

On-line syntactic and semantic influences in reading revisited

Joel Pynte

Université Paris-Descartes, CNRS

Boris New

Université Paris-Descartes, CNRS

Alan Kennedy

Université Paris-Descartes, CNRS and
University of Dundee

This study is a follow-up to Pynte, New and Kennedy (2008), *Journal of Eye Movement Research*. 2(1):4, 1-11. A new series of multiple regression analyses were conducted on the French part of the Dundee corpus, using a new set of syntactic and semantic predictors. In line with our prior study, quite different patterns of results were obtained for function and content words. We conclude that syntactic processing operations during reading mainly concern function words and are carried out ahead of semantic processing

Keywords: Reading, eye movements, Latent Semantic Analysis, syntactic constraint, semantic constraint

Introduction

Whether and how high-level processing operations, for example, those developed during sentence comprehension, translate in the eye-movement record has recently emerged as a major issue in the reading literature. Classically, this question has been investigated by manipulating the “predictability” of target words in a particular sentential context (e.g., as assessed by the Cloze task, Taylor, 1953). This approach, implemented in well-controlled laboratory experiments, has established that first-fixation, single-fixation and gaze durations are all sensitive to sources of influence exerted from above the word level (see Clifton, Staub and Rayner, 2007, for a discussion). Inspection times have also been shown to be influenced by the predictability of both the prior and following words in text (Kliegl, Nuthmann & Engbert, 2006), suggesting that different types of processing operations may be “distributed” over several adjacent words (see Drieghe, Rayner & Pollatsek, 2008; Kennedy, 2008; and Rayner, Pollatsek, Drieghe, Slattery & Reichle, 2007 for a discussion of possible artefacts). The pattern of results reported so far is quite complex,

however, with occasional inversions in the direction of the obtained effects, depending on the measure of visual inspection used in the analysis (e.g. single fixation vs. gaze duration, Kliegl, et al., 2006), or the class of the words involved (e.g., whether the target is a function word, Kliegl, 2007).

More recently, several studies have examined the sources of influence likely to contribute to context effects of this kind. For example, Pynte, New and Kennedy (2008a) used Latent Semantic Analysis (Landauer & Dumais, 1997; Landauer, Foltz & Laham, 1998) to assess the degree of semantic relatedness between target words and prior sentential contexts. Using the French part of the Dundee corpus (Kennedy, Hill & Pynte, 2003), the measured LSA index was found to be a significant determinant of inspection time, suggesting that visual inspection parameters are sensitive to a hypothetical semantic component of predictability. Equally, the possible contribution of syntax has been investigated using the English part of the Dundee corpus (Demberg and Keller, 2008) and the (German) Potsdam sentence corpus (Boston, Hale, Kliegl, Patil & Vasisht, 2008). In both studies, inspection time was found to be sensitive to the probability of target words in the syntactic context

within which they appeared (see the notion of “surprisal” developed by Hale, 2001).

It can be concluded that there are both semantic and syntactic determinants of obtained predictability effects, but this leaves open questions as to the precise mechanisms involved. Are several syntactic options pursued in parallel or does the human parser operate in strictly serial fashion? Does syntax come into play first, or do syntactic and semantic factors operate concurrently? These questions have been intensively debated in the Psycholinguistic literature, contrasting the “syntax-first” class of models (Frazier & Clifton, 1998) with the constraint satisfaction approach (McClelland, St. John, & Taraban, 1989; McDonald, 1994). We believe that multiple-regression modelling carried out over large corpora of the kind detailed above, can usefully contribute to this debate. Adopting this approach, the method used for assessing the degree of syntactic or semantic constraint for a given word is submitted can be relatively neutral. For example, the probability that a given word is attached in a particular way can be estimated with reference to the set of alternative syntactic options still open at that point in the sentence (using a syntactically annotated corpus). If inspection times are found to vary as a function of such a measure, this may suggest that alternative options were considered in parallel (see Levy, 2008 for a discussion). This need not to be the case, however, as the frequency of a particular construction itself (e.g., familiarity with a given situation) may be the critical determinant. Further clarification can be obtained by examining the time course of such sources of influence over several adjacent words. For example, Kliegl et al. (2006) show that increased reading times can be observed before a predictable target word is encountered and this is a clear indication that expectation production may be time consuming, in contrast to what a fully parallel approach of parsing would have predicted (see Levy, 2008 for a discussion).

It is possible, of course, that syntax and semantics exert a joint influence on processing time. In a recent study, again conducted on the French part of the Dundee corpus, both semantic and syntactic predictors were included in a regression analysis (Pynte, New & Kennedy, 2008b). The semantic index (SEM hereafter), based on Latent Semantic Analysis, provided an estimate of the degree to which contexts and target words were semantically related. A novel syntactic index was

derived (SYN hereafter), aimed at accounting for syntactic processing difficulty on a word by word basis and based on the syntactic descriptions of the corpus provided by Abeillé, Clément and Toussanel (2003). Both predictors were found independently to contribute to the goodness of fit of the model. However, quite different patterns of effects were found for function and content words. For function words, a main SYN effect was observed as early as first-fixation duration, whereas no clear syntactic influence could be found in the case of content words (the effect was only present at sentence beginning). The method for assessing semantic relatedness, namely Latent Semantic Analysis, does not permit an examination of the possible influences of this factor in the case of function words, because a particular function word can appear in innumerable different contexts. Nevertheless, the fact that early syntactic influences are observed on function words, together with the presence of semantic effects for content words suggests that function and content words might be submitted to different types of processing operations during reading. A similar suggestion has emerged from the extensive work on letter detection showing that letter-detection errors occur disproportionately on frequent function words (see Greenberg, Healy, Koriat & Kreiner, 2004, for a discussion). Koriat and Greenberg (1994) argued that function words are monitored on the basis of a shallow and rapid initial analysis and that this paves the way for the semantic integration of content words, in line with the syntax-first approach. A first aim of the present study was to further examine the time course of syntactic influences for function vs. content words. As first-fixation duration was found to be sensitive to syntactic influences in our previous study, this was preferred to single-fixation duration in the present study.

A second aim was to characterise semantic and syntactic influences as “early” or “late” in processing terms. There is one important methodological change with respect to our earlier study (Pynte, New and Kennedy, 2008a). The semantic index we employed did not control for the degree of semantic relatedness between the target word and the word located immediately to its left. This factor might have been responsible in part for the facilitation observed, through the operation of some kind of inter-lexical priming mechanism (Collins & Loftus, 1975; Meyer & Schvaneveldt, 1971). This could have led to an overestimation of sentence-level semantic influences

during the early stages of syntactic integration. To meet this objection, the procedure used to measure the SEM index in the present study excluded the immediately prior word. The new measure was thus independent of the degree of semantic relatedness between the target and prior words.

A further issue relates to the ability of the syntactic index we used to capture the type of syntactic property likely to exert an immediate influence during reading. Consider (1) below: (NP = Noun Phrase VP = Verb Phrase; PP = Prepositional Phrase; AP = Adjectival Phrase):

(1) [NP Pierre] [VP donne [NP une pomme [AP mure]]] [PP à [NP Marie]]]

(Peter gives an apple ripe to Mary)

(1a) [NP] [VP [NP [AP + P

(1b) [NP] [VP [NP [AP]]] [PP

The SYN value associated with the preposition “à” (underlined in the example) was obtained by computing the frequency with which (1a), representing the current state of the syntactic structure at the moment when the incoming preposition is encountered, leads to an outcome like (1b), representing the structure once the preposition has been integrated, (see Pynte, New and Kennedy, 2008b, Method section). Note that no brackets are represented to the left of the incoming word in (1a), as their presence depends on the way this word will be integrated in the structure. In (1b), a new PP is now attached as a complement of the verb, which involved closing the NP and the AP located in between. In other words, SYN provides an index of the probability of the specific continuation obtained *once the incoming word has been integrated*. Clearly, such an index is likely to tap into relatively late processing stages. For this reason we employed a second syntactic index (FAM hereafter), providing a measure of the familiarity of the initial syntactic configuration (e.g. the probability of (1a) in the corpus) at the moment the target word is encountered, irrespective of the specific continuation obtained at the end of the integration process. If FAM is found to affect the time spent inspecting content words, this would

indicate that some early syntactic processing can be carried out on content words, after all.

One final purpose of the present study was to further investigate Kliegl et al.’s (2006) notion of distributed processing. Distributed processing could explain the failure to find clear effects of SYN for content words if, for example, part of the processing operations concerning a given target word can be carried out while the eyes are fixating either the prior or the next word. For example, the time spent inspecting a given target word might be affected by the FAM value associated with the word located to its right, and still in the parafovea. This is theoretically possible if one admits that the syntactic category of a word can sometimes be obtained from the parafovea (see Pynte & Kennedy, 2007, for a discussion). As in the case of the SYN index, spill-over effects can be expected. As noted above, SYN is based on the probability of the syntactic continuation obtained once an incoming word has been integrated, which involves making a decision regarding all the constituents still open at that point. Deciding not to close the preceding AP in (1) above would lead to a noun-complement PP, as in (1b’), whereas leaving both the AP and the NP open would correspond to an attachment to the preceding AP (continuation 1b’), as in “une pomme mure à ravir”. Note, the word located to the right of the preposition is a noun in both (1b) and (1b’), whereas it is verb in (1b’). Making the correct choice may thus sometimes involve parafoveal processing in order to acquire information about the next word’s syntactic category. Alternatively, readers may decide to postpone the decision until the next word is fixated, leading to a spill-over effect.

(1b’) [NP] [VP [NP [AP]] [PP

(1b’’) [NP] [VP [NP [AP [PP

Method

Materials

The analyses were conducted on the French part (52,173 tokens and 11,321 types) of the Dundee Corpus (Kennedy, Hill & Pynte, 2003) which is based on extended articles taken from the French language newspaper *Le Monde*. Over a number of testing sessions, ten French-speaking participants read the texts presented

at a viewing distance of 500 mm from a display screen, five lines at a time. The set of articles presented to participants was selected from those used by Abeillé, Clément and Toussanel (2003) to construct their French tree-bank, and the syntactic indexes used in the present study were based on the syntactic descriptions provided by these authors.

Syntactic scores

Each word in the eye-movement corpus was associated with two distinct syntactic scores, namely FAM, defined as the probability of configurations such as (1a) in the corpus; and SYN, defined as the probability with which (1a) leads to an outcome like (1b). As noted above, FAM provides a measure of the familiarity of a specific syntactic-context + word-category configuration before the word has been integrated, whereas SYN provides a measure of the probability of the syntactic continuation obtained, once the target has been integrated. (a fuller account can be found in the Method section of Pynte, New and Kennedy, (2008b).

Semantic scores.

Each word in the eye-movement corpus was also associated with a measure of its semantic relatedness to the prior sentence fragment (SEM hereafter), using the Latent Semantic Analysis framework (Landauer, & Dumais, 1997; Landauer, Foltz, & Laham, 1998). In this framework, word meanings are represented as vectors in a high dimensional space (e.g., 300 dimensions) obtained from a representative sample of the target language, in our case, a large corpus of French novels and film dialogues corresponding to more than 30 million words (see Pynte, New and Kennedy, 2008a, Appendix A, for further details of the procedure). Importantly, any sentence fragment, whether present in the original corpus or not, can be represented as a vector in the same semantic space. The representation of a sentence fragment in LSA is simply the average of the vectors of the words it contains, independent of their order, and it is straightforward to compute the similarity between vectors by means of the cosine function. The context taken into account for the measure of the SEM index consisted of all the words located between the previous sentence terminator (e.g., full stop, question mark, etc) and the target word, *excluding the immediately prior word*. All words, at all steps of the procedure, were submitted to lemma transformation.

Selection criteria

Except when specified otherwise, for selection in the present analyses, a word (“word n” or “target word” hereafter) had to be reached by a saccade launched from the immediately prior word, and its length had to be between 2 and 12 characters. The first three words in each sentence were ignored. SEM scores are of little interest in the case of high-frequency function words such as determiners, prepositions, pronouns, etc., simply because such words can be found in any context. For this reason, function target words were excluded from the analyses involving the SEM index and were submitted to separate analyses.

Procedure

Syntactic and semantic influences on target-word first-fixation and gaze durations were assessed via a series of regression analyses. Syntactic and semantic indexes were successively added to a baseline model comprising a set of predictors known to influence inspection time. The contribution of these syntactic and semantic properties to the goodness of fit of the model was evaluated. The analyses were conducted in the linear-mixed effects model (*lme*) framework, using the *lme4* package (Bates, 2007) for the *R* system for statistical computing (R Development Core Team, 2006). Both readers and words were treated as random factors. Syntactic and semantic effects were estimated as varying across readers. All independent variables, except those expressed in number of characters were log transformed.

Baseline Model

In the baseline model, first-fixation and gaze durations were accounted for in terms of the length and frequency of the target word, the length, frequency and class (0 = content word, 1 = function word) of the prior word, and the length, frequency and class of the next word. In addition, the baseline model comprised three predictors whose purpose was to account for variation in inspection time arising as a function of landing position and preview benefit. These were: the size of the saccade entering the word, its relative landing position (landing position divided by word length), and the square of this latter measure (quadratic trend). To maintain compatibility with previous analyses (Pynte & Kennedy,

2006; 2007), measures of lexical frequency were based on the texts used in the Dundee Corpus.

Correlation between predictors

The correlation matrix is provided in the appendix. Apart from correlations between length and frequency (visible for words n-1, n and n+1), the highest value (+.74) was obtained between FAM and FAM+1, that is, the familiarity of the syntactic configuration involving the current and next words, respectively. All other values are relatively low (< .35).

Results

The results of the analyses conducted on function words are presented in Table 1, and those for content words in Table 2, with first-fixation durations on the left and gaze durations on the right. We start with a brief description of the baseline model, corresponding to predictors 1 to 11 in the tables. The contribution of syntax and semantics is subsequently examined in the next sections, as the corresponding predictors are successively added to the regression equation.

Baseline model

Most factors included in the baseline model are well-known determinants of inspection time. We only briefly mention them here. Unsurprisingly, first-fixation and gaze durations were longer for long and low frequency words. A spillover effect of prior-word frequency was present for both function and content words, whereas prior-word length only affected first-fixation duration for content words. First-fixation durations (either on content or function words) were shorter when the prior word was a function word. Backward influences (associated with next-word properties) are also visible on Tables 1 and 2. Shorter first-fixation and gaze durations were obtained as the length of the next word increased. An effect of next-word frequency is also apparent for gaze durations recorded on content words, in line with prior results (Kennedy and Pynte, 2005; but see also Drieghe, Rayner & Pollatsek, 2007, and Kennedy, 2008). An influence of the class of the next word was observed for function target words only, with shorter first-fixation and gaze durations when the next word was also a function word. The baseline model also comprised a set of predictors

aimed at accounting for inspection strategies. In line with prior results, inspection time was found to vary as a function of landing position (linear and quadratic trends), with longer fixation durations for landing points in the middle of the word (O'Regan, Pynte & Coeffé, 1986; Vitu, McConkie, Kerr & O'Regan, 2001). Longer first-fixation and gaze durations were also associated with longer incoming saccades (Radach & Heller, 2000; Vitu, et al., 2001; McDonald, 2005).

Table 1: Regression coefficients with associated standard errors from the analysis of function words

	Variance	
	First fixation	Gaze duration
Random effects:		
Item (Intercept)	547,35	876,28
Sub (Intercept)	165,79	321,49
Sub Syn-1	0,00	6,90
Reader Syn	0,29	0,00
Residual	4732,53	7659,40
Reader Fam	0,00	0,00
Reader Fam+1	0,00	0,00

	Estimate (Std. Error)			
	First fixation		Gaze duration	
Fixed effects:				
(Intercept)	193,75	(5,51)	265,23	(7,37)
Saccade	1,53	(0,12*)	1,59	(0,15*)
Landing	35,58	(11,09*)	-101,23	(14,11*)
Landing ²	-10,31	(8,24)	91,21	(10,48*)
Freq.-1	-2,30	(0,42*)	-3,01	(0,54*)
Length-1	-0,14	(0,26)	-0,44	(0,33)
Frequency	-3,20	(0,50*)	-2,97	(0,63*)
Length	2,59	(0,46*)	10,41	(0,58*)
Freq.+1	-0,32	(0,42)	-0,88	(0,54)
Length+1	-2,03	(0,28*)	-2,24	(0,36*)
Class-1	2,48	(1,75)	1,97	(2,23)
Class+1	-4,07	(1,85*)	-5,58	(2,35*)
Syn	-1,56	(0,63*)	-1,18	(0,77)
Syn-1	0,24	(0,71)	-2,22	(1,24)
Fam	0,16	(0,28)	0,45	(0,36)
Fam+1	-0,61	(0,41)	-0,71	(0,52)

	Estimate (Std. Error)			
	First fixation		Gaze duration	
Interactions				
Syn : Fam	-0,70	(0,36)	-0,92	(0,45*)

The SYN index

Adding the SYN index to the baseline model significantly improved the goodness of fit for first-

fixations recorded on function words ($X^2 = 6.42$, $p = 0.05$). In line with our prior study (Pynte et al., 2008b), first-fixation duration was found to decrease as the probability of the syntactic continuation increased ($t = -2.46$). No significant improvement was obtained for gaze duration, however, also in line with our prior study ($X^2 = 2.87$, $p > 0.23$, $t = -1.70$). This point will be discussed later in relation with the FAM index. In the case of content words, a significant improvement was obtained for both first-fixation and gaze duration ($X^2 = 10.99$ and 8.64 , $p < 0.02$, respectively), even though the corresponding regression coefficients were not significant ($t = -0.60$ and -0.82 , respectively). In our prior study, a similar pattern of results was found to be associated with a significant interaction with the position of the target word in the carrying sentence, a factor not included in the present analysis.

The FAM index

Contrary to our predictions, adding FAM as a potential predictor of inspection time did not result in any improvement in fit. This was true for both function words ($X^2 = 0.33$ and 1.63 , $p > 0.40$; $t = 0.58$ and 1.63 , for first-fixation and gaze durations, respectively) and for content words ($X^2 = 0.23$ and 1.92 , $p > 0.38$; $t = 0.53$ and 1.05 , for first-fixation and gaze durations, respectively). Importantly, however, an improvement in fit was obtained when the SYN:FAM interaction term was added to the model for function words. The improvement was close to significance for first-fixation duration ($X^2 = 3.82$, $p = 0.05$, $t = -1.95$), and significant for gaze duration ($X^2 = 4.15$, $p < 0.05$; $t = -2.05$). Longer inspection times were obtained in the condition associating a low SYN value with a high FAM value. That is, when a low-frequency attachment decision had to be made in a situation where the syntactic configuration the reader was confronted with, prior to attachment, looked familiar. This is illustrated in Figure 1 for first-fixation durations.

Spill-over effects

In order to examine possible spill-over effects, SYN scores associated with the prior word (SYN-1 hereafter) were subsequently included in the analysis. In the case of function words, no fit improvement was observed for first fixation duration ($X^2 = 0.01$, $p > 0.94$). The improvement in fit was significant for the measure of gaze duration ($X^2 = 8.83$, $p < 0.02$), but the corresponding regression

coefficient failed to reach significance ($t = -1.79$, n.s.). In contrast, the addition of SYN-1 clearly increased the goodness of fit of the model for both dependent variables in the case of content words ($X^2 = 7.13$ and 6.35 , $p < 0.05$ for first-fixation and gaze durations, respectively), and the corresponding regression coefficients were significant ($t = -2.21$ and -2.52 , respectively). Importantly, further improvement was obtained when the SYN-1:Class-1 interaction term was added to the regression equations for content words ($X^2 = 4.91$ and 7.93 , $p < 0.03$, $t = -2.22$ and -2.81 , respectively). This interaction was explored by separate analyses conducted for content words preceded by a function word and for content words preceded by another content word. A significant SYN-1 effect was obtained when the prior word was a function word ($B = -2.12$ and -4.04 ; $t = -2.99$ and -3.36 for first-fixation and gaze duration, respectively). Virtually no effect was present when the target word was preceded by another content word ($B = 0.06$ and 0.82 ; $t = 0.09$ and 0.71 , respectively). In other words, a spill-over was only observed when the prior word, the source of the effect, was a function word. This pattern of results confirms the importance of function words, as far as SYN effects are concerned.

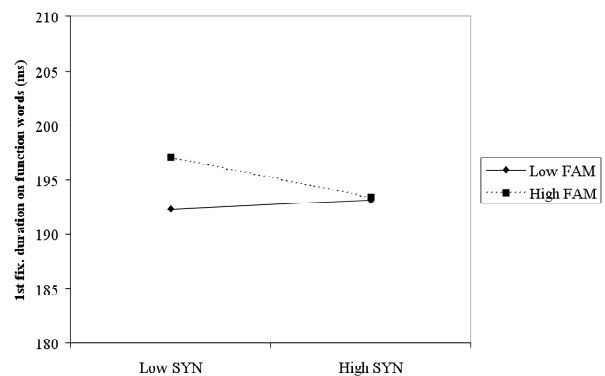


Figure 1. First-fixation duration for function words with low vs. high SYN scores and low vs. high FAM scores

Backward effects

In this section, we examine whether inspection times were affected by FAM scores associated with the next word (FAM+1 hereafter). FAM+1 could theoretically exert an influence if it is admitted that the syntactic category of the next word can be obtained from the parafovea (see Pynte and Kennedy, 2007, for a

discussion). Again, a different pattern of results was obtained for function and content words. No FAM+1 effect was apparent when the word being fixated was a function word. Adding FAM+1 did not result in any significant fit improvement ($X^2 = 2.25$ and 1.89 , $p > 0.30$, $t = -1.51$ and -1.38 for first-fixation and gaze duration, respectively). However, a clear contribution to both first-fixation and gaze duration was apparent for content words ($X^2 = 10.93$ and 23.83 , $p < 0.005$, $t = 2.89$ and 3.11 , respectively). It follows that content words were subject to contextual influences from both the prior word (as shown by the SYN+1 effect noted above) and next word. These results will be discussed in the General Discussion in relation with the notion of distributed processing (Kliegl, Nuthmann and Engbert, 2006).

Table 2: Regression coefficients with associated standard errors from the analysis of content words

Random effects:	Variance	
	First fixation	Gaze duration
Item (Intercept)	225,25	927,11
Sub (Intercept)	163,00	899,10
Sub Syn-1	0,83	0,00
Reader Syn	6,67	14,24
Residual	4722,09	13091,00
Reader Fam	0,01	0,44
Reader Sem	0,00	14,89
Reader Fam+1	0,18	2,06

Fixed effects:	Estimate (Std. Error)			
	First fixation		Gaze duration	
(Intercept)	182,75	(4,57)	293,80	(10,15)
Saccade	1,02	(0,10*)	0,86	(0,16*)
Landing	95,06	(6,51*)	-207,30	(10,93*)
Landing ²	-56,07	(5,49*)	168,98	(9,21*)
Freq.-1	-1,22	(0,31*)	-2,51	(0,53*)
Length-1	0,51	(0,21*)	-0,45	(0,36)
Frequency	-4,12	(0,32*)	-8,41	(0,55*)
Length	3,13	(0,20*)	14,15	(0,34*)
Freq.+1	-0,26	(0,31)	-1,46	(0,54*)
Length+1	-0,67	(0,22*)	-1,13	(0,39*)
Class-1	-4,61	(1,18*)	-1,24	(2,05)
Class+1	-1,55	(1,32)	-0,10	(2,28)
Syn	-0,60	(1,00)	-1,27	(1,56)
Syn-1	-1,19	(0,54*)	-1,96	(0,78*)
Fam	0,10	(0,18)	0,40	(0,38)
Sem	-1,21	(0,61)	-8,28	(1,63*)
Fam+1	0,92	(0,32*)	2,10	(0,68*)

Interactions	Estimate (Std. Error)			
	First fixation		Gaze duration	
Syn-1 : Class-1	-2,02	(0,91*)	-4,42	(1,57*)
Syn-1 / content	0,40	(0,74)	0,92	(1,19*)
Syn-1 / function	-1,71	(0,66*)	-3,52	(1,02*)
Syn-1 : Sem			-2,38	(1,14*)

The SEM index

In this section we examine the possible influence of semantic relatedness, as assessed by SEM scores. The analyses were restricted to content target words, because of the lack of relevance of Latent Semantic Analysis to the case of function words. It should be kept in mind that the word located immediately to the left of the target word was excluded from the context used to compute SEM scores in the present study. As we were interested in possible interactions between semantic relatedness and syntax, SEM scores were added to a model already containing the syntactic predictors discussed in the previous sections. This did not result in any improvement to the fit in the first-fixation case ($X^2 = 3.94$, $p > 0.13$). The goodness of fit was improved when the SEM index was added to the regression equation for gaze duration, however ($X^2 = 61.25$, $p < 0.000001$), with a 8.3 ms decrease per log increment of the SEM index ($t = -5.09$). Turning to the joint influence of syntax and semantics, adding the SYN:SEM interaction term to the regression equation did not significantly improve the goodness of fit of the model ($X^2 = 2.75$, $p < 0.10$ and $X^2 = 1.26$, $p > 0.25$ for first-fixation and gaze durations, respectively) and the corresponding regression coefficients failed to reach the significance level ($t = 1.65$ and 1.12 , respectively). Semantic relatedness apparently modulated the processing operations devoted to the syntactic integration of the prior word, however. That is, a significant fit improvement was obtained for gaze durations when the SEM:SYN-1 interaction term was added to the regression equation ($X^2 = 4.40$, $p < 0.04$, $t = -2.09$). As the interaction was observed for gaze duration only, it can be concluded that only relatively late processing operations were involved. As can be seen in Figure 2, no cross-over was involved. Shorter gaze durations were obtained when high SYN-1 scores were associated with high SEM scores.

Complementary analyses

The analyses presented so far were conducted on a restricted set of data, with the constraint that the entering saccade had to be launched from the immediately prior word. Prior words were frequently skipped however (67% of the cases when a content target word was preceded by a function word). A series of complementary analyses, devoted to this particular situation is presented in this section. All other cases (e.g., involving two adjacent content or function words, or a function word preceded by a content word) will be ignored, as no significant effects emerged. The regression analyses conducted for the function word + content word case are summarised in Table 3, with first-fixation duration on the left, and gaze duration on the right. The pattern of effects observed for SYN-1 (that is, SYN defined relative to word n-1) and SEM is similar to what was found in the no-skip condition, with a significant or marginally significant influence of SYN-1 for first fixation and gaze durations ($X^2 = 7.73$ and 6.56 , $p < 0.04$, $t = -2.66$ and -1.81 , respectively), and a significant SEM influence for gaze duration ($X^2 = 39.89$, $p < 0.000001$, $t = -6.32$). A significant influence of FAM on gaze duration ($X^2 = 10.54$, $p < 0.006$, $t = 3.25$), associated with an absence of a main effect of FAM+1 ($X^2 = 0.97$ and 1.16 , $p > 0.50$, $t = -0.98$ and -0.77) is also worth noting. This suggests that less attention may have been devoted to the next word, as compared to the no-skip situation.

Table 3 : Regression coefficients with associated standard errors from the complementary analysis of content words (prior function word skipped)

	Variance	
	First fixation	Gaze duration
Random effects:		
Item (Intercept)	142,55	738,61
Sub (Intercept)	213,31	1304,10
Sub Syn-1	0,19	4,14
Reader Syn	5,23	11,91
Residual	4562,27	14719,28
Reader Fam	0,25	0,00
Reader Sem	0,91	0,00
Reader Fam+1	0,00	0,47

	Estimate (Std. Error)			
	First fixation		Gaze duration	
Fixed effects:				
(Intercept)	231,91	(5,05)	303,96	(12,01)
Saccade	-0,35	(0,11*)	-0,30	(0,20)

Landing	41,40	(5,93)	-64,40	(10,73*)
Landing ²	-31,16	(5,46*)	83,39	(9,88*)
Freq.-1	-0,03	(0,40)	-3,56	(0,74*)
Length-1	-2,05	(0,53*)	0,07	(0,98)
Frequency	-1,21	(0,32*)	-7,93	(0,59*)
Length	0,32	(0,20)	12,92	(0,36*)
Freq.+1	-0,46	(0,25)	-1,25	(0,47*)
Length+1	-0,40	(0,20*)	-1,02	(0,38*)
Syn	0,57	(1,04)	2,90	(1,77)
Syn-1	-1,15	(0,43*)	-1,81	(1,00)
Fam	0,25	(0,24)	1,06	(0,33*)
Sem	-0,90	(0,70)	-7,94	(1,16*)
Fam+1	-0,33	(0,33)	-0,51	(0,66)

Interactions	Estimate (Std. Error)	
	Gaze duration	
Syn : Fam	1,62	(0,80*)
Syn : Fam+1	2,17	(0,74*)

The most striking result, however, is a paradoxical marginally significant SYN effect for gaze duration ($X^2 = 5.57$, $p < 0.07$, $t = 1.66$). High SYN values were associated with an *increase*, not a decrease, in gaze duration. This quite surprising result was qualified by a significant interactions with FAM ($X^2 = 4.08$, $p < 0.05$, $t = 2.03$) and by a significant interaction with FAM+1 ($X^2 = 8.55$, $p < 0.004$, $t = 2.93$). High SYN scores were associated with a gaze-duration increase in the high-FAM condition on the one hand, and in the high-FAM+1 condition on the other hand.

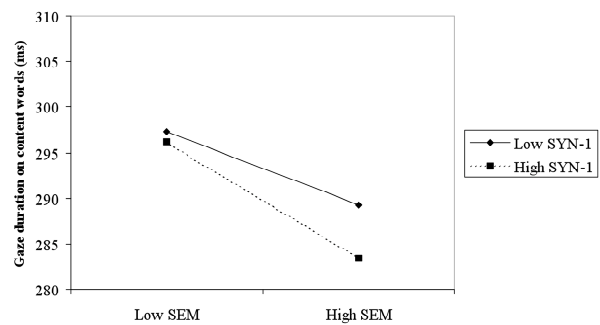


Figure 2. Gaze duration for content words with low vs. high SEM scores and low vs. high SYN-1 scores.

General discussion

In this paper, we present a new series of regression analyses conducted on the French part of the Dundee corpus, with the aim of further examining whether, and

how, semantic and syntactic sources of influence translate in the eye movement record, in terms of their temporal locus (early vs. late measures of visual inspection), class of word affected (function vs. content words), and possible impact on prior and following words.

In a prior study, inspection times were found to be affected by both syntactic and semantic sources of influence. Quite different patterns of results were found for function and content words, however. A main effect of SYN was found for first fixation durations recorded on function words, suggesting that syntax may exert quite an early influence. This conclusion had to be qualified for content words, however, as in this case a clear semantic influence was present in the absence of any equivalent clear syntactic effect. In an attempt to better track down early syntactic influences, a new index (FAM) likely to tap into early processing stages (e.g., before an incoming word has been integrated in the current structure), was used in the present study, together with SYN, already used in our prior study. Contrary to our expectations, no clear influence of FAM was found, either for content or function words. As with the SYN index, the results were consistent with what was initially found in our prior study. First-fixation durations recorded on function words were shorter when the prior sentence fragment was highly constraining in terms of syntactic attachment, that is, when the integration process ended up in a high-frequency syntactic continuation. Interestingly, sensitivity to the SYN effect was modulated by an interaction with FAM, which provides some support to the notion that several syntactic options may be taken into consideration before a final decision is made. As indicated in Figure 1, a clear SYN effect was only present for high-FAM values, that is, when the initial syntactic configuration was relatively frequent, and the parser probably confronted with several, relatively frequent, attachment options. Our data are not consistent with a fully parallel account, however. Rather, the observed interaction points to some kind of garden-path mechanism, with longer inspection times recorded when the integration processes ends up with a non-preferred continuation. From this perspective, the human parser must be seen as either ranked parallel or serial. In conclusion, adding a new syntactic predictor to the model did allow us to shed some light on the processing operations devoted to integrating function words. We failed to find any clear evidence of similar syntactic influences in the case of content words, however.

As demonstrated by Kliegl, Nuthmann and Engbert (2006), some of the processing operations involved in sentence comprehension may not always be time-locked to visual inspection. Could it be the case that some of the effects we are looking for (e.g., concerning content words) would only emerge on either the prior or next words? SYN can be thought of as reflecting syntactic difficulty at a relatively late processing stage. Spill-over effects could thus be expected in this case. In contrast, FAM, which was assumed to reflect quite early processes, could be expected to manifest itself as soon as a new word is available in parafoveal vision, leading to what Kliegl et al., 2006, called a “successor effect”. Let us consider successor effects first (referred to as backward effects in the Results section). Inspection times were sensitive to the FAM index associated with the following word, still in the parafovea, but this was only true when the word being fixated was a content word (see FAM+1 in Table 2). Moreover, the direction of the observed FAM+1 effect was opposite to what could be expected. Longer, not shorter, inspection times were observed for high FAM+1 scores, that is when the syntactic category of the parafoveal word could possibly be guessed, given the current syntactic context. This quite paradoxical outcome is reminiscent of Kliegl, Nuthmann and Engbert’s (2006) results, obtained with a classical measure of predictability. In their study, conducted on the Potsdam sentence corpus, Kliegl et al. observed inflated single fixation durations prior to entering a highly predictable word. They suggested that this might reflect the time necessary for sentence-level expectations (based on memory retrieval) to develop (see Kliegl, 2007, for further discussion). A similar mechanism, e.g., aimed at developing syntactic expectations is possibly responsible for the FAM+1 effect observed here.

With regard to spill-over effects (“predecessor effects” in Kliegl et al.’s terminology), the SYN index associated with the prior word (SYN-1 in the tables) did not contribute to inspection times in the function-word case. A SYN-1 effect was present for content words. However, the presence of an interaction with the class of the prior word indicated that the effect was only present when the content target word was preceded by a function word. What we observe here is thus a spill-over of the processing operations aimed at integrating a prior function word. This pattern of results suggests that syntactic parsing is not fully incremental, as fully parallel parsing models predict it should be. As illustrated by (1)

in the Introduction, the way an incoming word will eventually be integrated into the current syntactic structure may sometimes depend on the syntactic category of the following word (e.g. the presence of a verb to the right of the preposition in the example would have favoured an attachment to the preceding AP). In such circumstances, knowing the syntactic category of the next word clearly confers an advantage, and this could explain the FAM+1 effect discussed above. In the absence of any parafoveal information, however, part of the processing operations devoted to integrating the incoming word may have to be postponed, until the syntactic category of the following word can be determined, thus leading to a spill-over effect.

In such circumstances, the reader may also decide simply to skip the function word. In our data, 67% of function words followed by a content word were skipped during first-pass reading, and it can be assumed that a proportion of these correspond to a situation similar the one described above. Consistent with the notion of distributed processing, the time spent inspecting the (content) target word was still subject to a SYN-1 spill-over effect, suggesting that part of the processing operations devoted to integrating the (skipped) prior function word were carried out while the content target word was being fixated. In contrast, no clear “successor effect” was observed, apart from a SYN:FAM+1 interaction, with a paradoxical effect of SYN (longer gazes associated to high SYN values) when the next syntactic configurations looked familiar. We are unable to provide any explanation for this quite complex pattern of results. It may be worth noting that, overall, more time was spent on the target word in the skip condition, relative to the no-skip condition, as witnessed by the results of a combined analysis ($t = 32.31$ and 35.43 , for single-fixation and gaze durations, respectively). A possible explanation could be found in the necessity for the reader to integrate two words in parallel, namely the currently fixated word and the skipped word. In other words, it may be the case that, for reasons as yet unknown, the reader may carry out most of the processing operations devoted to integrating the function word while inspecting the content word.

We can now return to our original concern: disentangling syntactic and semantic sources of influence during reading. Two types of contrast are at our disposal: in terms of temporal locus (first-fixation vs. gaze

durations) on the one hand, and in terms of class of word affected (function vs. content words) on the other hand. Regarding the latter contrast, the pattern of results discussed above, taken as a whole, suggests that function words are the place where syntactic decisions are made, even though part of the processing operation may spill over onto a content word located to the right. Content words, on the other hand, seem to be mostly sensitive to semantic influences. SYN-1 and FAM+1 did exert an influence, but these predictors do not directly concern the word being fixated. Moreover, as far as SYN-1 is concerned, the effect was modulated by the degree of semantic support that the target word received from the prior context (significant SEM:SYN-1 interaction for gaze durations), thus providing further evidence of the importance of semantic factors in the case of content words.

There is a weakness in this reasoning, however. The analyses conducted on function words were confined to syntactic influences. Because the method used for computing the SEM index was not appropriate for function words we do not know whether semantic effects were present for function words. This is where the second contrast, opposing first-fixation to gaze durations, enters into play. As noted in the Introduction, part of the semantic effect found in our prior study could have had its origin at a lexical level, e.g., via some kind of priming mechanism. In an attempt to dissociate sentence-level contextual effects from purely local sources of influence, the SEM index used in the present study was rendered independent of the degree of semantic relatedness between the target word and the word located immediately to its left. With this new SEM index, an effect of semantic relatedness was only found in the analysis of gaze duration: shorter inspection times were associated with higher SEM values. In contrast to our earlier study no SEM effect was apparent for first-fixation durations, suggesting that semantic support becomes effective at a relatively late processing stage. As the SYN index manifested itself as early as on first fixations for function words, the possibility remains that syntax might be ahead of semantics, after all. In fact, it seems extremely unlikely that semantic influence would show up on first fixation for function words, whereas the SEM index only manifested itself on gaze durations for content words.

In conclusion, our results seem to be consistent with a version of syntax-first models in which syntactic processing operations mainly concern function words and are carried out ahead of semantic processing. This is partly in line with Koriat and Greenberg's (1994) suggestion that function words "pave the way" for the semantic integration of content words. However, given the nature of the SYN index used in the present study, we are tempted to conclude that function words are submitted to quite deep processing operations, in contrast with Koriat and Greenberg's suggestion that only shallow parsing is necessary.

Appendix: Correlations between predictors

	frl-	syn-	olg	frl	fam	syn	sem	olg+	frl+	Fam+
olg-1	-0,79	0,09	-0,20	0,26	-0,11	-0,16	-0,01	0,02	-0,04	0,00
frl-1		-0,14	0,26	-0,30	0,09	0,17	0,03	-0,01	0,05	-0,01
syn-1			-0,04	0,05	0,01	0,01	0,00	0,02	-0,03	-0,02
olg				-0,81	0,06	0,13	0,15	-0,22	0,30	-0,10
frl					-0,06	-0,17	-0,07	0,28	-0,34	0,08
fam						-0,09	-0,12	0,00	0,00	0,74
syn							0,00	-0,04	0,06	0,02
lsp								-0,04	0,09	-0,15
olg+1									-0,78	0,03
frl+1										-0,03

References

Abeillé A., Clément L., Kinyon A. (2003). Building a Treebank for French. In A. Abeillé (ed), *Treebanks: Building and Using Syntactically annotated corpora*. (pp.165-187). Dordrecht: Kluwer.

Bates, D.M. (2007). *lme4: Linear mixed-effect models using S4 classes*. R package version 0.995-2.

Boston, F.M., Hale, J., Kliegl, R., Patil, U., & Vasisht, S. (2008). Parsing costs as predictors of reading difficulty: an evaluation using the Potsdam sentence corpus. *Journal of Eye Movement Research*, 2(1):1,1-12.

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. V. Gompel, M. Fisher, W. Murray, & R.L. Hill (Eds.), *Eye Movements: A Window in Mind and Brain* (pp. 341-372). Elsevier.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic priming. *Psychological Review*, 82, 407-428.

Demberg, V., & Keller, F. (2008). Eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, 109, 193-210.

Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *The Quarterly Journal of Experimental Psychology*, 61, 1239-1249

Frazier, L., & Clifton, C. (1998). *Construal*. Cambridge, MA: MIT press.

Greenberg, S.N., Healy, A.F., Koriat, A., & Kreiner, H. (2004). The GO model: A reconsideration of the role of structural units in guiding and organizing text on line. *Psychonomic Bulletin & Review*, 11, 428-433.

Hale, J. (2001) A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics*. Pittsburgh, PA.

Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science*, 30(4). 609-642.

Kennedy, A. (2008). Parafoveal-on-foveal effects are not an artefact of mis-located saccades. *Journal of Eye Movement Research*, 2(1), 1-10.

Kennedy, A. & Pynte J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, 45, 153-158.

Kennedy, A., Hill, R., & Pynte, J. (2003). The Dundee corpus. *Poster presented at ECEM12: 12th European Conference on eye movements.*, Dundee, August 2003.

Kliegl, R. (2007). Towards a perceptual-span theory of distributed processing in reading: a reply to Rayner, Pollatsek, Drieghe, Slattery, & Reichle (2007). *Journal of Experimental Psychology: General*, 138, 530-537

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: the influence of past, present and future words on fixation durations. *Journal of Experimental Psychology: General*, 135, 12-35.

- Koriat, A. & Greenberg, S.N. (1994). The extraction of phrase structure during reading: Evidence from letter detection errors. *Psychonomic Bulletin & Review*, 1, 345-356.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259-284.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259-284.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106, 1126-1177.
- McClelland, J.L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed approach. *Language and Cognitive Processes*, 4, 287-336.
- McDonald, M.C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, 9, 157-201.
- McDonald, S. A. (2005). Parafoveal preview benefit in reading is not cumulative across multiple saccades. *Vision Research*, 45, 1829-1834.
- Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227-234
- O'Regan, K., Pynte, J., & Coëffé, C. (1986). Comment le regard explore un mot isolé. *Bulletin de Psychologie*, 39, 7-10.
- Pynte, J. & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading English and French. *Vision Research*, 46, 3786-3801.
- Pynte, J. & Kennedy, A. (2007). The influence of punctuation and word class on distributed processing in normal reading. *Vision Research*, 47, 1215-1227.
- Pynte, J., New, B., Pynte, J. (2008a). On-line contextual influences during reading normal text: a multiple-regression analysis. *Vision Research*, 48, 2172-2183.
- Pynte, J., New, B., Pynte, J. (2008b). A multiple regression analysis of syntactic and semantic influences in reading normal text. *Journal of Eye Movement Research, special issue*. 2(1):4, 1-11.
- R Development Core Team (2006). *R: A language and environment for statistical computing*. (version 2.3.1). R Foundation for Statistical Computing, Vienna, Austria
- Radach, R., & Heller, D. (2000). Relation between spatial and temporal aspects of eye movement control. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 165-192). Oxford: Elsevier.
- Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T.J., & Reichle, E.D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology: General*, 136, 520-529
- Taylor, W.L. (1953). Cloze procedure: a new tool for measuring readability. *Journalism Quarterly*, 30, 415-433.
- Vitu, F., McConkie, G. W., Kerr, P. & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: an inverted optimal viewing position effect. *Vision Research*, 41, 3513-3533.