

Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings

C. Frenck-Mestre

Centre National de Recherche Scientifique

Laboratoire Parole et Langage, Université de Provence

SUMMARY

The complex trace of saccades, fixations and regressions that the eyes make while taking in a line of text is unquestionably one of the richest accounts available as concerns the process of reading. Recording these jumps, stops and re-takes provides a to-the-letter, millisecond-precise report of the readers' immediate syntactic processing as well as revisions thereof. In addition, the influence of innumerable factors--from low-level visual conditions to high-level pragmatic cues--on the reading process can be measured via this method, thereby rendering possible the testing of various psycholinguistic models of parsing and comprehension. For all of these reasons, eye-movement recording has become an invaluable tool in the study of how readers process text. Interestingly, however, only a handful of eye-movement studies have been published as concerns reading in the second language. The goal of the present paper is to outline the findings of these second-language studies and, hopefully, to encourage future work in the field using this tool.

An overview of eye-movement recording

What exactly does one measure when recording eye-movements, and how? What basic factors need be taken into account before using eye-movements as the dependent variable in any psycholinguistic experiment? To answer these questions, we will provide a brief overview of the basic findings in the literature (for a more

complete description, cf. the chapters on eye-movement recording in monolingual psycholinguistic studies by Brysbaert & Vitu, 1998; Rayner & Sereno, 1994).

Figure 1 presents an example of an actual recording obtained in our laboratory from a native reader of French while she read this structurally ambiguous experimental sentence (cf. Frenck-Mestre, in press; forthcoming, for a theoretical description of how native and non-native readers process this ambiguity). Various different standard measures of eye-movements can be illustrated. Indeed, the trace left by the eyes as they scan a line of text (whether isolated or in context) can be broken down into several measures. Namely, as concerns the measurement of time, we can distinguish between "first pass" reading time, "second pass" reading time and "total" reading time (the sum of the former two measures), for any given region of interest (ROI). First pass reading can furthermore be broken into "first fixation", i.e. the first time the eyes land in a given region of interest (ROI), and "gaze duration", i.e. all fixations in the ROI from the first fixation until the eyes exit to the right or left. In addition to the measurement of time, fixations provide other indications of processing. Namely, one can tally the number of first-pass fixations as well as re-fixations in a region, and determine both the length of saccades between fixations and the probability of skipping over an ROI, all of which can be indicative of processing difficulty. In addition, the pattern of regressions, as well as the probability of re-fixating a region provide valuable information as concerns the eventual difficulty of text processing. To illustrate, let us consider the eye-movement record for the sentence in Figure 1. For the first segment, "L'atelier" (i.e. "The workshop"), we have a first fixation of 180 ms on the capital letter "L" and a second fixation on the letter "i" for 260 ms; hence in this ROI we have two first pass fixations, with a gaze duration of 440 ms. There are no second pass fixations in this region, hence total reading time is equivalent to first pass gaze duration. Consider, however, the region containing the adverb "pendant" (i.e. "while"), and the disambiguating region, comprised of the auxiliary "a" ("has") and the past participle "ouvert" ("opened"). Both regions show increased processing time, due to re-reading i.e. "second pass" fixations, accompanied or not by regressive saccades. The adverb

was initially processed in a single fixation (the eye being placed at the letter "p"), for 420 ms. It was re-read, however, for 510 ms, following a regressive saccade which was initiated at the disambiguating region. The disambiguating region itself was initially read for 300 ms in a single fixation (at the auxiliary "a"), and re-read in two second pass fixations (290 and 220 ms, at the space prior to "a" and at the letter "e" in "ouvert" respectively). Total reading time for the disambiguating region was thus over one second (1100 ms), and close to one second (950 ms) for the adverb adjacent to the area where the ambiguity began (at the word "refait" i.e. "redone", which could be either a past participle form, as in the example, or an adjective). Hence, both the disambiguating region and the region prior thereto show effects of processing difficulty. It is important to note that both of these regions were initially fixated, and within a standard time frame (an average fixation being roughly 240 ms), i.e. that re-fixations were not due to initial skipping of the region. Finally, we can note the last fixation, on the last word of the sentence ("tôt" i.e. "early"). As is clearly visible, this fixation was quite long (670 ms) and roughly two and a half times as long as the mean of all the other fixations in the sentence. This reflects the well-known "sentence wrap-up effect" (Just & Carpenter, 1980), whereby readers spend more time at sentence final words not because of the word itself but rather because the entire sentence can be comprehended at this point. As such, the end word of a sentence is generally a poor region to use as one's dependent measure. In like fashion, the very beginning of the sentence is a poor region for analysis as is the end of a line of text prior to a line break, as these regions are skipped far more frequently than other regions of the sentence.

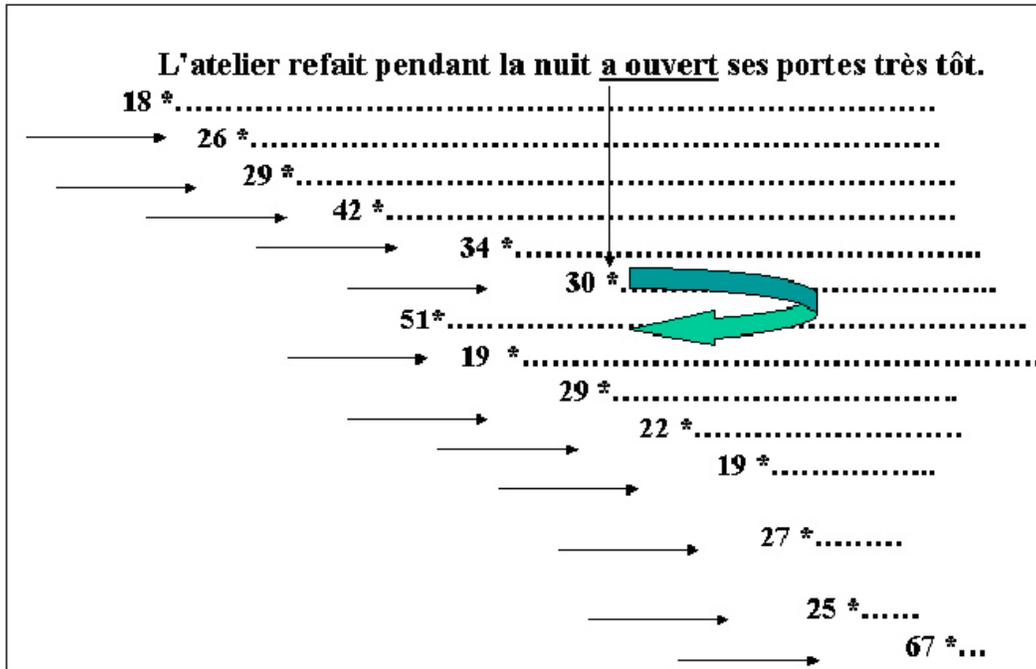


Figure 1. Eye movement record of reading a structurally ambiguous sentence.

The above illustration, based on an actual recording, clearly shows how eye-movement recordings can translate both the reader's initial and subsequent processing of a sentence. It is also very clear from this example that eye-movement data provide an extremely rich data set; numerous dependent variables are present, and may be used to determine *when* (during the first or second pass through a sentence) and *where* (at a disambiguating point, prior thereto, immediately thereafter, etc.) processing difficulty occurs, and *how* the reader deals with this difficulty (by fixating longer, by making a regressive saccade to an earlier point in the sentence, by re-reading a region). It is important to note, nonetheless, that eye-movement patterns (i.e. the percentage of forward to regressive saccades, the likelihood to either regress to a previous word or to skip an upcoming word, and the precision of a regression), as well as fixation durations (i.e. how long the reader keeps her eyes in a particular region) can only be used as an indication of processing difficulty when compared to an equivalent; experimentally controlled alternative condition. It is this issue that we will now turn to, i.e. what are some of

the factors to consider when constructing clean eye-movement materials. We will return to the issue of multiple processing measures momentarily.

Prior to using eye-movements to assess psychological aspects of reading one must first ensure that all physical aspects have been neutralized. That is, as in any experiment, all factors must be controlled in order for one to be able to make any useful conclusions from the data. While this may seem self-evident, there are several factors that can affect the progression of saccades across a sentence, the duration of fixations and the probability of skipping and/or refixating a particular word that may not be immediately obvious to the researcher whose main thrust is the study of syntactic manipulations. Indeed, the debate yet rages as to whether "low-level" factors, linked to physical properties of words and the ocular-motor system itself, or "high-level" linguistic factors play the major role in deciding where to send one's eyes and for how long during reading (cf. Brysbaert & Vitu, 1998; Rayner & Well, 1996, for very different viewpoints). For example, it is well documented that the probability of fixating a word "n" during the first pass through a sentence as well as the time spent on word n when fixated is influenced by several low-level factors such as how long the eye fixated on the word n-1 and how far away the eye was from word n (i.e. the launch site from the word n-1), where the eye landed in word n, and the length of word n itself. However, there is also substantial evidence that more "linguistic" variables play a role, such as lexical frequency and the predictability of the word in the sentence context. Ocular-motor and lexical factors may also interact. For example, short (2 to 3) letter words are often skipped (roughly 70% on average). This can be due to the fact that the average saccade length often precludes landing in such short regions, and/or to the fact that a short region can be read in the periphery if the eye stays long enough at the end of the prior word, and/or to the fact that highly predictable words are skipped more often in general. Hence, both psychophysical and psycholinguistic advocates would predict, and be able to explain, why short words are more likely to be skipped (and hence, why it would be wise to avoid using short words as ROIs). Let us simply remark that there is strong evidence for both

viewpoints, and any psycholinguistic researcher would do well to take heed of all "low-level" factors when setting up sentence materials.

Measurements of reading : Eye-movements vs. self-paced reading.

By far the most important difference between the measurements obtained via the recording of eye-movements and self-paced reading times is the level of definition. As outlined above, eye movements provide a multifaceted trace of the reader's processing of text. In comparison, self-paced reading generally provides but a single measure: the total reading time of a given segment. In eye-tracking, one can distinguish first pass measures, which are often associated with 'initial' parsing decisions, from second pass measures, which are more often taken to reflect re-analysis (although a note of caution is in order as there is not a systematic relationship). This is not the case of self-paced reading, where initial analysis and re-analysis cannot easily be distinguished. Note that if one records not only the time spent reading a segment, but, as an independent measure, the *pause* time between key-presses (i.e. following the key-release for segment A and prior to key-pressing for segment B) one can in fact render the self-paced reading measurement more fine-grained. In this way, the "spill over" effects linked to the processing of the prior segment can be estimated and effects that may not otherwise be observed may become apparent. Indeed, it has been shown that different types of readers differentially show effects on either reading or pause times (Lété, Zagar & Pynte, 1994). Self-paced reading is a quick and readily available tool, which has provided a wealth of information (cf. Mitchell, in press; Papadopoulou, this issue). However, we would like to suggest, as illustrated by the results of a recent bilingual study (Altarriba, Kroll, Scholl and Rayner, 1996), that self-paced reading time can potentially mislead the researcher as concerns second language processing.

In their study, Altarriba et al. (1996) recorded the eye-movements of Spanish-English bilinguals while they read either mixed-language or monolingual sentences. In both conditions, two factors were manipulated: the printed frequency

and the Cloze frequency (i.e. sentence constraint) of the target word in the sentence. Examples 1 through 4, below, illustrate the sentences used, where the underlined word indicates the within- and between-language target. The aim of the study was to determine which of two factors--contextual or lexical constraints--played the dominant role in facilitating the processing of the target word in the sentence. This debate has indeed been the subject of numerous monolingual studies of sentence processing (cf. Simpson, 1994, for an extended review). Altarriba et al. (1996) used their Spanish-English bilinguals' capacity (and reported natural tendency) to switch between their two languages to test the "conceptual" versus "lexical" hypothesis of sentence-context effects. Indeed, if the conceptual context provided by the sentence plays the dominant role in facilitating the processing of an upcoming word, then lexical properties, such as the co-occurrence value between two words in the sentence (i.e. "country" and "city", "Halloween" and "pumpkin" which have high co-occurrence, vs. "country" and "ciudad", "Halloween" and "calabaza" which do not) should not override it, and equal facilitation should be observed for monolingual and mixed-language sentences.

1. She moved from the country to the city / ciudad to find a better job (High constraint, high frequency)
2. We took a walk in the city / ciudad before we drove back home. (Low constraint, high frequency)
3. On Halloween the children carved an orange pumpkin/calabaza for the front steps. (High constraint, low frequency)
4. The market had a new variety of pumpkin / calabaza in the fall. (Low constraint, low frequency)

Altarriba et al. used several dependent variables to measure facilitated processing of the target word: the probability of skipping the target word (cf. Balota, Pollatsek & Rayner, 1985), first-fixation time on the target, and first-pass gaze duration on the target. Note that neither skipping probability nor first-fixation times are available in self-paced reading. This is noteworthy, as the significant results reported by Altarriba et al. involve exactly these measures. For both monolingual and mixed-language sentences, the bilingual readers skipped over the target word more often when it was placed in a highly constraining than an unconstrained sentence context. Hence, for this measure, sentential context apparently outweighed lexical constraints. However, the pattern of first fixations

revealed different patterns of results for within- and mixed-language sentences. For monolingual sentences, first-fixations were shorter for highly-constrained targets than for unconstrained targets and the effect was independent of the frequency of the target. Again, thus, sentence constraint outweighed lexical probability. For mixed-language sentences, however, first-fixation durations revealed an interaction between sentential constraint and frequency. For low frequency Spanish targets high-constraint sentences facilitated processing compared to low-constraint, whereas for high frequency Spanish target words the opposite was true. Altarriba et al. (1996) concluded from these results that both sentential context and lexical probability play a role in facilitating the processing of upcoming words. When highly constrained by the sentence context, a high-frequency word in the wrong language did not meet the lexical predictions and caused a penalty in initial reading times. Yet, when subjects took longer to initially read the target words, as was the case for low-frequency Spanish target words, the effect of lexical probability was outweighed by sentence context¹.

Two important conclusions can be drawn from the Altarriba et al. (1996) study. First, the bilingual population provided the ideal situation for pitting the "conceptual" versus "lexical" hypotheses of sentential context against each other. While no effect of lexical probability was obtained in the monolingual sentences, the data from the mixed-language sentences made it clear that this factor can play a strong role. Second, and in relation to our argument as concerns the sensitivity of measures, the recording of eye-movements provided a multifaceted record which allowed the authors to reach the above conclusions. Note, again, that no significant results were reported as concerns sentential constraint on *gaze duration*. As such, had the study been run using self-paced reading, where only gaze duration is available, the researchers would have come to very different conclusions than those they reported in their eye-movement study.

Another illustration of how eye-movements can provide a more complete picture of bilingual sentence processing is provided by the results of an experiment on syntactic ambiguity resolution (Frenck-Mestre & Pynte, 1997). The aim of the experiment was to establish whether lexical properties of the critical verb in the

sentence (italicized in the examples) would affect syntactic ambiguity resolution, and whether such would hold for non-native as well as native readers of French . Examples of the sentences used are provided in 5 and 6, below, with the ambiguous site underlined.

5. Il *rate* le train de peu / de midi et décide alors de chercher un hôtel.

(He misses the train by little / of noon and decided thus to look for a hotel.)

6. Il *accuse* l'ambassadeur d'Indonésie / d'espionnage mais il n'est pas certain des faits.

(He accuses the ambassador of Indonesia / of espionage but he isn't certain of the facts.)

In 5, the critical verb is a monotransitive verb and generally takes but one argument to fulfil its thematic grid. In 6, the critical verb is ditransitive and generally takes two arguments to fulfil its thematic grid. Frenck-Mestre and Pynte (1997) found that attachment preferences differed as a function of verb type, for both native French speakers and English-French bilinguals. For both groups, monotransitive verbs lead readers to prefer "low" attachment, i.e. they spent less time reading sentences in which the target segment modified the preceding noun ("de midi" in the example) than those in which the target segment necessitated modifying the preceding verb ("de peu" in the example). The exact opposite was true for ditransitive verbs, with "low" or NP attachment ("d'Indonésie" in the example) taking longer to process than "high" or VP attachment (d'espionnage" in the example). This interaction was apparent in the mean first pass gaze durations of both groups. However, it is noteworthy that the effect of verb type appeared earlier for the bilingual readers, reading in their second language, than in the group of native speakers. That is, for the English-French bilinguals, a significant interaction was found on mean first fixation duration, i.e. the from the very first time the eye landed in the ambiguous region of the sentence. This was not true of the group of native readers. While the latter result did not change the general conclusion, it did reveal differential processing between non-native and native readers, showing even greater sensitivity to verb constraints in the non-native group. Had the study been run with a less sensitive measure, this effect would not have been apparent.

In sum, the above two studies clearly demonstrate the advantage for recording eye-movements when examining the immediate influence of lexical- and sentence-level variables on comprehension. Further work with this measure is to be encouraged in the field of second language processing, especially given the present paucity of studies.

Studies of second language sentence processing: Eye-movements vs. event-related potentials.

We have argued that eye-movement recording is perhaps the richest record of a reader's linguistic behaviour as s/he unravels the information present in a piece of text. Yet, another on-line multidimensional measure has also been used in the last twenty years with great success to examine on-line syntactic (and semantic) processing in monolinguals and more recently in bilinguals. Indeed, the recording of event-related potentials (ERPs) during the visual (word-by-word) and/or auditory presentation of sentences provides a to-the-millisecond precise record of how readers' (brains) react to materials that are either ambiguous or anomalous from either a syntactic or semantic point of view (cf. Osterhout, McLaughlin, Kim, Greenwald & Inoue in press; Müller, this issue, for recent reviews). Furthermore, syntactic and semantic processing apparently provoke different electrophysiological responses (but see Osterhout et al., in press), which can aid in answering theoretical questions about which types of information are taken into account at what particular stage of processing.

Of interest to the present paper are the apparent discrepancies between data obtained via on-line reading measures on the one hand and ERPs on the other as concerns syntactic processing in the second language. Based on comparisons between native and non-native readers' ERP traces to syntactically *anomalous* sentences, it has been argued by some (Hahne & Friederici, 2001; Sanders & Neville, 2003; Weber-Fox & Neville, 1996) that non-native readers do not show the same degree of automaticity when processing such material. This has been argued due to the absence, even in proficient non-native readers, of "early" components (notably early anterior negativities) and, in some studies, to the

delayed onset of later components, notably the P600 (this component, first evidenced in monolinguals by Osterhout, 1990, and Osterhout & Holcomb, 1992; is quite often associated with syntactic processing and syntactic integration difficulty). This claim is at odds with the data from several eye-movement studies of second language syntactic processing, showing immediate use (as early as the first fixation on a target word of the sentence) of both lexical and sentence-level information during the processing of unambiguous as well as structurally ambiguous sentences (Altarriba et al., 1996; Frenck-Mestre, 1997; 2002, in press; Frenck-Mestre & Pynte, 1997). It is also at odds with the results from several self-paced reading studies of second-language sentence processing showing that, provided strong second language skills, the L2 reader behaves in a virtually indistinguishable manner from the native speaker when engaged in syntactic processing (Hoover & Dwivedi, 1998; Juffs, 1998; Juffs & Harrington, 1996, but see Felser, Roberts, Gross & Marinis, 2003; Marinis Roberts, Felser & Clahsen, 2003). Where might the differences lie?

We believe that the rift between ERP and on-line reading studies of second language syntactic processing may be (at least somewhat) lessened if two considerations are taken into account (see also Frenck-Mestre, in press). The first is a methodological issue, related to ERP components and what they tell us. The second is related to the types of structures that have been examined with the two methods. To begin, let us consider the ERP components themselves. A quick overview of the second language ERP literature reveals that by far the most consistently reported difference between native and second language syntactic processing involves the presence vs. absence of a distinct left-hemisphere dominant "early negativity" in the non-native group². This has been reported in several studies, in two different laboratories and using various native and second language pairings (Hahne, 2001; Hahne & Frederici, 2001; Hahne, Müller & Clahsen, 2003; Weber-Fox & Neville, 1996). In contrast, the P600 signature (i.e. between 500 and 800 ms post stimulus word, central, and larger posteriorly) has been consistently reported in the second as well as native language in studies of syntactic anomaly processing (Hahne, 2001; Hahne & Frederici, 2001; Hahne, et

al., 2003; Weber-Fox & Neville, 1996). It is noteworthy that only one of these studies reported a delayed onset of the P600 for proficient bilinguals compared to native speakers. We ourselves (Foucart & Frenck-Mestre, forthcoming; Frenck-Mestre, 2004) found that grammatical gender agreement errors between the noun and preceding definite article in a sentence context, such as illustrated in examples 7 and 8, produce a large P600 in both native French speakers and non-native (German) speakers.³ Results for the native French speakers, showing a large P600 response to French nouns preceded by an article with the wrong gender is presented in Figure 2. We did not find a delay in onset of the P600 in the non-native group.

7. Hier la neige était partout dans le jardin
(Yesterday the(fem) snow(fem) was all over the garden.)
8. Hier le neige* était partout dans le jardin
(Yesterday the(masc) snow(fem) was all over the garden.)

Moreover, although we did see some early negativity, we did not find evidence of a left-lateralized early negativity in response to grammatical gender violations in either native or non-native readers. Indeed, early negativities are not consistently reported for syntactic anomalies in the monolingual literature (the examples abound where P600 is found in the absence of early negativities) and there is a large current debate as to the significance of the P600 and "early negativities" as concerns syntactic processing. While some have argued that differences in "automatic" syntactic processing and more effortful "repair" processes can be indexed by these two components (cf. Friederici, Hahne & Mecklinger, 1996) this is not accepted by all. It would seem indeed that the lack of early negativities in non-native readers cannot be clearly interpreted until some consensus in the monolingual literature has been achieved.

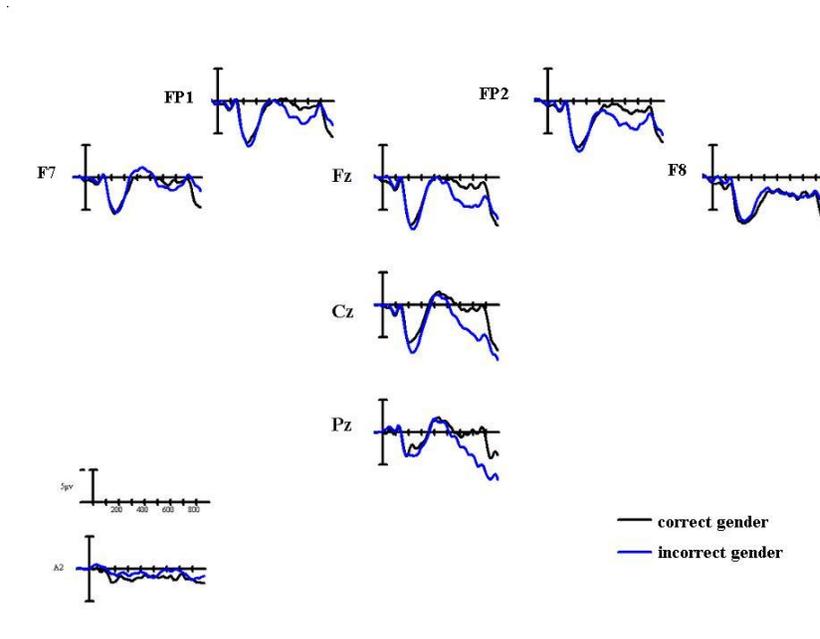


Figure 2. P600 response to grammatical gender error in French

The second issue is that of the type of syntactic processing that has been examined via reading studies and ERP studies. In all of the second language studies cited above, syntactic anomalies rather than ambiguities have been the subject of attention. The opposite is true of the majority of reading studies, where structural ambiguity has been used to test processing abilities. In the one second language ERP study that we are aware of where structural ambiguity and verb subcategorization information was at play (Kotz, 1991), the author reports use of the subcategorization information by the non-native readers (Spanish-English bilinguals) in like fashion to native readers, and equivalent P600 responses in the two groups when the ambiguous structure was not resolved in the manner expected. This concurs with the results from previous reading studies showing use of lexical information in the second language (Frenck-Mestre, in press; Frenck-Mestre & Pynte, 1997; Juffs, 1998). It would be of great interest to see more ERP studies of second language syntactic *ambiguity* processing, to determine whether or not the data concur with those from reading studies.

In line with the above, it would be of interest to compare ERP and eye-movement second language data for different types of ambiguities and/or anomalies. To date, only a handful of studies have examined this and in the native language only. In a recent ERP study of monolingual English readers, Osterhout and colleagues (Osterhout et al., 1994; Osterhout & Nicol, 1999) showed that sentences of the type illustrated in 9, below, elicit separable electrophysiological responses depending upon the type of anomaly.

9. The black widow spider likes to [hide/sigh/hiding/sighing] in dark places

Compared to semantic and syntactically correct verbs ("hide"), semantic (selectional restriction) violations ("sigh") produced an increased N400 response, whereas syntactic (verb tense) violations ("hiding") elicited a P600 effect. Quite interestingly, verbs that presented both types of violation ("sighing") produced both of these responses. These results suggest that how readers resolve anomalies can be tracked quite efficiently via the recording of ERPs. Eye-movement data as well can provide a window onto different processing types, although more inferences may be involved. Indeed, Ni and collaborators (Braze, Shankweiler, Ni & Palumbo, 2002; Ni, Fodor, Crain & Shankweiler 1998) examined processing of the same anomalies as illustrated in 9 via the recording of eye-movements. There was only weak evidence that initial (first pass) reading times distinguished the processing of these two types of anomalies. The pattern of regressions, however, did differ for the two types of violations. As such, the multi-faceted record provided by eye-movement recording can provide information about how a reader differentially resolves different processing types (although it must be said that ERPs are more revealing in this aspect, as quite distinct patterns are obtained for semantic and syntactic violations). Work along these lines in the second language of skilled bilinguals, or indeed as a function of processing skill in the second language, should provide extremely informative and interesting information.

The influence of the native language in second language syntactic parsing :
Evidence from eye-movements and self-paced reading

In several prior publications we have examined the on-line processing, via the recording of eye-movements, of a particular structural ambiguity, illustrated in 10 and 11. This particular ambiguity has been the subject of a plethora of papers involving the study of at least fifteen different languages (for recent reviews, see Cuetos, Mitchell & Corely, 1996; Frazier & Clifton, 1996; Frenck-Mestre & Pynte, 2000a, 2000b).

10. Jean saw the daughter of the woman who was leaving the shop.

11. Jean saw the daughter with the woman who was leaving the shop.

The reason for testing this structure across so many studies and languages lies in its unique sensitivity to linguistic as well as extra-linguistic variables. It has been shown that, even within a language, whether one prefers the subject of the relative clause to be the head noun (the daughter) or the second noun (the woman) depends upon the frequency of the two nouns (Pynte & Colonna, 2000), the length (or "weight") of the relative clause (Fodor, 1998; Pynte & Colonna, 2000), how the sentence is segmented (due assumedly to the effect of segmentation on prosody, cf. Gilboy, & Sopena, 1996 as well as Fodor, 1998, but also Carreiras & Clifton, 1999), training on a given structure (Frenck-Mestre & Pynte, 2000a) and whether one considers off-line measures such as questionnaire or corpus data as opposed to on-line measures of reading (De Vincenzi & Job, 1995; Mitchell, Brysbaert, Grondelaers, & Swanepoel, 2000). The preposition also plays a major role, with almost universal "N2" or "low" attachment of the relative preferred, both on and off-line, when the preposition has semantic content, as in 11 (De Vincenzi & Job, 1995; Frenck-Mestre & Pynte, 2000b; Gilboy, Sopena, Clifton & Frazier, 1995), while wide cross-linguistic variation is observed when the preposition is devoid of semantic content, as in the example 10 (cf. Mitchell, Brysbaert, Grondelaers & Swanepoel, 2000, for a review). Given

the number of factors that affect its resolution, no single model can account for how this ambiguity is resolved (see Cuetos et al., 1996; Frazier & Clifton, 1996; Gibson Pearlmutter, Canseco-Gonzalez & Hicock, 1996; Mitchell et al., 2000; for different proposals).

Of interest for the present purposes is how non-native readers treat this ambiguity. Three factors will be considered: the two languages of the reader, their level of experience in the second language, and the sensitivity of the measure used to examine processing. The language pairings that are used provide a test-bed for competing models of how this ambiguity is resolved. In like manner, examining readers of different levels of competence in their second language--measured by both numbers of years of using the L2, of formal instruction, and by standard test scores--allows one to pit models of processing against each other. Last, whether one uses eye-movements or self-paced reading to test how readers resolve this ambiguity will provide the investigator with different means with which to make their claims.

Consider first the data from a series of eye-movement studies (Frenck-Mestre, 1997; 2002), which examined resolution in French of the ambiguity presented in 10, for native English speakers on one hand and Spanish speakers on the other. Proficiency was measured primarily by years of residence in France (all subjects had roughly equivalent formal instruction in French in scholastic settings during adolescence and early adulthood), either "low" (1 year of residency) or "advanced" (5 years residency). These language pairings allow one to compare, on the one hand, two languages--Spanish and French--for which clear "N1" or "high" attachment preferences have been reported in monolinguals when 1) the nouns are of equal frequency, 2) the preposition is "of", 3) the relative clause contains several words, 4) no prior training or massive exposure to an alternative structure has been given, 5) no segmentation is present (cf. Carreiras & Clifton, 1993, 1999; Cuetos & Mitchell, 1988; Cuetos et al., 1996; Frenck-Mestre & Pynte, 2000b; Gilboy et al., 1995; Zagar, Pynte & Rativeau, 1997). The other pair of languages--English and French--provides a comparison of languages which do not coincide as concerns the immediate resolution of this ambiguous

structure. English shows a relatively systematic bias for "low" attachment of the relative clause, i.e. to the second noun of the complex NP (Carrieras & Clifton 1999; Dussias, 2001; Fernandez, 1999; Frazier & Clifton, 1996), although there is some debate as to how strong the N2 bias is in this language (Carrieras & Clifton, 1999; 1993)

The results from these eye-movement studies, reproduced in Figure 3, show that Spanish-French and English-French beginning bilinguals differ during the on-line processing of the NP1-Prep-NP2-RC ambiguity in French, as measured by first pass reading times. Both groups apparently apply the same strategy as used in their native language, that is, high attachment for Spanish and low attachment for English⁴. One could argue, in line with Gibson et al.'s (1996) recency/predicate proximity model, that two competing principles are available to resolve this ambiguity which have different weights cross-linguistically, and our bilingual readers used different principles. Spanish would be a "predicate proximity" language (i.e. take the noun closest to the main verb) whereas English would weight recency more heavily (i.e. take the noun most recently processed), and our non-proficient bilinguals acted accordingly when processing this ambiguity in their second language, independent of the weight of the two principles in the latter. This argument is bolstered by the data of more skilled English-French bilinguals, also reported in Figure 3 (rightmost bars). These readers, who had been living in France for a mean of 5 years, showed a change in preference, with an N1 preference akin to that found for the control group of native French. Again, in line with Gibson et al.'s model, as well as Mitchell and collaborators model of "linguistic tuning" (Brysbaert & Mitchell, 1996; Cuetos et al., 1996), one can account for the results by assuming that the L2 reader is now applying the weights assigned by his/her second language due to the prolonged exposure to this language (and the lessened exposure to the native language). A look at Figure 3 reveals that readers' attachment preferences were declared early, within the first 500 ms of processing, during the initial (first pass) reading of the segment that disambiguated the sentence. In addition, but not depicted in the graph, numerically more regressive saccades were launched from this region to

early parts of the sentence, as compared to the preferred structure (cf. Frenck-Mestre, 2002). Both of these measures indicate that readers experienced immediate difficulty when sentences were disambiguated in a manner that did not meet their expectations.

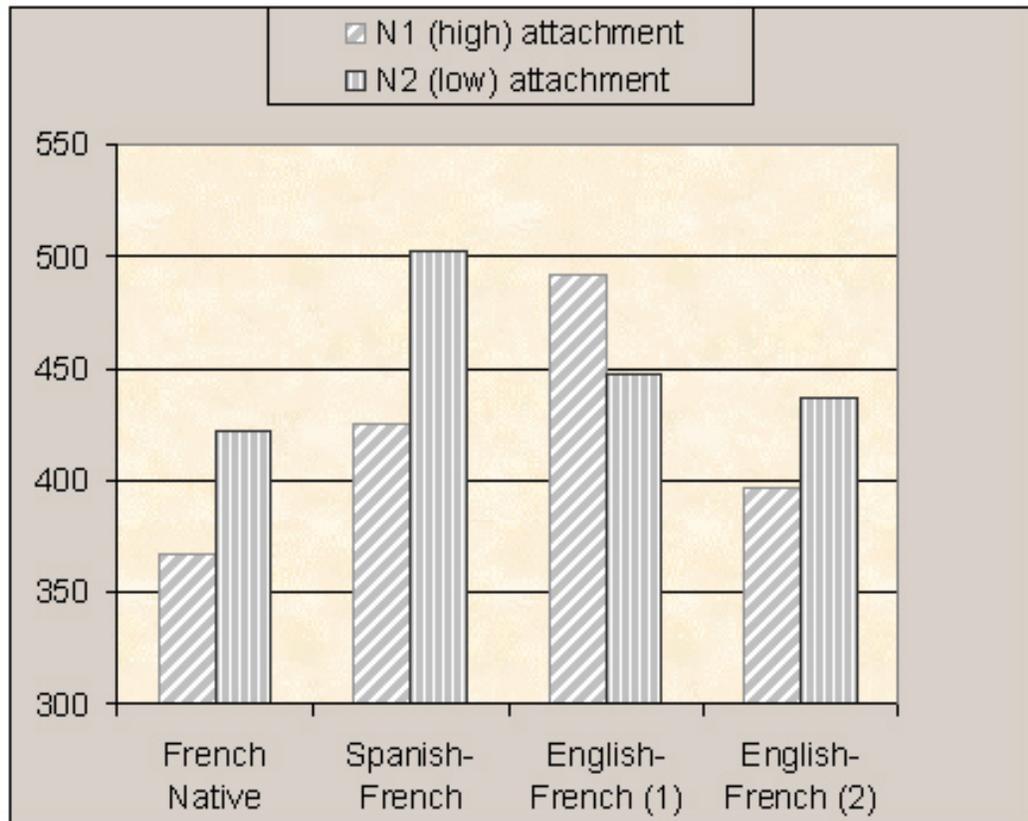


Figure 3. Mean first pass gaze duration on the disambiguating verb for NP1-Prep-NP2-Relative clause constructions (e.g. *Jean a vu [la gardienne des filles /les filles de la gardienne] qui revient de Paris*: "John saw the [nanny of the girls/ the girls of the nanny] who come back from Paris"), as a function of type of attachment (to NP1 or NP2) and type of reader (native French, beginning Spanish-French bilinguals, beginning English-French bilinguals (1), proficient English-French bilinguals (2)).

The results from self-paced studies of the same ambiguity (cf. examples 10 and 11) in non-native readers (Fernandez, 1999; Felser et al., 2003, Papadopoulou & Clahsen, 2003) have not always provided similar conclusions to those obtained in eye-movement studies, however. Felser et al. (2003) used self paced reading to look at German-English and Greek-English proficient bilinguals'

attachment preferences for this temporarily ambiguous structure, in English. Unlike our data, the authors found neither a clear "native language" strategy (which would have dictated N1 attachment) nor a significant "second language" strategy (where N2 should have been preferred). The authors argue that 'transfer' cannot explain their results. We believe, however, that the lack of any clear preference can be readily explained if one assumes that these readers were in a transitional stage, having moved beyond systematically applying the principle dominant in their L1 but not yet to a point of systematically applying that of their L2. Alternatively, it could be argued that some of the subjects in the two groups applied one strategy and others another, resulting in a null overall effect. Our own preference leans toward the former hypothesis, whereby (at least the majority of) subjects were in a state of flux. It would be highly interesting to see a follow-up study of these two subject pools after several years of residency in the U.K., to see if a stable preference would evidence itself. Indeed, the self paced reading data from Fernandez (1998) looking at Spanish-English and Japanese-English bilinguals treatment of this ambiguity show that the best predictor of attachment preference was proficiency in the second language based on self-reports (and more so than the age of acquisition of the language). Those who rated themselves as dominant in their native language showed effects of forward transfer whereas those who thought of themselves as dominant in the second language, English, showed preferences similar to those of native English speakers. The same argument has been forwarded by Dussias (2001), where she found that native Spanish speakers having lived in an English speaking environment for numerous years actually showed low attachment preferences (on off-line questionnaires) when processing *Spanish* NP1-de-NP2-RC ambiguities, whereas Spanish monolinguals showed high attachment preference. This upset of preference in the native language is indeed surprising, but in line with models which assume a dominant role of linguistic experience (cf. MacDonald, Pearlmutter, & Seidenberg; Mitchell, 1994). Note, nonetheless, that Papadopoulou & Clahsen (2003) have argued from self-paced reading data, obtained with proficient Russian, Spanish and German readers of Greek, that

fluctuating cross-linguistic processing preferences cannot explain the L2 processing of these subjects when reading ambiguities such as 10 or 11. They argue that L2 processing is driven by lexical properties, more than structurally based ones, whereas native language processing reflects an integration of the two (cf Marinis et al., 2003 for a similar stance). Their account is interesting, and indeed prior eye-movement studies have shown that semantic and lexical properties have immediate effects upon L2 processing (Altarriba et al., 1996, Frenck-Mestre & Pynte, 1997), and can disrupt analysis. Note, however, as outlined above, that self-paced reading may be less sensitive to the difference between initial analysis and re-analysis than eye-tracking. Given the quite long reading times reported in Papadopoulou & Clahsen (2003) for the disambiguating word (over 2 seconds), it is not implausible that both of these measures were included in their reading times, making it thus rather difficult to make any strong claims as concerns readers' immediate processing preferences.

Concluding remarks

In the present review we have made a strong case for the advantages of using eye-tracking as a means to study second language syntactic and semantic processing during sentence interpretation. We have also underlined the need to pay close attention to "low level" psychophysical factors when using this technique, the importance of which is often underestimated. Moreover, while we fully promote the use of eye-tracking, it is our belief that a complete picture of second language sentence processing can only be obtained via the combined use of several techniques. Notably, we have suggested a need to compare the results from the three major on-line tasks of reading--eye movement recording, self paced reading and event-related potentials--for equivalent structures and populations to determine the cause for possible differences across studies.

References

- Altarriba, J., Kroll, J.F., Scholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye-fixations and reading times. *Memory & Cognition*, 24, 477-492.
- Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal information in reading. *Cognitive Psychology*, 17, 364-390.
- Braze, D., Shankweiler, D., Ni, W., & Palumbo, L.C. (2003). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research*, 31, 25-44.
- Brysbaert, M., & Mitchell, D.C. (1996). Modifier attachment in sentence processing: evidence from Dutch. *Quarterly Journal of Experimental Psychology*, 49A, 664-695.
- Brysbaert, M., & Vitu, F. (1998). Word Skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* Oxford: Elsevier, pp. 125-147.
- Carreiras, M., & Clifton, C. (1993). Relative clause interpretation preferences in Spanish and English. *Language and Speech*, 36, 353-372.
- Carreiras, M., & Clifton, C. (1999). Another word on parsing relative clauses: Eye-tracking evidence from Spanish and English. *Memory and Cognition*, 27, 826-833.
- Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the late closure strategy in Spanish. *Cognition*, 30, 73-105.
- Cuetos, F., Mitchell, D. C., & Corley, M. M. B. (1996). Parsing in different languages. In M. Carreiras, J. Garcia-Albea, & N. Sabastian-Galles (Eds.), *Language Processing in Spanish*, pp 145-187. Mahwah, NJ: Erlbaum
- De Vincenzi, M. & Job, R. (1995) An investigation of Late Closure: the role of syntax, thematic structure and pragmatics in initial and final interpretation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, pp. 1303-1321.
- Dussias, P.E. (2001). Sentence parsing in fluent Spanish-English bilinguals. In J. L. Nicol (Ed.), *One mind, two languages*, pp. 159-176. Oxford: Blackwell Publishers.
- Fernandez, E. (1998). Language dependency in parsing : Evidence from monolingual and bilingual processing. *Psychologica Belgica*, 38, 197-230.
- Fernandez, E. (1999). Processing strategies in second language acquisition: Some preliminary results. In E.C. Klein & G. Martohardjono (Eds.), *The development of second language grammars: A generative approach*, pp. 217-240. Amsterdam: John Benjamins
- Felser, C., Roberts, L., Gross, R., & Marinis, T. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, 24, 453-489.
- Fodor, J.D. (1998). Learning to Parse. *Journal of Psycholinguistic Research*, 27, 285-319.
- Frazier, L., & Clifton, C. Jr. (1996). *Construal*. Cambridge, Massachusetts: MIT Press.
- Frenck-Mestre, C. (1997). Examining second language reading: An on-line look. In A. Sorace, C. Heycock & R. Shillcock (Eds.), *Language Acquisition, Knowledge Representation and Processing: GALA 1997*, pp. 444-448. HCRC, Edinburgh.
- Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. Heredia and J. Altarriba (Eds.), *Bilingual Sentence Processing*, pp. 217-236. Amsterdam: Elsevier, North Holland.
- Frenck-Mestre, C. (in press). Ambiguities and anomalies: What can eye-movements and event-related potentials reveal about second language sentence processing? In J. Kroll and A.M.B. de Groot (Eds.), *Handbook of Bilingualism*.
- Frenck-Mestre, C. (2004) Grammatical gender agreement in French for article-noun agreement in native and non-native speakers: Evidence from ERPs. *Cognitive Neuroscience Society*, San Francisco, USA, 18-20 April
- Frenck-Mestre, C., & Pynte, J. (1995). Lexical influences on parsing strategies: Evidence from eye movements. In J.M. Findlay, R.W. Kentridge & R. Walker (Eds.), *Eye movement research: Mechanisms, Processes and Applications*, pp. 433-444. Amsterdam: Elsevier, North Holland.
- Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology*, 50A, 119-148.

- Frenc-Mestre, C., & Pynte, J. (2000a). Romancing syntactic ambiguity: Why the French and the Italians don't see eye to eye. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a Perceptual Process*, pp. 549-564. Amsterdam: Elsevier, North Holland
- Frenc-Mestre, C., & Pynte, J. (2000b). Resolving syntactic ambiguities: Cross-linguistic differences? In M. deVincenzi & V. Lombardo (Eds.), *Cross-linguistic perspectives on language processing*, pp. 119-148. Kluwer Academic Press.
- Friederici, A.D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 5, 1219-1248.
- Gibson, E., Pearlmutter, N., Canseco-Gonzales, E. & Hickok, G. (1996). Cross-linguistic attachment preferences: Evidence from English and Spanish. *Cognition*, 59, 23-59.
- Gilboy, E., Sopena, J.M., Clifton, C. & Frazier, L. (1995). Argument structure and association preferences in Spanish and English complex NPs. *Cognition*, 54, 131-167.
- Gilboy, E., & Sopena, J.M. (1996) Segmentation effects in processing complex noun pronouns with relative clauses. In M. Carreiras, J. Garcia-Albea and N. Sebastien-Galles (Eds.) *Language Processing in Spanish*, Mahwah, NJ: LEA
- Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, 30, 251-66.
- Hahne, A., & Friederici, A.D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, 4, 123-142.
- Hahne, A., Müller, J., & Clahsen, H. (2003). Second language learners' processing of inflected words: Behavioral and ERP evidence for storage and decomposition. *Essex Research Reports in Linguistics*, 45, 1-42.
- Herschensohn, J. (1998). Minimally raising the verb issue. *Proceedings of the Twenty-second Annual Boston University Conference on Language Development*, A. Greenhill, M. Hughes, H. Littlefield & H. Walsh (eds.). Somerville, Mass.: Cascadilla Press, 325-336.
- Hoover, M.L., & Dwivedi, V.D. (1998). Syntactic processing by skilled bilinguals. *Language Learning*, 48, 1-29.
- Juffs, A. (1998). Main verb v.s reduced relative clause ambiguity resolution in second language sentence processing. *Language Learning*, 48, 107-147.
- Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language sentence processing. *Language Learning*, 46, 283-326.
- Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye-fixations to comprehension. *Psychological Review*, 87, 329-354.
- Kotz, S.A. (1991). Event-related brain potentials: A sensitive measurement of bilingual sentence comprehension? *Unpublished Master's thesis*. Tufts University.
- Lété, B., Zagar, D., & Pynte, J. (1994). A new paradigm for studying word recognition during reading: Explorations with word frequency effects. *International Journal of Psychology*, 29, 291-318.
- MacDonald, M.C., Pearlmutter, N.J., & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.
- Marinis T., Roberts, L., Felser, C., & Clahsen, H. (2003). Gaps in second language processing. *Essex Research Reports in Linguistics*, 45, 44-79.
- Mitchell, D.C. (1994). Sentence parsing. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*, pp. 375-409. New York: Academic Press.
- Mitchell, D.C. (in press). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. Clifton, Jr. (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERP and Beyond*. Brighton, UK: Psychology Press.
- Mitchell, D.C., Brysbaert, M., Grondelaers, S., & Swanepoel, P. (2000). Modifier attachment in Dutch: Testing aspects of construal theory. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a Perceptual Process*, pp. 493-518. Amsterdam: Elsevier, North Holland
- Ni, W., Fodor, J.D., & Crain, S. , & Shankweiler, D. (1998). Anomaly detection: Eye movement patterns. *Journal of Psycholinguistic Research*, 27, 515-539.
- Osterhout, L. (1990). Event-related brain potentials elicited during sentence comprehension. *Unpublished doctoral dissertation*. Tufts University.

- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785-806.
- Osterhout, L., Nicol, J., McKinnon, R., Ni, W., Fodor, J.D., & Crain, S. (1994). An event-related brain potential investigation of the temporal course of sentence comprehension. Poster presented at the 7th annual *CUNY Sentence Processing Conference*, City University New York, New York.
- Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, 14, 283-317
- Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (in press). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras & C. Clifton, Jr. (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERP and Beyond*. Brighton, UK: Psychology Press.
- Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, 25, 501-528.
- Pynte, J., & Colonna, S. (2000). Decoupling syntactic parsing from visual inspection: The case of relative clause attachment in French. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 529-547). Elsevier: Oxford.
- Rayner, K., & Sereno, S. (1994). Eye movements in reading: Psycholinguistic studies. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*, New York: Academic Press, pp. 57-81.
- Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, 3, 504-509.
- Sanders, L.D., & Neville, H.J. (2003). An ERP study of continuous speech processing II. Segmentation, semantics and syntax in non-native speakers. *Cognitive Brain Research*, 15, 214-227.
- Simpson, G. (1994). Context and the processing of ambiguous words. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*, New York: Academic Press, pp. 359-374.
- Weber-Fox, C., & Neville, H.J. (1996). Maturation constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231-256.
- Zagar, D., Pynte, J. & Rativeau, S. (1997). Evidence for early closure attachment on first-pass reading time in French. *Quarterly Journal of Experimental Psychology*, 50A, 421-438.

NOTES

1. The same pattern of results for both monolingual and mixed-language sentences was found in a second experiment, where another group of Spanish-English bilinguals were asked to name the target word aloud as quickly as possible upon its presentation in a rapid serial visual presentation (RSVP) paradigm.
2. Note that Sanders and Neville (2003) also claim for less automatic syntactic processing in late Japanese-English bilinguals, but using auditory materials and based on differences in N100 (linked to segmentation of the auditory stream) and "LAN" (left anterior negativity) for these bilinguals and native English speakers.
3. In this study, which will be reported in greater detail in Frenck-Mestre & Foucart (forthcoming), we manipulated the correspondence between the syntactic gender of nouns in the French and German languages. Indeed, as syntactic gender is arbitrary, the same inanimate noun can carry either the same or different genders across languages. Our results show that when the article preceding the noun in French carries incorrect gender, German-French bilinguals produce a reliable P600 and that the effect is found both when the French noun has the same gender in German (feminine = feminine, masculine = masculine) and when it has the opposite gender (feminine = masculine, masculine = feminine).
4. An alternative hypothesis for the English-French non-proficient bilinguals' attachment preference is available, but in our minds less plausible. That is, it is conceivable that these readers chose to attach the relative clause "low", to the second NP, in order to respect a general recency tendency, whenever possible. The data from the Spanish-French bilinguals lead us to reject this hypothesis as, were there a general tendency to prefer recent host sites in the second language, we should have observed the same trend in this group.

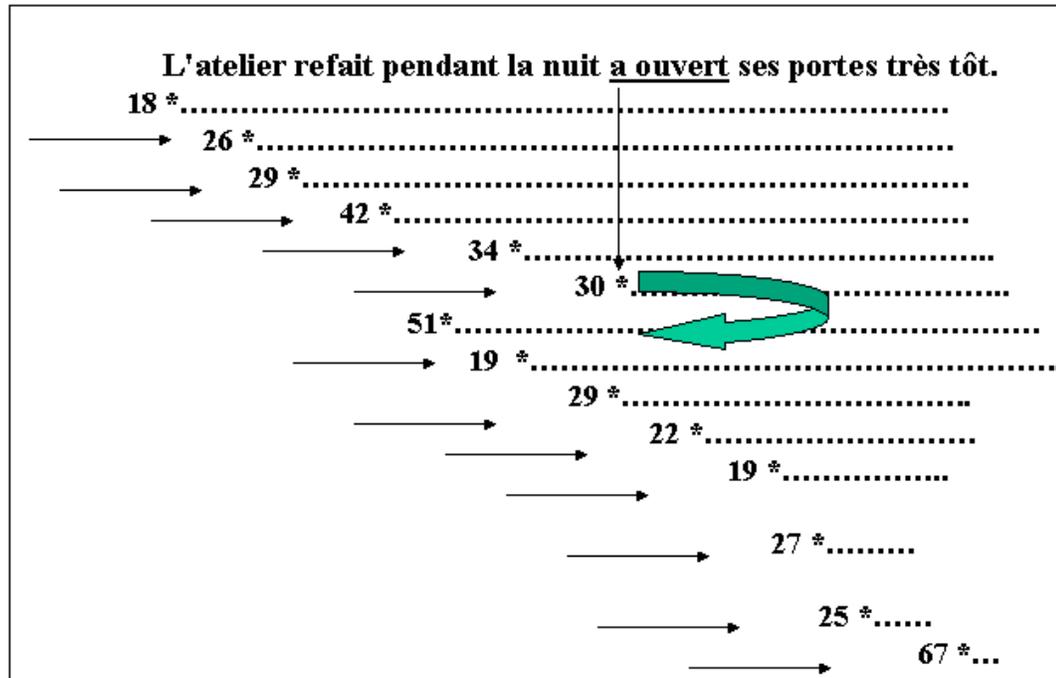


Figure 1. Eye-movements, depicted in hundredths of a second and at $.02^\circ$ precision, during reading, for the ambiguous French sentence "The factory fixed during the night opened its doors very early." At the disambiguation point "a ouvert" (opened) there is an initial fixation of 300 ms. and immediate regression to the word "pendant" (during) which was re-read for 510 ms., followed by second-pass fixations on the NP "la nuit" (the night) and the disambiguation point, and several further first-pass saccades/fixations on subsequent regions.

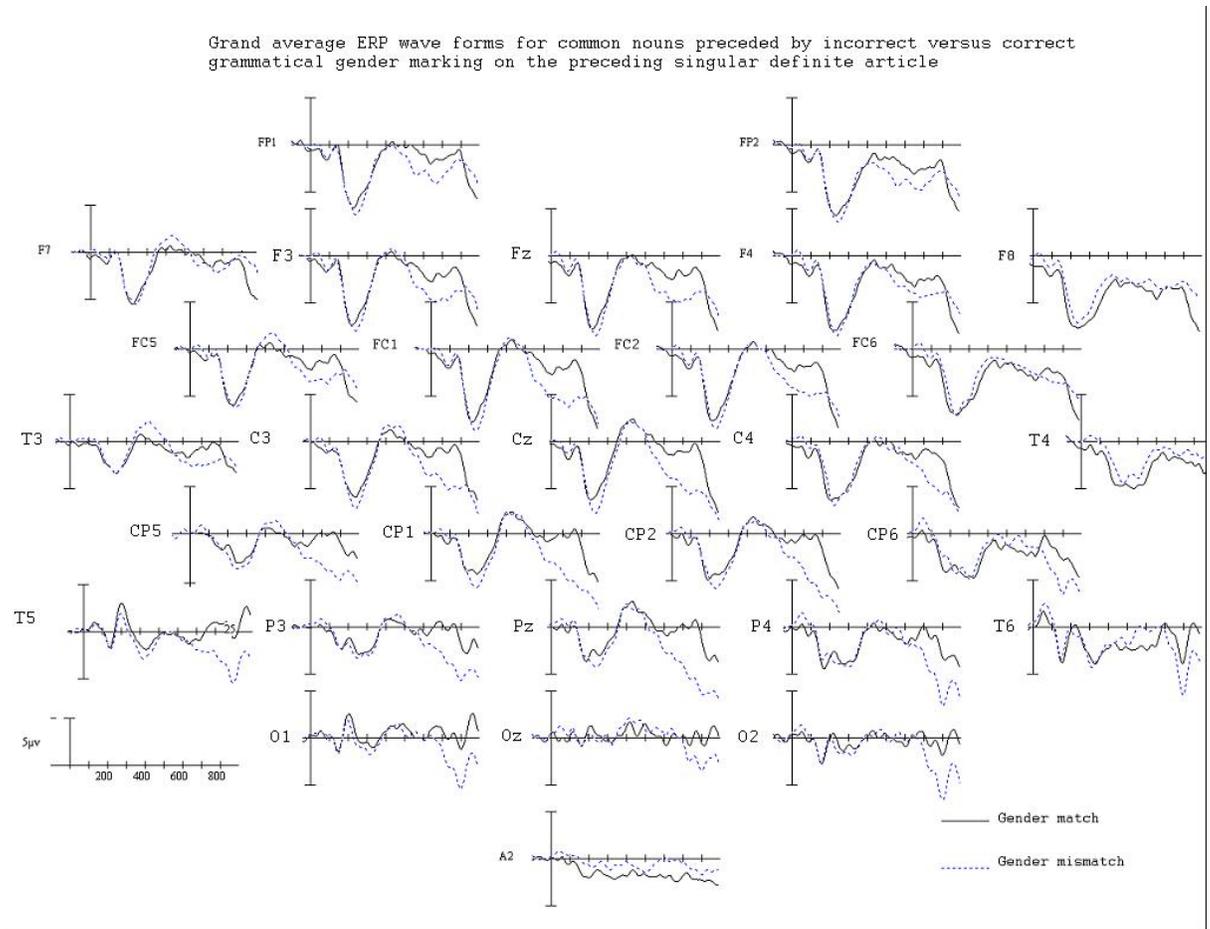


Figure 2. Grand average ERP responses, for native French speakers, when processing a common noun preceded by a singular definite article that correctly indicated the grammatical gender of the noun (gender match) or did not (gender mismatch). All words were presented within a short sentence context.

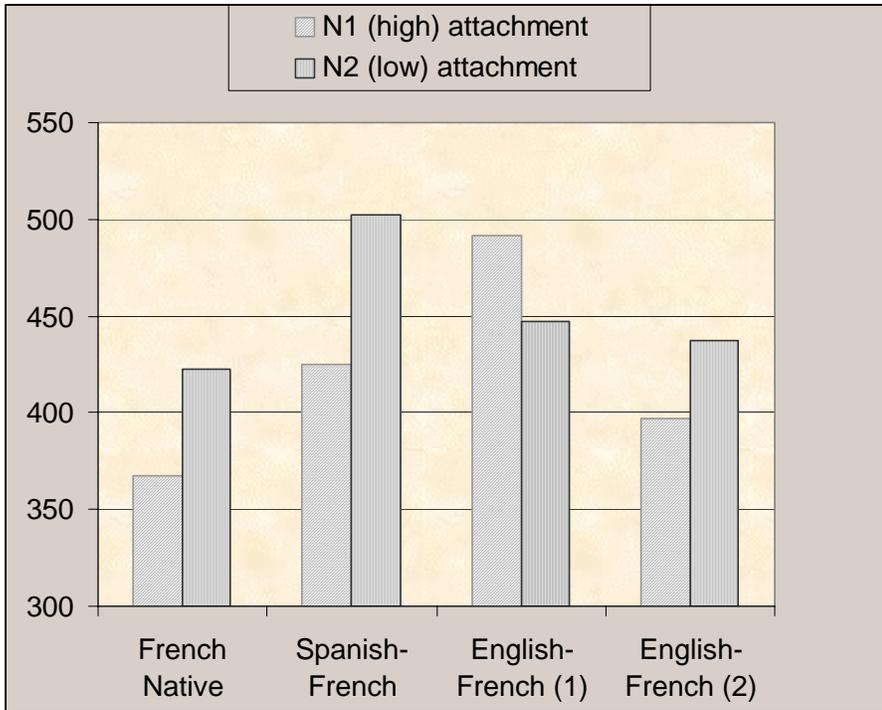


Figure 3. Mean first pass gaze duration on the disambiguating verb for NP1-Prep-NP2-Relative clause constructions (e.g. *Jean a vu [la gardienne des filles /les filles de la gardienne] qui revient de Paris*: "John saw the [nanny of the girls/ the girls of the nanny] who come back from Paris"), as a function of type of attachment (to NP1 or NP2) and type of reader (native French, beginning Spanish-French bilinguals, beginning English-French bilinguals (1), proficient English-French bilinguals (2)).