University of Essex

Referential Context Effects in L2 Ambiguity Resolution: Evidence from Self-Paced Reading

Hui-Yu Pan
University of Essex
Claudia Felser
University of Essex

Essex Research Reports in Linguistics

Volume 59

Number 1

13 Apr, 2010

Dept. of Language and Linguistics,
University of Essex,
Wivenhoe Park,
Colchester, Essex, UK,
CO4 3SQ

http://www.essex.ac.uk/linguistics/publications/errl/

Essex Research Reports in Linguistics present ongoing research activities of the members of the Department of Language and Linguistics.

The main purpose of these reports is to provide a quick publication outlet. They have 'pre-publication status', and most will subsequently appear in revised form as research articles in professional journals or in edited books.

If you have technical problems downloading a paper, or for further information about these reports, please contact the editor:

Doug Arnold: doug@essex.ac.uk.

Citation Information:

Hui-Yu Pan and Claudia Felser. 'Referential Context Effects in L2 Ambiguity Resolution: Evidence from Self-Paced Reading', Essex Research Reports in Linguistics, Vol. 59.1. Dept. of Language and Linguistics, University of Essex, Colchester, UK, Apr, 2010.
http://www.essex.ac.uk/linguistics/publications/errl/errl59-1.pdf

Referential Context Effects in L2 Ambiguity Resolution: Evidence from Self-Paced Reading

Hui-Yu Pan & Claudia Felser

University of Essex

Abstract

This study investigates the degree to which referential context information influences structural ambiguity resolution preferences in non-native sentence comprehension, using both an off-line questionnaire and an on-line self-paced reading task. The critical target sentences contained prepositional phrases (PPs) modifying either the verb phrase (VP) or the preceding noun phrase (NP), as in *Bill glanced at the customer with strong suspicion (with ripped jeans)*. These were embedded within short context paragraphs providing either one or two potential referents for the postverbal NP. The results showed that native Chinese-speaking learners of English and native English speakers were affected differently by referential information in the on-line task. The learners' reading times of the critical PP were influenced significantly by the referential context, with VP-modifying items being read faster than NP-modifying ones in a VP-supporting context, and the reverse pattern seen in NP-supporting contexts. The native speakers' ambiguity resolution preferences, on the other hand, were modulated by the referential context in the off-line task only. Our results indicate that non-native comprehenders are highly sensitive to extra-sentential discourse-level information during processing even at intermediate levels of proficiency, a finding that provides a challenge for 'processing capacity limitation' accounts for non-targetlike L2 performance.

1. Introduction

Real-time sentence comprehension involves the rapid evaluation and integration of multiple different cues to interpretation, including morphological, syntactic, semantic, discourse-level and probabilistic ones (Gibson and Pearlmutter, 1998). Using experimental psycholinguistic techniques to investigate how readers or listeners resolve structural ambiguities can help reveal whether, and to what extent, different types of information affect on-line sentence comprehension. Most previous studies of second language (L2) ambiguity resolution have focused on the role of sentence-internal cues to interpretation, such as the extent to which on-line parsing decisions are affected by verb argument structure information or other types of lexical biases, or on the question of whether non-native speakers are guided by the same kind of phrase-structure based parsing principles that have been proposed in the monolingual processing literature (see Frenck-Mestre, 2005; Papadopoulou, 2005, for reviews).

Very little is known about non-native speakers' sensitivity to sentence-external interpretation cues during L2 comprehension, however. L2 learners have been argued to be less well able than native (L1) speakers to integrate different types of information in real time, due to general processing capacity limitations that make it difficult for learners to simultaneously handle different kinds of information in situations of increased processing demands (compare e.g., Hopp, 2009; Kilborn, 1992). Investigating whether learners can utilize sentence-external context information during on-line processing can tell us something about their ability to integrate pragmatic 'top-down' information with 'bottom-up' cues to meaning. As the on-line integration of extra-sentential discourse information is contingent on the availability of sufficient processing or working memory (WM) resources (Just and Carpenter, 1992), then given that L2 processing is considered to be more resource demanding than native language processing (Harrington, 1992), we might expect learners' ability to use this kind of information to be affected more than native speakers' by task demands and individual differences in WM capacity.

To date, very few studies have investigated L2 learners' sensitivity to top-down information in processing tasks. Kilborn (1992), for example, found that only native speakers' – but not L2 learners' - performance in a word monitoring task improved if the target words appeared in a semantically and syntactically congruent sentence context, compared to a semantically anomalous 'syntactic prose' condition. This finding suggests that learners' ability to integrate syntactic and sentence-level semantic information during processing may be reduced compared to native speakers', supporting the hypothesis that "the limiting factor for L2 speakers is likely to be a lack of immediate, on-line control over relevant top-down information" (Kilborn, 1992: 340).

Results from other L2 processing studies, in contrast, indicate that learners' ability to use sentence-internal semantic cues to interpretation may be similar to native speakers', but that their sensitivity to morphosyntactic and phrase structure information might be reduced in comparison to the latter (see Clahsen and Felser, 2006, for review and discussion). Clahsen and Felser hypothesized that late L2 learners might be able to compensate for their grammatical processing problems by making efficient use of non-grammatical cues to interpretation during on-line comprehension, including semantic and discourse-level information. Evidence that L2 processing may indeed be more semantics-based than L1 processing comes from a study by Guo et al. (2008) using event-related potentials (ERPs). Guo et al. found that subcategorization violations in sentences such as **Joe's father didn't show him drive the car* elicited qualitatively different brain responses in English native

speakers and Chinese-speaking learners of English. While the former showed a positive ERP effect ('P600') of the kind that is thought to reflect syntactic processing difficulty, the learners showed a negative effect ('N400') instead, a brain response characteristically elicited by semantic anomalies. This rather striking L1/L2 difference in participants' brain responses suggests that learners may rely more on meaning-based than on syntactic processing strategies in L2 comprehension.

However, the question of whether or not non-native speakers' supposedly more meaning-based comprehension strategies also include the ability to process and integrate extra-sentential context information on-line is still largely unexplored. The few published studies that have examined the role of preceding discourse information in L2 processing include an eye-movement monitoring study by Roberts et al. (2008) on pronoun resolution in L2 Dutch, and a reading-time study by Hopp (2009) on the processing of German scrambling structures. Roberts et al. (2008) report that highly proficient native German and Turkish-speaking L2 learners' – but not the native Dutch-speaking controls' – reading of ambiguous pronouns was affected by the number of potential antecedents provided in the preceding context sentence.

Examining learners' sensitivity to information-structure constraints on German object scrambling, Hopp (2009) found that the reading times of both advanced and near-native L1 English and Russian speakers, although not those of L1 Dutch speakers, were affected by discourse-level constraints on scrambling in a native-like way. The fact that the less proficient L1 English-speaking participants did not show any sensitivity to discourse context information in a complementary off-line task indicates, according to Hopp, that learners may have difficulty integrating syntactic and discourse-level information under certain conditions.

While these findings suggest that learners at or near the top end of the proficiency scale may be able to process and integrate extra-sentential context information on-line, the extent to which discourse-level cues affect L2 structural ambiguity resolution, and possible effects of L2 proficiency or WM capacity on learners' ability to integrate discourse-level and bottom-up information, have not yet been systematically investigated. The current study is the first to examine the degree to which non-native comprehenders are sensitive to referential information provided by the preceding discourse context in L2 ambiguity resolution.

2. The role of referential context information in parsing

Our study investigates the resolution of prepositional phrase (PP) ambiguities in sentences such as *The policeman watched the spy with binoculars*. Here the PP *with binoculars* can

either be interpreted as modifying the verb phrase (= VP attachment, as indicated in (1a)) or the postverbal noun phrase (= NP attachment, as in (1b)).

(1) a. The policeman [$_{VP}$ watched [$_{NP}$ the spy] [$_{PP}$ with binoculars]]

b. The policeman [$_{VP}$ watched [$_{NP}$ the spy [$_{PP}$ with binoculars]]]

Earlier monolingual processing studies have shown that when sentences such as those in (1) are presented in isolation, native speakers of English tend to prefer the VP modification over the NP modification reading (e.g. Clifton et al., 1991; Frazier, 1979; Rayner et al., 1983 – but see Taraban and McClelland, 1988), especially in the case of action verbs or potentially ditransitive verbs (compare e.g. Britt, 1994; Schütze and Gibson, 1999; Spivey-Knowlton and Sedivy, 1995).

According to structural models of parsing, the preference for interpreting ambiguous PPs as VP rather than NP modifiers reflects the application of phrase structure-based 'least effort' principles such as MINIMAL ATTACHMENT (Frazier and Rayner, 1982), on the assumption that VP modification requires a less complex syntactic representation than NP modification, or PREDICATE PROXIMITY, which biases the parser towards attaching ambiguous modifiers as closely as possible to the head of the current clause or predicate (Gibson et al., 1996).

Discourse-sensitive theories such as the Referential Theory developed by Altmann and Steedman (1988) and Crain and Steedman (1985), on the other hand, claim that the parser's initial analysis will be influenced by referential context information, if available. For sentences containing ambiguous PPs, a preference for VP modification is argued to result from the processing system's preference for the analysis that requires the fewest possible pragmatic presuppositions. In sentences such as *The policeman watched the spy with binoculars* in which the postverbal NP is definite, the VP modification reading requires the presupposition of a single referent for this NP only. The NP modification reading, on the other hand, requires us to presuppose the existence of more than one referent for *the spy*, with the PP functioning as a restrictive modifier that identifies a unique referent from a set. According to the PRINCIPLE OF REFERENTIAL SUPPORT (2), a discourse context providing more than one potential referent for the postverbal NP will render the NP modification analysis more felicitous than it would be in the absence of NP-supporting context information (Altmann and Steedman, 1988).

(2) PRINCIPLE OF REFERENTIAL SUPPORT

An NP analysis which is referentially supported will be favored over one that is not.

There is a substantial body of evidence for the use of referential information in monolingual processing (Altmann and Steedman, 1988; Britt, 1994; Britt et al., 1992; Crain and Steedman, 1985; Kaiser and Trueswell, 2004; Papadopoulou and Clahsen, 2006; Spivey-Knowlton et al., 1993; Spivey-Knowlton and Tanenhaus, 1994; Spivey and Tanenhaus, 1998; van Berkum et al., 1999), although the question of whether or under what conditions referential information affects the initial stages of L1 parsing is still under debate.

Altmann and Steedman (1988), for example, had native speakers of English read syntactically ambiguous sentences such as (3) that were semantically disambiguated towards either VP modification (*with the dynamite*) or NP modification (*with the new lock*), and which were preceded by short context paragraphs.

(3) The burglar blew open the safe with the dynamite (with the new lock) and made off with the loot.

The referential context was manipulated in such as way so as to provide either one or two potential discourse referents for the postverbal NP, *the safe*. The results showed that participants' global reading times for sentences containing NP-modifying PPs were significantly shorter than for those containing VP-modifying ones if these were preceded by a two-referent ('NP-supporting') context such as (4) below.

(4) A burglar broke into a bank carrying some dynamite. He planned to blow open a safe. Once inside he saw that there was *a safe with a new lock* and *a safe with an old lock*.

As longer reading times are thought to reflect increased processing or comprehension difficulty, these results indicate that an NP modification reading is indeed preferred over a VP modification reading in an two-referent (NP-supporting) context. No reversal of this effect on participants' disambiguation preferences was seen in the VP-supporting context conditions, however. In a self-paced reading version of the experiment using segment-by-segment rather than whole-sentence presentation, the referential context did not reliably modulate participants' on-line disambiguation preferences. Instead, PPs in a VP-supporting context were generally read faster than those in an NP-supporting context, and NP

modification elicited shorter reading times than VP modification across both context conditions.

Other findings also suggest that the use of discourse-level information may be delayed in L1 ambiguity resolution, with referential context information affecting final interpretations but not initial parsing decisions. Evidence against the immediate use of referential information in L1 English has been reported, for instance, in a self-paced reading study on complement/relative clause ambiguities by Mitchell et al. (1992), and in an eye-movement study by Binder et al. (2001) examining main clause/reduced relative clause ambiguities. The results from eye-movement monitoring experiments investigating relative clause attachment ambiguities in French (Zagar et al., 1997) and Dutch (Desmet et al., 2002) also showed no evidence that participants' initial attachment decisions were influenced by referential information.

In contrast, evidence that preceding context information may have immediate effects on parsing has been found, for example, in an ERP study on the processing of complement/relative clause ambiguities in Dutch (van Berkum et al., 1999) and in a reading-time experiment on the processing of scrambled OVS structures in Finnish (Kaiser and Trueswell, 2004). Early referential context effects were also observed in a self-paced reading study on relative clause attachment ambiguities in Greek (Papadopoulou and Clahsen, 2006), although context information was not able to override any lexical biases towards local modification here. The lack of consistency regarding the relative timing of context effects in monolingual processing studies may be due to differences in the experimental tasks or materials used, or in the relative strength of the context manipulations compared to other factors affecting parsing.

Comparatively little is known about the extent to which non-native ambiguity resolution is influenced by referential context information. Ying (1996) examined advanced non-native speakers' interpretation preferences for ambiguous PPs in a series of untimed reading and listening based tasks. In his first experiment Ying asked learners of English from various L1 backgrounds to read ambiguous sentences such as *The man talked to the girl with a sense of humour* and indicate their preferred interpretation on an answer sheet. In experiment 2, participants read the same sentences again, which now appeared embedded within an NP-biasing context as shown in (5).

(5) There were two girls. One of them had a sense of humour, and the other did not. The man talked to the girl with a sense of humour.

The results revealed a preference for VP over NP modification in the no-context condition (experiment 1) that was reversed if the sentence were preceded by an NP-supporting context (experiment 2). As noted above, a preference for VP modification in the absence of an NP-biasing context could reflect either the application of structural least-effort principles or an attempt to minimize the number of pragmatic presuppositions – or possibly, a combination of both. The results from experiment 2 suggest that referential information is able to override any potential intra-sentential economy constraints on interpretation, however.

Although Ying's results are consistent with the predictions made by the Referential Theory, they can only tell us something about learners' ultimate interpretations, but nothing about the extent to which referential information affects their on-line ambiguity resolution preferences. The absence of a native control group moreover precludes any direct comparison between native and non-native speakers' degree of sensitivity to referential context information. Using participants from a range of typologically different language backgrounds also seems less than ideal as this makes it difficult to control for potential L1 effects on L2 ambiguity resolution. Finally, it is conceivable that the preference for VP modification seen in experiment 1 was at least partly due to the use of action verbs such as *hit*, *eat* or *strike* in about half of Ying's experimental items, which may have triggered an expectation for an 'instrument' PP.[1]

The current study uses experimental materials similar to those above to examine and compare how native and non-native readers are influenced by referential context information during processing, and how contextual biases affect their ultimate interpretation preferences. Our critical target sentences were of the type shown in (6a,b) below and contained only psych or perception verbs so as to minimize any lexical biases towards VP modification (see Spivey-Knowlton and Sedivy, 1995).

(6) a. Bill glanced at the customer *with strong suspicion.*

b. Bill glanced at the customer *with ripped jeans*.

Although both (6a) and (6b) are syntactically ambiguous, the italicised PP in (6a) is semantically disambiguated towards VP modification, whereas in (6b) it can only be

[1] Compare Frenck-Mestre and Pynte (1997), who showed that proficient English-speaking learners of French are guided by verb subcategorization information in a similar way to native French speakers when processing ambiguous prepositional phrases.

understood as modifying the preceding NP *the customer*. If L2 processing is generally guided more by discourse-level information than is the case in monolingual processing, we would expect L2 learners' ambiguity resolution preferences to be more strongly affected by biasing referential context than native speakers'. Using an on-line subject-paced reading task as well as an off-line task should furthermore allow us to assess whether referential information is used immediately during on-line comprehension or whether it affects participants' ultimate interpretations only.

3. Experiment 1: Sentence completion

To examine how referential context information affects native and non-native speakers' ultimate PP modification preferences, we carried out an off-line binary-choice sentence completion task with a group of Chinese-speaking learners of L2 English and a group of native English-speaking controls. Given the findings reported in earlier monolingual (e.g., Altmann and Steedman, 1988) and L2 studies (Ying, 1996), we would expect the proportion of participants' VP modification choices to decrease in an NP-supporting context. Note that in Chinese, PP attachment ambiguities of the type under investigation do not exist, as VP and NP modification each normally require a different constituent order.[2] This effectively precludes the possibility of the learners transferring any potential L1-specific ambiguity resolution preferences to their L2.

3.1. Method

3.1.1. Participants

A group of 30 Chinese-speaking learners of English as a second language (19 females; mean age: 24.6; range: 17-52; SD: 6.2) and a group of 30 native English-speaking controls (22 females; mean age: 30.86; range: 18-60; SD: 12.7) volunteered to participate in the off-line experiment. Participants were recruited from the University of Essex student and staff communities, and from universities and other higher education institutions in Taiwan. The non-native participants were all Taiwanese Chinese speakers whose native dialect was Mandarin. The Chinese-speaking participants were first exposed to English at the age of 11

[2] In Chinese, PP modifiers usually immediately precede the constituent they modify. For example, for VP modification the natural constituent order in Chinese is *Bill [with strong suspicion] glanced at the customer*, whereas NP modification requires the order *Bill glanced at the [with ripped jeans] customer*. Moreover, in the absence of any direct counterpart of prepositions like *with* in Chinese, the choice of suffix (adverbial *–di* versus adjectival *-de*) provides an additional grammatical disambiguation cue.

on average (range: 4-15; SD: 2.1), in a classroom setting. According to the scores for IELTS, TOFEL or GEPT ('General English Proficiency Test', a language proficiency test commonly adopted in Taiwan) provided by the Chinese participants, their general level of proficiency in L2 English ranged from IELTS 5.5 or equivalent ('upper intermediate') to IELTS 7 or equivalent ('proficient'), with the median corresponding to IELTS 6.5. Twelve of the Chinese participants had had no immersion in English at all, while the remaining 18 had spent, on average, about two years and ten months in an English-speaking environment at the time of testing (range: 11 months to 8 years). All of the participants were naïve with regard to the ultimate purpose of the experiment.

3.1.2. Materials

The materials for the off-line sentence completion task were adapted from Spivey-Knowlton and Sedivy (1995) and comprised 16 experimental and 16 filler sentences. The experimental items all contained a blank after the second noun phrase (e.g. *Bill glanced at the customer* _____), followed by two alternative options for filling the blank, one consistent with a VP modification reading (i.e., *with strong suspicion*) and the other one consistent with an NP modification interpretation (i.e., *with ripped jeans*). The matrix verbs used were either psych or perception verbs, and the postverbal noun phrase was always definite. The average string length of the entire PP modifier *(with + NP)* was matched across the two modification conditions, with VP-modifying PPs consisting of 17.19 and NP-modifying ones of 17.06 characters on average. In addition, the word form frequencies of the disambiguating nouns (e.g. *suspicion* vs. *jeans*) were also matched as closely as possible (VP modification: 35.56, NP modification: 36, according to the CELEX database).

Each experimental sentence appeared in two experimental conditions, preceded either by a VP-supporting or by an NP-supporting context, as illustrated by the examples shown in (7) and (8) below.

(7) VP-SUPPORTING CONTEXT
Bill walked into a shop that he knew the police were keeping an eye on. There was only one other customer in the shop. The customer was wearing old and filthy clothes, whereas the sales assistant was dressed very smartly.

(8) NP-SUPPORTING CONTEXT
Bill walked into a shop that he knew the police were keeping an eye on. There were two other customers in the shop. One customer was wearing old and filthy clothes, whereas the other one was dressed very smartly.

In the VP-supporting context condition (7), only a single discourse referent for the postverbal noun phrase *the customer* is provided, whereas in the NP-supporting context condition (8), two potential referents are introduced. According to the Referential Theory the presence of more than one potential discourse referent for *the customer* in (8) should increase the number of NP modification choices as compared to a neutral or VP-supporting discourse context. A complete list of our experimental stimulus items can be found in the Appendix.

The experimental items were distributed across two presentation lists using a Latin Square design, to ensure that each participant would see each experimental sentence only once, and with the order of VP and NP modification answer options counterbalanced in both lists. The filler items also consisted of short paragraphs containing a blank, followed by two answer options. Some of the fillers were pseudo-fillers that were similar to the experimental items in that they also needed to be completed by prepositional phrases, as in *George enjoys feeding the ducks ____ (a) in the morning / (b) in the pond.* The experimental sentences were mixed with the fillers and pseudo-randomised.

3.1.3. Procedure

The materials were presented in the form of a web-based written questionnaire.[3] Participants were given the link to the questionnaire website and were instructed, in writing, to read each paragraph carefully and to complete it by ticking one of the two answer options provided. Participants who clicked on the link to the questionnaire would first see the instructions, which also reminded them that their initial answer choices should not be changed later on. After reading through the instructions, participants had to click on a 'Next' button to proceed to the main task. The 32 test items were presented on four pages containing eight items each. Participants were asked to complete the questionnaire in a single uninterrupted session. Once they finished the task, clicking on a 'Submit' button at the bottom of the last page would submit their answers to the database. There was no time limit for this task. The questionnaire took about 15-20 minutes for native English speakers to complete, and around 25-30 minutes for Chinese speakers.

---

[3] For some discussion of the merits - and potential drawbacks - of this method of data collection, see Wilson and Dewaele (2010).

3.2. Results

A summary of the results is provided in Table 1. Although both groups showed an overall preference for VP modification, both the native and the non-native speakers showed a marked increase in their proportion of NP modification choices in an NP-supporting context.

Table 1. *Proportion of NP modification choices (in percent, SDs in parentheses) per group and condition.*

	English group	Chinese group
VP-supporting context	30 (18)	29 (16)
NP-supporting context	44 (21)	46 (22)

A mixed repeated-measures analysis of variance (ANOVA) with the within-subjects factor Context (*VP-supporting, NP-supporting*) and the between-subjects factor Group (English, *Chinese*) showed a significant main effect of Context (F_1 (1, 58) = 25.946, p_1 < .001; F_2 (1, 30) = 22.973, p_2 < .001), confirming that the number of participants' NP modification choices was significantly higher in the NP-supporting than in the VP-supporting context condition, and no interaction between Context and Group.

3.3. Discussion

The results from the sentence completion questionnaire show that our context manipulation was effective in both groups, such that the proportion of VP modification choices was significantly reduced for experimental items that were preceded by an NP-supporting context. This replicates the context effect observed by Ying (1996) and is in line with what the Referential Theory would lead us to expect. Our learners did not differ statistically from the native speakers in the degree to which their answer choices were affected by the referential context, or with regard to their absolute proportions of NP versus VP modification choices across the two experimental conditions. Note, however, that the results from the above off-line task only tell us something about participants' *ultimate* modification preferences, leaving open the possibility that their decision was - partly or wholly - based on a conscious (re-)evaluation of the discourse context following their initial reading of the stimulus items. Experiment 2 investigates whether referential information also affects learners' PP-ambiguity resolution preferences during on-line processing.

4. Experiment 2: Self-paced reading

To investigate whether the preceding discourse-pragmatic context would also affect non-native readers' *initial* disambiguation preferences, we carried out an on-line self-paced reading experiment. If referential information has an immediate effect on parsing as predicted by discourse-sensitive processing models, then the critical PP should be read faster if it is pragmatically congruent with the preceding context compared to when it is not. However, if the use of referential information in on-line ambiguity resolution is reduced or delayed, participants' reading times of the ambiguous PP region should not be modulated by our context manipulation.

Furthermore, if L2 learners have more difficulty than native speakers integrating extra-sentential discourse information in situations of increased processing demands, then only the native speaker controls, but not the Chinese-speaking participants, should show immediate sensitivity to referential context information. On the assumption that readers' on-line sensitivity to preceding discourse-level information is dependent on the availability of sufficient computational resources (Just and Carpenter, 1992), we might moreover expect participants with a relatively lower WM capacity to be less sensitive to extra-sentential referential cues compared to those with a higher WM capacity.

Conversely, if L2 learners generally rely on meaning-based processing strategies to a greater extent than native speakers do, and if these include the rapid integration of extra-sentential discourse-level cues, we might see the learners being affected by referential context information earlier and/or to a greater extent than the native speakers during processing. If a stronger focus on meaning is a general L2 processing strategy that helps learners compensate for grammatical processing problems, we might expect even less proficient learners to show sensitivity to referential context information, and little or no effects of individual differences in their L2 WM capacity.

4.1. Method

4.1.1. Participants

Thirty-two native speakers of English (23 female; mean age: 28.6; range: 18-61; SD: 12.1) and 36 Chinese-speaking learners of English (25 female; mean age: 27.3; range: 17-35; SD:

5.2) volunteered to participate in the experiment.[4] They were recruited from the same participant pools as those who took part in Experiment 1. The learners' mean age of first exposure to English was 11 (range: 5-14; SD: 1.8) in a classroom setting. Seventeen of the non-native participants had had no immersion in English at all, whilst the remaining 19 participants had spent 13 months, on average, in an English-speaking environment at the time of testing (range: 1 month to 5.4 years). All participants had normal or corrected-to-normal vision and were naïve with regard to the ultimate purpose of the experiment. The non-native participants were all Taiwanese Chinese speakers whose first language was Mandarin.

Besides the main experiment, the Chinese-speaking participants also completed the grammar part of the Oxford Placement Test (OPT; Allan, 2004) and a vocabulary test. The OPT results showed that the average level of the learners' general English grammar proficiency was that of 'proficient' users, with individual scores ranging from 'lower intermediate' to 'highly advanced' user levels (mean OPT score: 77/100; range: 61-94; SD: 9). The purpose of the vocabulary test was to ensure that the learners were familiar with the verbs, nouns and adjectives that were used in the critical target sentences.

To allow us to examine potential working memory effects on participants' processing performance, all participants additionally underwent a reading span test (L1: Daneman and Carpenter, 1980; L2: Harrington and Sawyer, 1992). These tests required participants to read increasingly larger sets of sentences and then to recall the final word of each sentence at the end of each set. The group mean reading span score was 3 (out of 6) for the native controls (range: 2 – 4.5, SD = 0.66) and 26 (out of 42) for the Chinese participants (range: 14 – 39, SD = 6.16).[5]

4.1.2. Materials

The same 16 experimental items (including both the referential contexts and the target sentences) that were used in the sentence completion experiment were also used in our reading-time experiment. The experiment had a 2×2 design with Context (*VP-supporting, NP-supporting*) and Attachment (*VP modification, NP modification*) as within-subjects

---

[4] Five of the learners had also previously taken part in Experiment 1. As the interval between the two experiments was more than one month, however, this is unlikely to have affected their performance in the on-line experiment.

[5] Note that Harrington and Sawyer's (1992) variant of Daneman and Carpenter's (1980) reading span test uses a slightly different scoring procedure. The scores obtained by either measure tend to be highly correlated, however (Whitney et al., 2001).

factors. Each target sentence was segmented into five presentation segments as indicated by the slashes in (9).

(9) Bill / glanced at / the customer / with strong suspicion (with ripped jeans) / and then walked away.

Segment-by-segment presentation (as in Altmann and Steedman's, 1988, second experiment) was chosen after some pilot testing which revealed that L1 Chinese speakers within the English proficiency range to be examined had difficulty comprehending the experimental stimulus texts when the target sentence was presented one word at a time, suggesting that word-by-word presentation would have resulted in an excessive number of comprehension errors and thus yielded an unacceptably large proportion of unusable data.

The 16 experimental items were distributed across four counterbalanced presentation lists using a Latin Square design, and then randomized with 24 filler texts. To help ensure that participants would read both the context paragraphs and target sentences carefully for meaning, all experimental and filler items were followed by a yes/no comprehension question. The end-of-trial comprehension questions probed the content of either the context paragraph or the target sentence, to help ensure that participants would read both of these carefully for meaning.

4.1.3. Procedures

All participants were tested individually in a quiet setting. The Chinese participants were tested in two separate sessions around one week apart. Session one included the main experiment and the reading span test, which together took the learners about 50 minutes to complete. The reading span test required participants to read sets of up to five sentences and provide a grammaticality judgement at the end of each. At the end of each set, they were asked to recall the last word of each sentence. The paper-and-pencil proficiency and vocabulary tests were completed in session two, which took around 40-45 minutes in total.

The self-paced reading experiment started with three practice items to allow participants to familiarize themselves with the task. Each trial began with the presentation of a context paragraph such as (7) or (8) above, which was shown as a whole. Participants were instructed to press a 'Continue' button on a Logitech PC game pad in order to proceed to the target sentences, which were presented one segment at a time using the non-cumulative moving-window technique (Just et al., 1982). Participants controlled the presentation speed by

pressing the 'Continue' button on the gamepad when they felt ready to receive the next segment, which then replaced the previous one on the screen.

The last segment was replaced by a comprehension question, which participants were asked to answer as quickly and as accurately as possible by pressing a designated 'Yes' or 'No' button on the gamepad. Halfway though the experiment participants were offered to take a short break. The stimulus texts were presented in 14-font Arial in white letters on black background, and the presentation of the stimuli and the recording of reaction times and responses was controlled by the DMDX experimental software package (Forster and Forster, 2003).

4.2. Results

The native speakers answered the end-of-trial questions correctly 97% (range: 87% - 100%) and the L2 learners 93% (range: 81% -100%) of the time, indicating that both native and non-native participants were reading the stimulus items properly and were paying attention to the task. Statistical analyses of the reading time data were carried out for correctly answered trials only. We also removed individual trials from the L2 data set that contained any unknown vocabulary items, according to the results of the vocabulary test, which affected 8.5% of the remaining trials for the Chinese group. Furthermore, individual outlier data points of 2.5 SDs or more beyond the group means per condition were eliminated from the data set prior to the statistical analysis. This procedure affected 4.2% of the reading times at the critical PP region and 6% at the post-critical region for the native speakers, and 6.8% of the reading times at the PP region and 6.1% at the post-critical region for the L2 group. Table 2 provides an overview of participants' mean reading times per segment for each of the four experimental conditions after data trimming.

We analysed participants' reading times for both the critical PP region (the point at which the experimental conditions started to diverge) and the postcritical sentence segment. To determine whether or not the two participant groups' reading-time patterns differed across the experimental conditions at the critical PP and/or postcritical region, we firstly ran preliminary repeated-measures ANOVAs with Context (*VP-supporting, NP-supporting*) and Attachment (*VP modification, NP modification*) as within-subject factors and Group (*English, Chinese*) as a between-subjects factor for each of the two regions of interest. For the critical PP region, we found a significant a two-way interaction of the factors Context and Attachment (F_1 (1, 65) = 9.838, $p_1 < .01$; F_2 (1, 30) = 6.471, $p_2 < .05$) as well as a three-way

Table 2. *Mean reading times in milliseconds and SDs (in brackets) per group and condition, Experiment 2*

Region	Subject *Bill*		Verb (+ P) *glanced at*		NP *the customer*		PP *with strong suspicion / with ripped jeans*		Final *and then walked away*	
	NSs	L2	NSs	L2	NSs	L2	NSs	L2	NSs	L2
VP context -	577	708	407	644	452	687	639	1471	798	1698
VP attachment	(120)	(151)	(91)	(216)	(129)	(156)	(200)	(313)	(227)	(409)
VP context -	563	712	410	594	462	717	646	1659	862	1697
NP attachment	(89)	(172)	(57)	(131)	(117)	(218)	(197)	(456)	(298)	(428)
NP context -	539	742	394	591	452	658	643	1694	801	1762
VP attachment	(96)	(186)	(72)	(105)	(133)	(162)	(213)	(467)	(254)	(440)
NP context -	579	711	404	623	455	666	639	1566	835	1827
NP attachment	(114)	(171)	(73)	(186)	(129)	(167)	(218)	(350)	(277)	(474)

interaction between Context, Attachment and Group (F_1 (1, 65) = 8.376, $p_1 < .01$; F_2 (1, 30) = 4.852, $p_2 < .05$). The fact that the L2 group generally read the critical PP more slowly than the L1 group was reflected in a significant main effect of Group (F_1 (1, 65) = 256.309, $p_1 < .001$; F_2 (1, 30) = 252.173, $p_2 < .001$). A main effect of Group was also found at the postcritical region (F_1 (1, 65) = 173.091, $p_1 < .001$; F_2 (1, 30) = 74.259, $p_2 < .05$), as well as a marginally significant Context × Group interaction in the analysis by participants (F_1 (1, 65) = 2.901, $p_1 = .093$; F_2 (1, 30) = 2.562, $p_2 = .120$).

As the results of the preliminary analyses were indicative of different processing patterns in the two groups, we subsequently analysed the data from each group separately using repeated-measures ANOVAs with the factors Context and Attachment. For the native speakers, these showed no main effects or interaction at the critical PP region but a significant main effect of Attachment (F_1 (1, 31) = 5.145, $p_1 < .05$; F_2 (1, 15) = 5.339, $p_2 < .05$) at the post-critical region, reflecting the fact that the final region of VP-disambiguated items was read faster than the final region of NP-disambiguated ones, irrespectively of context. The analysis of the Chinese group's reading times, on the other hand, revealed a significant Context × Attachment interaction at the critical PP region (F_1 (1, 34) = 10.518, $p_1 < .01$; F_2 (1, 15) = 7.467, $p_2 < .05$). Subsequent paired *t*-tests (one-tailed) confirmed that the L2 learners read VP-disambiguated PPs faster than NP-disambiguated ones in a VP-supporting context ($t_1 = 2.202$, $p_1 < .05$; $t_2 = 1.802$, $p_2 < .05$), and NP-disambiguated PPs faster than VP-disambiguated ones in an NP-supporting context, in the analysis by participants ($t_1 = 1.713$, $p_1 < .05$; $t_2 = .638$, $p_2 = .267$). No significant effects or interaction were found for the post-critical region here.

4.3. Further analyses

Additional factors that might potentially have affected participants' reading-time patterns include individual differences in WM capacity, L2 proficiency, and whether or not they had spent time in an immersion setting. To examine whether the observed L1/L2 differences in participants' sensitivity to referential information might be linked to differences in participants' WM capacity (as measured by the reading span tests), repeated-measures ANOVAs with Reading Span as a covariate were carried out for both the native and the non-native speakers' reading times of the critical PP region. The results for the native control group showed a significant main effect of

Reading Span (F (1, 30) = 6.066, p <.05), reflecting the fact that native speakers with relatively lower reading spans generally read more slowly than those with higher reading spans, and a significant Context × Reading Span interaction (F (1, 30) = 4.945, p < .05).

To explore the source of this interaction, we divided the native speakers up into two WM subgroups. Those with reading span scores lower than the median score of 3.0 were grouped together as the Low WM subgroup (n=6), whereas those with scores higher than 3.0 formed the High WM subgroup (n=9). Participants with reading span scores equal to the median (n=17) were excluded. Subsequent t-test revealed that the difference between the collapsed reading times for the VP-supporting and NP-supporting context conditions was marginally significant by participants for the Low WM subgroup (t_1= 2.040, p_1= .097; t_2= 1.208, p_2= .247), reflecting the fact that their reading times were generally shorter for target sentences appearing in a VP-supporting compared to those in an NP-supporting context (721 vs. 789 ms). For the High WM subgroup, on the other hand, reading times were shorter for the NP-supporting than for the VP-supporting context conditions (500 vs. 540 ms), a difference that was marginally significant in the analysis by items (t_1= 1.762, p_1= .116; t_2= 1.780, p_2= .095). However, in the absence of any interactions with the factor Attachment, these results merely suggest that high and low-span native speakers' reading times of the critical PP might have been affected differently be the preceding context, but without our context manipulation affecting either the high or the low-span readers' disambiguation preferences. A parallel ANOVA on the L2 learners' reading times showed no significant effects of, or interactions with, the factor Reading Span for this group.

As our non-native participant group was not particularly homogeneous with regard to their general L2 grammar proficiency (as measured by the OPT), to determine whether individual differences in the learners' L2 proficiency had any effects on their processing patterns, we ran a corresponding analysis with OPT scores as a covariate for the learners' reading times at the PP region. This revealed a main effect of Attachment (F (1, 33) = 9.530, p < .01), a marginal main effect of OPT score (F (1, 33) = 3.226, p = .082), as well as a significant interaction of Attachment × OPT Score (F (1, 33) = 9.888, p < .01). To examine the source of this interaction, we divided the learners into two proficiency subgroups based on their median OPT score. The High Proficiency subgroup included learners with a mean OPT score of 85.5

(range: 79-94) and the Low Proficiency subgroup those with a mean score of 69.8 (range: 61-78).[6] Subsequent paired *t*-tests showed that learners with relatively lower L2 proficiency showed a weak trend of reading NP-disambiguated items faster than VP-disambiguated ones (t_1 = 1.946, p_1 = .068; t_2 = .607, p_2 = .553), whilst the High Proficiency subgroup showed the opposite pattern (t_1 = 3.157, p_1 < .01; t_2 = 1.662, p_2 = .117). Crucially though, individual differences in L2 grammar proficiency did not modulate the learners' sensitivity to referential context information.

Finally, we divided the L2 participants into two subgroups according to whether to not they had spent time in an immersion setting. A mixed repeated-measures ANOVA with Context and Attachment as within-subjects factors and Immersion as a between-subjects factor showed a main effect of Immersion (F (1, 33) = 4.173, p < .05), reflecting the fact that participants who were immersed in English at the time of testing tended to be faster readers than those without immersion, and a Context × Attachment interaction (F (1, 33) = 9.907, p = .003) that was not modulated by the factor Immersion.

Together, the results from the above analyses show that the observed L1/L2 differences in participants' on-line sensitivity to referential context information cannot be accounted for by individual differences in WM capacity or L2 proficiency, or whether or not the learners had spent any time immersed in their L2.

5. Discussion

While referential context information affected both participant groups' ultimate modification preferences in Experiment 1 in essentially the same way, clear L1/L2 differences were observed in participants' reading-time patterns in Experiment 2. The results from the self-paced reading experiment indicate that the L2 learners' processing of the critical PP was affected by referential information, such that VP-modifying PPs were easier to process in a one-referent context and NP-modifying ones easier in a two-referent context. There is no evidence in our data to suggest that the use of referential context information is reduced or delayed in L2 relative to L1 ambiguity resolution.

---

[6] The Lower Proficiency subgroup included all learners who fell within the 'lower to upper intermediate' proficiency bands, while the Higher Proficiency subgroup included only 'proficient to highly proficient' learners, according to the OPT scale.

The native speakers, in contrast, showed some evidence of a general preference for VP modification that was not modulated by the referential context, at the region following disambiguation. The results from our native group resembled those from Altmann and Steedman's (1988) self-paced reading experiment in that no reliable context by attachment interactions were found in their participants' reading times of the critical PP segment, either, and are consistent with the results from other studies which show that context effects may be delayed in monolingual processing or affect ultimate interpretations only (compare, e.g., Desmet et al., 2002; Mitchell et al., 1992; Zagar et al., 1997).

The reading-time patterns seen in the learners, on the other hand, are in line with the predictions made by the Referential Theory, which claims that discourse-pragmatic information can affect on-line parsing. In what follows, we will consider possible explanations for the observed L1/L2 processing differences.

5.1. Computational and subject-specific factors

Let us consider first the possibility that the observed L1/L2 differences in processing syntactically ambiguous PP might have been due to general computational factors such as slower processing speed, or the increased WM demands associated with processing a non-native language. As is usually the case in L2 processing studies, our learners were indeed significantly slower to read the experimental stimuli in the on-line task than were the native speakers. Nevertheless, it was the learners who showed immediate sensitivity to referential context information whereas the native speakers did not – the opposite of what might be expected if slower processing speed or other processing capacity limitations reduce the ability to integrate top-down and bottom-up information during on-line comprehension.

Recall further that, even though L2 participants who had not been immersed in English tended to read our stimulus materials more slowly than those with immersion, the two subgroups' on-line ambiguity resolution preferences were affected by our context manipulation in the same way. That is, while the increased amount of L2 practice and exposure provided in an immersion setting seemed to be associated with faster L2 reading speed, immersion did not measurably influence our learners' sensitivity to discourse information.

Although non-native language processing is likely to require more computational resources than does processing one's native language, possible WM shortages on the

part of our Chinese participants cannot account for our findings, either. Recall that individual differences in WM capacity as measured by Harrington and Sawyer's (1992) reading span test did not affect the learners' reading-time pattern across the experimental conditions. The lack of any statistical WM effects aside, it is difficult to see why potential capacity limitations should have led to our learners being guided *more* by sentence-external pragmatic information than native speakers during on-line ambiguity resolution. Note that in contrast to our L2 learners, young children, whose computational or WM resources are also likely to be more limited than mature native speakers', have been found to show no or reduced sensitivity to referential context information when processing ambiguous PPs (Hurewitz et al., 2000; Trueswell et al., 1999).

For the native control group in our Experiment 2, additional WM analyses showed some evidence that high and low span participants' reading times of the critical PP might have been affected differently by the preceding context, with the low-span subgroup patterning with the participants in Altmann and Steedman's (1988) segment-by-segment reading time experiment in showing a general (numerical) reading time advantage for PPs embedded in a VP-supporting context. While high-span participants showed a numerical trend in the opposite direction in our study, individual differences in reading span cannot account for the lack of context effects on our native speakers' on-line ambiguity resolution preferences. There is no evidence in our data to suggest that the L1 Chinese participants' processing pattern resembled that of either high or low-span native speakers.

Individual differences in L2 proficiency as measured by the OPT, on the other hand, had a small effect on learners' processing patterns such that more proficient learners tended to read VP-disambiguated items faster than NP-disambiguated ones, whereas less proficient ones showed the opposite tendency. Crucially, however, differences in L2 grammar proficiency did not affect the learners' sensitivity to referential context information, either. That VP modification should become relatively more felicitous with increased proficiency could either be due to an increase in learners' ability to apply structural parsing principles such as Minimal Attachment during processing, or could reflect greater sensitivity to the pragmatic bias towards VP modification bias triggered by the definiteness of the preceding NP. While this may be an issue worth investigating further by systematically comparing learners from a wider range of proficiency levels, there is no indication in the current results

that learners' reliance on extra-sentential disambiguation cues during L2 comprehension either grows or diminishes with increasing proficiency.

In summary, the L1/L2 differences in participants' reading-time patterns seen in Experiment 2 cannot obviously be accounted for by general processing or subject-specific factors such learners' slower reading speed, WM shortages, or lack of L2 proficiency or exposure.

5.2. The role of discourse-level information in L2 processing

The results from the current study confirm and extend those reported by Ying (1996) by demonstrating that PP ambiguity resolution in L2 English is highly sensitive to referential context information not only in off-line but also in on-line tasks. They furthermore extend previous findings by Roberts et al. (2008) and Hopp (2009) from learners at or near the top end of the proficiency scale by showing that even learners at intermediate proficiency levels, and learners who had not spent any time at all in an immersion setting, are highly sensitive to extra-sentential discourse-level information during on-line L2 processing.

Our finding that immediate effects of the referential context were seen in the learners but not in the native speaker controls moreover support the hypothesis that on-line processing in an L2 may be guided more strongly by semantic and pragmatic cues to interpretation compared to native language processing (e.g., Clahsen and Felser, 2006). This hypothesis is also supported by Roberts et al.'s (2008) finding that only the L2 learners, but not the native speaker controls, were influenced by extra-sentential discourse-level information when resolving ambiguous pronouns.

Greater sensitivity to discourse-level cues in non-native compared to native processing has also been reported in an eye-movement monitoring study investigating L1 German speakers' processing of English reflexives (Felser and Cunnings, submitted), and advanced Greek-speaking learners of English were found to be more strongly (mis-)guided by pragmatic congruence than native speakers when processing garden-path sentences such as *While the band played the song pleased all the customers* (Roberts and Felser, in press).

One possible way of accounting for these findings is that proficient non-native speakers may in fact be better able to integrate bottom-up and top-down information during processing than native speakers (contra Kilborn, 1992). Even though this possibility should perhaps not be dismissed entirely out of hand, it seems rather

counter-intuitive, and difficult to reconcile with the absence of any proficiency or WM effects on learners' ability to utilize extra-sentential context information in the current study.

Alternatively, it may be that non-native speakers generally tend to apply more meaning-based processing strategies than native speakers, whose processing may be more grammar-based instead (Guo et al., 2008). This could be the case if grammatical processing is particularly 'hard' in a non-native language, thus biasing L2 comprehenders towards focusing more on semantic and pragmatic information instead when processing the L2 input (Clahsen and Felser, 2006; Gass, 1989; among others). Under this view, our finding that even less proficient learners and learners without any immersion were guided by extra-sentential discourse-level information during L2 ambiguity resolution is not particularly surprising.

The hypothesis that L2 processing is generally more meaning-based than L1 processing raises the question of whether L2 learners' grammatical processing abilities can ever become native-like. While there is some evidence from previous L2 processing studies suggesting that learners' reliance on contextual cues might decrease with increasing L2 proficiency (Dekydtspotter and Outcalt, 2005), and that highly proficient learners may be native-like in their ability to integrate discourse-level and syntactic information (Hopp, 2009), this question clearly warrants further investigation.

6. Conclusion

Our results provide evidence for the immediate use of referential context information in L2 ambiguity resolution. Unlike many previous L2 processing studies which identified domains in which learners showed reduced processing ability or cue sensitivity compared to native speakers, the results from the current study suggest that learners' sensitivity to discourse-level pragmatic cues to interpretation may in fact be stronger than native speakers'. An enhanced ability to efficiently exploit non-structural cues to interpretation during processing may allow learners to compensate for potential grammatical processing difficulties, thus helping ensure successful L2 comprehension. Future research will show whether our findings generalize to other linguistic phenomena and other L1/L2 combinations, and how learners' on-line sensitivity to discourse-level information might change across a broader range of proficiency levels.

References

Allan, D., 2004. The Oxford Placement Test [2nd Edition]. Oxford University Press, Oxford.

Altmann, G., Steedman, M., 1988. Interaction with context during human sentence processing. Cognition 30, 191-238

Binder, K., Duffy, S., Rayner, K., 2001: The effects of thematic fit and discourse context on syntactic ambiguity resolution. Journal of Memory and Language 44, 297-324.

Britt, M., 1994. The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. Journal of Memory and Language 33**,** 251–283.

Britt, M., Perfetti, C., Garrod, S., Rayner, K., 1992. Parsing in discourse: Context effects and their limits. Journal of Memory and Language 31**,** 293–314.

Clahsen, H., Felser, C., 2006. Grammatical processing in language learners. Applied Psycholinguistics 27, 3-42.

Clifton, C., Speer, S., Abney, S., 1991. Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. Journal of Memory and Language 30**,** 251–271.

Crain, S., Steedman, M., 1985. On not being led up the garden path: the use of context by the psychological parser. In: Dowty, D., Karttunen, L., Zwicky, A. (Eds.), Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives. Cambridge University Press, Cambridge, pp. 320–58.

Daneman, M., Carpenter, P., 1980. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior 19, 450-466.

Dekydtspotter, L., Outcalt, S., 2005. A syntactic bias in scope ambiguity resolution in the processing of English-French cardinality Interrogatives: Evidence for informational encapsulation. Language Learning 55, 1-36.

Desmet, T**.,** De Baecke, C., Brysbaert, M., 2002. The influence of referential discourse context on modifier attachment in Dutch. Memory & Cognition 30, 150-157.

Felser, C., Cunnings, I., submitted. Processing reflexives in English as a second language: The role of structural and discourse-level constraints. Ms. University of Essex.

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

Frazier, L., 1979. On Comprehending Sentences: Syntactic Parsing Strategies. Indiana University Linguistics Club, Bloomington, IN.

Frazier, L., Rayner, K., 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. Cognitive Psychology 14, 178-210.

Frenck-Mestre, C., 2005: Ambiguities and anomalies: What can eye-movements and event-related potentials reveal about second language sentence processing? In:

Kroll, J., De Groot, A. (Eds.), Handbook of Bilingualism. Oxford University Press, Oxford.

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

Gass, S., 1989. How do learners resolve linguistic conflicts? In: Gass, S., Schachter, J. (Eds.), Linguistic perspectives on second language acquisition. Cambridge University Press, Cambridge, pp. 183-199.

Gibson, E., Pearlmutter, N., 1998. Constraints on sentence comprehension. Trends in Cognitive Sciences, 2, 262-268.

Gibson, E., Pearlmutter, N., Canseco-Gonzales, E., Hickok, G., 1996. Recency preferences in the human sentence processing mechanism. Cognition 59, 23-59.

Guo, J., Guo, T., Yan, Y., Jiang, N. and Peng, D., 2008. ERP evidence for different strategies employed by native speakers and L2 learners in sentence processing. Journal of Neurolinguistics 22, 123-134.

Harrington, M., 1992. Working memory capacity as a constraint on L2 development. In: Harris, R. (Ed.), Cognitive Processing in Bilinguals. North-Holland, Amsterdam, pp. 123–136.

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

Hopp, H., 2009. The syntax-discourse interface in near-native L2 acquisition: Off-line and on-line performance. Bilingualism: Language and Cognition 12, 463-483.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L., Trueswell, J., 2000. One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. Journal of Psycholinguistic Research 29, 597-626.

Just, M., Carpenter, P., 1992. A capacity theory of comprehension: Individual differences in working memory. Psychological Review 99, 122-149.

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

Kaiser, E., Trueswell, J., 2004. The role of discourse context in the processing of a flexible word-order language. Cognition 94, 113-147.

Kilborn, K., 1992. On-line integration of grammatical information in a second language. In: Harris, R. (Ed.), Cognitive Processing in Bilinguals. Elsevier, Amsterdam, pp. 337–350.

Mitchell, D., Corley, M., Garnham, A., 1992. Effects of context in human sentence parsing: Evidence against a discourse-based proposal mechanism. Journal of Experimental Psychology: Learning, Memory and Cognition 18, 69-88.

Papadopoulou, D., 2005. Reading-time studies of second language ambiguity resolution. Second Language Research 21, 98-120.

Papadopoulou, D., Clahsen, H., 2006. Ambiguity resolution in sentence processing: the role of lexical and contextual information. Journal of Linguistics 42, 109–138.

Rayner, K., Carlson, M., Frazier, L., 1983. The interaction of syntax and semantics during sentence processing: eye movements in the analysis of semantically biased sentences. Journal of Verbal Learning and Verbal Behavior 22, 358-374.

Roberts, L., Felser, C., in press. Plausibility and recovery from garden-paths in second-language sentence processing. *Applied Psycholinguistics.*

Roberts, L., Gullberg, M., Indefrey, P., 2008. On-line pronoun resolution in L2 discourse: L1 influence and general learner effects. Studies in Second Language Acquisition 30, 333-357.

Schütze, C., Gibson, E., 1999. Argumenthood and English prepositional phrase attachment. Journal of Memory and Language 40, 409-431.

Spivey, M., Tanenhaus, M., 1998. Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. Journal of Experimental Psychology: Learning, Memory, and Cognition 24, 1521-1543.

Spivey-Knowlton, M., Sedivy, J., 1995. Resolving attachment ambiguities with multiple constraints. Cognition 55, 227-267.

Spivey-Knowlton, M., Tanenhaus, M., 1994. Referential context and syntactic ambiguity resolution. In: Clifton, C., Frazier, L., Rayner, K. (Eds.), Perspectives on Sentence Processing. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 415–39.

Spivey-Knowlton, M., Trueswell, J., Tanenhaus, M., 1993. Context effects in syntactic ambiguity resolution. Canadian Journal of Experimental Psychology 47, 276-309.

Taraban, R., McClelland, J., 1988. Constituent attachment and thematic role expectations. Journal of Memory and Language 27, 597-632.

Trueswell, J., Sekerina, I., Hill, N., Logrip, M., 1999. The kindergarden-path effect: Studying on-line sentence processing in children. Cognition 73, 89-134.

van Berkum, J., Brown, C., Hagoort, P., 1999. Early referential context effects in sentence processing: Evidence from event-related brain potentials. Journal of Memory and Language 41, 147-182.

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

Wilson, R., Dewaele, J.-M., 2010. The use of web questionnaires in Second Language Acquisition and bilingualism research. Second Language Research 26, 103-123.

Ying, I., 1996. Multiple constraints on processing ambiguous sentences: Evidence from adult L2 learners. Language Learning 46, 681-711.

Zagar, D., Pynte, J., Rativeau, S., 1997. Evidence for early-closure attachment on first-pass reading times in French. The Quarterly Journal of Experimental Psychology 50A, 421-438.