Bridging Mechanisms of Reading, Viewing and Working Memory during Attachment Resolution of Ambiguous Relative Clauses

Eunice Fernandes (fernandese@campus.ul.pt)[a]
Armanda Costa (armandacosta@letras.ulisboa.pt)[a]
Moreno I. Coco (micoco@psicologia.ulisboa.pt)[b]
[a] Department of Linguistics and [b] Department of Psychology, University of Lisbon,
Alameda da Universidade, 1600-214 LISBOA - PORTUGAL

Abstract

Cognitive mechanisms sustaining reading, viewing and working memory are mostly independently examined. In this study, we investigate their interaction in a visual-world priming task. Participants first read relative clauses (RC) morphologically disambiguated for high-(to *NP1*, HA) or low-attachment (to *NP2*, LA) (e.g., *The helper|helpers of the bakers|baker who* ***will*** $_{[sg]}$ *deliver the bread* ***has|have*** *arrived*), and then heard a spoken temporarily ambiguous RC (e.g., *the father of the baby who will drink the beer|baby bottle is tall*) while presented with a visual context (i.e., VWP). Using linear-mixed effects models, we predict anticipatory fixations to the visual referents associated with NP1 and NP2 as a function of: (i) *second-pass* time observed during their reading and (ii) individual working memory scores. We demonstrate that high-capacity individuals anticipate more the (non-primed) visual referent when they reread more its associated NP (e.g., anticipate the visual referent 'father' when reread more often NP1, *the helpers*). We suggest that working memory capacity allows individuals to maintain alternative syntactic analyses of the sentence, and evaluate them upon a subsequent visual context. These findings provide support to the constraint-capacity theory, and shed new light on the cross-modal mechanisms underlying syntactic ambiguity resolution.

Keywords: reading; visual-world paradigm; working memory; relative clause attachment; eye-movements.

Introduction

As we read, words are attended and integrated into sentences (Marslen-Wilson & Tyler 1980). Similar processes of integration are at work, when spoken language occurs concurrently with a visual context (VWP, Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). Working memory (WM) is also actively involved during sentence processing, either by constraining the system principles (e.g., Frazier & Rayner (1982)'s garden-path theory) or by determining the extent to which multiple interpretations can be maintained in parallel (e.g., Just & Carpenter (1992)'s capacity theory). Some accounts of the VWP also assume WM to be an essential component (e.g., Huettig, Olivers & Hartsuiker, 2011). In this study, we investigate how reading, and viewing during situated language understanding, depend on WM capacity.

To preview our work: we utilize data from a structural priming task (detailed in section *Methods*), where participants were screened for verbal (reading and backward-digit spans), and non-verbal (spatial span) WM tasks. Then, in an eye-tracking experiment, they first read temporary ambiguous relative clause sentences (RC), disambiguated for high-(HA) or low attachment (LA), and were subsequently presented with another spoken RC, but this time concurrently with a visual context. Building on the capacity theory, we hypothesized that high WM individuals can maintain in memory multiple analyses of the sentences read and, therefore, evaluate them on the visual context once attachment ambiguity needs to be resolved during situated language understanding. To test our hypothesis, we examine anticipatory looks to visual objects corresponding to each of the primed attachments (HA, LA), as a function of associated reading time to NP1 and NP2, and to the scores of working memory tests. We find high-WM individuals to anticipate more the visual object associated to the alternative RC (non-primed) interpretation, when they reread for longer the corresponding NP. These results give support to the capacity constraint-theory, and go beyond it, by uncovering the shared cognitive mechanisms bridging reading, viewing and working memory.

Before going into the details of our study, however, we contextualize our research hypothesis within the literature on reading and situated understanding of structurally ambiguous sentences.

Background

Structural ambiguity has always played a pivotal role for research on sentence processing and theories of syntactic parsing. Frazier and Rayner (1982), for example, analysed eye-movements while participants read sentences with a temporary ambiguous post-verbal noun phrase, such as (a) *Since Jay always jogs a mile this seems like a short distance to him* and (b) *Since Jay always jogs a mile seems like a very short distance to him*. The results show longer reading times in (b) compared to (a). The *garden-path* model was proposed to explain that (b) *a mile* is initially interpreted as the direct object of the verb *jogs* (i.e., late closure principle): an incorrect interpretation that needs to be revised by reanalysing the ambiguous constituent. Constraint-based models instead assume that the sentence processor activates multiple interpretations in parallel to resolve ambiguity, and that this process is mediated by different sources of information like lexical frequency or pragmatic plausibility (refer to MacDonald & Seidenberg, 2006, for a review). In particular, Just and Carpenter (1992) propose that all comprehenders can construct multiple interpretations when encountering an ambiguity, each having an activation level dependent on inherent linguistic properties of the structure. Furthermore, WM capacity is assumed to constrain the duration the sentence processor can maintain multiple interpretations: high-span readers can hold more efficiently in memory multiple representations, while low-span ones have difficulties maintaining more than one interpretation, and therefore abandon the less preferred one.

Important to the present work is the study by Traxler (2007), which investigates how RC attachment resolution (i.e., the syntactic structure examined in our study) is sensitive to WM capacity. Participants, assessed on WM reading span, were eye-tracked as they read sentences such as *The writer of the letter that had blonde hair arrived this morning*, where the RC must be high-attached (HA), i.e., the NP antecedent of *that* is *the writer* (and cannot be *the letter*); and compared it with a low-attached (LA) version of the same sentence, *The letter of the writer that had blond hair arrived this morning*. Among other corroborating measures, *total time* in and *regressions* from a post-disambiguating region (*arrived this*) decreased with increases in WM in HA compared to LA (i.e., easier processing)[1]. This result is taken to support the capacity theory of Just and Carpenter (1992), as high-span readers should retain more easily in memory the representation of earlier constituents in the sentence.

Evidence for attachment preferences of RCs on eye-movement can also be found in situated language understanding. Kamide (2012), for example, shows that attachment preference for RCs such as *The uncle of the girl who will ride the motorbike (OBJ1) /carousel (OBJ2) is from France* can be anchored to the sex identity of the speaker's voice; whereby, if a male talker is always associated to HA, and a female voice, instead, always to LA, then anticipatory eye-movement to either *the motorbike* or *the carousel* (respectively for male and female talkers) are observed when the verb *ride* is heard. Anticipatory looks during this linguistic region to the visual referents of OBJ1 or OBJ2 inform on the interpretation pursued (HA or LA).

Eye-movements are a sensible index of the interpretative strategies adopted as ambiguous RC are either read or understood in a visual context. Moreover, WM can be a key indicator of readers' capacity to retain and evaluate multiple analyses of attachment resolution, as well as an important component for situated language understanding. To the best of our knowledge, attachment resolution of RCs has been studied independently in reading and during situated language processing. Reading of RCs may be informative of attentional patterns occurring when these structures are understood, by the same individual, in a visual context. In addition, WM capacity can be involved on acquisition of structural information during reading and on its pro-active re-use during situated understanding.

The present study

We investigate how reading times on temporary ambiguous RCs (HA and LA) can predict anticipatory eye-movement during spoken understanding of the same structure situated in a visual context. Most importantly, we examine how WM capacity mediates the transfer of structural information from reading to viewing in a visual-world paradigm (VWP) task.

Portuguese speakers have a preference for low-attachment of RCs (Maia, Fernández, Costa & Lourenço-Gomes, 2007); therefore, more difficulties are expected for high-attachment (for corroborating results refer to next section). This could be reflected, for example, by more regressions into the antecedent NP1. However, low-memory readers may not retain the attachment information long enough to reuse it on a subsequent visual context, especially when that interpretation can be abandoned. On the contrary, high-span readers should be able to activate and retain for longer both alternatives. So, in sentences disambiguated for LA, we expect reading of NP1 (HA interpretation) by high-span readers to predict stronger anticipatory looks to its associated visual object when spoken ambiguity needs to be resolved, i.e., upon hearing the ambiguous pronoun. This result would indicate that WM directly modulates the capacity to activate and retain information associated with the alternative syntactic attachment, and pro-actively utilize such information in a subsequent visual-world context.

Method

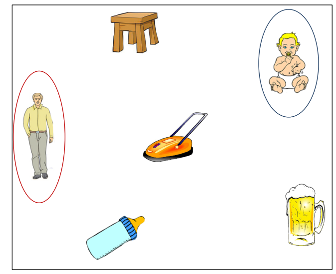
We utilize a subset of data collected in a previous study which investigates syntactic priming of RCs (Fernandes, Coco & Branigan, 2014). In particular, participants (N=24) read aloud two Portuguese RC sentences (1a-b), both disambiguated for high- or low-attachment[2] (prime: HA, LA), or a filler and (1b), thus manipulating the number of primes (1 vs. 2) read within the trial. Subsequently, they listened to another temporarily ambiguous spoken RC sentence, like (2), while concurrently viewing a clip-art object array depicting the possible antecedents of the ambiguous pronoun (S1 and S2, the visual correspondents of NP1 and NP2), as well as the objects of the relative clause, and two other distractors (please refer to Table 1 for an example of sentences and visual context, and to Figure 1 for a trial run). In addition to the 48 experimental items, 32 sentences and 64 visual-world trials, with no RCs

[1] But see Swets, Desmet, Hambrick & Ferreira (2007) for inconsistent evidence on WM's mediation of ambiguous RCs' attachment in an offline task.

[2] In Portuguese *will* is morphologically marked for number and therefore disambiguates the attachment.

Table1: Example of sentences of a critical trial

(1)	**Written primes**
(a)	**High attachment: to NP1 (Low attachment: to NP2)**
	O irmão (irmãos) dos herdeiros (do herdeiro) que vai ler o testamento é (são) de França
	The brother (brothers) of the heirs (heir) who will [sg] read the will is (are) from France
(b)	**High attachment: to NP1 (Low attachment: to NP2)**
	O ajudante (ajudantes) dos padeiros (do padeiro) que vai distribuir o pão chegou (chegaram)
	The helper (helpers) of the bakers (baker) who will [sg] deliver the bread has (have) arrived
(2)	**Visual-World Targets**
	High attachment: to NP1 (Low attachment: to NP2)
	O pai do bebé que vai beber a cerveja (o biberão) é alto
	The father of the baby who will [sg] drink the beer (the baby bottle) is tall



(ditransitive verbs and conjoined constructions), were fillers, interleaved with the experimental items. From this dataset, we only consider the subset of data where two sentences with the same structure[3] were read (i.e., 2 primes condition). The same participants were also screened for verbal (reading and backward-digit spans) and non-verbal (spatial span) working memory capacity prior/after (counterbalanced between-participants) the eye-tracking session.

A pre-test on our target sentences corroborated the Portuguese preference for low-attachment. Participants (N=15) read 96 sentences (48 sentences in both HA and LA versions, presented in two random orders), and were asked to tell who was the agent of the action (i.e., *the father* or *the baby* in (2)), and score the acceptability of the sentence (on a Likert scale form 1 (totally unacceptable) to 7 (totally acceptable). A low-attachment preference was substantiated by higher accuracy in the LA condition (mean 0.95, sd = 0.21) compared to the HA condition (mean 0.58, sd = 0.49). Sentence acceptability, however, did not depend on its attachment type: ratings for LA (mean 5.1, sd = 1.69) were not significantly higher than HA (mean 4.4, sd = 1.78), as indicated by a Wilcoxon Signed-Ranks Test ($Z = -1.24$, $p = 0.22$). Additional evidence for this preference is found also in reading. In particular, an analysis of regressions into NP1 and NP2 as a function of AOI (NP1 vs. NP2) and sentence type (HA vs. LA), shows main effects of sentence type and AOI, whereby readers regressed more often in HA sentences and into NP1 ($\beta = -0.09$ and $\beta = -0.33$, both $p < 0.05$).

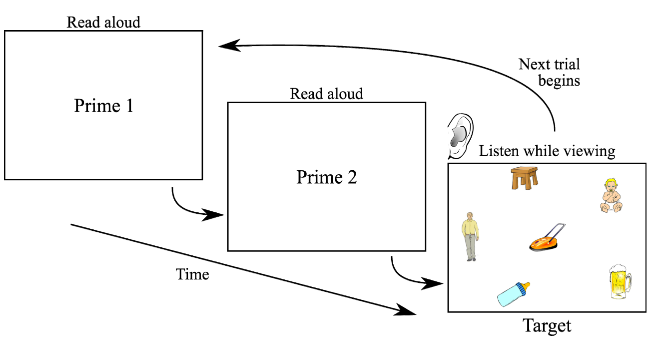


Figure 1: Illustration of a critical trial's time course

Eye-movement Monitoring Procedure

Participants' eye-movements were recorded using an SMI IVIEW X™ HI-SPEED eye-tracker at a sampling rate of 1250 Hz on a 21" screen (1024 x 768 px. image resolution). Viewing was binocular but only the participant's dominant eye was tracked (determined by a prior parallax test). Connected to the participant PC was a satellite speaker and subwoofer system for auditory presentation. A 5-point calibration was done before the experiment began and was repeated every 4 critical trials (20 sentences) or whenever the experimenter found it necessary (half of these included a validation process where we accepted the calibration for angle deviations smaller than 0.5/1 for x and y respectively).

Working Memory Assessment

Participants' working memory was assessed through verbal (reading and backward-digit) and non-verbal (spatial) WM span tasks (Swets et al., 2007; Waters and Caplan, 2003). Each task had 70 items in sets of increasing size (5 sets of 2 to 5 items in the reading and spatial tasks, and 2 sets of 2 to 8 items in the backward-digit task).

In the reading task, the sentences were presented for 5 seconds. After each sentence, participants had to make a grammaticality judgment and, after the end of the set, they had to recall the final word of each sentence read, in the correct order of presentation[4]. In the backward-digit task, the sets consisted of digits from 1 to 9 (in sequences drawn from a table of random numbers), presented one at a time for 700ms. At the end of each set, digits should be recalled in reverse order of presentation. In the spatial span task, participants were presented with sets of letters, one at a time, for 3 seconds. Each letter could be normal or mirror-

[3] We do so in order to have a more reliable index of reading, by computing the mean of reading measures across the 2 sentences with an identical structure.

[4] We used Gaspar and Pinto (2001)'s materials in the adaptation of Daneman and Carpenter's (1980) task to European Portuguese.

imaged, and appear in different orientations (i.e., rotated in 7 possible angles). Participants had to decide, after each letter, if it was normal or mirror imaged; and, after each set, recall the original orientation of each letter of the set in the correct order of presentation.

In all tasks, a recall prompt signaled that the set had ended and that the participant should recall the items. The duration of this prompt increased as a function of the set size. For each task, the total number of correctly recalled items (ranging from 0 to 70) by one participant was taken as that participant's score on the task. For each participant, we computed a composite measure based on the average of the scores of the three tasks, and standardized it into *z*-scores. This procedure increases reliability of the scores, and their generalizability across tasks (Salthouse, 1994).

Analyses

Ambiguity resolution is differently characterized by eye-movement in reading and during situated understanding. During reading, eye-movement responses index processes of attachment interpretation and re-analysis (e.g., re-reading previous ambiguous regions[5]). During situated understanding, instead, they index anticipatory mechanisms evaluating the object in the context which most likely resolves the attachment (e.g., fixation to S1 for an HA interpretation upon hearing *who*).

From the reading data, as AOIs, we consider the two noun phrases referred by the pronoun, NP1 (*The helper|The helpers*) and NP2 (*of the bakers|of the baker*). From the VWP data, as AOI, we consider the corresponding depicted referents, S1 (The father) and S2 (the baby), refer to Table 1 for example material. We focus on these linguistic and visual AOIs to establish clear links between the referents read, and their counterparts in the visual context.

Reading Measures: We consider *second-pass* on NP1 and NP2, as the sum of all fixations duration, beginning with the first reentry in the region and ending when the reader leaves it, in any direction. Length-residualized values, i.e., the differences between predicted (by linear regression as function of word length) and observed values, were computed for each subject, to account for word variability (Trueswell, Tanenhaus, & Garnsey, 1994). As stated in the *Method* section, each reading trial has two prime sentences with the same structure (both HA or LA). In order to have a unique reading measure to use as predictor of fixation, we consider the mean *second pass* between the first and second sentences. If the eye-movement of a trial did not record two or more of seven AOIs considered (indicated by the slashes in *The helper* | *of the bakers* | *who will* | *deliver* | *the bread* | *has arrived*), we removed it, assuming mis-calibration. This procedure removed 148 of 1150 trials (12.8%).

Visual-World Measures Six visual objects were defined in the array by drawing annotated polygons labelled as: S1 and S2 (the antecedents of the pronoun for HA and LA, respectively), O1 and O2 (the objects acted upon by the subjects), Central (accounting for center bias) and Distractor (an unrelated object). Fixation coordinates from the tracker output were mapped onto these areas.

For each trial, fixations were aligned, item-by-item, to the onset of the pronoun *who*, creating a time window up to 400ms after it[7] (with 200ms added to account for oculo-motor programming). Fixation points were aggregated in 50ms bins, and the percentage of fixations on each visual object was calculated relative to the total amount of fixation on every other object. We took the mean percentage of fixations across 8 bins (i.e., 400ms), and obtain a unique measure of fixation to correlate with the *second-pass* observed in the same trial (refer to Figure 1 for the trial run).

Our dependent variable is percentage of fixations on S1 (and S2), predicted as a function of reading NP1 (and NP2), prime type (HA vs. LA) and WM z-scores in a maximal-random mixed effect models (Barr, Levy, Scheepers & Tily, 2013). The predictors are centered, as a precaution, to avoid co-linearity between predictors. Participant (24) and Item (48) are the random effects, entered as intercept (e.g., (1 | Participant)), as well as uncorrelated slopes for the predictors (e.g., (0 + Prime | Participant, in `R` syntax).

Results and Discussion

In Figures 2 and 3, we scatter percentage of fixations to S1 and S2, as a function of *second-pass* readings on NP1 and NP2 respectively, for high and low WM groups[8] in the two priming conditions (HA - left panels, LA - right panels). We observe that fixations to S1 and S2 are differently modulated by *second-pass* reading times on NP1 and NP2, the type of Prime, and WM. In particular, starting with fixation to S1, we observe a significant two-way interaction between *prime* and *second-pass*, whereby stronger anticipation to S1 is observed with longer *second-passes*, especially when the prime is LA (two-way interaction *prime*:*secondpass*, refer to Table 2 for the model coefficients). Moreover, this effect is especially strong in individuals with high WM (*prime*:*secondpass*:*WM*, visualized in Figure 2, right LA panel). This result shows that memory capacity enables comprehenders to encode and recall the HA alternative when primed with the preferred

[5] This is a general prediction of garden-path theories that are not explicit concerning where to (NP1 or NP2), in such RCs, readers regress when encountering ambiguity (cf. Frazier & Rayner, 1982).

[7] See Arnold, Eisenband, Brown-Schmidt & Trueswell (2000) for evidence of rapid anticipatory eye-movement as indexes of pronoun resolution in a visual-world task.

[8] Note the WM grouping (considering two halves of the distribution) is only for purposes of visualization; the WM scores were introduced as a continuous measure in the model.

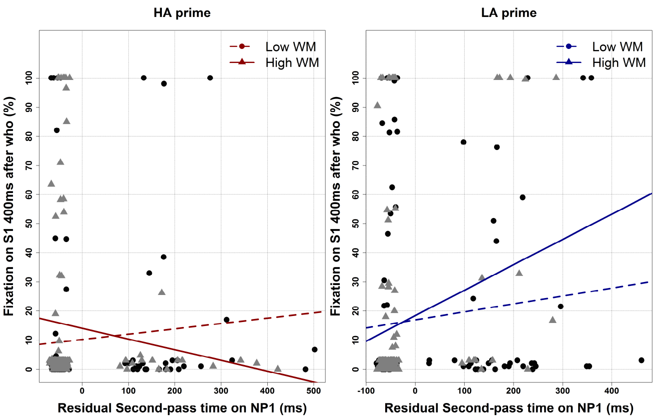


Figure 2: Scatter plot of Percentage of fixation on S1 as a function of residual Second-pass time on NP1 in HA (left) vs. LA (right) prime and low (dashed) vs. high (solid) WM.

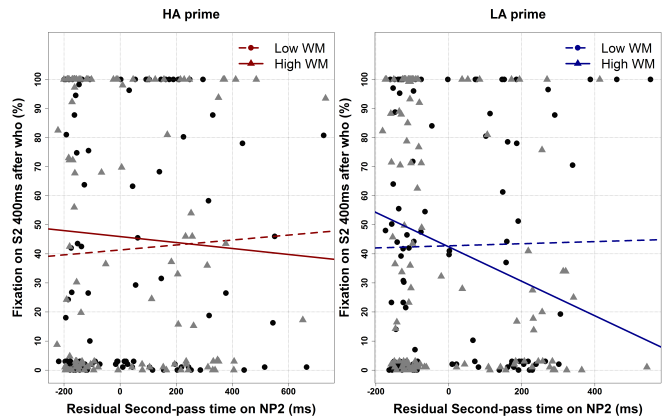


Figure 3: Scatter plot of Percentage of fixation on S2 as a function of residual Second-pass time on NP2 in HA vs. LA prime and low vs. high WM.

Table 2: Summary of maximal models fitted to Fixations on S1 (left) and S2 (right) as a function of *Second-Pass* Residual Reading times, Prime and WM. For each effect, we report the coefficient estimate (in percentage of fixation), its standard errors and significance tests (t and p-values). All factors are centered to reduce co-linearity; Prime (HA = -0.5; LA = 0.5), Reading (ranging from -64 to 501ms (NP1); -212 to 719 (NP2)) and WM (ranging from -2 to 1.6) are continuous measures.

Residual ***Second-pass*** time									
Percentage of Fixation on S1 (400ms after who)					Percentage of Fixation on S2 (400ms after who)				
	Estimate	**SE**	**t**	**p**		**Estimate**	**SE**	**t**	**p**
(Intercept)	13.378	2.042	6.552	< 0.0001	*(Intercept)*	42.400	3.570	11.878	< 0.0001
prime	5.325	3.084	1.727	0.084	*prime*	-1.774	3.811	-0.465	0.642
secondpass (NP1)	0.022	0.014	1.568	0.117	*secondpass* (NP2)	-0.006	0.011	-0.524	0.600
WM	1.037	2.040	0.508	0.611	*WM*	-0.336	3.529	-0.095	0.924
prime:WM	-1.706	3.204	-0.532	0.594	*prime:WM*	-1.285	3.962	-0.324	0.746
prime:secondpass	**0.070**	**0.028**	**2.530**	**0.011**	*prime:secondpass*	-0.034	0.021	-1.617	0.106
secondpass:WM	-0.001	0.013	-0.074	0.941	*secondpass:WM*	-0.013	0.011	-1.133	0.257
prime:secondpass:WM	**0.050**	**0.025**	**2.023**	**0.043**	***prime:secondpass:WM***	**-0.042**	**0.021**	**-1.961**	**0.050**

LA structure. Most importantly, this happens when they reread for longer NP1, indicating that they evaluated an HA interpretation of the sentence. Fixations to S2 corroborate this result. Here, we observe less anticipatory looks to S2, for longer *second-pass* reading NP2 in LA primes, for high-WM individuals. Our results can be motivated by the capacity-theory of sentence processing, where WM holds multiple possible representations for the same sentence.

The interaction, however, was stronger on fixations to S1 with respect to the *second-pass* readings on NP1. This result might reflect the preference of Portuguese readers for LA. In fact, percentage of fixations to S2 (LA interpretation, Figure 3) is overall higher than anticipation of S1 (HA interpretation, Figure 2), reflecting a bias towards LA. In Just & Carpenter's theory, the higher WM is, the more the individuals can construct multiple representations when faced with an ambiguity, which levels of activation are determined by other factors such as the frequency of occurrence in the language. In our data, high-WM readers, as expected, show less difficulty with the dispreferred high-attachment; they do not have to re-inspect NP1 in reading HA to further anticipate S1 (additional analysis on residual *second-pass* on NP1 as a function of WM and Prime revealed a main effect of WM ($\beta = -24.61$, $p = 0.036$), confirming that HA sentences are not so demanding for high-capacity readers). Yet, the capacity model assumes that the less frequent alternative (here HA) should have an a priori lower level of activation, even when high-memory allows for its activation. Why, then, should high-WM readers strongly anticipate S1, rather than S2, when presented with a LA sentence? We suggest that memory capacity not only allows for encoding of multiple alternatives but also leads to strategic strengthening of less frequent representation, to avoid future prediction errors, as proposed by learning accounts of priming (e.g., Chang, Dell & Bock, 2006). This strategy involves a greater allocation of attention to the syntactic elements associated with alternative interpretation, a process allowing transfer of structural information when needed to be recalled in a visual context.

Conclusion

Important advances in theories of sentence processing have been made by investigating visual attention during reading (e.g., Frazier & Rayner, 1982). More recently, the VWP has been used to uncover the referential mappings of syntactic analyses in a visual context. As noted by Huettig, Rommers & Meyer (2011), VWP research focused on sentence processing, but the dependent measures concern visual attention, and might demand the involvement of memory.

In the present study we provide, for the first time, linking evidences between online reading measures of sentences and online measures of their incremental processing when situated in a visual context[9]. Furthermore, we aimed to reconcile the dynamics of these processes with working memory capacity. We did it by showing that ambiguity resolution strategies at reading can directly inform on anticipatory mechanisms during language understanding, and are mediated by WM capacity.

Our results lend support to parallel processing accounts of syntactic ambiguity resolution and, in particular, to Just and Carpenter (1992)'s capacity theory. Moreover, they directly contribute to advance such theory by showing that translational mechanisms are at place when comprehenders engage into reading and situated language understanding.

Acknowledgments

Fundação para a Ciência e Tecnologia under awards number SFRH/BD/72307/2010 to EF and SFRH/BDP/88374/2012 to MIC, is gratefully acknowledged.

References

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76, B13-B26.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255-278.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113, 234–72.

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

Fernandes, E., Coco, M. I., & Branigan, H. P. (2014). Saturating the Anticipatory Mechanisms of Syntactic Priming during Situated Language Understanding. *AMLaP XX*, Edinburgh, Scotland, 4–6 September 2014.

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

Gaspar, N. & Pinto, A. da C. (2001). Amplitude de memória a curto prazo e operatória para dígitos e palavras com a adaptação da tarefa de Daneman e Carpenter (1980). *Psicologia, Educação e Cultura*, V (2), 217-236.

Huettig, F., Olivers, C. N. L., & Hartsuiker, R. J. (2011). Looking, language, and memory: Bridging research from the visual world and visual. *Acta Psychologica*, 137, 138-150.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137, 151-171.

Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, 24, 56-98.

Kamide, Y. (2012). Learning individual talkers' structural preferences. *Cognition*, 124(1), 66-71.

Knoeferle, P., Urbach, T. P., & Kutas, M. (2011). Comprehending how visual context influences incremental sentence processing: Insights from ERPs and picture-sentence verification. *Psychophysiology*, 48, 495–506.

MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. Traxler & M. Gernsbacher (Eds.), *The handbook of psycholinguistics* Amsterdam: Elsevier.

Maia, M., Fernández, E. M., Costa, A., & Lourenço-Gomes, M. C. (2007). Early and late preferences in relative clause attachment in Portuguese and Spanish. *Journal of Portuguese Linguistics*, 6, 227-250.

Marslen-Wilson, W. & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1-71.

Salthouse, T. A. (1994). The Aging of Working Memory. *Neuropsychology*, 8 (4), 535-543.

Swets, B., Desmet, T. Hambrick, D., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, 136(1), 64-81.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632-1634.

Traxler, M. J. (2007). Working memory contributions to relative clause attachment processing: A hierarchical linear modelling analysis. *Memory & Cognition*, 35, 1107-1121.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285–318.

Waters, G. S. & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers*, 35, 550-64.

[9] See Knoeferle, Urbach & Kutas (2011) for linking evidence of congruency effects measured by ERPs (while reading sentences following pictures) and offline verification answers.