English vs. French: Determinants of Eye Movement Control in Reading

Sébastien Miellet, Cyril Pernet, Patrick J. O'Donnell, and Sara C. Sereno

Department of Psychology, University of Glasgow

UNIVERSITY *of* GLASGOW

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Glasgow Language Processing

Introduction

Eye movement (EM) control through text is determined by lower-level perceptual and attentional processes as well as higher-level psycholinguistic properties of the text itself. Three current points of debate concern:

(1) the relative extent to which EM behavior is affected by low-level oculomotor factors *vs.* higher-level cognitive processes;

(2) the relative depth to which readers process information from the parafoveal word; and

(3) the reliability of these effects across different languages.

Determinants of eye movement control in reading

The variables of word length, word frequency, and contextual predictability typically produce robust effects on EM behavior. These factors are used as the main predictors in models of EM control in reading such as E-Z Reader (Pollatsek et al., 2006) or SWIFT (Engbert et al., 2005).

However, few studies have explored the relative weights of a larger set of factors on the spatial and temporal aspects of oculomotor behavior in reading. ***Repeated-measures multiple regression*** (rmMR), as described by Lorch and Myers (1990), provides measures of effect magnitude across a large number of factors. It involves a regression analysis on the entire data set, with each individual observation constituting a separate case.

Parafoveal-on-foveal effects

Using rmMR with a German corpus, Kliegl et al. (2006) studied whether fixation time on a target was affected by the word length, frequency, and predictability of the words *before* and *after* the target. They found parafoveal-on-foveal effects (i.e., the characteristics of word *n+1* affected the fixation time on word *n*), and concluded that more than one word can be processed in parallel.

Present study

- **To study a large set of oculomotor and psycholinguistic variables and to estimate their relative weights on EM behavior.**
- **To examine whether parafoveal-on-foveal effects could be obtained in new datasets.**
- **To determine if these findings can be generalized across languages.**

Method

Participants: 20 native English and 20 native French speakers.

Apparatus: Fourward Technologies Dual-Purkinje Eyetracker (Gen 5.5).

Procedure: Participants read sentences *in their native language* while their eye movements were monitored, with Y/N questions on 1/3 of the trials.

Materials and Design:

80 experimental sentences (in each language; translations), each containing an Adjective-Noun (English) or a Noun-Adjective (French) target pair:

wd1 **wd2**

As expected, the **amateur cyclist** has lost the race.

Comme prévu, le **cycliste amateur** a perdu la course.

Adjectives and nouns were selected as either high (H) or low (L) frequency, and all combinations were used. Noun phrases (NPs) were either plausible or implausible. Note that implausible NPs were only implausible up to the NP; they were made plausible by the remainder of the sentence (e.g., *Benjamin cut off the* ***left head*** *of the mutant in the video game.*).

The conditions were as follows (2 x 2 x 2), with 10 items per condition:

Plausible: HH, HL, LH, LL Implausible: HH, HL, LH, LL

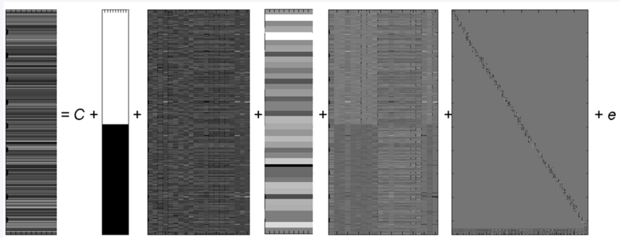
Frequency values were obtained from the British National Corpus (1995) for English, and Lexique3 (New et al., 2001) for French.

Although conditions were dichotomous (e.g., H vs. L frequency; plausible vs. implausible), *exact frequency, predictability, and plausibility values for each item were used for analyses*.

Predictors used in the global regression model:

Abbrev.	Predictor	Description	Interaction * Lang FFD	SFD	GD	TFT
Lang	language	English or French				
InFreq	In(freq of word)	In of word frequency		*		*
FixLoc	fixation location	location of first fixation	*	*	*	*
Launch	launch site	distance from previous fix		*		*
Length	length of word	number of letters word	*		*	*
Plaus	plausibility	values from a rating task			*	*
LengthM	length of NP mate	number of letters NP mate				*
InFreqM	In(freq of NP mate)	In of NP mate frequency				
Pred	predictability	values from a Cloze task				
Wd#	word number	word1 or word2 in NP	*		*	

Determinants of EM control: Statistical model matrices with R^2

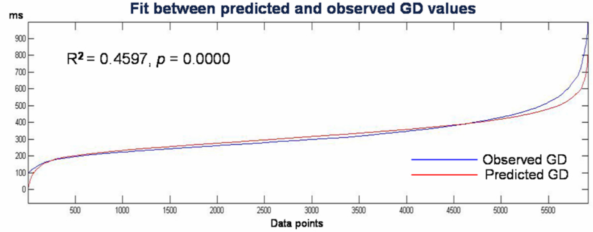


Y	= C +Lang+	Predictors	+ Subs	+Lang*Predictors	+ Subs*Predictors + e
FFD	0.0018*	0.0685*	0.0783*	0.0120*	0.2938
SFD	0.0005	0.0855*	0.0869*	0.0130*	0.3423
GD	0.0000	0.1131*	0.1847*	0.0102*	0.2647
TFT	0.0000	0.1572*	0.0590*	0.0178*	0.2644

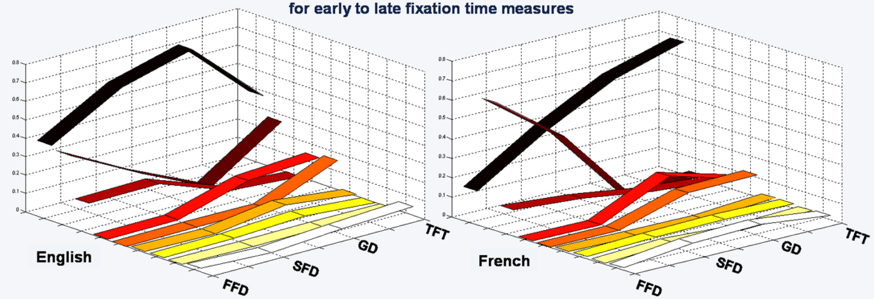
- **No effect of Language on Single Fixation Duration (SFD), Gaze Duration (GD), or Total Fixation Time (TFT)**
- **Strong effects of Predictors**
- **Interaction between Language and Predictors**

  FixLoc has a stronger relative influence on the early fixation time measures in French than in English (probably due to the importance of number and gender inflections in French).

Fit between predicted and observed GD values



Relative part of variance explained by predictors in English and French for early to late fixation time measures



Parafoveal-on-foveal effects on GD: Regression analysis on word1 of NP

English:

	R^2	F	p	beta
InFreq	**0.0277**	**41.3820**	**0.0000**	**-11.8069**
LengthM	**0.0030**	**4.4642**	**0.0014**	**-4.1132**
Pred	**0.0020**	**3.0045**	**0.0128**	**183.2258**
Length	0.0006	0.8531	0.6179	3.6934
Launch	0.0004	0.5386	0.8859	2.5037
PredM	0.0002	0.3586	0.9756	-14.5816
InFreqM	0.0002	0.2670	0.9938	-1.1236
FixLoc x InFreqM	0.0000	0.0741	1.0000	-1.3050
FixLoc	0.0000	0.0737	1.0000	5.7326
FixLoc x PredM	0.0000	0.0016	1.0000	0.0074
Plaus	0.0000	0.0001	1.0000	-0.0253

French:

	R^2	F	p	beta
InFreq	**0.0157**	**26.0791**	**0.0000**	**-7.7123**
Launch	**0.0037**	**6.1959**	**0.0002**	**7.3310**
Length	**0.0016**	**2.7156**	**0.0210**	**5.6613**
FixLoc x PredM	**0.0015**	**2.5709**	**0.0271**	**0.0061**
Plaus	**0.0014**	**2.4044**	**0.0366**	**-3.1937**
PredM	0.0007	1.1175	0.4037	39.4500
InFreqM	0.0006	1.0044	0.4887	1.4365
FixLoc	0.0002	0.3462	0.9791	-4.2206
FixLoc x InFreqM	0.0001	0.1155	0.9999	3.1878
Pred	0.0001	0.0909	1.0000	-59.7966
LengthM	0.0000	0.0235	1.0000	0.3125

English: No effect of wd2's frequency or predictability on wd1. Effect of wd2's length on GD of wd1.

French: No effect of wd2's frequency or predictability on wd1. Interaction between Pred of wd2 and FixLoc on wd1.

Discussion

- Despite no significant global differences in fixation durations in English and French, there does seem to be underlying processing differences as evidenced by the different patterns obtained in the regression analyses by Language.
- The relative weights of 10 predictors (both oculomotor and psycholinguistic) on fixation durations in reading English and French were obtained. These findings can help inform models of eye movement control in reading.
- On some fixation time measures, plausibility accounts for a larger amount of variance than predictability. Further investigations are needed to directly compare of the relative weightings of different contextual constraint indices (e.g., predictability, plausibility, or transitional probability) in order to select the best index for the models.
- Parafoveal-on-foveal effects were examined via a regression model on word1's fixation time. Although there were no clear lexical parafoveal-on-foveal effects, there was evidence that certain characteristics of word2 affected EM behavior on word1.
- The weights of the variables used in models of EM control in reading should be adapted to the language that is being read.

References and Acknowledgments

British National Corpus. (1995). http://www.natcorp.ox.ac.uk/.

Engbert, R., Nuthmann, A., Richter, E.M., & Kliegl, R. (2005). A dynamical model of saccade generation during reading. Psych Rev, 112, 777-813.

Kliegl, R. Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. JEP:Gen, 135, 12-35.

Lorch, R.F., & Myers, J.L. (1990). Regression analyses of repeated measures data in cognitive research. JEP:LMC, 16, 149-157.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE (http://www.lexique.org/). L'Année Psychologique, 101, 447-462.

Pollatsek, A., Reichle, E.D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. Cog Psych, 52, 1-56.

This work was supported by an ESRC Research Grant to S.C. Sereno and P.J. O'Donnell. Corresponding author: **miellet@psy.gla.ac.uk**