctor argued *e'_i* that the rude patient had angered *e_i* is refusing to work late.

b. **EXTRACTION, NP**

The nurse *who_i* the doctor's argument about the rude patient had angered *e_i* is refusing to work late.

c. **NON-EXTRACTION, VP**

The nurse thought the doctor argued that the rude patient had angered the staff at the hospital.

d. **NON-EXTRACTION, 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 G&W. 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 misanalyses 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 towards taking a sentential complement. Six verbs (*claim, argue, prove, suggest, conclude, and decide*) were selected from Garnsey, Pearlmutter, Myers & Lotocky's (1997) list of sentential complement verbs, and a further six were tested independently for their complement bias. To this end, ten native speakers of English were given both a free and a forced-choice sentence completion task, 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 non-extraction conditions differed from G&W's in that they contained exactly the same number of words (up to the embedded verb) as the sentences in the corresponding extraction conditions. By way of avoiding any asymmetry between the extraction and corresponding non-extraction conditions with respect to the degree of structural complexity, we added a further level of embedding to the sentences in the non-extraction 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 pseudo-randomised and mixed with the filler sentences, and were 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 their contents.

Procedure

The pretests (OPT & Questionnaire) and the on-line task were administered in two separate sessions, with approximately one week in between. Reading-time and comprehension accuracy data were collected using the non-cumulative 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 NESU software package (Baumann, Nagengast & Klaas, 1993). The stimulus sentences were presented in a segment-by-segment fashion, in white letters (Arial 24pt) on a black background in the centre of a 17" monitor. The experimental sentences were divided into six segments as indicated in example (11) below.

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

Participants were instructed to read each segment as quickly as possible for comprehension and then to press a pacing button as soon as they were ready to receive the next segment. In this way, a step-by-step record can be obtained of the parse as it unfolds. The underlying rationale is that increased reaction times to a specific segment (relative to the corresponding segment in a control condition) indicate a relatively higher processing difficulty at this region of the sentence.

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 of 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 at the left and one on the right-hand side. For half of the questions the correct answer was the one on the left-hand side of the screen, and for the other half the one on the right-hand side was the correct one. 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 thought was the correct one. After the end of each trial, a message appeared on the screen instructing participants to press a dedicated key on the keyboard in order to trigger the start of the next trial. All participants completed the on-line task in approximately 30 minutes.

4. Results

4.1 Accuracy

All participant groups scored highly in answering the comprehension questions that followed the experimental sentences. The native speakers' 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. non-extraction) and Phrase Type (VP vs. NP) as within-subjects factors and Language (English, Chinese, Japanese, German, Greek) as 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 native speakers with respect to their ability to comprehend the experimental sentences. Moreover, neither the presence vs. absence of a filler-gap dependency nor the type of the intervening phrase (VP vs. NP) appeared to have influenced the participants' accuracy scores.

4.2 Reading times

The relevant segments for determining whether or not intermediate syntactic gaps were postulated during processing are segments 3 and 5. Segment 3 in the Extraction-VP condition (10a) contains the complementiser *that*, which is hypothesized to trigger the reactivation of the filler *who* at this position, whereas in the corresponding non-extraction condition (10c) the complementiser merely indicates the beginning of an embedded clause. Thus if the participants postulate an intermediate gap at this point during processing in sentences such as (10a), we expect segment 3 to elicit longer RTs in the Extraction-VP condition than in the Non-Extraction-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 subcategoriser 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 G&W, 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 *wh*-dependencies in the same way as native speakers, we expect to find no statistical differences between the five participant groups. If, however, properties of the first language have an impact on the way long *wh*-dependencies are processed in L2 English, then 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 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 native speakers and the learners, but not among the individual learner groups.

Data trimming

Following standard practise in this type of experiment, we included only reading times (RTs) from correctly answered trials in the statistical analysis. Prior to the analysis of the data, responses were screened for trials whose total reading time exceeded a time-out of 20,000 ms for the native speakers and 25,000 for the learners.⁴ This affected 0.79% of the data from the English group, 1.67% of data from the Chinese group, 1.04% of the data from the Japanese group, 0.47% of the data from the German group, and 1.04% of data from the Greek group. In addition, we screened the participants' RTs to each segment for outliers, and eliminated individual data points beyond 2 SD from the mean RTs for each condition per subject and item. This affected 3.86% of the data from the native speakers, 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 from one Chinese and two German-speaking participants were excluded because their data sets were incomplete. The remaining data from 32 Chinese, 26 Japanese, 22 German and 30 Greek-speaking learners of English, and from 24 native speaker controls were included in the statistical analysis.

Between-groups analyses

Table 3 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. non-extraction) 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.

At Segment 3, we found a main effect of Extraction ($F(1, 129) = 8.412, p < 0.01$; $F(1, 94) = 6.321, p < 0.05$), a main effect of Phrase Type in the items analysis that was approaching significance in the subjects analysis ($F(1, 129) = 3.081, p = 0.085$; $F(1, 94) = 8.814, p < 0.01$), as well as a main effect of Language ($F(1, 129) = 2.443, p = 0.05$; $F(4, 94) = 8.082, p < 0.001$). The analysis also revealed an interaction of Phrase

Table 3: Mean RTs (in milliseconds) per segment and condition

		1	2	3	4	5	6
English	Extraction-VP	753	1069	825	1268	1075	1359
	Extraction-NP	717	1507	833	1366	1307	1343
	Non-extraction-VP	712	1195	729	1237	811	985
	Non-extraction-NP	718	1099	657	1066	820	1073
Chinese	Extraction-VP	1585	2665	1062	2155	1635	3065
	Extraction-NP	1434	3227	814	2492	1838	2822
	Non-extraction-VP	1723	2691	836	1963	1373	1876
	Non-extraction-NP	1783	2329	857	1676	1551	1996
Japanese	Extraction-VP	1547	2652	956	2053	1473	2994
	Extraction-NP	1507	3052	1126	1972	1874	2859
	Non-extraction-VP	1470	2512	955	1675	1314	1929
	Non-extraction-NP	1590	2282	918	1643	1358	1833
German	Extraction-VP	844	1562	977	1628	1609	1599
	Extraction-NP	845	1801	935	1351	1374	1343
	Non-extraction-VP	958	1659	925	1196	959	1273
	Non-extraction-NP	968	1533	753	1265	925	1214
Greek	Extraction-VP	851	1636	838	1757	1167	1786
	Extraction-NP	909	2170	837	2022	1217	1682
	Non-extraction-VP	1173	2015	875	1634	1004	1209
	Non-extraction-NP	1288	2052	664	1417	945	1262

Type and Language that was significant in the items analysis ($F(4, 129) = 0.718, p > 0.1$; $F(4, 94) = 2.534, p < 0.05$). Additionally, a three-way interaction between Extraction, Phrase Type and Language was found to be approaching significance ($F(4, 129) = 2.418, p = 0.05$; $F(4, 94) = 2.181, p = 0.077$). These results suggest that there were differences among the five participant groups with respect to how they processed segment 3.

The analysis of segment 5 revealed a main effect of Extraction ($F(1, 130) = 49.011, p < 0.001$; $F(1, 94) = 56.882, p < 0.001$), a main effect of Language ($F(4, 130) = 4.130, p < 0.001$; $F(4, 94) = 30.554, p < 0.001$), and an interaction of Phrase Type and

Language in the items analysis ($F(4, 130) = 2.004, p = 0.098$; $F(4, 94) = 3.088, p < 0.05$). This interaction indicates that there are differences between the language groups with respect to Phrase Type. As the above results showed interactions with the factor Language at both segments 3 and 5, we went on to analyse the data from each of the five participant groups separately.

Native speakers

We performed separate two-way ANOVAs for segments 3 and 5 with the factors Extraction and Phrase Type. The analysis of the native speakers' RTs to segment 3 revealed a main effect of Extraction ($F(1, 23) = 4.578, p < 0.05$; $F(1, 19) = 9.672, p < 0.01$), reflecting the fact that RTs for the extraction conditions were significantly slower than those for the non-extraction conditions, as well as a main effect of Phrase Type in the items analysis ($F(1, 23) = 0.844, p > 0.1$; $F(1, 19) = 4.967, p < 0.05$). Like G&W, we found no significant interaction between Extraction and Phrase Type, however.

At Segment 5, the two extraction conditions again produced longer reading times than the non-extraction conditions, but with the Extraction-VP condition being read more quickly than the Extraction-NP condition. The ANOVA for this segment showed a main effect of Extraction ($F(1, 23) = 11.054, p < 0.01$; $F(1, 19) = 20.355, p < 0.001$), a main effect of Phrase Type ($F(1, 23) = 4.759, p < 0.05$; $F(1, 19) = 4.715, p < 0.05$) and an interaction between Extraction and Phrase Type ($F(1, 23) = 4.994, p < 0.05$; $F(1, 19) = 4.364, p = 0.05$). Subsequent pairwise comparisons revealed significant differences between the Extraction-VP and Non-Extraction-VP conditions ($t(23) = 2.560, p < 0.05$; $t(19) = 2.389, p < 0.05$) and between the Extraction-NP and Non-Extraction-NP conditions ($t(23) = 3.551, p < 0.01$; $t(19) = 4.322, p < 0.001$), which reflect the additional processing cost associated with integrating a filler with its subcategoriser in the extraction conditions. Crucially, the difference between the Extraction-VP and Extraction-NP conditions also proved significant ($t(23) = 2.441, p < 0.05$; $t(19) = 2.220, p < 0.05$), suggesting that filler integration was comparatively less costly in the Extraction-VP condition. The fact that the RTs for the two non-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 native speakers essentially replicate G&W's findings. Like G&W, we observed elevated reading times at the intervening clause boundary in the

Extraction-VP condition compared to the corresponding non-extraction condition, and shorter RTs to the segment containing the filler's subcategoriser for the Extraction-VP than for the Extraction-NP condition. These results support the hypothesis that native speakers of English postulate intermediate gaps during the processing of long *wh*-dependencies, which facilitates the filler's integration with its subcategoriser.

L2 learners

As before for the native speakers, we carried out separate two-way ANOVA 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 ($F(1, 21) = 3.925, p = 0.061$; $F(1, 18) = 10.569, p < 0.01$), and a main effect of Extraction in the subjects analysis only ($F(1, 21) = 4.388, p < 0.05$; $F(1, 18) = 2.186, p > 0.1$). The analysis of the Greek learners' data showed a main effect of Phrase Type in the items analysis that was approaching significance in the subjects analysis ($F(1, 21) = 3.593, p = 0.068$; $F(1, 19) = 6.315, p < 0.05$), but no effect of Extraction. Unlike the native speakers, then, none of the learners group 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 reading times in the extraction conditions than in the non-extraction conditions. The analyses of variance revealed a main effect of Extraction for all groups (Chinese group: $F(1, 31) = 6.069, p < 0.05$; $F(1, 19) = 4.157, p = 0.056$; Japanese group: $F(1, 25) = 8.162, p < 0.01$; $F(1, 19) = 5.216, p < 0.05$; German group: $F(1, 21) = 15.175, p = 0.001$; $F(1, 19) = 27.241, p < 0.001$; Greek group: $F(1, 29) = 9.149, p < 0.01$; $F(1, 19) = 12.415, p < 0.01$). In addition, we found a main effect of Phrase Type for the Japanese group ($F(1, 25) = 5.031, p < 0.05$; $F(1, 19) = 7.482, p = 0.01$) that was approaching significance in the subjects analysis for the Chinese group ($F(1, 31) = 3.129, p = 0.086$; $F(1, 19) = 0.868, p > 0.1$) and in the items analysis for the German group ($F(1, 21) = 2.551, p > 0.1$; $F(1, 19) = 4.278, p = 0.053$). No interactions between Extraction and Phrase Type were observed in any of the learner groups, however. The Extraction effects suggests that the learners integrated the filler with its subcategoriser 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 native speakers was the interaction between Extraction and Phrase Type observed at segment 5. In the sentences involving extraction, the Extraction-VP condition elicited shorter reading times than the Extraction-NP condition, whereas no such difference was found between the corresponding non-extraction conditions. This pattern indicates that native speakers 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 learners groups showed any such interaction or 'intermediate gap' effect, suggesting that L2 learners process long *wh*-dependencies differently from native speakers.

5. Discussion

The purpose of the current study was to investigate whether or not both native speakers and L2 learners of English make use of intermediate gaps during the processing of long *wh*-dependencies. The main results of the self-paced reading experiment can be summarised as follows:

- Both the native speakers and the learners proved equally good at comprehending sentences involving long-distance *wh*-dependencies.
- All participants showed an effect of extraction at segment 5, indicating that they attempted to integrate the displaced *wh*-phrase with its subcategorising verb at this point during the parse (*filler-integration effect*).
- For the native speakers, filler integration was facilitated by the presence of an intermediate gap in sentences involving extraction across a clause boundary (*intermediate gap effect*).

- We found no intermediate gap effect in the L2 data, suggesting that independently of their L1 background, the learners failed to postulate any intermediate syntactic structure when processing sentences containing long *wh*-dependencies.

In the following, we will discuss the above findings and their implications for theories of L1 and L2 sentence processing.

Intermediate gaps in L1 processing

Several studies of the processing of filler-gap dependencies have shown that native speakers of English reactivate a displaced *wh*-constituent at the position of its associated syntactic gap. As the purported *wh*-gaps are located immediately after the verb or other lexical subcategoriser in English, however, results from these studies usually do not provide unambiguous evidence for the Trace Reactivation Hypothesis, but instead can also be explained by direct lexical association. That is, if the processing of a potential subcategoriser results in the parser's automatically reconstructing its lexical argument structure (Pickering & Barry, 1991; Sag & Fodor, 1994), then the observed filler-reactivation effects might be the result of lexically-driven or 'backward' gap-filling rather than reflecting the parser's use of empty categories. By way of eliminating this potential confound, we followed G&W in using sentences that involved long-distance dependencies. According to syntactic theories such as Chomsky's (1995, and later) minimalist framework, sentences involving long *wh*-movement also contain *intermediate* syntactic gaps.

Comparing sentences with and without extraction across either VPs or NPs, we found that English native speakers postulate such intermediate gaps during real-time processing (thus replicating G&W's findings). The native speaker group showed elevated RTs to the complementiser *that* in the condition involving an intermediate gap relative to the same word in the corresponding non-extraction condition, which may reflect the increased processing load associated with the parser's consulting a memory representation of the filler at this point and linking it to the gap. As in G&W'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 by themselves do not provide any conclusive evidence for the mental reality of intermediate gaps. According to G&W, the main effect of extraction observed at this segment reflects the additional processing cost associated with storing the filler in working memory in the two

extraction conditions. It is also conceivable that although extraction from a noun phrase containing a 'specified subject' is illicit, the parser attempts to analyse the *wh*-filler as the object of the preposition *about* at this point during processing (compare also Pickering, Barton & Shillcock, 1994). If the parser does indeed consider the preposition *about* as a potential host for the *wh*-filler, though, 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 subcategoriser. The elevated RTs elicited by the two extraction conditions compared to the non-extraction conditions can be taken to reflect the cost of filler integration at this point. Importantly, however, RTs in the condition involving an intermediate gap were significantly shorter than in the condition involving extraction across a noun phrase. This effect cannot be attributed to any differences in subject-verb distance between the two extraction conditions as no such difference in reading times was observed between the two non-extraction conditions. Given the assumption that the processing load associated with filler integration increases with the distance between the filler and its subcategoriser (Gibson, 1998, among others), the shorter RTs observed in the Extraction-VP condition can be explained by the availability of an intermediate landing site that serves to break up the long dependency into two shorter ones. Taken together, these results provide evidence that native speakers of English postulate intermediate gaps during the processing of long *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 subcategoriser.

Native speakers vs. L2 learners

The results from Williams et al.'s (2001) reading-time study suggest that L2 learners also employ a filler-driven strategy when processing *wh*-dependencies. Like native speakers, their learners appeared to try and integrate a fronted *wh*-phrase as soon as possible with a potential subcategoriser. Both Williams et al.'s and Juffs & Harrington's (1995) results furthermore indicate that L2 learners may have more difficulty than

native speakers recovering from an initial misanalysis. Recall, however, that the elevated reading times observed by both Williams et al. and Juffs & Harrington at the reanalysis region are consistent both with an empty category-based account of gap-filling such as the TRH, and with direct lexical association. By way of separating out these two possibilities, we used materials which according to syntactic theory also contain intermediate syntactic gaps. We found that the presence of such gaps facilitated a *wh*-filler's integration with its subcategoriser for the native speakers 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 subcategoriser irrespective of the availability of an intermediate landing site. Gap-filling in L2 processing, then, appears to be lexically-driven rather than driven by requirements of the grammar such as the subadjacency constraint.

The extraction effect observed for the L2 learners at segment 5 - the region containing the subcategorising verb - is in accordance with the predictions made by the DAH, whereas the interaction between phrase type and extraction that we found for the native speakers indicates that they postulated intermediate empty categories during processing. In short, there is no evidence in our results to suggest that L2 learners employ a filler-driven strategy when processing sentences containing long *wh*-dependencies, or indeed that they postulate any syntactic gaps at all. Note that our results are not necessarily incompatible with Williams et al.'s (2002) findings, given that the effects they observed also occurred at or after a potential subcategoriser. 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 complementiser or preposition, and that the *wh*-filler was not mentally reactivated at this point.

Interestingly, their failure to make use of intermediate syntactic gaps did not seem to compromise the learners' ability to understand the experimental sentences. This observation suggests that the comprehension of sentences containing long *wh*-dependencies is not contingent upon successive-cyclic movement. While the native speakers and the L2 learners were equally good at comprehending the experimental sentences, the two groups appeared to employ different processing mechanisms in order to achieve this result. English native speakers made use of intermediate syntactic gaps, which led to an increase in processing cost at the clause boundary but facilitated filler integration further downstream. The L2 learners, in contrast, attempted to establish a direct dependency between the fronted *wh*-phrase and its lexical subcategoriser, a

strategy that was applied regardless of the possibility of breaking down the *wh*-dependency into a series of smaller steps.

Our results are compatible with other studies indicating that second language learners under-use syntactic information during L2 processing, compared to native speakers. While there is evidence from various studies that L2 learners, like native speakers, 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, Roberts, Gross & Marinis (2003), Papadopoulou & Clahsen (2003) and Roberts (2003) have shown that L2 learners do not seem to apply any phrase-structure based locality principles when processing temporarily ambiguous sentences. The hypothesis that L2 processing differs from L1 processing with respect to the role of phrase structure information in on-line processing is further supported by the results from ERP studies. Hahne (2001), Hahne & Friederici (2001) and Isel (2002) found qualitative differences in first-pass parsing between native speakers and L2 learners of German for sentences that were syntactically ill-formed. While sentences containing phrase-structure violations elicited two different ERP components associated with the processing of syntactic information in the native speakers, 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.

L1 transfer in L2 processing

The role of L1 transfer in non-native language processing is controversial. Differences in linguistic performance between native speakers and L2 learners may be due to differences between the target grammar and the learners' L1 or interlanguage grammar, to differences in their parser, or possibly, both. While some studies have found evidence of processing transfer in L2 sentence comprehension (e.g. Frenck-Mestre, 1997; 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, two of which from *wh*-movement backgrounds (German and Greek) and two from *wh*-in-situ backgrounds (Chinese and Japanese), with similar levels of proficiency in L2 English. If properties of the learners' L1 grammar influence the way they process sentences from the L2, then we might have expected the German and Greek-speaking participants to pattern with the native

speakers, and differently from the Chinese and Japanese-speaking participants. Our results, however, indicate that all L2 groups processed the experimental sentences in essentially the same way but differently from native speakers. Specifically, none of the learner groups appeared to postulate any intermediate gaps during real-time processing, irrespective of whether or not 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 *wh*-dependencies in L2 English. Thus our results suggest that the (successive-cyclic version of the) active filler strategy is not transferred from the L1 to the L2. The absence of such transfer effects may be at least partially explained by the above hypothesis that L2 learners' sensitivity to syntactic information during real-time processing is more limited than that of native speakers. Instead, L2 learners seem to rely more on lexical-semantic and other non-syntactic cues to sentence interpretation.

6. Conclusion

Our results show that native speakers but not L2 learners apply a successive-cyclic version of the active filler strategy when processing long *wh*-dependencies in English. The learners' failure to make use of intermediate syntactic structure during parsing proved independent of whether or not their L1 instantiated successive-cyclic *wh*-movement. This finding lends further support to the hypothesis that L2 learners' sensitivity to syntactic information during L2 processing is restricted relative to that of native speakers. The observed dissociation between the learners' comprehension abilities and on-line sentence processing shows that although learners are able to comprehend sentences containing long *wh*-extractions, they do not use native-like, phrase structure-based processing mechanisms in order to achieve this goal.

Acknowledgements

This research was supported by the Leverhulme Trust (Grant No. F/00 213B to Harald Clahsen, Claudia Felser, and Roger Hawkins), which is gratefully acknowledged. Parts of the research reported here have previously been presented at the 24th DGfS Conference, Mannheim, EUROSLA 12, Basle, the 27th Boston University Conference on Language Development, and EUROSLA 13, Edinburgh. We thank the audiences for their comments and suggestions, Ted Gibson and Tessa Warren for making their 1999 manuscript available to us, and Ritta Husted and Michaela Wenzlaff for helping with the data collection.

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.
- 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* (273-317). Dordrecht: Kluwer Academic Publishers.
- Crocker, M. (1996). *Computational Psycholinguistics*. Dordrecht: Kluwer.
- Cuetos, F., Mitchell, D., & Corley, M. (1996). Parsing in different languages. In M. Carreiras, J. Garcia-Albea, & N. Sebastian-Galles (Eds.), *Language Processing in Spanish* (pp. 145-187). Mahwah, NJ: Lawrence Erlbaum Associates.
- Featherston, S. (2001). *Empty Categories in Sentence Processing*. Amsterdam: John Benjamins.
- Felser, C. (in press). *Wh*-copying, phases, and successive cyclicity. *Lingua*.
- Felser, C., Roberts, L., Gross, R., & Marinis T. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, 24, 453-489.
- Fiebach, C. J., Schlewsky, 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.
- Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, 4, 93-126.
- Frazier, L., & Rayner, K. (1988). Parameterizing the language processing system: Left- vs. right branching within and across languages. In J. Hawkins (Ed.), *Explaining Language Universals* (pp. 247-279). Oxford: Blackwell.
- Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology*, 50A, 119-148.

- 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.
- Gibson, E. (1998). Syntactic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1-75.
- Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Cross-linguistic attachment preferences: Evidence from English and Spanish. *Cognition*, 59, 23-59.
- Gibson, E., & Warren, T. (1999). The psychological reality of intermediate linguistic structure in long-distance dependencies. Unpublished manuscript, MIT, Cambridge, MA.
- 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.
- Harrington, M. (1987). Processing transfer: Language-specific processing strategies as a source of interlanguage variation. *Applied Psycholinguistics*, 8, 351-378.
- 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 DGfS Conference, Mannheim, Germany.
- Juffs, A. (1998). Main verb vs 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.
- 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.
- 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.

- Mazuka, R., & Lust, B. (1990). On parameter setting and parsing: Predictions for cross-linguistic differences in adult and child processing. In L. Frazier & J. de Villiers (Eds.), *Language Processing and Language Acquisition* (pp. 163-205). Dordrecht: Kluwer.
- MacWhinney, B. (1997). Second Language Acquisition and the Competition Model. In A. De Groot & J. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 113-142). Mahwah, NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. (2002). Extending the competition model. In R. Heredia & J. Altarriba (Eds.), *Bilingual Sentence Processing*. New York: Elsevier.
- McCloskey, J., 2001. The morphosyntax of *wh*-extraction in Irish. *Journal of Linguistics*, 37, 67-100.
- 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 & Shillcock, R. (Eds.), *Cognitive Models of Speech Processing: The Second Sperlonga Meeting* (pp. 321-350). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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., & 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). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pollard, C., & I. A. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Roberts, L. (2003). *Syntactic Processing in Learners of English*. Unpublished Ph.D. Dissertation, University of Essex, Colchester.
- Ross, J. (1967). *Constraints on variables in syntax*. Ph.D. dissertation, MIT, Cambridge, MA.
- Sag, I. A., & Fodor, J. D. (1994). Extraction Without Traces. In *Proceedings of the Thirteenth Annual Meeting of the West Coast Conference on Formal Linguistics* (pp. 365-384). Stanford: CSLI Publications.
- Stowe, L. (1986). Parsing *wh*-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, 1, 227-245.
- Swinney, D., Ford, M., Frauenfelder, U., & Bresnan, J. (1988). On the temporal course of gap-filling and antecedent-assignment during sentence processing. Unpublished Manuscript.

- Thornton, R. (1990). *Adventures in long-distance moving: The acquisition of complex wh-questions*. Unpublished Ph.D. dissertation, University of Connecticut, Storrs.
- 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'Neill (Eds.), *The Generative Study of Second Language Acquisition*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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.
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.
- 1b The manager who the secretary's claim about the new salesman had pleased will raise company salaries.
- 1c The manager thought the secretary claimed that the new salesman had pleased the boss in the meeting.
- 1d 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.
- 2b The student who the headmaster's thoughts about the clever teacher had surprised does not like doing homework.
- 2c The student believed the headmaster thought that the clever teacher had surprised everybody at school last week.
- 2d 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.
- 3b The nurse who the doctor's argument about the rude patient had angered is refusing to work late.
- 3c The nurse thought the doctor argued that the rude patient had angered the staff at the hospital.
- 3d 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.
- 4b The witness who the lawyer's proof about the evil criminal had confused does not want to testify.
- 4c The witness said the lawyer proved that the evil criminal had confused the judge during the trial.
- 4d 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.
- 5b The actress who the journalist's suggestion about the talented writer had inspired will go on stage tonight.
- 5c The actress thought the journalist suggested that the talented writer had inspired everybody with the new play.
- 5d 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.

- 6b The customer who the receptionist's statement about the lazy cleaner had annoyed will not pay his bill.
- 6c The customer thought the receptionist stated that the lazy cleaner had annoyed the manager of the hotel.
- 6d 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.
- 7b The farmer who the builder's thoughts about the dedicated worker had amazed will give everybody extra money.
- 7c The farmer said the builder thought that the dedicated worker had amazed the new boss last week.
- 7d 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.
- 8b The singer who the musician's statement about the drunken guitarist had offended will not perform this evening.
- 8c The singer thought the musician stated that the drunken guitarist had offended the drummer after the performance.
- 8d 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.
- 9b The schoolboy who the teacher's proof about the aggressive child had distressed will complain at the meeting.
- 9c The schoolboy said the teacher proved that the aggressive child had distressed the class at school yesterday.
- 9d 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.
- 10b The girl who the policeman's conclusion about the nasty boy had frightened has stopped going to school.
- 10c The girl said the policeman concluded that the nasty boy had frightened the children at the school.
- 10d 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.
- 11b The coach who the manager's decision about the violent footballer had annoyed will cancel the match today.
- 11c The coach said the manager decided that the violent footballer had annoyed his fans at the match.

- 11d 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.
- 12b The politician who the minister's statement about the TV journalist had upset will not give an interview.
- 12c The politician thought the minister stated that the TV journalist had upset the president on the programme.
- 12d The politician thought the minister's statement about the TV journalist had upset the president on the programme.
- 13a The chef who the cook argued that the head waitress had bothered wants to find another job.
- 13b The chef who the cook's argument about the head waitress had bothered wants to find another job.
- 13c The chef said the cook argued that the head waitress had bothered the manager of the restaurant.
- 13d 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.
- 14b The director who the agent's suggestion about the unpleasant dancer had disappointed will cancel the performance tonight.
- 14c The director said the agent suggested that the unpleasant dancer had disappointed the others in the ballet.
- 14d 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.
- 15b The film star who the interviewer's suggestion about the horrible photographer had embarrassed will not answer any questions.
- 15c The film star said the interviewer suggested that the horrible photographer had embarrassed the editor of the newspaper.
- 15d The film star said the interviewer's suggestion about the horrible photographer had embarrassed the editor of the newspaper.
- 16a The man who the customer thought that the shop assistant had amused was trying not to laugh.
- 16b The man who the customer's thoughts about the shop assistant had amused was trying not to laugh.
- 16c The man believed the customer thought that the shop assistant had amused everybody in the store yesterday.
- 16d 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.

- 17b The therapist who the patient's dream about the strange woman had fascinated is writing a new book.
- 17c The therapist said the patient dreamed that the strange woman had fascinated the members of the group.
- 17d 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.
- 18b The man who the detective's conclusion about the dangerous thief had distressed will buy a new alarm.
- 18c The man thought the detective concluded that the dangerous thief had distressed the people in the neighbourhood.
- 18d The man thought the detective's conclusion about the dangerous thief had distressed the people in the neighbourhood.
- 19a The captain who the officer decided that the young soldier had displeased will write a formal report.
- 19b The captain who the officer's decision about the young soldier had displeased will write a formal report.
- 19c The captain said the officer decided that the young soldier had displeased the colonel at training today.
- 19d 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.
- 20b The tourist who the guide's claim about the hotel manager had angered wants to return home now.
- 20c The tourist believed the guide claimed that the hotel manager had angered everybody in the holiday party.
- 20d The tourist believed the guide's claim about the hotel manager had angered everybody in the holiday party.

Endnotes

- ¹ Theodore Marinis is now working at the Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, and Leah Roberts at the Max-Planck-Institute for Psycholinguistics, Nijmegen.
- ² The term 'subjacency' refers to the requirement that non-local 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 The reason for using different cut-off points for the learners and for the native speaker controls was that the learner's reading times were slower overall than those of the native speakers.