Probing for Intermediate Traces in the Processing of Long-Distance *Wh*-Dependencies in English as a Second Language

Laurent Dekydtspotter and A. Katherine Miller
Indiana University

1. Movement traces and second language processing

A body of research argues that unlike native speakers (NSs), non-native speakers (NNSs) cannot access a structural reflex during sentence processing and rely instead on shallow structures dependent on lexical-thematic information and context (Clahsen & Felser, 2006a, b; Papadopoulou & Clahsen, 2003; Felser, Roberts, Gross, & Marinis, 2003; Marinis, Roberts, Felser & Clahsen, 2005; Felser & Roberts, 2007). Thus, in experiments examining ambiguous relative clause attachment in structures of the type N1 *of* N2 RC, Dussias (2001, 2003) for second language (L2) Spanish, Papadopoulou and Clahsen (2003) for L2 Greek, and Felser, Roberts, Gross, and Marinis (2003) for L2 English, noted that their L2 learner groups showed no attachment preferences in the reading times of the disambiguating segment unlike the native groups. They argued that lack of L1 effects or convergence revealed an inability of NNSs to rely on syntactic information in disambiguation. In contrast, Frenck-Mestre (2002) found L1 effects and convergence on target in English-French and Spanish-French comparisons. Focusing on interactions of context and prosody with syntax, Dekydtspotter, Donaldson, Edmonds, Liljestrand Fultz and Petrush (2008) found evidence of autonomous reflexes in the processing of RC. This challenges the Shallow Structure Hypothesis.

Marinis, Roberts, Felser, and Clahsen (2005) and Felser and Roberts (2007) point to the presence/absence of movement trace as a sort of trace-theoretic litmus test in adjudicating the nature of L2 sentence processing. Indeed, a body of processing evidence points to traces in L1 processing of long distance dependencies (Bever & McElree, 1988; Frazier & Clifton, 1989; Gibson & Warren, 2004; Love & Swinney, 1996; Nicol & Swinney, 1989). Psycholinguistic research reveals evidence suggesting that a long-distance dependency between a filler and a gap inside an embedded clause as in (1a) requires cyclic movement crucially implicating traces, i.e. copies of moved expressions, as in (1b) (Gibson & Warren, 2004).

(1) a. Who did the consultant claim that the proposal had pleased?
b. [Who did the consultant claim [<who> that the proposal had pleased <who>]]

Intermediate traces at clause edge are the presumed signatures of a processing mechanism highly specific to language design, which incorporates a grammatical apparatus that requires traces at the edge of phases as in standard generative theory. If L2 processing can be shown to require intermediate traces at clause edge (in particular), then this would suggest that L2 processing follows from a universal processing mechanism.

Thus, Marinis, Roberts, Felser, and Clahsen (2005) devised an experiment, based on Gibson and Warren (2004) with four groups of Chinese, Japanese, Greek and German speakers of English as a second language and a group of English NSs. The learners were similar in age, had generally started learning English in their eleventh or twelfth year (apart from Greek learners with a mean age 8.67), and seemed to be within the same general proficiency range. They were given a grammatical

 Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009), ed. Melissa Bowles et al., 113-124. Somerville, MA: Cascadilla Proceedings Project.

questionnaire with 20 items as in (2) in order to show that learners could indeed understand the type of sentences under study. All groups tested above 90%.

(2) 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 respondents were tested on sentences such as (3a-d). (3a, b) involved the movement of a *wh*-expression *who*. (3c, d) did not involve movement. (3a, b) differed in the following way: (3a) involved cyclic movement with a trace, a silent copy of the moved item (<who>), at the edge of the embedded clause. (3b) did not.

(3) a. The manager who/ the secretary claimed/ <*who*> that/ the new salesman/ had pleased <*who*>/ will raise company salaries.
b. The manager who/ the secretary's claim/ about/ the new salesman/ had pleased <*who*>/ will raise company salaries.
c. The manager thought/ the secretary claimed/ that/ the new salesman/ had pleased/ the boss in the meeting.
d. The manager thought/ the secretary's claim/ about/ the new salesman/ had pleased/ the boss in the meeting.

Marinis, Roberts, Felser, and Clahsen's (2005) NS results replicated those obtained by Gibson and Warren (2004). NSs produced longer reading times for *that* in (3a) versus (3c), but shorter reading times on the verb segment *had pleased* in (3a) versus (3b). In (3a), reading slowed down on the complementizer as a result of the need to accommodate the trace at the edge of the embedded clause, but this provided a processing advantage for later trace integration in thematic position. As pointed out by Gibson and Warren (2004) following (Gibson, 1998, 2000), processing the sentences in (3a, b) requires accommodating a referent in memory during the processing of the *filler*—thereby decreasing the filler activation. Intermediate traces by re-activating fillers reduce computational load at the time of integration by restoring the activation-level of the referent. In stark contrast, learners did not show these effects. Their processing of verbs in extraction contexts took significantly longer than in non-extraction contexts.

Even if L1 and L2 sentence processing obey the same principles, there is no guarantee that NNSs' and NSs' reading will reflect the very same abstract computational moments as they read a given segment, given processing lags induced by under-learned lexical access routines *inter alia*. We refer readers to Dekydtspotter, Schwartz and Sprouse (2006) for discussion of factors that may confound Marinis, Roberts, Felser and Clahsen's (2005) conclusion. In short, if activations are weak in L2 processing by virtue of slower lexical access in the limits of fixed resources, then the relative enhancement induced by the reactivation provided by the trace in (3a) may not be sufficient to have a measurable impact on the integration of the filler with the thematic verb. Hence, the sentence analyzer could deploy grammatical representations of the classical kind, but speed of access interacting with computational complexity could produce different effects.

For their part, Felser and Roberts (2007) used a cross modal priming task to detect sentence processing that is reliant on a trace-theoretic parse. Such processing includes dedicated reactivation moments in contradistinction to processing that is shallow which lacks these structurally determined reactivations. Their experiment focused on the ability of traces to prime their antecedents. Sitting in front of a computer monitor, advanced Greek learners of English listened to sentences as in (4a-d). As respondents listened, a picture of the antecedent [SQUIRREL] (4a, b) or of an unrelated object [TOOTHBRUSH] (4c, d) appeared on a computer screen either in gap position (4a, c) where the silent trace of *to which* is expected, or in control position (4b, d) before this point. Respondents pressed a button to indicate whether the picture depicted something alive or not. Periodic comprehension checks verified that participants were actively listening, and that they understood the sentence structure.

(4) Fred chased the squirrel to which the nice monkey explained…
a. the game's difficult rules [SQUIRREL] in class last Wednesday.
b. the game's [SQUIRREL] difficult rules in class last Wednesday.
c. the game's difficult rules [TOOTHBRUSH] in class last Wednesday.
d. the game's [TOOTHBRUSH] difficult rules in class last Wednesday.

The results were compared with those from four groups of NSs (high and low working memory (WM) adults and children) in a study by Roberts, Marinis, Felser, and Clahsen (2007) on the same stimulus. High WM NSs reacted significantly faster to pictures related to the filler than to pictures unrelated to it, although crucially not in pre-gap positions. Non-native speakers did not produce differences between gap and pre-gap positions, reacting more quickly overall to related than to unrelated probes. In native speakers' processing, the facilitation on filler-related probes was limited to the gap position, suggestive of reactivation at this point only. This is consistent with the presence of a trace which reactivates the filler. For Felser and Roberts (2007), learners' processing shows priming evidence consistent with the Direct Association method (Pickering & Barry, 1991; Sag & Fodor, 1995; Traxler & Pickering, 1996; Pickering, 1999) expected with shallow representations: facilitations on filler-related probes are insensitive to positions because the filler must be kept active in working memory until which point it is resolved with the verb. This explains the nature of this asymmetry. Unfortunately, this is not the only explanation for the learners' behavior. Listening comprehension often carries a certain level of difficulty as learners must attune to talker, speech rate etc., which may have encouraged learners to deploy task-specific strategies. In this task, the related/not related dimension of probes matched the alive/not alive distinction so that while faced with the need to classify probes as alive or not alive, very cautious learners could have actively scanned for referents in the sentence, coupling alive/not alive decisions with match/mismatch verifications. Mismatches are expected to show inhibitions, deriving the results. This response strategy also produces indiscriminate facilitations on alive/related probes in learners (cf. Clahsen & Featherton, 1999).

2. Picture classification during reading

In a cross-modal priming paradigm, respondents listen to sentences. A picture appears on a computer screen at selected moments. Respondents make simple categorization decisions, and their RTs are analyzed. A categorization decision that is concomitant with the processing of a trace that renders a corresponding referent active in memory is expected to be speedier than a categorization decision that is dissociated from such processing. Speedier classification of a filler-related probe at a processing point implicating a trace of that filler with respect to an earlier control position goes contrary to general complexity expectations based on distance. The fact that facilitation is obtained in spite of the greater distance provides compelling evidence for a syntactic reflex. This effect is often subject to working memory (WM) capacity, with low working memory individuals often not exhibiting the effect.

The task presented here follows the same logic. However, since typical instructed learners receive input through the written medium, reading might provide NNSs with greater support than listening and would therefore allow for the testing of respondents at lower levels of proficiency. Thus, our task required respondents to classify a probe while reading a sentence presented word by word on a computer screen in a low voice to themselves. At given moments, pictures appeared at the center of the screen for 650 milliseconds. Respondents classified these pictures as human or non-human as quickly as possible. RTs were recorded. Reading speed was controlled using DMDX (Forster & Forster, 2003) and calibrated so that reading was fluent. Non-native presentation was slightly slower than that for NSs: word segments appeared for 30 ms longer, and had an increased inter-item duration of 100 milliseconds. The instrument included 40 items involving cyclic movement across a clause as in (5).

(5) Harry is who$_i$ Mary said on Monday [t_i that the headmaster congratulated t_i at the assembly].

These sentences were given a context that rendered them felicitous. Respondents were told that they were to read a classroom log written by a substitute teacher. As a prank, children in the class had systematically switched names. The substitute teacher recorded a series of events to help the regular teacher sort things out. Respondents read the log and classified pictures as human or not.

Critical sentences were organized in a Latin square: Four conditions crossed (target/control) positions with filler-related probes matching the antecedents of the *wh*-pronoun *who* or non-filler-related probes matching the intervening NP as in Figure 1. The sentences included 10 girl-boy pairs and 10 boy-girl pairs as well as 10 pictures of boys and girls equally divided among them. There were 20 distracter items involving animals (9 primates and 11 others). The control positions were located 600 milliseconds before target positions, occurring within a prepositional phrase that offered a protective zone from any trace temporarily associated with the complement position of the verb.

Condition 1: target position, filler-related probe
Harry / is / who / Mary / said / on / Monday / that / [probe: boy] / the headmaster / congratulated / at the assembly.
Condition 2: target position, non-filler-related probe
Harry / is / who / Mary / said / on / Monday / that / [probe: girl] / the headmaster / congratulated / at the assembly.
Condition 3: control position, filler-related probe
Harry / is / who / Mary / said / on / [probe: boy] / Monday / that / the headmaster / congratulated / at the assembly.
Condition 4: control position, non-filler-related probe
Harry / is / who / Mary / said / on / [probe: girl] / Monday / that / the headmaster / congratulated / at the assembly.

Figure 1. Sample test item

The 20 probes used for the critical sentences were pictures of boys and girls obtained from Microsoft Office Clipart and adjusted to similar size using Photoshop. They appeared uniformly at the center of the screen. There were four versions of the task so that the pictures of young males and females cycled through the four conditions across the experiment. The names used in the experiment were common and unambiguously male or female: Harry, Adam, Matt, Kevin, Jason, Mark, Bill, Tom, Dave and John for the boys, and Mary, Beth, Amy, Sue, Karen, Julie, Jill, Ann, Kate and Jane for the girls. Respondents were tested to make certain that they knew the gender associated with each name. Results show that the gender of the names were known to non-native participants. The probes occurred in environments where a common noun was disfavored. While the form *that* is ambiguous between a treatment as a complementizer and as a distal demonstrative, the matrix verbs were selected so that a propositional complement was expected, following Marinis, Roberts, Felser, and Clahsen (2005). Additionally, the need to integrate the *wh*-expression *who* when any matrix-clause options have been exhausted, strongly militates against the probability of the parser having committed to a common noun analysis as the probe was introduced. The probe in control position after the proposition *on* appeared in a position where a common noun is disfavored but likewise not excluded as *on girl/boy night out*.

Figure 2. Examples of probes

Modulo the effect of WM limitations, a typical priming profile should be as follows: Classification of pictures related to directly semantically antecedents at clause edge (Condition 1) will be facilitated with respect to pictures related to non-antecedents at clause edge (Condition 2). Classification of pictures directly semantically related to antecedents at clause edge (Condition 1) will be facilitated with respect to pictures related to antecedents in control position (Condition 3). Finding such patterns in a group of respondents would constitute priming evidence for traces. Crucially, priming expectations depend on the spreading of activations such that activating a concept sends a wave of activations through its semantic super-ordinate categories, and then possibly to its dependents. Activating 'Harry' makes the concepts 'Male" and 'Human' more salient, as they are likely to be relevant to understanding the broader context. This said, it could also be that activation spreads not only from the individual concept 'Harry' to the super-ordinate categories, which we assume to be the strongest line of spreading, but also to the sister category 'Female' under 'Human', which would also be activated. Still, finding filler-related priming at clause-edge versus control positions would, when the relevant control conditions are more stable, present priming evidence for traces. The exact contour of the priming effect largely depends on the strength of spreading through the semantically related concepts on which the probes rely.

Crucially, a facilitation profile that relies on activation spreading is not the only possible indicator of reliance on movement traces. In priming experiments with college-aged English NSs using newly acquired technical vocabulary (e.g. *accipiter,* the genus of goshawks and sparrow-hawks) as primes, inhibitions arose on related terms such as *eagle* in lieu of facilitations (Dagenbach, Carr & Barnhardt, 1990). This is explained by a dampening mechanism on related semantic categories associated with preserving a weak activation due to an under-routinized lexical access (Carr & Dagenbach, 1990). Similarly, the expectations of the 'priming' profile are reversed, given the under-routinized lexical access in (intermediate) L2 learners (Favreau & Segalowitz, 1983; Segalowitz & Segalowitz, 1993). In the event of weak activation of 'Harry', it becomes locally more strategic to dampen the super-ordinate concepts 'Male" and 'Human,' in order to focus resources on the weakly activated 'Harry'. In other words, the local need to preserve the activation takes precedence over the benefits of spreading activation, since if the activation dies all is lost. Thus, as the individual concept 'Harry' is re-activated at clause edge, the super-ordinate concept 'Male' is dampened (Condition 1), classifications of Male-pictures will be inhibited with respect to Female-pictures and vice versa (Condition 2). This is because inhibitions will presumably affect more categories in the super-ordinate relations (direct semantic line) than categories in sister lines, although a neighborhood effect is possible. Classification of Male pictures related to Male antecedents reactivated at clause edge (Condition 1) will be inhibited with respect to Male-pictures in control position (Condition 3). This is because inhibitions of the super-ordinate categories will be increased at reactivation points. Similarly, inhibition effects for non-filler related pictures matching the subject NPs are possible in Condition 4 due to locality of the activation, again due to dampening of super-ordinate categories.

After filling out a background questionnaire providing biographical information, respondents did a short picture classification task as a pretest, during which the pictures that appeared in the main experimental task were presented to the respondents, who classified them as human or non-human. This pretest provided an individual baseline. For each respondent, we kept track of which (sets of) pictures would appear in which conditions in the main experiment, so that comparisons could be made between picture classification alone and picture classification during language processing. Respondents also did a version of Harrington & Sawyer's (1992) WM reading span test modeled for L2 learners of English after Daneman and Carpenter (1980). Respondents read aloud 42 sentences organized in 12 sets increasing in size from 2 to 5 sentences per set by groups of 3. They needed to indicate whether they had detected ungrammaticalities and recall the last word of each sentence. After each set, respondents recorded the last word of each sentence in order on an answer sheet. WM scores are the total number of recalled words out of 42, and the number of completely correct sets out of 12. For NSs, the sentences were longer by 5 to 7 words, making them comparable in length to Daneman and Carpenter's (1980) task. After completing the WM reading span test, respondents did the main experimental task: a picture classification task during reading. An investigator verified their knowledge of the gender associated with each name and verified whether speed was appropriate. We collected the results of an in-house placement test that measured listening and reading comprehension, supplemented with a short essay. Brown's (1980) cloze-test was administered as an external proficiency measure. However, some subjects did not come back for the second session to do this test. These results, nevertheless, confirmed the intermediate proficiency level of the subject pool.

NS and NNS respondents received class credit for their participation. A control group of 45 undergraduate university students of similar age provided a validation group for our instrument. They are enrolled in introductory linguistics courses and had no knowledge of the structure of clauses of the type that were investigated. Their average WM word score was 36/42, with range from 24 to 42. Their average WM set score was 8/12, with range from 2 to 12. Respondents were split into a lower WM group and a higher WM group with above average scores. We grouped respondents based on set scores since the distribution in the two groups aligns closely. We tested 61 undergraduate NNSs in their early twenties, enrolled in English language support courses for matriculated students: Korean (n = 38), Mandarin (13), Cantonese (04), Indonesian (02), Japanese (02), Hindi (1), Arabic (1). The entirety of the data set was reported in Dekydtspotter, Miller, Chang, Kim, & Schaefer (2008) since priming often requires large numbers of subjects. We now focus on the 17 Chinese learners and 38 Korean learners for which there is sufficient data. They were of intermediate proficiency. Chinese learners had an average placement score 58/80, with a range of 41-70 on a battery of tests. The average WM score per word was 36/42, with a range from 28 to 42. The average WM score per set was 8/12, with a range of 3-12. Korean learners had an average placement score 62/80, with a range of 51-72 on a battery of tests. The average WM score per word was 33/42, with a range from 18 to 42. The average WM score per set was 7/12, with a range of 2-12. These NNSs are developing English proficiency, and a goal is in part to determine the type of processing that learners bring to the learning task.

The following analytical procedures were followed. The pretest was used to examine RTs for picture classification alone. For each respondent, we obtained a mean value for the pictures used in each of the four conditions and can therefore examine the difference between pretest means and means in the main task, obtaining values corrected for effects due to the pictures themselves. Only correct classifications were considered. Measures two standard deviations beyond the means were discarded. Missing values were replaced by means. For each respondent, mean RTs per condition were then obtained. Based on the predictions above, we carried a two-tailed *t*-test per hypothesis. For planned *t*-tests, the α-level is set at .05. Thus, we examined RTs in Condition 1 (C1) versus Condition 2 (C2) as well as RTs in Condition 1 (C1) versus Condition 3 (C3). A significant asymmetry in C1 vs. C3 for instance, when RT differences in the counterparts conditions C2 vs. C4 are for all intents and purposes flat would implicate the presence of a trace. On the best scenario, a mixed-design repeated measures ANOVA on RTs with WM group as a between-subject factor and (target/control) position and (related/non-related) probe as within-subject factor will exhibit an interaction of position and probe.

Given the spreading and dampening mechanisms related to activations of referents, it is paramount to examine individual patterns of distribution, seeking to determine whether learners exhibit patterns of behavior consistent with these mechanisms. Although the breadth of the spreading and dampening in the web of semantic relations creates a factor that can create disunity, RTs are predicted to co-vary on probes directly semantically related to recently (re)activated referents as in C1 and C4 reflecting either activation spreading to or dampening on the super-ordinate category. Likewise, probes not directly semantically related to recently (re)activated referents as in C2 and C3 should also exhibit similar response time tendencies.

3. Results

The results for the pretest by L1 groups are presented in Table 1. RTs for pictures appearing in all four conditions were extremely flat. This is as expected, since the pictures rotated across all four conditions. Any asymmetry appearing across conditions in the experimental task can be attributed to sentence processing. These RTs allowed us to examine individual differences: For each respondent, we took into consideration individual variation in responses to pictures and obtained corrected measures. There was no difference between NSs and NNSs given the fact that no language processing is involved.

Table 1
RTs in milliseconds to pictures appearing in the four conditions of experimental task

	NSs (n = 45)	Chinese NNSs (n = 17)	Korean NNSs (n = 38)
Pictures in C1	503 (54)	495 (71)	497 (60)
Pictures in C2	500 (58)	484 (70)	499 (57)
Pictures in C3	488 (52)	478 (65)	505 (53)
Pictures in C4	495 (57)	489 (63)	490 (47)

(Standard deviations in parentheses)

The results of the experimental task by L1 group and WM are provided in Table 2. An ANOVA revealed a main effect of position, $F(1, 43) = 17.602$, $p = .0005$. There was no significant effect of working memory. Pooled results are given in Figure 3. NS data produced a theoretically relevant facilitation in C1 (filler-related probes in test position) (538 ms, SD = 53) versus C3 (559 ms, SD = 57) (filler-related probes in control position): $t(44) = 3.767$, $p < .0005$. RT asymmetries in C2 (non-filler-related probes in test position) (537 ms, SD = 53) versus C4 (non-filler-related probes in control position) (547 ms, SD = 44) did not reach statistical significance, $p = .078$. The effect cannot therefore fully be attributed to position, but suggests facilitation associated with the postulation of a trace, evidence of priming.

Table 2
RTs in milliseconds in the task of picture classification during reading

	NSs		Chinese NNSs		Korean NNSs	
Condition	Lower WM (n=26)	Higher WM (n=19)	Lower WM (n=8)	Higher WM (n=9)	Lower WM (n=22)	Higher WM (n=16)
C1	542 (58)	533 (47)	593 (50)	603 (32)	577 (57)	569 (59)
C2	538 (51)	535 (57)	573 (68)	564 (41)	581 (65)	563 (54)
C3	558 (64)	560 (49)	572 (70)	586 (52)	593 (50)	562 (61)
C4	550 (45)	544 (44)	593 (69)	592 (36)	588 (58)	571 (61)

(Standard deviations in parentheses)

For the Chinese learners, however, an ANOVA revealed a crucial interaction of position with probe, $F(1, 15) = 9.217$, $p = .008$, with no effect of working memory. Pooled results are again in Figure 3. In the test position, t-tests revealed inhibitions with filler-related probes (C1) (598 ms, SD = 40) versus non-filler-related probes (C2) (568 ms, SD = 54), $t(16) = 3.303$, $p = .004$. For non-filler-related probes, there was also a significant inhibition in the closer control position C4 (592 ms, SD = 53) relative to the test position C2 (568 ms, SD = 54), $t(16) = 3.495$, $p = .003$. As evident in Figure 3, the statistics for Korean learners were all clearly non-significant: C1 (574 ms, SD = 57), C2 (573 ms, SD = 61), C3 (580 ms, SD = 56, C4 (581 ms, SD = 59).

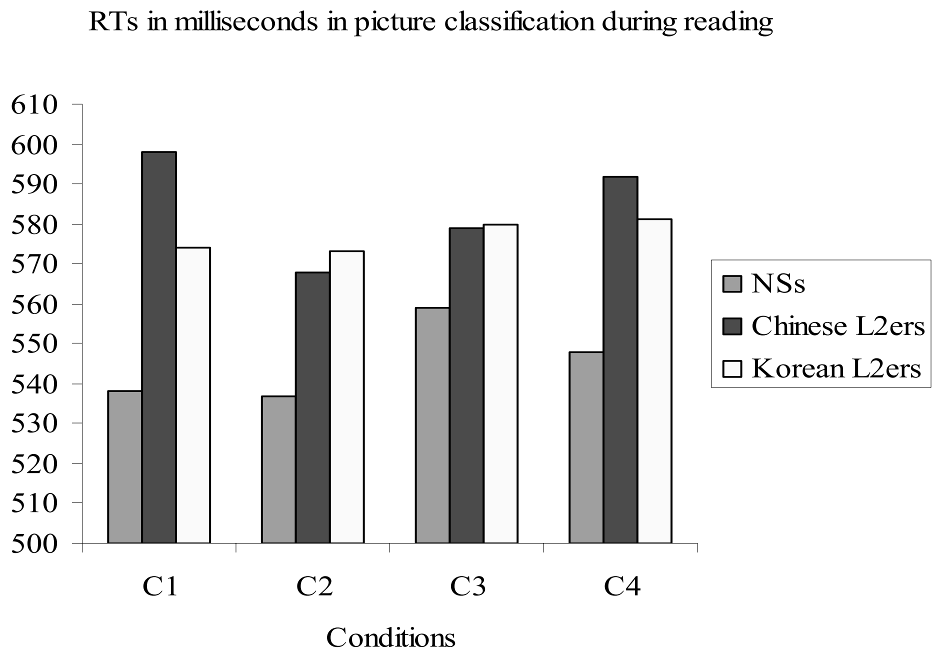

Figure 3. Response times by conditions by language groups

The corrected measures are in Table 3. For the NSs, an ANOVA revealed a main effect of position, $F(1, 43) = 26.722$, $p < .0005$ as well as a marginal interaction of position and probe, $F(1, 43) = 3.828$, $p = .057$. A highly statistically significant contrast was again observed between C1 (56 ms, SD = 8) and C3 (66 ms, SD = 10), $t(44) = 5.566$, $p < .0005$; This effect was significantly greater than the marginal contrast between C2 (37 ms, SD = 59) and C4 (52 ms, SD = 52), $t(44) = 1.946$, $p = .058$. In *t*-tests, Chinese learners revealed theoretically significant inhibitions with filler-related probes (C1) (103 ms, SD = 73) versus non-filler-related probes (C2) (84 ms, SD = 69), $t(16) = 1.891$, $p = .077$. Again, statistics for Korean learners did not reach significance: C1 (76 ms, SD = 62), C2 (74 ms, SD = 51), C3 (75 ms, SD = 66, C4 (91 ms, SD = 59). Results are maintained.

Table 3
Corrected RTs in milliseconds in the task of picture classification during reading

	NSs		Chinese NNSs		Korean NNSs	
Condition	Lower WM (n=26)	Higher WM (n=19)	Lower WM (n=8)	Higher WM (n=9)	Lower WM (n=22)	Higher WM (n=16)
C1	33 (60)	37 (50)	102 (78)	104 (75)	67 (62)	88 (62)
C2	26 (60)	51 (55)	83 (72)	85 (72)	73 (56)	75 (44)
C3	65 (66)	79 (66)	110 (32)	94 (71)	83 (64)	63 (70)
C4	44 (53)	64 (48)	101 (66)	104 (54)	93 (60)	87 (60)

(Standard deviations in parentheses)

NSs produced facilitations, whereas Chinese learners produced inhibitions. Korean learners did not reveal any asymmetries in group results. We therefore considered Korean learners' individual results for confirmation of tendencies suggestive of two activation patterns. Inhibition/facilitations are predicted to co-vary in crucial ways. In C1 and C4, RTs on probes related to recently (re)activated referents will reflect either spreading or dampening of activations affecting super-ordinate categories.

In C2 and C3, probes unrelated to recently (re)activated referents should exhibit similar patterns. The distribution of subjects provided in Figure 4 examines the relationship between RTs in Condition 1 (the x-axis) and in C4 (the y axis). Pearson's product-moment correlation statistics $r = .8$ indicated a solid correlation that is statistically significant, $p < .0005$. Such correlations were found for the Chinese learners (n = 17), Pearson's $r = .7$, $p = .004$, and for native speakers (n = 45), Pearson's $r = .7$, $p < .0005$. Likewise, Figure 5 shows RTs in Condition 2 (the x-axis) and Condition 3 (the y axis) to be solidly linearly correlated: $r = .8$, $p < .0005$. Again, solid correlations were found for Chinese learners (n = 17), Pearson's $r = .9$, $p < .0005$, and for native speakers (n = 45), Pearson's $r = .8$, $p < .0005$. These patterns are theoretically significant.

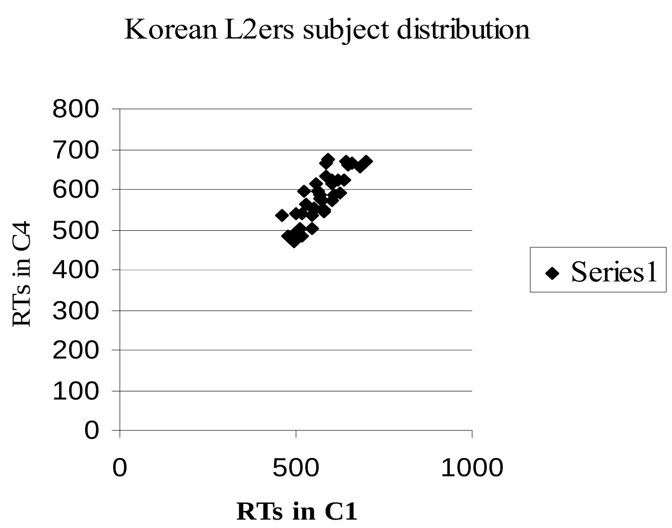

Figure 4. Probes related to recently (re)activated referents

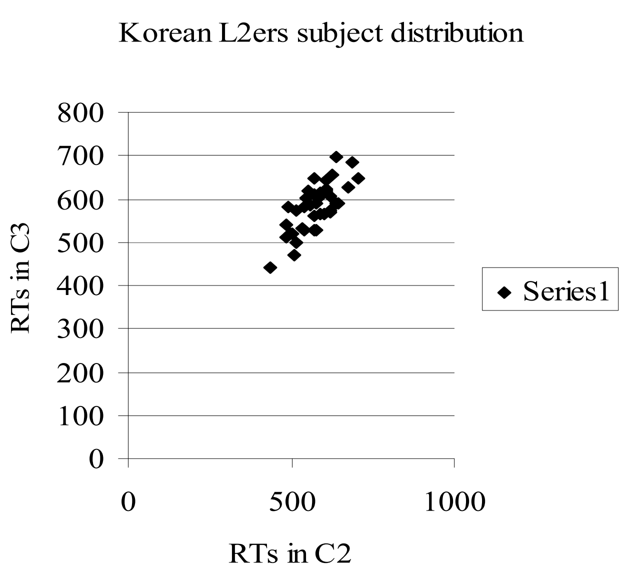

Figure 5. Probes unrelated to recently (re)activated referents

4. Weak activation and L2 representations

In summary, ANOVAs examining results from each language group revealed no effect of WM within the language groups. Native speakers exhibited highly significant facilitation of filler-related probes in test positions, suggestive of priming by movement traces. Perusal of individual data also showed individuals who exhibited inhibitions, although the facilitation profile dominated in the aggregate data. Chinese learners provided inhibitions of filler-related probes in test positions. In the corrected RTs, controlling for individual variation in responses to pictures, a crucial interaction of position with probe was obtained. This interaction points to specific computations at clause-edge that are naturally explained under a trace-theoretic parse. The difference between the (expected) facilitation profile and the obtained inhibition profile in the Chinese learners is suggestive not of major representational differences, but of major differences in the predominant activation mechanism: dampening of activations on (immediate) super-ordinate categories to focus resources on the referent itself. The aggregate data of Korean learners were non-instructive. However, the individual patterns were revealing: Correlation betweem RTs on C1 and C4 in which probes were directly semantically related to recently (re)activated referents and likewise on C2 and C3 in which probes were not directly semantically related to recently (re)activated also point to structurally dependent processing. The entirety of the data set is explained, if (re)activations of referents are accompanied by two activation mechanisms that induce distinct facilitation/inhibition profiles. This explanation is elegant, we believe, and motivated by independent mechanisms of lexical access. Activation spreading to (immediate) super-ordinate categories (yielding facilitations) allows for speedy access to contextual inferences. Activation dampening on (immediate) super-ordinate categories (yielding inhibitions) focuses resources on the primary activation of the referent. In the light of such results, it is unfortunate that mechanisms of activation have been excluded from the theoretical discussion thus far. This is a significant oversight, as results on a priming task are dependent on them. In fact, the low WM children in Roberts *et al.* (2007) produced inhibition on filler-related probes in test position. The Chinese-English learners in the current study seem to be behaving similarly to certain groups of NSs.

An alternative account could point to the distance between NP and probe related to that NP and or to the relative salience of the referent, i.e. first mention or last mention effects. The expectation is that distance induces inhibitions as the activation decays as new referents are introduced. This is consistent with inhibitions of filler-related probes versus non-filler-related probes in test positions, but inconsistent with inhibitions of non-filler related probes in control versus test positions. In terms of salience, we expect that a more salient referent would induce facilitations. In control position, the referent to which the *wh*-expression is linked seems more salient than the closest referent, although the intervening NP should have decreased the activation of the filler. The last-mentioned NP is least salient locally as the asymmetry between C2 and C4 shows. It is possible to entertain this as salience modulated by distance, but without any specific reason why this would be the case, this does not reach the level of a hypothesis. Furthermore, correlations between RTs in C1 and C4 (involving probes directly semantically related to recently activated referents) as well as between RTs in C2 and C3 (involving probes not directly semantically related to recently activated referents) in all groups—irrespective of facilitation or inhibition trends and including a group of learners that does not exhibit any group preference at all—speak loudly to specific syntactically motivated processing moments. Thus, activation mechanisms, including a dampening mechanism used to maintain weakly activated representations and a spreading mechanism used to make contextual inferences, best fit the entirety of the empirical evidence.

5. Perspectives on weak activations

Priming relies on the fact that activating a concept sends a wave of activations through its semantic network. Activation spreading helps establish the global context. However, the need to preserve representations that are weakly activated, presumably as a result of under-routinized lexical access, requires dampening related semantic concepts. When activation spreading is replaced by dampening, facilitations are replaced by inhibitions. The full patterns of asymmetries and correlations produced by L2 learners follow elegantly. In Marinis, Roberts, Felser and Clahsen's (2005) experiment, the cost of integrating a moved object with the thematic verb, independently of the presence of an intermediate trace in the representation, follows from the fact that the benefits afforded by reactivation at clause edge are limited by weak activation.

Weak activation of referents provides a simple and constrained working hypothesis. Given general activation mechanisms, complex behavior in the processing of *wh*-dependencies in a second language follows from differences in lexical access (Favreau & Segalowitz, 1983; Segalowitz & Segalowitz, 1993). In contrast, Clahsen and Felser's (2006a, b) Shallow Structures Hypothesis is extremely broad, claiming the absence of a syntactic reflex in L2 sentence processing caused by a fundamental epistemological difference between L1 and L2 grammars. This latter claim is at odds with a solid and massive body of evidence on the UG-constrained nature of L2 grammars (Dekydtspotter 2009, Slabakova 2008, White, 2003). In contradistinction to weak activation, the Shallow Structures Hypothesis does not predict the interactions and correlations that we found. Our results suggest that research on the processing of *wh*-dependencies in sentence processing must give full consideration to activation mechanisms or else a significant piece of the puzzle is missing.

Acknowledgements

The data were collected in collaboration with Yufen Chang, Ock-Hwan Kim, and Vance Schaefer. This study was made possible in part with support from the Department of Second Language Studies. This paper reflects comments we received at GASLA for which we are grateful. All remaining errors are ours.

References

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

Brown, J.D. (1980). Relative merits of four methods for scoring cloze tests. *Modern Language Journal, 64,* 311–317.

Carr, T., & D. Dagenbach (1990). Semantic priming and repetition priming from masked words: evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology, 16,* 341-350.

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

Clahsen, H., & C. Felser (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27,* 107-126.

Clahsen, H., & C. Felser (2006b). Grammatical processing in language learning. *Applied Psycholinguistics, 27,* 3-42.

Dagenbach, D., T. H. Carr, & T. M. Barnhardt (1990). Inhibitory semantic priming of lexical decisions due to failure to retrieve weakly activated codes. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 16, 328-340

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

Dekydtspotter L., & R.A. Sprouse (2003). L2 Performance: Interlanguage representations, computations, and intuitions. In J.M. Liceras, H. Zobl, & H. Goodluck (eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002): L2 Links* (pp. 45-54). Somerville, MA: Cascadilla Press.

Dekydtspotter, L., B. Donaldson, A.C. Edmonds, A. Liljestrand Fultz, & R.A. Petrush (2008). Syntax and prosody in English-French processing of relative clause attachment. *Studies in Second Language Acquisition, 30,* 453–480.

Dekydtspotter, L, B.D. Schwartz, & R.A. Sprouse (2006). The comparative fallacy in L2 processing research. In M. Grantham O'Brien, C. Shea, & J. Archibald (eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006): The Banff Conference* (pp. 33-40). Somerville, MA: Cascadilla Press.

Dekydtspotter, L., K. Miller, Y Chang, O.-H. Kim, & V. Schaefer (2008). Clause-edge reactivations of fillers in processing English as a second language. Talk delivered at Second Language Research Forum 2008, October 17, University of Hawaii, Honolulu.

Dekydtspotter, L. (2009). Second language epistemology. *Studies in Second Language Acquisition,* 31, 291-321.

Dussias, P. E. (2001).Sentence Parsing in fluent Spanish-English bilinguals. In J. Nicol (Ed.), One Mind, Two languages: Bilingual Language Processing (pp.159-176). Cambridge: Blackwell.

Dussias, P. E. (2003). The effect of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition, 25,* 529-557.

Favreau, M., & N. Segalowitz (1983). Automatic and controlled processes in the first and second language reading of fluent bilinguals. *Memory & Cognition, 11,* 565–574.

Felser, C., & L. Roberts (2007). Processing *wh-* dependencies in a second language: A cross-modal priming study. *Second Language Research, 31,* 9-36.

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

Forster, K.I., & J.C. Forster (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments and Computers, 35*, 116-124.

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

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R.R. Heredia & J. Altarriba (eds.), *Bilingual sentence processing* (pp. 217-235). Amsterdam: Benjamins.

Frenck-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research, 21,* 175-198.

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

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition* 78:1-76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In *Image Language and Brain,* (Eds.) Y. Miyashita, A. Marantz and W. O'Neil, (pp. 95-126), Cambridge, Mass.: MIT press.

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

Harrington, M., & M. Sawyer (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14,* 25-38.

Love, T., & D. Swinney (1996). Coreference processing and levels of analysis in object-relative constructions: Demonstration of antecedent reactivation in the cross-modal priming paradigm. *Journal of Psycholinguistic Research, 25,* 5-24.

Marinis, T., L. Roberts, C. Felser, & H. Clahsen (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27,* 53-78.

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., & H. Clahsen (2003). Parsing strategies in L2 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition, 25,* 501-528.

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

Pickering, M., & G. Barry (1991). Sentence Processing without empty categories. *Language and Cognitive Processes, 6,* 229-59.

Roberts, L., T. Marinis, C. Felser, & H. Clahsen (2007). Antecedent priming at gap positions in children's sentence processing. *Journal of Psycholinguistic Research*, *36*, 175-188.

Sag, I., & J. Fodor (1995). Extraction without traces. *Proceedings of the 13th West Coast Conference on Formal Linguistics, 13,* 365–384.

Segalowitz, N., & S. J. Segalowitz (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics, 14,* 369–385.

Slabakova, R. (2008). *Meaning in the Second Language,* Studies in Language Acquisition Series, New York: Mouton de Gruyter.

Traxler, M.J., & M.J. Pickering (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language, 35*, 454–475.

White, L. (2003). *Second language acquisition and Universal Grammar.* New York: Cambridge University Press.

Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)

edited by Melissa Bowles, Tania Ionin, Silvina Montrul, and Annie Tremblay

Cascadilla Proceedings Project Somerville, MA 2009

Copyright information

ISBN 978-1-57473-434-8 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Dekydtspotter, Laurent and A. Katherine Miller. 2009. Probing for Intermediate Traces in the Processing of Long-Distance *Wh*-Dependencies in English as a Second Language. In *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, ed. Melissa Bowles et al., 113-124. Somerville, MA: Cascadilla Proceedings Project.

or:

Dekydtspotter, Laurent and A. Katherine Miller. 2009. Probing for Intermediate Traces in the Processing of Long-Distance *Wh*-Dependencies in English as a Second Language. In *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, ed. Melissa Bowles et al., 113-124. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2262.