Processing wh-dependencies in a second language: a cross-modal priming study

Claudia Felser and Leah Roberts

Second Language Research 2007; 23; 9
DOI: 10.1177/0267658307071600

Second Language Research 23,1 (2007); pp. 9–36

Processing *wh*-dependencies in a second language: a cross-modal priming study

Claudia Felser and **Leah Roberts** *University of Essex*

This study investigates the real-time processing of *wh*-dependencies by advanced Greek-speaking learners of English using a cross-modal picture priming task. Participants were asked to respond to different types of picture target presented either at structurally defined gap positions, or at pre-gap control positions, while listening to sentences containing indirect-object relative clauses. Our results indicate that the learners processed the experimental sentences differently from both adult native speakers of English and monolingual English-speaking children. Contrary to what has been found for native speakers, the learners' response pattern was not influenced by individual working memory differences. Adult second language learners differed from native speakers with a relatively high reading or listening span in that they did not show any evidence of structurally based antecedent reactivation at the point of the indirect object gap. They also differed from low-span native speakers, however, in that they showed evidence of maintained antecedent activation during the processing of the experimental sentences. Whereas the localized priming effect observed in the high-span controls is indicative of trace-based antecedent reactivation in native sentence processing, the results from the Greek-speaking learners support the hypothesis that the mental representations built during non-native language processing lack abstract linguistic structure such as movement traces.

I Introduction

Sentences containing unbounded dependencies present a challenge for the human sentence comprehension mechanism as a dislocated

Address for correspondence: Claudia Felser, Department of Linguistics, University of Essex, Colchester, CO4 3SQ, UK; email: felsec@essex.ac.uk

 10.1177/0267658307071600

constituent must be retained in short-term memory until it can be linked to its subcategorizing head or other licenser, which often does not appear until much later on. In native (L1) sentence processing, encountering a dislocated constituent (or 'filler') such as the fronted *wh*-phrase *which book* in sentences like *Which book did Mary ask her students to read?* is thought to trigger the prediction of a lexical head to license it, or of a corresponding syntactic 'gap' (Frazier and Clifton, 1989; Frazier and Flores d'Arcais, 1989; Gibson, 1998). Keeping a filler active in working memory (WM) incurs a processing cost that has been found to increase with distance (see, among others, Gibson, 1998). Studies using event-related brain potentials (ERPs), for example, have found evidence for 'maintained activation' in L1 processing, in the shape of an ERP component (a sustained left-anterior negativity, or LAN) that has been claimed to index working memory cost, including the memory cost associated with temporarily storing a filler in WM (King and Just, 1991; Kluender and Kutas, 1993; King and Kutas, 1995; Kluender and Münte, 1998; Fiebach *et al.*, 2002). Once a potential gap has been identified, the filler will be retrieved from WM and integrated into the emerging sentence representation. Evidence from ERP studies of filler-gap dependencies suggests that in L1 processing, WM storage and filler integration are distinct mental processes, reflected in qualitatively different brain signals. While LAN effects have been associated with WM storage, the so-called P600, or 'syntactic positive shift', has been claimed to reflect, *inter alia*, the cost of filler integration (Kaan *et al.*, 2000; Fiebach *et al.*, 2002; Felser *et al.*, 2003).

Exactly how filler integration is accomplished is, however, still controversial. According to the Direct Association Hypothesis (DAH) (e.g. Pickering and Barry, 1991; Pickering, 1993), dislocated constituents are linked directly to their lexical subcategorizer when the latter is encountered, and semantically integrated into the subcategorizer's argument structure or thematic grid. Alternatively, it has been suggested that filler integration is mediated by empty syntactic categories known as 'traces' in generative-transformational grammar (e.g. Chomsky, 1981; 1995), which form part of the grammatical representations built during parsing (see, for example, Bever and McElree, 1988; Nicol and Swinney, 1989; Gibson and Hickok, 1993; Gorrell, 1993; Love and Swinney, 1996). The Trace Reactivation Hypothesis (TRH) claims that a filler is retrieved from working memory – that is, 'reactivated' – when the parser has identified a potential syntactic gap, and irrespective of the position of its lexical subcategorizer. These two hypotheses are notoriously difficult to dissociate empirically in head-initial languages like English, especially since many studies have investigated dependencies involving dislocated direct objects, whose hypothesized traces are located immediately after the subcategorizing verb. More persuasive evidence in support of trace-based reactivation in L1 sentence processing comes from studies on verb-final languages (Clahsen and Featherston, 1999; Miyamoto and Takahashi, 2002; Nakano *et al.*, 2002; Fiebach *et al.*, 2004) and from studies investigating the processing of indirect-object dependencies (Nicol, 1993; Roberts *et al.*, 2007), subject-relative clauses (Swinney and Zurif, 1995; Lee, 2004), or dependencies spanning more than one clause (Gibson and Warren, 2004; Marinis *et al.*, 2005).

Few studies have investigated the way second-language (L2) learners process filler–gap dependencies in real time, however. While there is evidence that learners are able to link dislocated constituents to their lexical subcategorizers when processing *wh*-dependencies in the L2 (Williams *et al.,* 2001; Juffs and Harrington, 1995; Juffs, 2005; Marinis *et al.*, 2005), the results from a reading-time study by Marinis *et al.* (2005) suggest that even highly proficient learners of L2 English process such dependencies differently from English native speakers. Marinis *et al.* investigated the processing of long-distance *wh*-dependencies in complex sentences such as (1) below, in both native English speakers and advanced adult L2 learners of English from different language backgrounds.

1) a. The nurse $[_{CP}$ who$_i$ the doctor argued $[_{CP}$ e_i' that the rude patient had angered $e_i]]$ is refusing to work late.
   b. The nurse $[_{CP}$ who$_i$ the doctor's argument about the rude patient had angered $e_i]$ is refusing to work late.

For the native speakers, filler integration at the embedded verb *angered* was found to be facilitated by the availability of an intermediate syntactic gap (marked e_i') at the clause boundary in (1a), compared to sentences of the same length that did not contain an intermediate gap such as (1b) (see also Gibson and Warren, 2004). No such facilitation effect was observed for the L2 learners, however, regardless of whether or not the subjacency constraint was operative in their native language. This finding indicates that intermediate gaps do not form part of the mental representations constructed during L2 processing. Note that the learners had no particular difficulty understanding sentences such as (1) above, however, and that their reading profiles showed evidence of filler integration at the point at which they encountered the subcategorizing verb. In short, Marinis *et al.*'s learners seemed to process long-distance *wh*-dependencies in accordance with the DAH but not the TRH.

The absence of any intermediate gap effects in Marinis *et al.*'s L2 data supports the hypothesis that the grammatical representations constructed during L2 processing are shallower than those built during L1 comprehension and lack abstract elements such as empty syntactic categories (Clahsen and Felser, 2006). As regards the processing of filler-gap dependencies in the L2, the Shallow Structure Hypothesis predicts that, although learners may be able to keep a fronted constituent in short-term memory and semantically associate it with an appropriate lexical head further downstream, filler integration will not be mediated by any structurally defined gaps. Note that the results from Marinis *et al.*'s (2005) reading-time study have provided only indirect evidence for the presence of syntactic gaps non-adjacent to the subcategorizing verb in native speakers' representations, and for their absence in L2 processing. The present study aims to test the above prediction more directly, by investigating antecedent reactivation effects at structural gap positions during the processing of indirect object dependencies in L2 English, using a cross-modal priming task. Potential effects of individual working memory differences on L2 processing will also be examined.

II Antecedent priming in L1 sentence processing

The cross-modal lexical priming technique provides a useful tool for examining whether dislocated constituents are mentally reactivated at particular structural positions. In this task, participants are required to make a word-based (e.g. a word/nonword discrimination) decision to visually presented targets while listening to stimulus words or sentences spoken at normal speed (Swinney *et al.*, 1979). If dislocated constituents are reactivated at gap positions, then participants' responses to targets semantically related or identical to the antecedent should be facilitated at the point of a gap, relative to non-gap (control) positions. This prediction is based on the well-documented phenomenon of automatic priming, the observation that the processing of visual targets is facilitated if they are presented immediately after the auditory presentation of an identical or semantically related word, or 'prime' (Neely, 1991).

Antecedent-priming effects at gap locations have previously been observed in both adult native speakers (Nicol and Swinney, 1989; Nicol, 1993; Swinney and Zurif, 1995; Love and Swinney, 1996; Clahsen and Featherston, 1999; Nakano *et al.*, 2002) and monolingual children (Love and Swinney, 1997; Hestvik *et al.,* 2005; Roberts *et al*., 2007). Nicol and Swinney (1989), for example, report the results from a cross-modal priming experiment investigating how mature native speakers of English processed *wh*-dependencies in sentences such as (2).

2) The policeman saw *the boy* who the crowd at the party accused _____ of the crime.

Participants' response times to target words semantically related to *boy* were shorter than response times to unrelated ones at the test position immediately following the subcategorizing verb, *accused*, but not at an earlier (preverbal) control position. The antecedent priming effect observed at the point of the gap indicates that the antecedent was retrieved from short-term memory, or 'reactivated', at its canonical direct object position. As the above results are compatible both with the DAH and the TRH, however, they do not provide any unequivocal evidence for trace-based reactivation. One possible way of dissociating lexically based and structurally based reactivation effects is to examine the processing of filler–gap dependencies in head-final languages. Cross-modal priming studies on languages such as Japanese or German have found evidence for structurally based antecedent reactivation even before the subcategorizing verb was processed. Investigating the processing of long-distance object scrambling sentences in Japanese, Nakano *et al.* (2002), for example, found antecedent priming effects at the preverbal object gap but not at an earlier control position, and Clahsen and Featherston (1999) report similar results for object scrambling constructions in German. Antecedent reactivation effects have also been observed at the end of sentences (Balogh *et al.*, 1998), suggesting that some memory representation of the antecedent is retained even after gap-filling, and is re-accessed during end-of-sentence interpretation (or 'wrap up') processes.

As the memory cost incurred by temporarily storing a dislocated constituent in WM is thought to increase with distance, we may expect antecedent reactivation to be affected by individual WM differences. There is some evidence from L1 processing studies that this is indeed the case. In Nakano *et al.*'s (2002) study, only participants with a relatively high reading span showed priming effects at the position of the gap. Low-span participants, on the other hand, seemed unable to retain the filler in working memory for long enough to be able to retrieve it at the gap site. Working memory effects were also observed in Roberts *et al.*'s (2007) study with English-speaking adults and children. Using a cross-modal picture priming task (McKee *et al.*, 1993; Love and Swinney, 1997), Roberts *et al.* investigated antecedent reactivation in sentences such as (3).

3) John saw *the peacock* to which the small penguin gave the nice birthday present ___ in the garden last weekend.

Participants were asked to make an 'alive'/'not alive' decision to picture targets presented either at the gap site (i.e. after the direct object *the nice birthday present*) or at an earlier control position. As the participants' performance in this task was influenced by individual WM differences (reading span for adults, listening span for children), Roberts *et al.* divided both the children and the adults up into 'high-span' and 'low-span' participants, according to their median scores in the WM tests.[1] A summary of the four participant groups' results in the cross-modal priming task is provided in Table 1.

For both high-span children and adults, reaction times (RTs) to identical targets were faster than those to unrelated targets at the gap position, whereas there was no advantage at all for identical targets at the earlier control position. This RT pattern is expected if the aliveness decision task is facilitated by the presence of a *wh*-gap at the later test point, but not

[1]Adult NSs were split up into two subgroups based on their scores in Daneman and Carpenter's (1980) reading-span test, and the children on the basis of their performance in Gaulin and Campbell's (1994) listening-span test for children.

Table 1 Summary of Roberts *et al.*'s (2007) results from high-span (HS) and low-span (LS) children and adults (mean RTs for unrelated targets minus mean RTs for identical targets, in milliseconds)

	HS adults ($n = 22$)	HS children ($n = 19$)	LS adults ($n = 32$)	LS children ($n = 25$)
Pre-gap position	−2	−87	19	−36
Gap position	31	53	−5	−147

during the processing of other (gap-free) sentence regions. Low-span participants, on the other hand, did not show any facilitation for identical targets at either of the two test points. The low-span children actually showed a lexical interference effect, with RTs to identical targets (e.g. a picture of a peacock) being longer than those to unrelated ones at both test points, even though the word *peacock* had previously been mentioned in the auditory stimulus sentence (for some discussion of this finding, see Roberts *et al.*, 2007). Note that the high-span participants' RT pattern cannot be explained by maintained (or 'residual') activation, which would predict the facilitation effect for identical targets to be either similar across different test points, or reduced (or absent) at later test points, as the mental representation of the antecedent will gradually fade from WM over time.[2] In short, only the high-span native speakers in Roberts *et al.*'s study showed the predicted antecedent priming effect at the position of the indirect object gap. Roberts *et al.*'s results further indicate that children as young as five are capable of building representations that include syntactic gaps when processing their native language, and confirm earlier findings showing that in L1 sentence processing, antecedent reactivation at gap sites is influenced by individual working memory differences.

The observation that the low-span participants in both Nakano *et al.*'s and Roberts *et al.*'s studies were still able to understand the experimental sentences, though, indicates that successful comprehension does not necessarily depend on participants' constructing representations that include abstract linguistic structure such as movement traces.[3] From the

[2]Numerically at least, the latter pattern can be seen in the low-span adults. The low-span children's RT pattern may reflect some degree of lexical interference between the auditory input and the visual target's semantic features, or could result from their being confused by the dual-task demands of the experiment (for further discussion, see Roberts *et al.*, 2007).
[3]Using a (less resource-demanding) cross-modal naming task, Hestvik *et al.* (2005) found antecedent reactivation effects at object gaps even in low working memory span children.

point of view of processing economy, reconstructing a filler at its gap site in addition to integrating it with its subcategorizer (and into the emerging semantic or discourse representation) may even seem a waste of computational resources. There is indeed evidence from the L1 processing literature that native speakers will, under certain conditions, perform shallow, or incomplete, parses only (e.g. Ferreira *et al.*, 2002; Sanford and Sturt, 2002). We will return to the concept of shallow processing in relation to *wh*-dependencies in the Discussion section below. An obvious advantage of building grammatical representations that include syntactically defined gaps, however, is that this allows us to parse even semantically incongruous or implausible sentences correctly. In the processing of head-final languages, trace-based gap-filling will enable comprehenders to begin building the verb phrase before the verb has been received, by allowing them to link dislocated arguments to their preverbal base positions prior to the processing of their lexical subcategorizer (Crocker, 1996). Furthermore, cyclically reactivating dislocated constituents at intermediate gap sites during the processing of long-distance dependencies helps ensure that the filler does not simply fade from memory before it is needed, thus facilitating filler integration further downstream (compare, e.g., Frazier and Clifton, 1989; Marinis *et al.*, 2005; Gibson and Warren, 2004). In short, postulating syntactically defined gaps may benefit comprehension even if reactivating a filler at gap sites incurs a certain amount of extra processing cost.

III The current study

The present study has two main aims:

- to investigate whether or not filler integration is mediated by syntactic gaps in L2 processing; and
- to examine whether non-native processing of *wh*-dependencies is influenced by individual WM differences.

To this end, we tested the processing of indirect object dependencies by advanced Greek-speaking learners of L2 English using Roberts *et al.*'s (2007) cross-modal picture priming (CMPP) task. Although this task has not, to our knowledge, previously been used with non-native speakers, it seems to offer several advantages over other methods for studying on-line sentence processing in language learners. First, this experimental technique allows for the stimulus materials to be presented uninterrupted and at a normal speech rate, thus rendering it a more natural than, for example, self-paced reading or listening. Second, the use of picture rather than word targets eliminates any potential processing delay caused by the learners' having to identify target words written in a non-native script and, third, using a picture-based decision task reduces the possibility of the participants focusing consciously on the structure of the auditory stimulus sentences, thus minimizing the likelihood of their responses being influenced by their metalinguistic knowledge of the L2 grammar.

Note that Greek, like English, is a *wh*-movement language and head-initial. Indirect-object relatives are formed in essentially the same way in Greek as in English (compare, e.g., Holton *et al.*, 1997), as the following translation of sentence (3) above illustrates.[4]

4)	O	Yiannis	ide	to	pagoni	ston	opoio	o	mikros	pinguinos	edose	to
	The	John	saw	the	peacock	to	which	the	small	penguin	gave	the

oreo	doro	genethlion	ston	kipo	to	perasmeno	Savvato-Kyriako.
nice	present	birthday	in the	garden	the	last	weekend.

As in English, prepositional indirect objects canonically follow the direct object in Greek. Given that the two languages pattern very similarly in the domain under investigation, we might expect that proficient Greek-speaking learners of English show the same kind of antecedent priming effects in their L2 that have previously been found for English native speakers. To the extent that antecedent reactivation depends on the availability of sufficient WM resources, we might further expect their performance pattern to be influenced by individual WM differences. According to the Shallow Structure Hypothesis for L2 processing, on the other hand, which claims that syntactically defined gaps, or 'traces', are absent from the mental representations constructed during L2 sentence processing, we would not expect to find any such localized priming

[4]Alternative ways of forming indirect-object relatives in Greek include using a genitive-marked *wh*-phrase (*tou opoiou*) instead of a prepositional one (*ston opoio*), or the use of the complementizer *pu* 'that' instead of a *wh*-relativizer (Holton *et al.*, 1997).

effects, even for sentences that are derived in a similar way in the learners' L1.

IV Method

1 Participants

Twenty-four Greek-speaking learners of English (mean age: 25.17; range: 20–31), all of whom students at the University of Essex, UK, participated in our study. Their performance was compared with that of 54 mature native speakers (NSs) of English (mean age: 22.8; range: 19–42) and 44 monolingual English-speaking children (mean age: 6.25; range: 5–7) from Roberts *et al.*'s (2007) study. All participants had normal hearing and normal or corrected-to-normal vision, and were not informed of the ultimate purpose of the main experiment. They received a small fee for their participation.

The Greek-speaking participants had first been exposed to English aged between 6 and 11 in a classroom setting or during private lessons, and none of them considered themselves bilingual. They had been resident in the UK for an average of 2.9 years. To allow us to determine their general L2 proficiency level at the time of testing, the learners completed the Oxford Placement Test (Allan, 1992), a standardized English proficiency test. All learners scored at or above 164/200 points in the OPT, indicating that all of them were advanced learners of English. The learners' bio-data and proficiency scores are summarized in Table 2.

As working memory capacity had been found to be a predictor for native speakers' performance in several cross-modal priming studies, the learners also underwent a reading span test (Harrington and Sawyer, 1992). This test involved the participants' reading sets of sentences, first

Table 2 Summary of the learners' bio-data and proficiency scores (n = 24)

	Age[a]	Age of first exposure[a]	Time spent in UK[a]	OPT scores[b]
Mean	25.17	8.63	2.9	178.79
Range	20–31	6–11	0.8–8.0	164–196
SD	3.06	1.46	2.5	6.51

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

a set of two, then three, up to a maximum of five. At the end of each set, the final word of each sentence had to be recalled and written down in the correct order. Scores indicate the maximum number of words that were correctly recalled (out of a total possible 42). The learners scored an average of 27.08 (range: 9–42; SD: 8.82) in this task, with all but one participant scoring above 13. Recall that in Roberts *et al.*'s study, only participants with a relatively high WM score had shown the expected antecedent priming effect at the gap position, though. To make the L2 group as homogeneous as possible with respect to their reading span, and better comparable to the high-span controls, we decided to exclude the data from the one learner who only achieved a very low score of 9 (2 SDs below the group mean) in the Harrington and Sawyer (1992) test from the analyses of the main experiment. Excluding this participant raised the remaining participants' average WM score to 27.87 (range: 13–42; SD: 8.11).[5]

2 *Materials*

The materials for the cross-modal priming task were the same as in Roberts *et al.*'s study, and comprised 20 experimental sentences containing indirect–object relatives, plus 60 filler sentences similar in length to the experimental ones. As indirect prepositional objects canonically follow the direct object in English, the hypothesized gap in our experimental sentences is not directly adjacent to the subcategorizing verb. Note that only the Trace Reactivation Hypothesis – but not the Direct Association Hypothesis – predicts that we should find antecedent priming effects at the position of the indirect object gap, thus allowing for the two hypotheses to be empirically dissociated. To prevent participants from focusing their attention on specific points during the sentences, 12 of the fillers were structurally similar to the experimental sentences but with visual targets presented at positions other than the

[5]Like Gaulin and Campbell's (1994) listening span test for children, the Harrington and Sawyer (1992) reading span test is a derivative of Daneman and Carpenter's (1980) original reading span test. Although Harrington and Sawyer's test uses a slightly different scoring procedure from the latter (total number of words recalled vs. highest level), there is evidence that the scores obtained by either measure in Daneman and Carpenter's test and its analogues are highly correlated (see, e.g., Whitney *et al.*, 2001).

critical test points. The remaining 48 fillers included constructions of different types.

All 80 sentences were read by a female native speaker of English, with natural intonation, and pre-recorded on a digital tape recorder. Pictures of animals and inanimate objects (all but one taken from the set of pictures in Snodgrass and Vanderwart, 1980) were used as visual targets. For each experimental sentence we selected two visual targets: an 'identical' picture target showing the referent of the indirect object noun and a picture showing an unrelated object. For each such identical vs. unrelated pair, the nouns depicted were matched for syllable length and lemma frequency (based on Francis and Kucera, 1982).[6] The pictures were scanned and their presentation linked to the experimental sentences such that the pictures would appear at one of two different test points:

- at the offset of the direct object NP; and
- at a pre-gap control position 500 ms earlier.

This 2×2 design yielded four different experimental conditions, as illustrated in (5a)–(5d).

5) Fred chased the squirrel to which the nice monkey explained…
   a. *Identical, gap position*:
      … the game's difficult rules [SQUIRREL] in the class last Wednesday.
   b. *Identical, pre-gap position*:
      … the game's [SQUIRREL] difficult rules in the class last Wednesday.
   c. *Unrelated, gap position*:
      … the game's difficult rules [TOOTHBRUSH] in the class last Wednesday.
   d. *Unrelated, pre-gap position*:
      … the game's [TOOTHBRUSH] difficult rules in the class last Wednesday.

A complete list of experimental sentences and targets is provided in Appendix 1. The experimental sentences were distributed across four counterbalanced presentation lists so as to ensure that each participant listened to each sentence only once, and with all presentation lists including an equal number of identical and related targets.

[6]Note that using identical rather than semantically related targets provides the most direct way of testing for antecedent priming effects. Any facilitation due to the referential identity of filler and target will be factored out by comparing response times to the same target pictures at the gap position with those at the earlier control position (see also Clahsen and Featherston, 1999).

The experimental sentences in each list were pseudo-randomized and mixed with the fillers.

3 Procedure

Each participant was tested individually in one of our experimental laboratories. They were seated in front of a 17″ PC monitor and asked to listen carefully to the pre-recorded sentences over headphones, and to watch the screen for pictures that would appear at some point during each sentence. The participants were told that whenever a picture appeared on the screen they had to decide as quickly as possible whether the animal or object in the picture was alive or not alive, by pushing either the left or the right-hand button of a dual push-button box. Participants' response times were measured from the point at which the picture appeared on the screen to their pressing one of the response buttons. Stimulus presentation and the recording of RTs was controlled by the NESU software package (Baumann *et al.*, 1993). To ensure that the participants made an active effort to comprehend the stimulus sentences, they additionally had to answer a total of 38 auditory comprehension questions randomly interspersed throughout the experiment. To allow the participants to familiarize themselves with the cross-modal aliveness decision task, the main experiment was preceded by a short practice phase. The experiment was presented over two brief sessions of no more than 15 minutes each, with a short break in between.

V Results

1 Accuracy

The learners answered 91.8% (adult NSs: 96%; children: 86%) of the end-of-trial comprehension questions correctly in the on-line task. Participants were also highly accurate in the aliveness decision task, with the learners correctly identifying 96.3% (adult NSs: 94%; children: 97%) of the picture targets as either 'alive' or 'not alive'. The data from one L2 participant who performed close to chance (at only 65% correct) on the picture targets were excluded from all subsequent analyses. Removing this participant's data raised the remaining learners' average accuracy scores in the comprehension task to 91.9% (range: 78–100%; SD: 4.9), and to 97.7% (range: 90–100%; SD: 4.0) in the aliveness decision task. These results demonstrate that the learners had no difficulty comprehending the experimental sentences, or coping with the dual-task demands of the cross-modal priming experiment.

2 *Reaction times*

Only trials that were responded to correctly were included in the analysis of the reaction time data. Applying the same data trimming criteria as in Roberts *et al.*'s study, we also removed trials that exceeded the set time-out of 2000 ms from the learners' data set (affecting 1.2% of the data), as well as individual outliers beyond 2 SD from each participant's mean RT per condition (affecting a further 5.9% of the L2 data). Statistical analyses were performed on the remaining RT data from 22 learners, and on the data from 22 high-span and 32 low-span adult NSs, and 19 high-span and 25 low-span children, from Roberts *et al.*'s study.

The learners' response times were shorter to identical than to unrelated targets at both the pre-gap (53 ms advantage for identical targets) and the gap position (57 ms advantage for identical targets). Figure 1 shows the learners' mean RTs to the visual targets at the two test points. To determine whether the learners' performance in the cross-modal priming task was affected by individual WM differences, we carried out a preliminary ANOVA with the within-participants factors Position

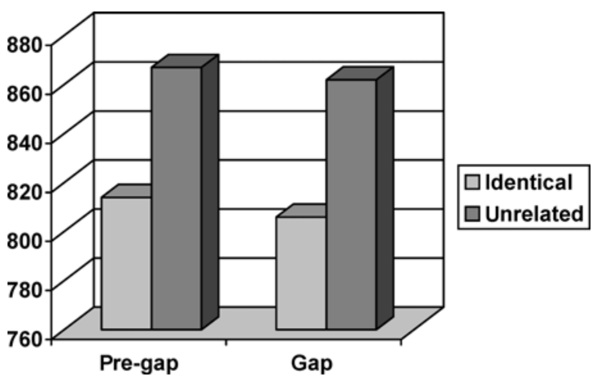

Figure 1 L2 learners' (n = 22) mean RTs to picture targets

(gap vs. control) and Target Type (identical vs. unrelated), and Reading Span as a covariate on the L2 data. The factor Reading Span did not interact with either Position ($p > .8$) or Target Type ($p > .1$), nor was there a significant three-way interaction ($p > .5$), indicating that individual WM differences (as measured by the test by Harrington and Sawyer, 1992) did not affect the learners' performance in the aliveness decision task. Although all our learners had demonstrated a relatively high level of general proficiency in English, to see whether the learners' performance was influenced by individual proficiency differences, we carried out a parallel ANOVA with Proficiency (as measured by the OPT) as a covariate. Again, there were no significant interactions with either experimental factor (Proficiency × Position, $p > .5$; Proficiency × Target Type, $p > .9$) nor was there a three-way interaction ($p > .2$), suggesting that individual differences in L2 proficiency did not affect the learners' RT pattern, either.

Recall that for the high-span participants in Roberts *et al.*'s study, RTs to identical targets were faster than those to unrelated targets at the gap position, whereas there was no such advantage for identical targets at the earlier control position. This RT pattern is expected if the aliveness decision task is facilitated by the presence of a *wh*-gap at the later test point. Low-span NSs, on the other hand, had shown no facilitation for identical targets at all. To determine whether the learners' performance pattern resembled that of either the high-span or the low-span adult NSs, or of the high-span or low-span children, we went on to compare the L2 group with each of the four subgroups from Roberts *et al.*'s study separately.

a L2 learners vs. adult NSs: To compare the learners with the high-span adult NSs, we carried out a mixed 2*2*2 ANOVA with the between-participants factor Group (high-span NSs, L2 learners), and the within-participants factors Position (gap vs. control) and Target Type (related vs. unrelated). This analysis revealed significant main effects of Group (F_1 (1,42) = 7.598, p = 0.009; F_2 (1,38) = 87.749, p = 0.000), reflecting the fact that the learners' RTs were higher overall than the NSs', and a main effect of Target Type (F_1 (1,42) = 10.963, p = 0.002; F_2 (1,38) = 10.393, p = 0.003). We moreover found a significant Target Type × Group interaction in the analysis by items that also approached significance in the by-participants analysis (F_1 (1,42) = 3.773, p = 0.059; F_2 (1,38) = 4.195, p = 0.047), and which indicates that the learners' RT pattern differed from the pattern seen in the high-span NSs. The two groups differed in that only the NSs showed a position-specific advantage for identical targets at the point of the gap. Subsequent pairwise comparisons confirmed that for the L2 group, identical targets elicited significantly shorter RTs than unrelated ones both at the gap (t_1 (21) = 2.415, p = 0.025; t_2 (19) = 2.659, p = 0.016) and at the pre-gap control position (t_1 (21) = 2.570, p = 0.018; t_2 (19) = 1.646, p = 0.116). In other words, the learners could identify (and, thus, classify) pictures showing the referent of a *wh*-filler more easily than pictures that were unrelated to any of the sentence's participants, but the size of this facilitation effect was not affected by the structural position at which these pictures were presented. In a parallel ANOVA, we also compared the L2 group with the low-span NSs. This analysis also showed significant main effects of Group (F_1 (1,52) = 4.907, p = 0.031; F_2 (1,38) = 51.620, p = 0.000) and Target Type (F_1 (1,52) = 11.054, p = 0.002; F_2 (1,38) = 9.360, p = 0.004), as well as a significant Target Type × Group interaction (F_1 (1,52) = 6.521, p = 0.014; F_2 (1,38) = 4.879, p = 0.033), confirming that the learners did not pattern with the low-span NSs, either. Specifically, these two participant groups differed in that the L2 learners but not the low-span NSs showed significantly shorter RTs to identical than to unrelated targets.

b L2 learners vs. children: Two further 2*2*2 ANOVAs were performed to compare the learners' RT pattern with those of the high-span and low-span English-speaking children. When comparing the learners with the high-span children, we found a significant main effect of Group (F_1 (1,39) = 24.588, p = 0.000; F_2 (1,38) = 184.371, p = 0.000), a significant interaction between Target Type and Position (F_1 (1,39) = 4.976, p = 0.032; F_2 (1,38) = 5.863, p = 0.02), a marginal interaction between Target Type and Group (F_1 (1,39) = 4.064, p = 0.051; F_2 (1,38) = 4.003, p = 0.053), and a three-way interaction between Target Type, Position and Group in the analysis by participants that also approached significance in the analysis by items (F_1 (1,39) = 4.663, p = 0.037; F_2 (1,38) = 3.685, p = 0.062). The comparison with the low-span children again revealed a significant main effect of Group (F_1 (1,44) = 47.782, p = 0.000; F_2 (1,38) = 473.351, p = 0.000), an interaction between Target Type and Position in the participants analysis (F_1 (1,44) = 4.546, p = 0.039; F_2 (1,38) = 1.458, p = 0.235), a significant interaction between Target Type and Group (F_1 (1,44) = 11.867, p = 0.001; F_2 (1,38) = 16.719, p = 0.000), and a significant interaction between Target Type, Position and Group in the analysis by participants (F_1 (1,44) = 4.874, p = 0.033; F_2 (1,38) = 2.482, p = 0.123). The main effects of Group reflect the fact that both the high-span and low-span children's RTs were longer overall than the learners', and the observed interactions with the factor Group shows that the learners' RT pattern differed reliably from those of both high-span and low-span children. The high-span children pattern with the high-span adult NSs rather than with the L2 group in that they showed shorter RTs to identical targets at the point of the gap only, whereas the low-span children did in fact show the exact opposite pattern to that found for the L2 group – longer RTs to identical than to unrelated targets at both test points (compare Table 1 above).[7]

3 Summary

Summarizing, we found that the L2 group behaved differently from all four NS subgroups examined in Roberts *et al.*'s study. The learners differed from the high-span native adults and children in that they did not show a localized antecedent priming effect at the point of the indirect object gap, and from the low-span NS groups in that they responded to identical targets more quickly than to unrelated ones. Crucially though, this facilitation effect occurred irrespective of the structural position at which the target pictures appeared. In the following, we will discuss these results in the light of previous research on

[7]Although the learner's performance was not influenced by their L2 reading span, they may nevertheless have formed a more heterogeneous WM group than did each of the four NS subgroups who served as controls. To test whether the learner's RT pattern resembled that of high-span and low-span NSs grouped together, we carried out additional statistical comparisons between the learners and all of the children, and between the learners and all of the adult NSs. Again, both of these comparisons revealed significant interactions with Group (p's < 0.01), confirming that the learners also behaved differently from both the two NS groups combined.

filler-gap processing and different hypotheses about non-native language processing.

VI Discussion

The results from the native speakers corroborate those from earlier studies reporting antecedent reactivation effects at gap sites in L1 sentence processing (e.g. Nicol, 1993; Clahsen and Featherston, 1999; Nakano *et al.*, 2002). The antecedent priming effect observed for the high-span NSs at the position of the indirect object gap support a trace-based account of the processing of *wh*-dependencies. Note that only the TRH – but not the DAH – predicts such an effect at the later test point. That is, upon identifying a syntactic gap, the L1 parser reconstructs the antecedent at this position even if it is not adjacent to the antecedent's lexical subcategorizer, and at a greater distance from the point at which the antecedent was first mentioned than the pre-gap control position. In cross-modal priming tasks, such structurally based antecedent reactivation is expected to be reflected in a significantly larger RT advantage for identical targets at the point of the gap than at earlier (pre-gap) control positions. Recall that Roberts *et al.* (2007) found evidence for-trace-based reactivation at indirect object gaps even in 5–7-year-old monolingual children, lending support to the 'continuity of parsing' view in L1 development (Crain and Wexler, 1999). Our L2 learners, in contrast, did not show any structurally determined antecedent priming effect of the kind that was observed in the high-span English-speaking children and adults. Instead, the learners' RT pattern differed reliably from those seen in any of our four NS control groups in that only the L2 group showed shorter RTs to identical than to unrelated targets at both test points. The results from the learners indicate that although they retained fronted *wh*-phrases in working memory during the processing of the experimental sentences, they did not retrieve them from WM (i.e. reactivate them) at structurally defined gap sites. In short, our L2 data show evidence of maintained activation, but not of reactivation.

Possible reasons for learners' non-native-like performance in real-time sentence processing tasks include:

- the lack of relevant L2 knowledge;
- negative L1 transfer;

- delayed or slowed-down processing; and
- the unavailability of certain processing routines in the L2 (compare also Clahsen and Felser, 2006).

Given the learners' high level of proficiency in English and their native-like accuracy scores on the end-of-trial comprehension questions, it seems unlikely that the absence of any gap-specific priming effects in the L2 data should reflect a mere knowledge deficit. Since indirect object *wh*-dependencies are formed similarly in Greek and English, the learners' on-line performance cannot obviously be accounted for by negative grammatical or processing transfer, either. Slowed processing also fails to provide a plausible explanation for our findings. While the learners did indeed show somewhat higher RTs overall than the adult NSs in the aliveness decision task, their RTs were still considerably shorter than the children's. For the native speakers examined by Roberts *et al.* (2007), individual WM differences rather than differences in processing speed between children and adults proved to be the crucial predictor for antecedent reactivation. Although the children's response latencies were almost twice as long on average as the adult NSs', the high-span children nevertheless showed the same antecedent priming effect as did the high-span adults. Recall that for the L2 learners, on the other hand, recognition of identical targets was equally facilitated at both test points, relative to unrelated ones. Had they merely required more time to process the indirect object trace, the facilitation effect for identical targets should still have been larger at the later test point than at the earlier one, irrespective of any differences in general processing speed between the learners and the native speaker controls.

The results from the L2 group are best compatible with the hypothesis that non-native comprehenders are unable to apply some of the parsing routines that are used in L1 comprehension. Recall that native speakers have been found to reconstruct the filler at structurally defined gap sites when processing filler-gap dependencies, in accordance with the TRH. Existing L2 processing studies, however, have thus far failed to find any evidence for trace-based gap-filling in non-native sentence processing. The results from the present study are consistent with Marinis *et al.*'s (2005) finding that proficient learners of L2 English do not postulate any intermediate traces when processing long-distance *wh*-dependencies in their L2. Together, the results from these studies provide converging evidence from different experimental tasks and different syntactic structures suggesting that, although L2 learners are able to interpret sentences containing filler-gap dependencies correctly, L2 comprehension does not involve any structure-driven gap-filling of the kind that has been observed in NSs. A reduced reliance on phrase structure based parsing routines has also been reported in studies examining L2 ambiguity resolution (Felser *et al.*, 2003; Papadopoulou and Clahsen, 2003). Investigating the processing of ambiguous relative clauses in L2 Greek, Papadopoulou and Clahsen (2003), for example, found that in the absence of relevant lexical cues to interpretation, advanced learners from different L1 backgrounds consistently failed to show any disambiguation preferences at all, even though Greek NSs' ambiguity resolution preferences were the same as those attested in the participants' L1s. Similar results were obtained by Felser *et al.* (2003) for L2 English.

By way of providing a unified account of these findings, Clahsen and Felser (2006) suggested that L2 learners typically perform partial or 'shallow' parses only, that is, they construct syntactic representations that lack deep hierarchical structure, and abstract elements of phrase structure such as movement traces. The concept of shallow parsing is familiar from computational approaches to language processing (e.g. Abney, 1991) and refers to 'the task of recovering only a limited amount of syntactic structure from natural language sentences' (Hammerton *et al.*, 2002: 552). According to Hammerton *et al.*, shallow parsing typically involves:

- identifying parts of speech;
- segmenting the input string into meaningful chunks (i.e. phrasal or clausal units); and
- determining what relations (e.g. subject, object, etc.) these chunks bear to the main verb.

Depending on comprehension goals or task demands, assigning a full hierarchical representation to an input string may often be unnecessary. There is evidence from L1 processing studies suggesting that native speakers sometimes rely on lexically or meaning-based comprehension heuristics, or compute incomplete syntactic representations that are just 'good enough' for the purpose at hand (Christianson *et al.*, 2001; Ferreira *et al.*, 2002; Sanford and Sturt, 2002). Sentences that express highly implausible propositions, for example, such as the passive sentence *The dog was bitten by the man* are frequently misinterpreted by adult NSs, suggesting that interfering pragmatic information may override the parser's syntactic analysis, or even prevent it from being carried out in full. Evidence from ERPs suggests that strongly plausible predicate–argument combinations may lead the parser to pursue an incorrect syntactic analysis (Kim and Osterhout, 2005), and native speakers' misinterpretations of certain types of garden-path sentence indicate that thematic roles once assigned tend to persist (Christianson *et al.*, 2001). These observations are in line with processing models that assume that L1 comprehension normally involves both the application of lexically based comprehension heuristics and full syntactic analyses (e.g. Townsend and Bever, 2001), and indicate that highly plausible and/or strong canonical meaning or form patterns may sometimes block correct, syntax-derived interpretations.

According to the Shallow Structure Hypothesis for L2 processing, 'late' L2 learners differ from native speakers in that they are largely restricted to shallow parsing. While (sufficiently proficient) learners may well appear indistinguishable from NSs in some domains, we would expect their reduced ability to carry out full syntactic analyses in real time to be reflected, for example, in non-native-like processing of unbounded dependencies. Learners' apparent failure to postulate syntactic gaps observed in both Marinis *et al.*'s (2005) and the current study confirm this prediction, and indicate that learners do not recover complete configurational structures from the input. One possible reason why L2 learners do not perform 'deep' parses even in situations where native speakers do so may be that (part of) their L2 grammatical knowledge is of a form that makes it unsuitable for use in real-time parsing, that is, it may be 'explicit' rather than 'implicit' knowledge (compare, e.g., Paradis, 2004).[8] As a result, learners may be forced to rely on lexical and pragmatic information to a larger extent than NSs in L2

[8]One should bear in mind, however, that a clear-cut separation of grammatical knowledge and processing routines may not in fact be possible. Although the Shallow Structure Hypothesis does not have anything to say about the possible neurophysiological bases of 'shallow' vs. 'deep' processing, it is broadly compatible with Ullman's (2001) claim that post-puberty L2 learners predominantly rely on declarative rather than procedural brain memory systems for the representation and processing of their L2. For more discussion, see Clahsen and Felser (2006).

comprehension. Learners' sensitivity to argument structure, thematic and plausibility information during L2 sentence processing is well attested (see, among others, Juffs and Harrington, 1995; Frenck-Mestre and Pynte, 1997; Juffs, 1998; 2004; Williams *et al.*, 2001; Felser *et al.*, 2003; Papadopoulou and Clahsen, 2003; Felser and Roberts, 2004), and may help compensate for their reduced ability to parse the L2 input in a native-like way.

Alternatively, it is conceivable that L2 learners resort to shallow processing because they lack sufficient WM resources to carry out full syntactic analyses of the input (compare, e.g., Harrington, 1992; Ardila, 2003). This question was not addressed in Marinis *et al.*'s (2005) study, and few published studies exist that have examined the possible influence of WM differences on L2 sentence processing. Although Juffs (2004) found some indication that digit span – but not reading span – affected learners' processing of temporarily ambiguous sentences in the L2, Juffs (2005) did not find any reliable influence of either of these WM measures on the processing of *wh*-dependencies in L2 English. These results fit with our observation that L2 reading span did not affect the participants' performance in the cross-modal priming task, either. Antecedent priming in native speakers, on the other hand, has been found to be influenced by individual WM differences. While the results from Roberts *et al.*'s low-span group are difficult to interpret, the results from the high-span NSs provide evidence for structurally determined gap-filling in L1 processing. Recall that our learners' performance pattern differed from both the high-span and the low span NSs' patterns. The learners' shorter RTs to identical pictures at both test points suggests that they were able to keep the filler active in short-term memory but without reactivating it at the gap site, independently of individual WM capacity as measured by Harrington and Sawyer's (1992) reading-span test.

VII Conclusions

Similar to Marinis *et al.*'s (2005) finding that L2 learners showed no evidence of postulating intermediate syntactic gaps during the processing of long-distance *wh*-dependencies, our results support the hypothesis that the representations constructed during L2 processing lack such abstract grammatical ingredients as movement traces. Contrary to (high-span) NSs, advanced Greek-speaking learners showed evidence of maintained activation but not of structurally determined antecedent reactivation. The lack of any interactions with L2 reading span moreover suggests that the learners' failure to postulate movement traces during real-time processing cannot be attributed to a shortage of WM resources. In both Marinis *et al.*'s and in the current study, the learners had no obvious difficulty understanding complex sentences of the types under investigation, however. We argued that this observation can be accounted for by assuming that learners are able to compensate for their relatively shallower grammatical analyses of the L2 input by exploiting lexical, pragmatic and other non-structural cues to interpretation.

Acknowledgements

Leah Roberts is now working at the Max Planck Institute for Psycholinguistics, Nijmegen. This research was supported by the Leverhulme Trust (Grant No. F/00 213B to H. Clahsen, C. Felser and R. Hawkins). We thank Harald Clahsen and two anonymous *Second Language Research* reviewers for their constructive comments, and Nikolitsa Stathopoulou for providing helpful information about Greek.

VIII References

Abney, S. 1991: Parsing by chunks. In Berwick, R., Abney, S. and Tenny, C., editors, *Principle-based parsing*, 257–78. Kluwer.

Allan, D. 1992: *The Oxford Placement Test*. Oxford University Press.

Ardila, A. 2003: Language representation and working memory with bilinguals. *Journal of Communication Disorders* 36, 233–40.

Balogh, J., Zurif, E., Prather, P., Swinney, D. and **Finkel, L.** 1998: Gap-filling and end-of-sentence effects in real-time language processing: Implications for modeling sentence comprehension in aphasia. *Brain and Language* 61, 169–82.

Baumann, H., Nagengast, J. and **Klaas, G.** 1993: New Experimental Setup (NESU). Ms. Max Planck Institute for Psycholinguistics, Nijmegen.

Bever, T. and **McElree, B.** 1988: Empty categories access their antecedents during comprehension. *Linguistic Inquiry* 19, 35–43.

Chomsky, N. 1981: *Lectures on government and binding*. Foris.

—— 1995: *The minimalist program*. MIT Press.

Christianson, K., Hollingworth, A., Halliwell, J. and **Ferreira, F.** 2001: Thematic roles assigned along the garden path linger. *Cognitive Psychology* 42, 368–407.

Clahsen, H. and **Featherston, S.** 1999: Antecedent-priming at trace positions: evidence from German scrambling. *Journal of Psycholinguistic Research* 28, 415–37.

Clahsen, H. and **Felser, C.** 2006: Grammatical processing in first and second language learners. *Applied Psycholinguistics* 27, 3–42.

Crain, S. and **Wexler, K.** 1999: Methodology in the study of language acquisition: a modular approach. In Richie, W. and Bhatia, T., editors, *Handbook of child language acquisition*, 387–425. Academic Press.

Crocker, M. 1996: *Computational psycholinguistics*. Kluwer.

Daneman, M. and **Carpenter, P.** 1980: Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19, 450–66.

Felser, C., Clahsen, H. and **Münte, T.** 2003: Storage and integration in the processing of filler-gap dependencies: an ERP study of topicalization and *wh*-movement in German. *Brain and Language* 87, 345–54.

Felser, C. and **Roberts, L.** 2004: Plausibility and recovery from garden paths in second language sentence processing. Poster presented at AMLaP, Aix-en-Provence.

Felser, C., Roberts, L., Gross, R. and **Marinis, T.** 2003: The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics* 24, 453–89.

Ferreira, F., Bailey, K. and **Ferraro, V.** 2002: Good enough representations in language comprehension. *Current Directions in Psychological Science* 11, 11–15.

Fiebach, C., Schlesewsky, M. and **Friederici, A.** 2002: Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German *wh*-questions. *Journal of Memory and Language* 47, 250–72.

Francis, N. and **Kucera, H.** 1982: *Frequency analysis of English usage: lexicon and grammar*. Houghton Mifflin.

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

Frazier, L. and **Flores d'Arcais, G.** 1989: Filler-driven parsing: the study of gap-filling in Dutch. *Journal of Memory and Language* 28, 331–44.

Frenck-Mestre, C. and **Pynte, J.** 1997: Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology* 50A, 119–48.

Gaulin, C. and **Campbell, T.** 1994: Procedure for assessing verbal working memory in normal school-age children: some preliminary data. *Perceptual and Motor Skills* 79, 55–64.

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

Gibson, E. and **Hickok, G.** 1993: Sentence processing with empty categories. *Language and Cognitive Processes* 8, 147–61.

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

Gorrell, P. 1993: Evaluating the direct association hypothesis: a reply to Pickering and Barry, 1991. *Language and Cognitive Processes* 8, 129–46.

Hammerton, J., Osborne, M., Armstrong, S. and **Daelemans, W.** 2002: Introduction to special issue on machine learning: approaches to shallow parsing. *Journal of Machine Learning Research* 2, 551–58.

Harrington, M. 1992: Working memory capacity as a constraint in L2 development. In Harris, R., editor, *Cognitive processing in bilinguals*, 123–134. Elsevier.

Harrington, M. and **Sawyer, M.** 1992: Working memory capacity and L2 reading skill. *Studies in Second Language Acquisition* 14, 25–38.

Hestvik, A., Schwartz, R., Tornyova, L. and **Datta, H.** 2005: Picture-naming shows children reactivate antecedents at trace positions. Poster presented at GASLA, Siena, Italy, September 2005.

Holton, D., Mackridge, P. and **Philippaki-Warburton, I.** 1997: *Greek. A comprehensive grammar of the modern language*. Routledge.

Juffs, A. 1998: Main verb vs. reduced relative clause ambiguity resolution in second language sentence processing. *Language Learning* 48, 107–47.

—— 2004: Representation, processing, and working memory in a second language. *Transactions of the Philological Society* 102, 199–225.

—— 2005: The influence of first language on the processing of *wh*-movement in English as a second language. *Second Language Research* 21, 121–51.

Juffs, A. and **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.

Kaan, E., Harris, T., Gibson, E. and **Holcomb, P.** 2000: The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15, 159–201.

Kim, A. and **Osterhout, L.** 2005: The independence of combinatory semantic processing: evidence from event-related potentials. *Journal of Memory and Language* 52, 205–25.

King, J. and **Just, M.** 1991: Individual differences in syntactic processing: the role of working memory. *Journal of Memory and Language* 30, 580–602.

King, J. and **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, 376–95.

Kluender, R. and **Kutas, M.** 1993: Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5, 196–214.

Kluender, R. and **Münte, T.** 1998: Subject/Object asymmetries: ERPs to grammatical and ungrammatical *wh*-questions. Poster presented at the 11th Annual CUNY Conference on Human Sentence Processing, Rutgers University.

Lee, M.-W. 2004: Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research* 33, 51–73.

Love, T. and **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* 20, 5–24.

—— 1997: Real time processing of object relative constructions by pre-school children. Poster presented at the 10th Annual CUNY Conference on Human Language Processing, Santa Monica.

Marinis, T., Roberts, L., Felser, C. and **Clahsen, H.** 2005: Gaps in second language sentence processing. *Studies in Second Language Acquisition* 27, 53–78.

McKee, C., Nicol, J. and **McDaniel, D.** 1993: Children's application of binding during sentence processing. *Language and Cognitive Processes* 8, 265–90.

Miyamoto, E. and **Takahashi, S.** 2002: Antecedent reactivation in the processing of scrambling in Japanese. *MIT Working Papers in Linguistics* 43, 127–42.

Nakano, Y., Felser, C. and **Clahsen, H.** 2002: Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research* 31, 531–71.

Neely, J. 1991: Semantic priming effects in visual word recognition: A selective review of current findings and theories. In Besner, D. and Humphreys, G., editors, *Basic processes in reading: visual word recognition*, 264–336. Erlbaum.

Nicol, J. 1993: Reconsidering reactivation. In Altmann, G. and Shillcock, R., editors, *Cognitive models of speech processing: the second Sperlonga meeting*, 321–50. Erlbaum.

Nicol, J. and **Swinney, D.** 1989: The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18, 5–20.

Papadopoulou, D. and **Clahsen, H.** 2003: Parsing strategies in L1 and L2 sentence processing: a study of relative clause attachment in Greek. *Studies in Second Language Acquisition* 24, 501–28.

Paradis, M. 2004: *A neurolinguistic theory of bilingualism*. John Benjamins.

Pickering, M. 1993: Direct association and sentence processing: a reply to Gibson and Hickok. *Language and Cognitive Processes* 8, 163–96.

Pickering, M. and **Barry, G.** 1991: Sentence processing without empty categories. *Language and Cognitive Processes* 6, 229–59.

Roberts, L., Marinis, T., Felser, C. and **Clahsen, H.** 2007: Antecedent priming at gap positions in children's sentence processing. *Journal of Psycholinguistic Research*, 36(3).

Sanford, A. and **Sturt, P.** 2002: Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences* 6, 382–86.

Snodgrass, J. and **Vanderwart, M.** 1980: A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory* 6, 174–215.

Swinney, D. and **Zurif, E.** 1995: Syntactic processing in aphasia. *Brain and Language* 50, 225–39.

Swinney, D., Onifer, W., Prather, P. and **Hirshkowitz, M.** 1979: Semantic facilitation across sensory modalities in the processing of individual words and sentences. *Memory and Cognition* 7, 159–65.

Townsend, D. and **Bever, T.** 2001: *Sentence comprehension: the integration of habits and rules*. MIT Press.

Ullman, M. 2001: The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition* 4, 105–22.

Williams, J., Möbius, P. and **Kim, C.** 2001: Native and non-native processing of English *wh*-questions: parsing strategies and plausibility constraints. *Applied Psycholinguistics* 22, 509–40.

Whitney, P., Arnett, P., Driver, A. and **Budd, D.** 2001: Measuring central executive functioning: what's in a reading span? *Brain and Cognition* 45, 1–14.

Appendix 1 Experimental sentences used in the cross-modal priming task

1. John saw the peacock to which the small penguin gave the nice birthday present in the garden last weekend.
   PEACOCK *CARROT*
2. Sue called the spider to which the big ostrich showed the small pretty picture at his house yesterday evening.
   BALLOON *SPIDER*
3. Jane loved the tiger to which the black beetle offered the sweet strawberry cake at the party last week.
   TIGER *PAINTBRUSH*
4. James phoned the giraffe to which the gray hippo told the scary ghost story in his room before bedtime.
   VIOLIN *GIRAFFE*
5. Sue phoned the zebra to which the old rhino sent a short thankyou letter at his house yesterday morning.
   ZEBRA *HAMMER*
6. George fed the panda to which the large leopard explained the difficult new game in the garden last Friday.
   APPLE *PANDA*

7. Fred chased the rabbit to which the brown eagle gave the small chocolate biscuit in the park last Monday.
   RABBIT *NEEDLE*
8. Bob loved the monkey to which the fat squirrel showed his excellent new trick in the playground last month.
   TOOTHBRUSH *MONKEY*
9. Ben liked the lobster to which the young camel offered the delicious melon at the beach last Saturday.
   LOBSTER *BUTTON*
10. Jack knew the donkey to which the nice tortoise told his most naughty secret by the river last week.
    ONION *DONKEY*
11. Jo knew the ostrich to which the black spider explained the difficult problem at school last Monday.
    OSTRICH *BALLOON*
12. John called the beetle to which the fat tiger showed his favorite photographs in the playground yesterday afternoon.
    PAINTBRUSH *BEETLE*
13. Sue saw the hippo to which the tall giraffe gave the sweet tasty orange in the jungle yesterday afternoon.
    HIPPO *VIOLIN*
14. Jane knew the rhino to which the big zebra told a really funny joke at school yesterday morning.
    HAMMER *RHINO*
15. James hit the leopard to which the old panda offered a very large ice-cream in the cinema after his lunch.
    LEOPARD *APPLE*
16. George liked the eagle to which the brown rabbit gave the biggest piece of cake at the party last Saturday.
    NEEDLE *EAGLE*
17. Fred chased the squirrel to which the nice monkey explained the game's difficult rules in the class last Wednesday.
    SQUIRREL *TOOTHBRUSH*
18. Bob fed the camel to which the pink lobster showed his new computer game at his office on Monday morning.
    BUTTON *CAMEL*
19. Ben saw the tortoise to which the grey donkey gave the small expensive gift at the party last weekend.
    TORTIOSE *ONION*
20. Jack liked the penguin to which the bright peacock sent a nice Christmas present in the post last year.
    CARROT *PENGUIN*