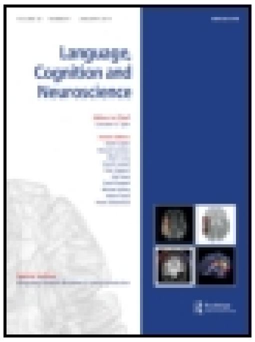

Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: http://www.tandfonline.com/loi/plcp21

Prominence in Spanish sentence comprehension: an eye-tracking study

Carolina A. Gattei, Yamila Sevilla, Ángel J. Tabullo, Alejandro J. Wainselboim, Luis A. París & Diego E. Shalom

To cite this article: Carolina A. Gattei, Yamila Sevilla, Ángel J. Tabullo, Alejandro J. Wainselboim, Luis A. París & Diego E. Shalom (2017): Prominence in Spanish sentence comprehension: an eye-tracking study, Language, Cognition and Neuroscience, DOI: 10.1080/23273798.2017.1397278

To link to this article: http://dx.doi.org/10.1080/23273798.2017.1397278

LANGUAGE, COGNITION AND NEUROSCIENCE, 2017
https://doi.org/10.1080/23273798.2017.1397278

Routledge
Taylor & Francis Group

Check for updates

Prominence in Spanish sentence comprehension: an eye-tracking study

Carolina A. Gattei [a,c], Yamila Sevilla[b], Ángel J. Tabullo[a], Alejandro J. Wainselboim[a], Luis A. París[a] and Diego E. Shalom[c]

[a]Instituto de Ciencias Humanas, Sociales y Ambientales - CONICET, Mendoza, Argentina; [b]Instituto de Lingüística, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina; [c]Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and IFIBA, CONICET, Ciudad Autónoma de Buenos Aires, Argentina

ABSTRACT

We report an eye-tracking experiment that examined argument linking and the role of prominence in Spanish sentence comprehension by testing the interplay between word order and verb type. Previous evidence from a self-paced reading study (Gattei, Dickey, Wainselboim, & París, 2015). showed that comprehenders use morphosyntactic information to form predictions about the thematic structure of the upcoming verb. In this study we focussed on the time course of this process. Results showed an interaction between verb type and word order for late eye movement measures but not for early eye movement measures. Participants regressed more and for longer time when word order did not match the canonical order for each verb class. This interaction is observed from the verb region onwards, independently of word order. We interpret that these effects take place due to the misinterpretation of the prominence status of the preverbal argument, leading to differential reading strategies.

ARTICLE HISTORY
Received 24 June 2016
Accepted 6 October 2017

KEYWORDS
Prominence; word order; verb class; eye-tracking; spanish

Introduction

The understanding of language involves the integration of multiple types of linguistic information. Over the past decades, special attention has been paid to the linking of syntactic and semantic information, since it is through this process that speakers understand "who did what to whom" in an event. To exemplify, consider the sentence "Mary responded to Ana." In order to understand this sentence, speakers need not only to understand the syntactic structure of the sentence; they also need to recover the information that describes a situation where someone responded to someone else; retrieve who the participants involved in the event are and assign them a thematic role (e.g. Agent, Patient, Experiencer, Recipient). Furthermore, speakers need to weigh these two participants against each other by building a hierarchical representation based on the morphosyntactic and semantic features of the sentence constituents. This hierarchy among the arguments of a sentence is known as *prominence* (Bornkessel & Schlesewsky, 2006). Several linguistic features are considered by the parser in order to establish the degree of prominence of a constituent relative to another constituent: thematic role, morphological case marking, argument position, animacy, person and definiteness are the most salient ones. These features are usually conceptualised in terms of "hierarchies that (…) rank animates over inanimates, definites over indefinites, first and second person over third and agents over patients. The higher an element's rank on the hierarchy, the more prominent it is considered to be." (Kretzschmar, Bornkessel-Schlesewsky, Staub, Roehm, & Schlesewsky, 2012).

The question about whether the computation of prominence is relevant for the incremental interpretation of sentential arguments and for the prediction of the type of event that will be described has been raised by the investigation of how verb-final constructions are processed. Crucially for the purposes of the present work, many of the studies related to this issue support the claim that prominence information is rapidly extracted in order to set predictions about the lexico-semantic structure of the upcoming main verb in a sentence. The evidence shows that a misanalysis of prominence information leads to differential neural and electrophysiological activity (Bornkessel, Schlesewsky, & Friederici, 2003; Bornkessel, Zysset, Friederici, von Cramon, & Schlesewsky, 2005; Dröge, Maffongelli, & Bornkessel-Schlesewsky, 2014; Gattei, Tabullo, París, & Wainselboim, 2015; Wang, Schlesewsky, Philipp, & Bornkessel-Schlesewsky, 2012, among others), longer reading (Kretzschmar et al., 2012) and response times, and higher error rates (Gattei, Dickey, Wainselboim, & París, 2015).

CONTACT Carolina A. Gattei carolina.gattei@conicet.gov.ar
Supplemental data for this article can be accessed http://dx.doi.org/10.1080/23273798.2017.1397278.

However, it is a well-established fact that the linguistic features that are relevant for argument interpretation may differ from language to language (see Bornkessel-Schlesewsky & Schlesewsky, 2009; Lamers & De Swart, 2012, for a discussion on the issue). For example, word order is a relevant cue for establishing arguments' prominence in languages with little or no case marking inflection (e.g. English, Chinese), but it is only considered a last-resort strategy in languages with a rich case marking system (e.g. Finnish, Japanese, Russian). Similarly, prominence computation may rely on subject dropping (e.g. Spanish), or on the relative animacy of arguments (e.g. Fore language, and see Wang, Schlesewsky, Bickel, & Bornkessel-Schlesewsky, 2009; Wang et al., 2012, for two studies on the role of animacy in Chinese). These language-specific differences call for the investigation of how the interaction of different linguistic cues that are related to the computation of prominence may affect incremental processing, and raise the question about whether there are common neuroanatomical, neurophysiological and cognitive correlates of prominence processing among languages. The present study aims to provide further evidence about the role of prominence information for the incremental interpretation of arguments during natural reading in Spanish, a language that has been scarcely addressed in this regard. More precisely, we investigate the eye movement patterns associated to the comprehension of sentences that require a reversal of argument prominence hierarchization. In the following section, we briefly describe the linguistic features of Spanish relevant for the establishment of prominence.

Argument hierarchization in spanish

Argument prominence scales in Spanish result from the interaction of word order, verb type, case marking, animacy and definiteness (Gutiérrez-Bravo, 2007; Primus, 2012). In this work, we focus on the interplay between the first three linguistic features. Consider sentences in (1).

(1) a. María le responde a Ana.
Mary$_{[NOM]}$ clitic$_{[DAT]}$ respond$_{[3rd-SG]}$ to Ana$_{[DAT]}$
"Mary responds to Ana"
b. A Ana le responde María.
To Ana$_{[DAT]}$ clitic$_{[DAT]}$ respond$_{[3rd-SG]}$ Mary$_{[NOM]}$
"Mary responds to Ana."

The verb in (1) expresses an activity. Example (1a) represents a subject-initial sentence; in contrast, in (1b) the dative object precedes the subject. Both sentences mean "Mary responds to Ana." Even though (1b) is a structure in which the Undergoer linearly precedes the Actor, speakers understand in both cases that it is Mary that responds to Ana and not the other way round through case marking (the dative pronoun "le"), clitic doubling (co-reference between "le" and "Ana") and the preposition "a," that functions as particle marking indirect object.[1]

Regarding the role of verb type in the computation of prominence, identical syntactic configurations may instantiate different semantic hierarchizations for its constituents, as it is shown in sentences in (2).

(2) a. María le teme a Ana.
Mary$_{[NOM]}$ clitic$_{[DAT]}$ fear$_{[3rd-SG]}$ to Ana$_{[DAT]}$
"Mary fears Ana."
b. María le encanta a Ana.
Mary$_{[NOM]}$ clitic$_{[DAT]}$ love$_{[3rd-SG]}$ to Ana$_{[DAT]}$
"Ana loves Mary."

Both sentences are syntactically alike and they require a similar semantic structure. One of its constituents is assigned the role of Experiencer and the other one is assigned the role of Theme/Undergoer. However, in sentence (2a) the role of Experiencer is linked to the subject of the sentence (i.e. "Mary"), while in (2b), the role of Experiencer is associated with the object ("Ana"). An important aspect of this distinction is that although sentence (2b) follows the canonical SVO word order of the language (Contreras, 1976; Hernanz & Brucart, 1987; Ocampo, 1995; Suñer, 1982, among others), its prominence hierarchy follows a non-canonical order (i.e. the Theme/Undergoer precedes the more actor-like participant). Now consider a sentence like (3).

(3) A Ana le encanta María.
To Ana$_{[DAT]}$ clitic$_{[DAT]}$ love$_{[3rd-SG]}$ Mary$_{[NOM]}$
"Ana loves Mary."

The sentence realises a marked OVS word order. However, the order of arguments parallels the canonical order established by the lexico-semantic structure of the verb "love," as exemplified in (4), in which "love" is a state predicate of emotion and "x," the left-most argument of this structure, is associated to the Emoter or Experiencer, and "y" is associated to the Target, Stimulus or Theme (Van Valin Jr, 2005, p. 45).

(4) love (x,y)

This mismatch between syntactic and semantic canonical linear order is particular of a class of verbs known as dative object-experiencer psychological verbs (hereafter ObjExp psych verbs). This group of verbs will be used to test the incremental interpretation of prominence scales in the current study. In particular, we use word order and morphological case marking to test at what stage comprehenders extract prominence information that comes from these linguistic features during natural reading of SVO sentences in Spanish.

Previous evidence from this language comes from a self-paced reading study performed by Gattei, Dickey, et al. (2015) that shows that prominence information indeed plays a role on the relative cost of on-line sentence comprehension in this language. The authors found that the cost of integrating a second argument in a simple SVO sentence was increased whenever understanding the sentence required a reversal of its thematic structure. The authors explained that even in non-ambiguous simple sentences, comprehenders rapidly use morphosyntactic information (i.e. word order and case marking) to form predictions about the thematic structure of the upcoming verb and to assign a thematic role to the preverbal argument. These predictions are based on the assumption that there is a preference for more prominent arguments to precede less prominent arguments in terms of linear order (Bornkessel et al., 2005; Grewe et al., 2006; Haupt, Schlesewsky, Roehm, Friederici, & Bornkessel-Schlesewsky, 2008; Wolff, Schlesewsky, Hirotami, & Bornkessel-Schlesewsky, 2008). The appearance of a verb that required a correction of this assumption resulted in longer reading times at the region of the second argument of the sentence. Furthermore, accuracy and response times were also increased for sentences that presented a non-canonical order of their arguments, showing that the effects of not respecting the canonical order established by the thematic hierarchy (i.e. "The Actor precedes the Undergoer," Van Valin & LaPolla, 1997) are so robust for comprehension that they may persist even after all integrative processes are completed.

Nevertheless, the methodological paradigm used in this task presents three main drawbacks: First, participants must press a button to reveal each new word. This results in unnaturally slow reading times (often 400–600 ms per word), even in short -and usually skipped during natural reading- functional words. Staub & Rayner (2007) argue that this lag between actual reading time and the motor task of pressing a key "may make it difficult, if not impossible, to determine exactly when a particular factor has its effect, and it may be impossible to detect real, if short-lived, syntactic misanalyses" (Staub & Rayner, 2007, p. 334). In addition, the moving window display does not allow participants to re-read previous regions in the sentence. Thus another question that remains unanswered in Gattei et al.'s study is what readers *actually do* in order to resolve the misleading interpretation of arguments' prominence. This type of paradigm does not provide any information about what participants use the extra time they spend in difficult conditions for, or if they look for any type of linguistic information to repair the incorrect prediction. Finally, mean comprehension accuracy in this study is relatively low for SVO sentences with ObjExp psych verbs(around 60%) and OVS sentences with activity verbs (around 70%), even when they comprise grammatical sentences of the language. The authors ask whether the use of a more ecologically valid method would improve overall comprehension.

Hypotheses and predictions

The current study attempts to tap into the aforementioned methodological issues by using eye-tracking measures during natural sentence reading. As Vasishth, von der Malsburg, & Engelmann (2013) point out, "eye-tracking has the potential to inform us about when an event occurs in the parser (timing); what the parser does when it encounters difficulty (parsing events); and how attention, the eyes, and the parser interact (the eye-parser link)" (Vasishth et al., 2013, p. 125).

In accordance with previous studies, we expect object-initial structures to be less costly with ObjExp psych verbs and subject-initial sentences to be easier to process with dative-marked activity verbs. Regarding the region and point at which prominence misanalysis exerts its influence, we expect to find a dissociation between early and late eye-movement measures. Following Staub & Rayner (2007), "careful examination of the point in the eye movement record at which the effect of some linguistic manipulation first appears can be highly informative about the nature of the underlying cognitive processes involved" (Staub & Rayner, 2007, p. 329). Since we are testing cognitive processes related to the integration of syntactic and semantic information, we expect that the manipulation of prominence hierarchies has no effects at early eye-movement measures, usually related to lexical properties of words, like length and frequency (but see Kretzschmar et al., 2012, for an eye-tracking study where prominence scales referentiality and syntactic ambiguity were manipulated). Regarding late eye-movement measures, we expect that the appearance of an unexpected verb (an ObjExp psych verb in subject-initial sentences and an activity verb in object-initial sentences) leads to slower regression latencies and higher regression probabilities from the verb region onwards.

Furthermore, previous studies have shown that when readers make regressive eye movements, they do not do so randomly (Frazier & Rayner, 1982; Meseguer, Carreiras, & Clifton, 2002). Instead, readers regress to the point at which their initial, incorrect analysis diverged from the correct one. Thus, we expect a higher probability of regressing into the regions previous to the disambiguating verb in both conditions that require a thematic reanalysis.

Finally, recall that in contrast to the moving window paradigm, in this task readers have the opportunity to spend as long as they need reading the whole sentence and re-reading the regions that they find more difficult to process. We expect that by using this method, performance at the comprehension task is improved. We will also examine what readers do in order to accurately resolve arguments' interpretation.

Experiment 1

Methods

Participants

Thirty-three native Spanish speakers (20 women; age range 20–39 years; M=24.7, SE=4.32) participated in a half-hour reading experiment. All participants had normal or corrected-to-normal vision and had no history of prior neurological disease, drug or alcohol abuse, psychiatric disorders, developmental speech/language disorders, or learning disabilities. All of them provided written consent prior to the study. Twenty-eight of the participants entered the final data analysis, the remaining five having been excluded on the basis of equipment-related artefacts and/or insufficient accuracy in the comprehension task (an error rate higher than 40% in the critical conditions). All participants were compensated with 30 Argentinian Pesos after finishing the experiment.

Materials

The experimental sentences for this study were constructed on the basis of two verb lists, consisting of 24 items for each type of verbs. Verbs were selected according to the Role and Reference Grammar characterisation of activity and state predicates (Van Valin & LaPolla, 1997; Van Valin Jr, 2005 and see Gattei, Dickey, et al. 2015 for a detailed theoretical discussion on how these types of verbs entail different types of linking and processing according to their thematic hierarchization). On the one hand, we used activity verbs that marked the Actor as the nominative argument, and contained a dative Object. On the other hand, we selected state verbs denoting feelings that marked the experiencer as a dative object. Most of the psych verbs used as data (19 out of 24) have accusative Experiencer versions as well. In each case, the accusative Experiencer realisation is interpreted as causative, in contrast to the dative Experiencer realisation which does not lead to a causative interpretation (see Belletti & Rizzi, 1988; Dowty, 1991; Franco, 2013, for different approaches to psych verbs in Romance languages). In order to focus on possible reanalysis effects that could arise from thematic reassignment, we only used the dative-marked version of psych verbs that allowed both accusative and dative case marking.

The two groups of verbs were matched in length (determined by the number of letters; ObjExp psych verb verbs: M=6.8, SE=.31; Activity verbs: M=6.3, SE=.25) and Log Frequency (ObjExp psych verb verbs M=4.32, SE=.17; Activity verbs: M=4.47, SE=.11) according to the LEXESP database (Davis & Perea, 2005). An independent-samples t-test showed no significant differences between groups: Length: $t(46) = -1.35$, p>.05; Log Frequency $t(46) = .71$, p>.05.

Verbs were placed in semantically reversible sentence frames that consisted of 48 pairs of proper nouns matched for length and frequency and counterbalanced for gender. Data of names' approximate frequency was collected by conducting an advanced Google search in the domain of the Facebook website from Argentina. We reasoned that Facebook would reflect the actual frequency distribution of the proper names that participants were exposed to during the experiment. The search was conducted within the same day during the preparation of materials. The name counts (number of hits returned in the search) were logged for statistical analysis. Table 1 shows the mean, standard errors, t-scores and p-values for frequency and length of the two groups of names.

The total number of 192 experimental sentences thus constructed was divided into four lists of 48 sentences (12 per condition) so that participants would see each verb twice, each time with a different word order and framed by a different pair of names. Finally, 48 additional Prepositional Phrase (PP) or Adverbial Phrase (AdvP) were added at the end of each block of sentences. These additional phrases could be attached to both NPs when placed after the verb, and they were semantically neutral, so that they did not facilitate any semantic interpretation rather than the one provided by the role assignment required by the verb.

Table 1. Mean values, standard errors, t-values and p-values for the control variables of length and *log* frequency of the proper nouns used in the current eye-tracking experiment, according to an advanced search on google in the domain of the facebook website in Argentina.

Variable	Group	M	SE	t	p
Length	NP1	6.02	0.18	−0.08	0.93
	NP2	6.04	0.15		
Frequency	NP1	9.64	0.12	−0.02	0.98
	NP2	9.65	0.12		

Notes: NP1, first noun phrase; NP2, second noun phrase.

Table 2. Critical sentences used in the current eye-tracking study.

Condition	Example	Question example
(a) ObjExp SVO	María le encanta a Ana porque siempre fue muy amable. *"Ana loves Mary because she was always very kind."*	¿Es María quien le encanta a alguien?
(b) ObjExp OVS	A María le encanta Ana porque siempre fue muy amable. *"Mary loves Ana because she was always very kind."*	*Is it Mary who is loved by someone?*
(c) Act SVO	María le responde a Ana porque siempre fue muy amable. *"Mary responds to Ana because she was always very kind."*	¿Es María quien le responde a alguien?
(d) Act OVS	A María le responde Ana porque siempre fue muy amable. *"Ana responds to Mary because she was always very kind."*	*Is it Mary who responds to someone?*

Notes: ObjExp, object experiencer psych verbs; act, activity verbs; SVO, subject-verb-object, OVS, object-verb-subject.

In addition, a set of three practice trials and 145 sentences that belonged to two other different experiments were used as fillers. These sentences were different in syntactic complexity and length than the critical trials, so that participants would not be aware of the aim of the experiment.

One question for each practice item, critical trial and 55 of the filler sentences was prepared to test comprehension. The questions were formulated in such a way that the participants had to judge whether it correctly described the content of the preceding experimental sentence or not. Half of the questions required the answer "yes" and half of them required the answer "no." In the case of critical items, half of the questions involved the subject of the sentence and half of them asked about the object. Table 2 shows an example of the critical sentences used in the current experiment. A complete list of the experiment materials may be found at Appendix A.

Equipment

Participants were seated in front of a 19-inch screen (Samsung SyncMaster 997MB, 1024 × 768 pixels resolution, 100 Hz refresh rate) at a viewing distance of 65 cm. A chinrest aligned with the centre of the screen prevented head movements. An EyeLink 1000 eye-tracker (SR Research Ltd.) was used to record gaze locations of both eyes during reading at a sampling rate of 1 kHz Nominal average accuracy was 0.5° and space resolution was 0.01° root mean square (RMS), as given by the manufacturer. The participant's gaze was calibrated with a standard 13-point grid for both eyes. All recordings and calibration were binocular. Only left eye data were used for the analysis.

All eye movements were labelled as fixations, saccades and blinks by the eye-tracker software using the default thresholds for Cognitive experiments (30°/sec for velocity, 8000°/sec for acceleration, and 0.1° for motion) (Cornelissen, Peters, & Palmer, 2002). Stimuli presentation was developed using Matlab (http://www.mathworks.com/, Massachusetts, United States) and Psychophysics Toolbox Version 3.

Procedure

All sentences were displayed on a single line and were presented in Courier New Bold font. At a distance of 65 cm, each letter subtended 0.44° of visual angle laterally. Subjects were instructed to read the sentences at their own rate. No instructions were given to suppress eye blinks. Before the eye-tracking experiment began, they had a practice session of six sentences. At the beginning of each trial, a dot appeared at the left edge of the screen and after participants fixated on this dot, the sentence appeared. The first letter of the sentence was located at the position of the dot. Participants were instructed to look at a second dot at the bottom right corner of the screen to indicate they had finished reading. The total reading time of each trial was measured starting from when participants triggered the appearance of the sentence by fixating on the left dot until they fixated on the bottom right dot and the sentence disappeared. Comprehension questions appeared after every critical stimuli and after 38% of the filler sentences. Participants responded by mouse-clicking on one of two possible answers ("Yes" or "No") displayed horizontally. Half of the times, the correct answer was positioned over the left. The order in which it appeared at this position was randomised. Response time was measured starting from the appearance of the question until participants clicked on one of the possible responses. A calibration procedure was performed at the beginning of the eye-tracking experiment. Experimental sessions lasted approximately 30 min.

Data analysis

Eye movement data from the 33 participants was screened for blinks and track losses. Fixations shorter than 50 ms and longer than 1000 ms were removed from the analysis. After this screening process, fixations were assigned to their respective word. Eye-tracking measures were computed using em2 package for R language for statistical computing (Logacev & Vasishth, 2013, version 3.0.2).

Table 3. Regions of interest used for the statistical analysis of the current eye-tracking experiment according to word order.

	1	2	3	4	5	6	7	8
SVO		María	le	responde|encanta	a	Ana	PP1	PP2
		Mary	*clitic$_{[DAT]}$*	*responds|loves*	*to*	*Ana$_{[DAT]}$*	*PP1*	*PP2*
OVS	A	María	le	responde|encanta		Ana	PP1	PP2
	To	*Mary$_{[DAT]}$*	*clitic$_{[DAT]}$*	*responds|loves*		*Ana*	*PP1*	*PP2*

Notes: SVO, subject-verb-object; OVS, object-verb-subject; PP1, first word of the prepositional phrase; PP2, second word of the prepositional phrase.

For the purpose of analysis, we divided the sentences into eight regions that consisted of the first eight words of each sentence, as shown in Table 3. Note that in order to facilitate statistical analysis and visual presentation of the results, we aligned the critical regions that comprised the proper names (regions 2 and 6), the clitic (region 3) and the verb (region 4). The region of the preposition has been labelled as (5) in subject-initial sentences, and (1) in object-initial sentences. The regions "PP1" and "PP2" correspond to the first and the second word of the prepositional phrase following the second noun phrase.

For each fixated word, we computed the following measures: (1) *First Fixation Duration* (FFD; the duration of the first fixation on the word); (2) *First Pass Reading Time* (FPRT; the sum of all fixation durations on the word before any other word was fixated); (3) *Right-Bounded Reading Time* (RBRT; the sum of all first-pass fixation durations on the word before another word to the right is fixated); (4) *Regression Path Duration* (RPD; also known as *go-past time*, it is the sum of all first-pass fixation durations on the word and all preceding words in the time period between the first fixation on the word up to the point where the reader leaves the critical region with a progressive saccade; (5) *Right-Bounded Regression Count* (RBRC; the number of regressions from the word before any word further to the right has been fixated); (6) *Total Fixation Time* (TFT; the sum of all fixations durations on a word); and (7) *Total Incoming Regressions* (TIR; the number of regressions to a specific word). Measures 1–3 are typically considered early measures, whereas measures 4–7 are assumed to reflect later processing stages (see Clifton, Staub, & Rayner, 2007; Vasishth et al., 2013, for a review on this discussion).

The data analysis was conducted in the R programming environment (R Core Team, 2013). For measures comprising reading or response time (i.e. Comprehension Task Response Time, FFD, FPRT, RBRT, and TFT) a linear mixed-effects model was used. Linear mixed-effects models are available in the package lme4 (Bates, Mächler, Bolker, & Walker, 2014; Pinheiro & Bates, 2000). For the accuracy measure, the data was fit to a generalised linear mixed-effects model with a binomial function. This function is adequate for analysing data measured on a dichotomous scale, namely "Correct" and "Incorrect" response. Count data (RBRC and TIR), on the other hand, was analysed with a generalised mixed-effects model with Poisson link function, which is appropriate for counts of events in a fixed time window (Baayen, 2008, p. 322).

All regression models included Verb Type and Word Order as fixed effects and Subject, and Item as random effects. Log Frequency and inverse length of each word were included as factors in every region except for regions 1 and 5 (preposition "a") and region 3 (clitic), the rationale being that these two variables may explain a significant part of the variability in reading times and amount of fixations on these regions (Just & Carpenter, 1980; Kliegl, Grabner, Rolfs, & Engbert, 2004; Rayner & Raney, 1996).

A maximal random-effects structure was included in both LMMs and GLMMs, as linear mixed-effects models that do not consider random intercepts and slopes involve the risk of Type I error inflation (Barr, Levy, Scheepers, & Tily, 2013). When models either did not converge or the correlation between variance components could not be estimated, the random effects structure was simplified by removing the correlations. For large samples, the t distribution approximates the normal distribution and an absolute value of t larger than 2 indicates a significant effect at $\alpha = 0.05$. For all the models presented in the study, covariates that involved reading times were scaled and centred.

Finally, we used an orthogonal contrast coding to test the interaction between verb type and word order. For the verb type contrast, sentences with activity verbs were coded as −1 and sentences with ObjExp psych verbs were coded as 1. For the word order contrast, SVO sentences were coded as −1 and OVS sentences were coded as 1.

Results

Comprehension task

Question accuracy. Mean accuracy for all comprehension questions was 88.57%. This indicates that participants were paying attention to the content of the sentences. Mean accuracy of critical sentences was 90.12%. Figure 1(a) shows mean accuracy according to condition. Differences in accuracy according to verb type and word order were analysed with a generalised linear mixed-

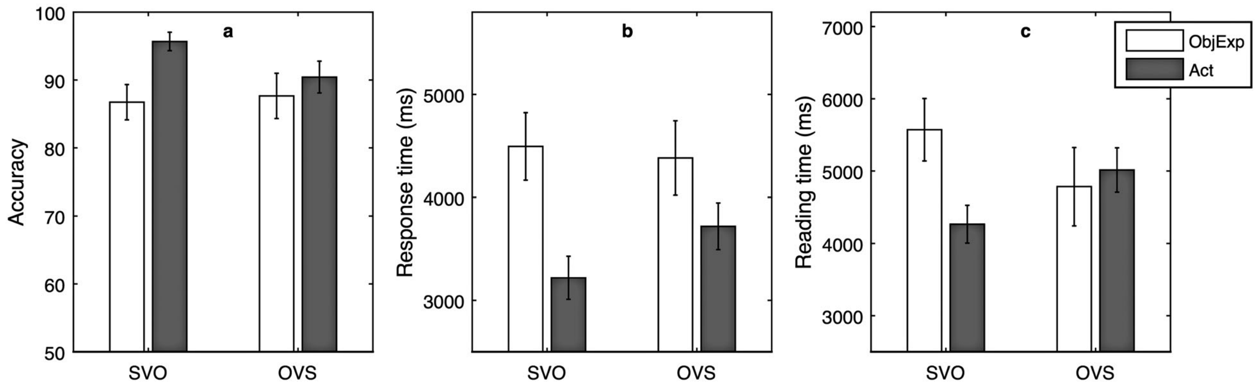

Figure 1. Percentage of accurate answers (panel a), mean response times for the comprehension question (panel b), and mean total reading time for the critical sentences (panel c) in the current eye-tracking study according to verb type (ObjExp psych verb vs. Act) and word order (SVO vs. OVS). Error bars correspond to Standard Error of the Mean. ObjExp psych verb, object experiencer psych verb; Act, activity verb; SVO, subject-verb-object; OVS, object-verb-subject.

effects model. The analysis revealed a significant effect of verb type. On average, participants responded more accurately after reading sentences with activity verbs ($M = 93.15\%$) than after reading sentences with ObjExp psych verbs ($M = 86.16\%$); $\beta = -0.424$, *SE*=0.161, *z*=2.63. An interaction between verb type and word order was found; $\beta = 0.256$, *SE*=0.112, *z*=2.29. Resolving this interaction revealed that accuracy was significantly higher for questions about sentences with Activity verbs and SVO word order than for the other three conditions (ActSVO - ActOVS; $\beta = 0.9271$;, *SE*=0.357, *z*=2.59, *p*=0.044; ActSVO - ObjExpSVO; $\beta = 1.361$, *SE*=0.418, *z*=3.255, *p*=0.006; ActSVO - ObjExpOVS; $\beta = 1.263$, *SE*=0.420, *z*=3.004, *p*=0.013).

Response times. Figure 1(b) shows mean response times (RTs) according to condition. Analysis of differences in RTs between verb type and word order revealed a main effect of verb type. On average, response times were significantly higher for questions about sentences with ObjExp psych verbs (*M*=4025 ms) than for questions about sentences with activity verbs (*M*=3075 ms); $\beta = 0.102$, *SE*=0.016, *t*=6.242. Interactions between both factors verb type and word order were significant; $\beta = -0.0369$;, *SE*=0.013, *t*=−2.839. Tukey post-hoc test revealed that participants answered comprehension questions for sentences with activity verbs significantly faster when sentences followed the SVO word order (*M*=3856 ms) than when they followed OVS word order (*M*=4053 ms); $\beta = -0.111$, *SE*=0.037, *z*=−3.023, *p*=0.013. However, this difference between OVS and SVO sentences was not significant for comprehension questions for sentences with ObjExp psych verbs; $\beta = -0.036$, *SE*=0.037, *z*=−0.992, *p*=0.752.

Total reading time. Figure 1(c) shows the average total reading time for the critical sentences used in the current eye-tracking experiment. The statistical analysis revealed an interaction between verb type and word order; $\beta = -0.090$, *SE*=0.013, *t*=−6.962. Resolving this interaction showed that participants spent significantly longer time reading sentences with activity verbs when they followed the OVS word order than when they followed the SVO order; $\beta = 0.170$, *SE*=0.036, *z*=4.654, *p*<0.001. Conversely, sentences with ObjExp psych verbs were read significantly faster when they contained OVS word order; $\beta = 0.189$, *SE*=0.036, *z*=5.192, *p*<0.001. Verb type also affected reading times of sentences with different word orders. The multiple comparisons test also revealed that when reading subject-initial sentences, participants spent significantly longer time reading sentences with ObjExp psych verbs than sentences with activity verbs; $\beta = 0.234$, *SE*=0.040, *z*=5.894, *p*<0.001. Reading of OVS sentences followed the opposite pattern: Participants took significantly longer time to read sentences with activity verbs than sentences with ObjExp psych verbs; $\beta = 0.125$, *SE*=0.040, *z*=3.145, *p*=0.008.

Eye-tracking measures

Following Kliegl, Nuthmann, & Engbert (2006), we excluded the first word of all sentences for all reading duration analysis since the first word was positioned at the location of the fixation spot that triggered the appearance of the sentence. For subject-initial sentences, the first word was labelled as "Region 2" and for object-initial sentences, the first word was labelled as "Region 1." We thus inspected the reading measures FFD, FPRT, RBRT, RBRC and RPD from Region 3 onwards. However, at Region 2, we did analyse the probability of regressing to the first word from other regions and the total fixation time.

Figure 2 summarises the contrast between sentences with activity verbs and sentences with ObjExp

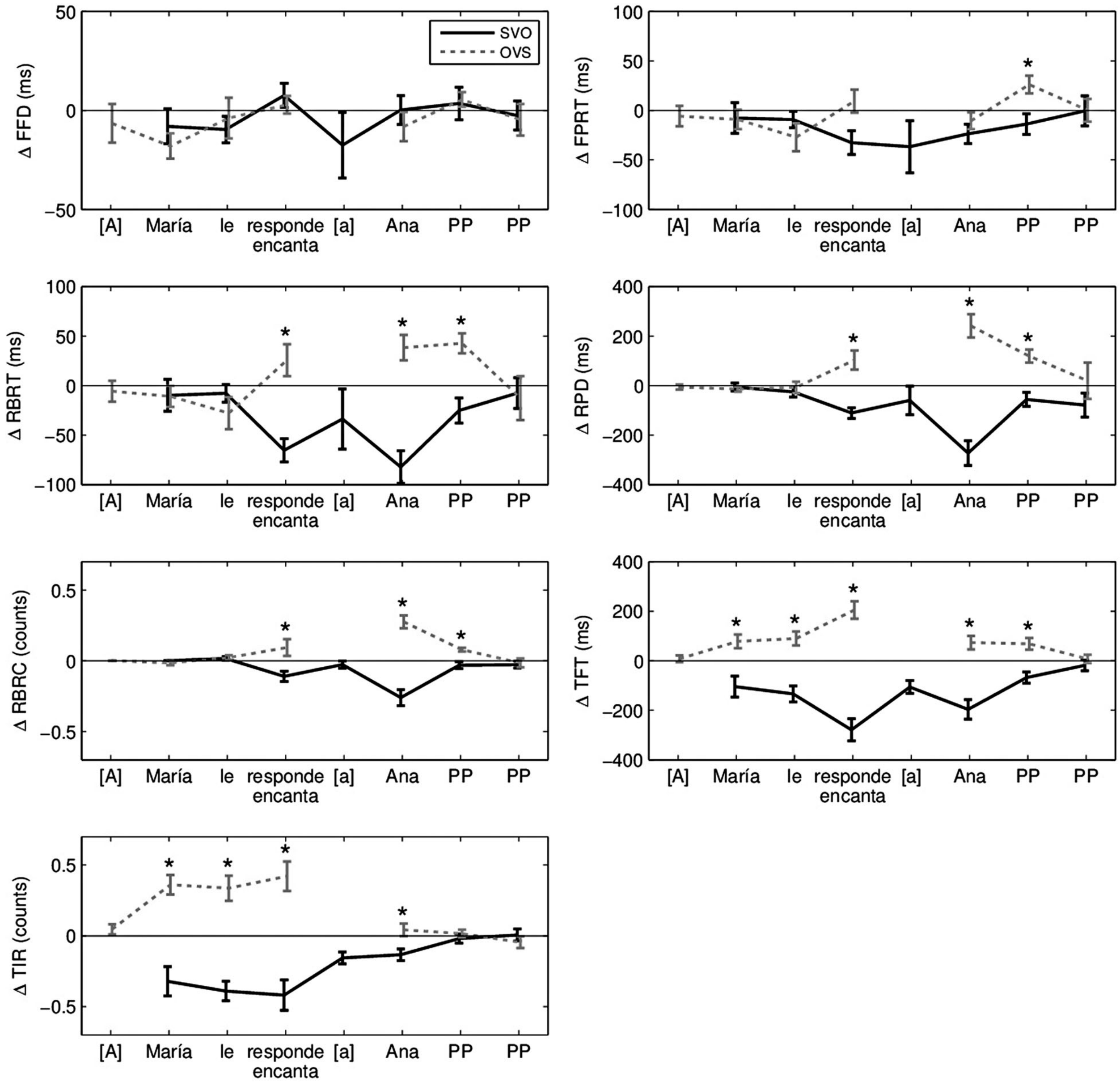

Figure 2. The Figure shows the difference (Δ) in mean fixation times (ms) and the amount of regressive saccades (counts) between conditions with Activity Verbs and conditions with ObjExp psych verbs according to the sentence word order. Error bars correspond to Standard Error of the Mean. Eye-tracking measures: FFD, first fixation duration; FPRT, first pass reading time; RBRT, right-bounded reading time; TFT, total fixation time; RPD, regression path duration; RBRC, right-bounded regression count; TIR, total incoming regressions. Word Order: SVO, subject-verb-object; OVS, object-verb-subject. The asterisk shows that the interaction between Word Order and Verb Type was significant.

psych verbs according to both word orders (SVO and OVS). Positive values mean that reading time and regression counts are higher for sentences with activity verbs than for sentences with ObjExp psych verbs. A positive grey dashed line (OVS conditions) and a negative black line (SVO conditions) correspond to an interaction between Verb Type and Word Order. This representation makes the interaction visually clear. The asterisks show the regions where the interaction was significant.

We now provide the analysis of regions of interest for both the early and late measures mentioned in the "Data analysis" section. For each region, we first present the analysis of the interaction between Verb Type and Word Order, since it is the contrast of interest of the current study. We then provide the relevant results of the multiple comparisons test whenever was needed. Finally, we report main effects of Verb Type or Word Order. Appendix B shows the final converging models for each measure at each region. A table with all statistical coefficients, standard errors and t values may be found at Appendix C.

Region 1 (Case marking preposition "a"). Analysis of the probability of regressions into this region ("TIR") revealed no significant differences between the probability of regressing to this area after reading an activity verb or after reading an ObjExp psych verb; $\beta = -0.534$, SE=0.377, z=−1.416.

Region 2 (First proper name). Analysis of late eye-movement measures showed a significant interaction between Verb Type and Word Order for Total Fixation Time (TFT) and Total Incoming Regressions (TIR); TFT: $\beta = -0.079$, SE=0.015, t=−5.095; TIR: $\beta = -0.228$, SE=0.033, z=−6.974). Resolving these interactions revealed that for subject-initial sentences, the probability of regressing to this region and the total fixation time are significantly higher when the sentence contains an ObjExp psych verb than when it contains an activity verb (TFT: $\beta = 0.148$, SE=0.045, z=3.304, p=0.005; TIR: $\beta = 0.493$, SE=0.099, z=4.967, p<0.001). Conversely, for object-initial sentences, the probability of regressing into this region and the total fixation time are significantly higher when the sentence contains an activity verb than when it contains an ObjExp psych verb (TFT: $\beta = 0.174$, SE=0.045, z=3.883, p<0.001; TIR: $\beta = 0.417$, SE=0.085, z=4.921, p<0.001).

Analysis of the probability of regressions into this region (TIR) also revealed a main effect of word order. On average, participants regressed to this word significantly more in object-initial sentences than in subject-initial sentences; $\beta = 0.155$, SE=0.033, z=4.710.

Region 3 (Clitic). Analysis of late eye-movement measures showed a significant interaction between verb type and word order for total fixation time, and for the probability of regressions into this region (TFT: $\beta = -0.128$, SE=0.020; t=−6.515; TIR: $\beta = -0.281$, SE=0.035, z=−7.945). This interaction follows the same direction as the interaction found on Region 2. For subject-initial sentences, participants regressed and fixated on this region significantly more when the sentence contained an ObjExp psych verb than when it contained an activity verb (TFT: $\beta = 0.147$, SE=0.044, z=3.304, p=0.005; TIR: $\beta = 0.592$, SE=0.099, z=5.967, p<0.001). In object-initial sentences; participants fixated and regressed to this region significantly more when the sentence contained an activity verb (TFT: $\beta = 0.174$, SE=0.004, z=3.883, p<0.001; TIR: $\beta = 0.531$, SE=0.101, z=5.274, p<0.001).

Region 4 (Disambiguating verb). Analysis of this region showed an interaction between verb type and word order for all late eye-movement measures (RPD: $\beta = -0.075$, SE=0.015, t=−4.897; RBRC: $\beta = -0.173$, SE=0.051, z=−3.380; TFT: $\beta = -0.132$, SE=0.0209, t=−6.318; TIR: $\beta = -0.247$, SE=0.031, z=−7.992). Tukey post-hoc test showed that in subject-initial sentences, participants fixated and regressed significantly more times to this region when they read an ObjExp psych verb than when they read an activity verb after the clitic (TFT: $\beta = 0.308$, SE=0.057, z=5.353, p<0.001; TIR: $\beta = 0.464$, SE=0.108, z=4.308, p<0.001). For object-initial sentences, participants fixated this region and regressed to previous regions before continuing reading for longer time, when they longer time, when they read an activity verb than when they found an ObjExp psych verb (RPD: $\beta = -0.130$, SE=0.040, z=−3.298, p=0.005; TFT: $\beta = 0.343$, SE=0.057, z=5.967, p<0.001). This difference followed the same direction for measures RBRC and TIR, showing that participants regressed significantly more times from and to this region when they read an activity verb instead of an ObjExp psych verb (RBRC: $\beta = -0.400$, SE=0.151, z=−2.639, SE=0.111, z=−4.844, p<0.001).

An interaction between Verb Type and Word Order was also found for early measure RBRT, $\beta = -0.052$, SE=0.0138, t=−3.807. However, the multiple comparisons test revealed that there were no significant differences between SVO conditions according to the verb type nor between OVS conditions, p>.1.

Region 5 (Case marking preposition "a"). Analysis of this region showed a significant main effect of Verb Type for Total Incoming Regressions; $\beta = 0.267$, SE=0.074; z=3.631. On average, participants regressed significantly more times to this region when the sentence contained an ObjExp psych verb than when it contained an activity verb.

Region 6 (Second proper name). Analysis of this region showed an interaction between Verb Type and Word Order for all late eye-movement measures (RPD: $\beta = -0.178$, SE=0.019, t=−9.186; RBRC: $\beta = -0.353$, SE=0.047, z=−7.556; TFT: $\beta = -0.124$; SE=0.0152; t=−8.141; TIR: $\beta = -0.155$, SE=0.053, z=−2.923). Resolving these interactions revealed that in SVO sentences, participants fixated for longer time at this and previous regions and regressed significantly more times from and to this region when the sentence contained an ObjExp psych verb than when it included an activity verb (RPD: $\beta = 0.380$, SE=0.055, z=6.867, p<0.001; RBRC: $\beta = 0.671$, SE=0.130, z=5.163, p<0.001; TFT:

$\beta = 0.343$, *SE*=0.043, *z*=7.914, *p*<0.001; TIR: $\beta = 0.456$, *SE*=0.146, *z*=3.124, *p*=0.009).

In OVS sentences, participants spent longer time on this and previous regions and regressed significantly more times from this region when the sentence contained an activity verb than when it contained an ObjExp psych verb. This difference was not significant for the probability of regressing into this region (RBRC: $\beta = 0.740$, *SE*=0.134, *z*=5.519, *p*<0.001; RPD: $\beta = 0.332$, *SE*=0.054, *z*=6.118; TIR: $\beta = -0.165$, *SE*=0.154, *z*=−1.069, *p*=0.707).

Analysis of early measures showed a significant interaction between Verb Type and Word Order for RBRT, $\beta = -0.074$, *SE*=0.013, *t*=−5.768. The multiple comparisons test showed that in SVO sentences, participants fixated for significantly longer time before continuing reading when the sentence contained an ObjExp psych verb, than when they had previously read an activity verb, $\beta = 0.214$, *SE*=0.0368, *z*=5.805, *p*<0.001. However, this difference was not significant for OVS sentences, $\beta = 0.084$, *SE*=0.0361, *z*=2.318, *p*=0.09.

Analysis of both early and late eye-movement measures also revealed main effects of Verb Type and Word Order. On average, participants spent longer time reading this region during first-pass reading (FPRT) when the sentence contained an ObjExp psych verb and followed the SVO word order (Verb Type: $\beta = 0.031$, *SE*=0.012, *t*=2.589; Word Order: $\beta = -0.038$, *SE*=0.012, *t*=−3.177). The multiple comparisons test shows that the word order effect is motivated by the longer reading time for SVO sentences with ObjExp psych verbs when compared to both OVS conditions (ObjExpSVO - ActOVS: $\beta = 0.139$, *SE*=0.034, *z*=4.081, *p*<0.001; ObjExpSVO - ObjExpOVS: $\beta = 0.089$, *SE*=0.034, *z*=2.643, *p*<0.04). The same effects were found for Right-bounded Reading Time and Total Fixation Time (RBRT Verb: $\beta = 0.033$, *SE*=0.013, *t*=2.522; Word Order: $\beta = -0.036$, *SE*=0.013, *t*=−2.839; TFT Verb: $\beta = 0.047$, *SE*=0.015, *t*=3.106; Order: $\beta = -0.053$, $\beta = -0.053$, *SE*=0.015, *t*=−3.443).

Region 7 (First word of the Spill-over region). Analysis of this region showed that the interaction between Verb Type and Word Order was significant for almost all late eye-movement measures and two early eye-movement measures (FPRT: $\beta = -0.028$, *SE*=0.011, *t*=−2.438; RBRT: $\beta = -0.050$, *SE*=0.0124, *t*=−3.967; RPD: $\beta = -0.089$, *SE*=0.019, *t*=−4.792; RBRC: $\beta = -0.367$; *SE*=0.104, *z*=−3.547; TFT: $\beta = -0.071$; $SE = 0.016$,t=-4.352). The multiple comparisons test showed that in subject-initial sentences, participants fixated longer at this region during first-pass reading and right-bounded reading when the sentence contained an ObjExp psych verb than when it included an activity verb (FPRT: *β*=0.1036, *SE*=0.0322, *z*=3.224, *p*=0.007; RBRT: $\beta = 0.214$; *SE*=0.038, *t*=5.596, *p*<0.001). Regression path duration and total fixation time were marginally higher for sentences with ObjExp psych verbs than for sentences with activity verbs (RPD: $\beta = 0.133$, *SE*=0.053, *z*=2.498, *p*=0.06; TFT: $\beta = 0.108$; *SE*=0.047, *z*=2.309, *p*=0.09).

In object-initial sentences, readers spent marginally longer time at this and significantly longer time at previous regions when the sentence contained an activity verb than when there was an ObjExp psych verb (FPRT: $\beta = -0.076$, *SE*=0.032, *z*=−2.393, *p*=0.078; RPD: $\beta = 0.222$, *SE*=0.052, *z*=4.301, *p*<0.001; TFT: $\beta = -0.178$, *SE*=0.046, *z*=−3.862, *p*<0.001). The analysis of regression counts (RBRC) showed that participants also regressed significantly more times from this sentences when it contained an activity verb than when it included an ObjExp psych verb; $\beta = 1.09$; *SE*=0.314, *z*=3.473, *p*=0.002. This difference was not significant for right-bounded reading time; $\beta = -0.084$; *SE*=0.037, *z*=−2.239, *p*=0.11.

Analysis of this region also showed a main effect of word order for both early measures, namely, FFD and FPRT and a few late measures. Participants fixated this region significantly longer and regressed to other regions before continuing reading for longer time when they read SVO sentences than when they read OVS sentences (FPRT: $\beta = -0.024$, *SE*=0.011, *t*=−2.114; FFD: $\beta = -0.031$, *SE*=0.009, *t*=−3.326; RBRT: $\beta = -0.033$, *SE*=0.014, *t*=−2.403; RPD: $\beta = -0.040$, *SE*=0.019, *t*=−2.160).

Region 8 (Second word of the Spill-over Region). No main effects nor interactions were found at this region.

Discussion

The purpose of the current study was to shed light on how different types of linguistic information are integrated during incremental parsing of Spanish sentences that required a reversal of thematic role assignment. A substantial amount of evidence from studies run in German (Bader & Bayer, 2006; Bornkessel, Schlesewsky, & Friederici, 2002; Bornkessel et al., 2003, 2005; Bornkessel-Schlesewsky & Schlesewsky, 2009; Frisch & Schlesewsky, 2001) and English (Weckerly & Kutas, 1999) has shown that the parser does not wait until verbal information is available in order to make predictions about the thematic structure of the verb to come. These predictions are based on morphosyntactic (i.e. case morphology, word order) and semantic cues (i.e. animacy, definiteness) provided by the arguments of the sentence. Cross-linguistic evidence suggests that the grade of relevance that each cue has for triggering these predictions may vary from language to language. Spanish is a language in which interpretation of prominence information depends on both word order and case morphology. In a previous self-paced reading study run in this language, Gattei, Dickey, et al. (2015) found that participants slowed down at the second argument of declarative sentences in which the order of semantic arguments did not parallel the hierarchically ordered arguments of the lexical-semantic structure of the verb. The authors also found that answering questions about subject-initial sentences was significantly more difficult than answering questions about object-initial sentences. This difference was triggered by the response accuracy at SVO sentences with ObjExp psych verbs, which was around 60%, very low for participants who were reading grammatical sentences. Two main issues remained unanswered from that experiment: First, due to the characteristics of the method used, it is not possible to disentangle at which region in the sentence and at which stage during reading, prominence exerts an influence over sentence processing in this language. Second, the self-paced reading method does not allow us to understand what participants actually *do* in order to retrieve the correct interpretation of the sentence. Recall that the moving-window paradigm impedes participants from reading previous words on the sentence. Hence, the results of the comprehension task of Gattei et al.'s study reflect what participants could achieve with the information provided by a single left-to-right reading, which is not the way people naturally read. Evidence has shown that even when reading simple sentences regressions to previous words, longer fixations on infrequent or long words and skipping functional and short words take place (Brysbaert, Drieghe, & Vitu, 2005; Rayner & McConkie, 1976). Both problems were tapped into by resorting to eye-tracking, a more ecological method for analysing reading strategies.

Results of Experiment 1 show that arguments' degree of prominence exerts its influence with the appearance of a single argument. Participants perceived that there was a mismatch between the morphosyntactic and semantic information of the argument that they had already read and the lexico-semantic representation of the verb as soon as they read the critical verb. When reaching this region, participants fixated their gaze longer and regressed more to previous regions when word order did not coincide with the canonical arguments order of the verb's lexical-semantic representation. Apart from supporting the account proposed by Gattei, Dickey, et al. (2015), this finding shows that with only one piece of syntactic and semantic information – as it is the case for OVS sentences – the parser establishes the relative degree of prominence of the initial argument and predicts the type of verb that will appear in the sentence. Recall that most of the studies done on this issue come from SOV languages, or languages that allow this construction (but see Dröge et al., 2014, for an experiment in Italian), and most of them used unnatural reading paradigms like the word-by-word visual presentation, frequently used for most event-related potentials (ERPs) studies. Results of the current experiment support the findings provided by Gattei, Dickey, et al. (2015) and comprise a good piece of evidence on how greedy (Crocker, 2012) the parser is when processing arguments' prominence.

These results are also informative about the temporal dynamics related to the processing of prominence information. Analyses of the different eye-movement measures have shown a dissociation on the degree and the region of influence of each of these measures during sentence reading. On the one side, early measures (FFD and FPRT) did not reflect any additional cognitive cost related to the repair of wrongly computing an argument's prominence status, neither at the verb's region nor at the regions of the second argument. These measures, and particularly FPRT, were only affected at the first region of the Prepositional Phrase, this is, once the sentence argument structure had already been confirmed. This is not an unexpected result, since the modulation of early eye-movement measures has been related to superficial aspects of words like frequency or length (Staub & Rayner, 2007), and both factors have been controlled for during the experimental design of this study. However, the modulation of FPRT once the syntax-to-semantics *linking* had been confirmed shows that this measure may be affected by the integration of more complex linguistic information, as it occurs when the parser tries to retrieve the meaning of an event with inverse linking.

Two effects that need to be taken into account, though, are the effects of verb type and word order found at Region 6 (Second proper name) for FPRT. These are indeed important effects, also discussed in Gattei, Dickey, et al. (2015). As of the verb type effect, recall that verbs in this experiment only differ in the way that their semantic content is syntactically realised, that is, in the way syntax-to-semantic linking is accomplished. While in activity verbs the nominative argument bears the "Actor" macrorole, in ObjExp psychological verbs, the nominative argument is assigned the "Undergoer" macrorole (Van Valin & LaPolla, 1997; Van Valin Jr, 2005). Our hypothesis is that the effect of verb type is the consequence of the computational cost of constructing the conceptual representation of events where only the "Undergoer" macrorole is assigned, and not the "Actor," as in sentences with ObjExp psych verbs. As evidenced in the response times in the comprehension task, difficulty still persists after most integrative processes are complete, also when using a more ecological paradigm as it is eye-tracking.

The word order effect seems motivated by the long reading time of this region in SVO sentences with ObjExp psych verbs in comparison to both OVS conditions (as evidenced in a multiple comparisons test). The fact that no difference takes place when SVO or OVS conditions are compared between each other, suggests that there is something particular about SVO sentences with ObjExp psych verbs. Subject-initial sentences with ObjExp psych verbs contain a cross-over linking between the thematic roles and the lexical-semantic structure of the verb. Object-initial sentences with activity verbs also use this type of linking. However, the integrative processes for comprehension of SVO order in ObjExp psych verbs require reanalysing a nominative argument into Undergoer. In contrast, reanalysis of OVS sentences with activity verbs involves the assignment of the Actor macrorole to a nominative constituent. The latter linking "Actor/nominative" is the typical linking in Spanish, whereas the "Undergoer/nominative" pairing is more marked (Arnaiz, 1998; Bakovic, 1998; Gutiérrez-Bravo, 2007). The multiple comparisons test indeed shows that there is a significant difference in first-pass reading time between both types of sentences. Since this effect was only found once the lexico-semantic structure of the event was complete and not at the verb region (when participants can already realise that the sentence denotes a different type of event than the one predicted) we hypothesise that the effect is a reflex of a later stage of processing, that is posterior to the computation of prominence itself.

On the other hand, late eye-movement measures show that the interaction between Verb Type and Word Order follows the direction of the results found in Gattei, Dickey, et al. (2015). When the canonical order of the arguments in the verb's lexical-semantic structure did not coincide with the morphosyntactic features of the argument presented as input, there was a higher probability that participants regressed more and spent longer time in previous regions. They also spent longer time reading the disambiguating verb before continuing reading. Finding these effects for late eye-movement measures also replicates previous findings from an eye-tracking study with similar sentences in German (Kretzschmar et al., 2012). It also comprises evidence in favour of the computation of prominence as one of the factors that the parser takes into account during sentence processing. Evidence in studies with ERPs shows that this computation exerts its influence at late stages of processing (between 400 and 900 ms: Dröge et al., 2014; Frisch & Schlesewsky, 2001; Gattei, Tabullo, et al. 2015; Roehm, Schlesewsky, Bornkessel, Frisch, & Haider, 2004; Weckerly & Kutas, 1999). In particular, two studies investigated this issue in a Romance language. Dröge et al. (2014) found that subject-initial sentences with ObjExp psych verbs and object-initial sentences with activity verbs elicited a P600 component that reflected, among other possibilities, a conflict monitoring or a well-formedness mismatch due to the thematic mismatch. Furthermore, the authors found that in object-initial sentences, the appearance of an activity verb elicited a biphasic N400 component as well. The authors explained the N400 effect as the interplay of two different factors: the thematic hierarchy and lexical predictability, and they argued that, in contrast to the high predictability for a dative object experiencer verb following an initial dative, an initial subject does not lead the processing system to anticipate a dative active verb. Hence, the lower degree of lexical predictability leads to an N400 increase in SVO sentences with activity verbs, thus masking the thematic N400 effect in SVO sentences with Object experiencer verbs.

Gattei, Tabullo, et al. (2015) used almost identical sentences to the ones used in the current experiment in order to investigate the neural correlates involved in the computation of prominence in Spanish. The authors also explored the role of arguments' span in the computation of prominence in this language, an aspect that had not been studied by the Italian study. All verbs used in the experiment required the use of a dative clitic immediately before the verb. This clitic either introduced a new piece of argument information (as in SVO sentences), or replicated the information of the initial argument (as in OVS sentences). Results showed that the appearance of the critical verb triggered different effects according to the sentence word order. When participants were shown a subject-initial sentence, the appearance of an ObjExp psych verb engendered a widely distributed positivity with a peak at 600 ms. (P600). Conversely, when participants read an object-initial sentence, the appearance of an Activity verb triggered a centro-parietal negativity with a peak at 400 ms. (N400). A cloze task showed that there was a qualitative difference in the type of responses given for each word order. In almost every case, participants completed subject-initial constructions with an Activity verb, although there was a great variability in the amount of responses given (N=58). On the other hand, half of object-initial constructions were completed with an inflected form of the verb "gustar" ("*to like*"). The rest of the responses (*N*=36) were divided among different types of verbs, from ObjExp psych verbs, to activity and state verbs. All events retrieved followed also a thematic structure that began with an argument with the highest prominence status. These results suggest that in the ERP study participants predicted a verb that ranked the first argument as the most prominent, thus triggering a late event related potential when this prediction was not correct. Crucially, they also showed that the amount of morphosyntactic information related to the argument structure of the event that was available before the verb appeared would modulate the level of uncertainty regarding either the thematic structure of the verb (more uncertainty for OVS sentences than for SVO sentences), or the lexical piece expected (more uncertainty for SVO constructions than for the OVS ones). The authors explained that the N400 could partly be accounted for by the thematic reanalysis required when the verb did not match the prominence given to the first argument, but also, by the cost of integrating of a lexical item different to the one expected. Conversely, the elicitation of the P600 effect in SVO sentences could be grounded on the appearance of a verb that required prominence reversal, but which was also thematically incompatible with the participants' expectations (an ObjExp psych verb instead of an Activity verb).

Although it is hard to establish absolute timing between the time course of the different effects found at both the current eye-tracking study and the two aforementioned ERP studies (see Pulvermüller, Shtyrov, & Hauk, 2009; Sereno & Rayner, 2003, for discussions on the incompatibility of the time course of different effects related to word recognition at each methodology), the current experiment shows that the effects related to prominence miscomputation become more evident when analysing late eye-movement measures.

On a similar note, the current experiment cannot disentangle whether the effects found in the different eye-movement measures are a reflex of difficulties resulting from prediction mismatches regarding the lexical or thematic structure of the verb as opposed to a possible thematic reanalysis occurring once the verb is encountered. First, prediction mismatches and reanalysis tasks have been shown to modulate the same eye-movement measures. Evidence reports an increase on first pass reading time, and on amount and duration of regressions when the sentence requires a syntactic reanalysis (Frazier & Rayner, 1982; Kretzschmar et al., 2012; Meseguer et al., 2002; von der Malsburg & Vasishth, 2013), but also when there is a mismatch between the word predicted and the word actually read (Ashby, Rayner, & Clifton, 2005; Balota, Pollatsek, & Rayner, 1985; Rayner, Ashby, Pollatsek, & Reichle, 2004). We do not mean, however, that differences between both processes could not be measured by an eye-tracking experiment, but in order to do this, there should be an appropriate manipulation of materials that took into account the cloze probability of each verb after the appearance of a certain noun phrase or prepositional phrase. This was not part of the aims of the current experiment, but it is matter of future research to unravel the eye-movement characteristics related to both prediction mismatch and thematic reanalysis processes during the computation of prominence (see Staub & Clifton Jr, 2006, for a discussion on eye-movement effects arising from lexical predictability and not from syntactic reanalysis even when reanalysis is needed).

Results of Experiment 1 have also shown what participants actually *do* when they need to solve the task of retrieving "who did what to whom" correctly. When higher cost of processing takes place during reading, observing the regressions *into* a region (TIR) may be informative of what type of hypothesis the readers formulate about where to solve the problem. The statistical analysis shows that the pattern of regressions into a region is not the same for subject-initial than for object-initial sentences. In fact, the multiple comparisons tests of the regions where there was a significant interaction between Verb Type and Word Order show that there are qualitative differences regarding what participants did in order to understand these types of sentences. In SVO sentences, the probability of regressing into the region of the second name was higher when the sentences included an ObjExp psych verb. This difference was not significant when OVS sentences were compared. In object-initial sentences, the probability of regressing to the first name was similar between both non-canonical sentences. However, TIR measure at Region 2 showed a main effect of Word Order. Participants regressed more to this region when the sentence was object-initial than subject-initial. This difference is driven by the high probability of regressing to this region when the sentence included an activity verb. This dissociated behaviour during the reading of SVO and OVS sentences shares something: Participants regress significantly more to the region that belongs to the name that comprises the sentential Object, this is, the name that was preceded by preposition "a." Literature on eye-tracking has shown that a great percentage of functional words and short words are skipped but still processed during reading, and that the shorter the word, the higher the chance of skipping it on a first fixation (Brysbaert et al., 2005; Rayner & McConkie, 1976; Rayner, Sereno, & Raney, 1996). Thus, although differences for TIR were found at the preposition region only in SVO sentences, readers may have processed the syntactic information that the preposition provided parafoveally (see Schotter, Angele, & Rayner, 2012, for a review on parafoveal processing), landing the saccade on the proper name. It is possible that in those sentences that required a revision of the arguments structure, they specially regressed to this region because it comprises the visually most informative region regarding the syntactic status of the argument. Re-reading it might ease the new interpretation of the event. If we consider that there were no semantic cues that could privilege one interpretation over the other (both arguments were equally animate), resorting to a visual cue of the argument's syntactic structure is the only possible strategy to retrieve the correct interpretation of the event.

In other words, when the reader can resort to language-specific morphosyntactic cues that are readily available in the sentence (as it is the case of preposition "a" in Spanish), and the reading paradigm facilitates the possibility of returning to them without restrictions (as it normally occurs during natural reading), the parser takes advantage of them to accurately understand the event denoted in the sentence. This hypothesis calls for further investigation, since other confounds (like the fact that the probability of regressing to initial regions in the sentence is higher than the probability of regressing to final words) may also modulate the amount of incoming regressions. We believe, however, that this experiment provides evidence consistent with prior findings. An extensive review on the role of prominence computation for sentence comprehension has shown that cues to which readers resort to for the computation of prominence may differ from language to language, although similar neural and cognitive correlates may take place when there is a prominence misanalysis (Bornkessel & Schlesewsky, 2006). For instance, in a language with an extensive case marking system like German, Kretzschmar et al. (2012) found that in order to recover from their initial misanalysis of prominence status, readers refixated both case ambiguous noun phrases.

Finally, the results of the comprehension task show an interaction between Verb Type and Word Order in the percentage of accurate responses when participants were prompted to identify who had performed the activity or felt something in the sentence read. This interaction was motivated by the high accuracy in questions about sentences with SVO word order and activity verbs. The statistical analysis shows that there was no difference among the other three conditions. This is not surprising if we consider that in the SVO sentences with activity verbs both semantic and syntactic canonical orders coincide; the other three conditions are versions of this sentence with either a syntactic or a semantic alteration. It is then possible that when both canonical orders coincide comprehension is enhanced.

Crucially to one of the purposes of re-running Gattei, Dickey, et al. (2015)'s experiment using the eye-tracking method, the percentage of accurate answers for those sentences that required a reinterpretation of the arguments structure of the verb was about 80% in both conditions (SVO with ObjExp psych verbs and OVS with activity verbs). This means that when readers face a more natural reading context and they can resort to their own reading strategies, they could understand "who did what to whom" correctly most of the times, task that was largely limited after reading the sentence with a moving window paradigm in a self-paced reading task.

Response times and total reading times may also enlighten us about two issues: On the one hand, questions response times are related to the cost of sentence processing. Recall that questions did not differ in syntactic structure, so participants would read similar questions for each set of sentences. The only difference between questions for ObjExp psych verbs sentences and for Activity verbs sentences was that questions included the verb read in the previous sentence. Hence, differences in response times could only be caused by either differences in processing the previous sentence (i.e. retrieving the correct thematic structure to solve "who did what to whom"), or a combination between the cost of processing the previous sentence and the cost of processing a specific verb type included in both the sentence and question. Consequently, the differences in response time cannot be attributed only to the question itself.

In a nutshell, response times were significantly higher for questions about sentences with ObjExp psych verbs than for questions about sentences with activity verbs, suggesting a greater cost for accurately retrieving psych verb events than for activity events. This effect is in line with the outcome of previous studies which have shown that ObjExp psych verbs lead to greater processing difficulty than activity verbs, as evidenced by increased reading times (Brennan & Pylkkänen, 2010; Cupples, 2002; Gattei, Dickey, et al. 2015; Gennari & MacDonald, 2009), and decreased comprehension accuracy in patients with Alzheimer's disease (Manouilidou, de Almeida, Schwartz, & Nair, 2009) and aphasia (Beretta et al., 2001; Piñango, 2006; Thompson & Lee, 2009). Results from previous ERP studies (Bornkessel et al., 2003; Gattei, Tabullo, et al. 2015) and neuroimaging techniques (Bornkessel et al., 2005) also found differential brain activity after the presentation of sentences with each type of verb. It is matter of future research to determine whether the same differences arise when comprehension of sentences with verbs of the same class that differ in the prominence status of their nominative argument (e.g. SubjExp vs. ObjExp psych verbs) is tested.

Response times have also revealed a significant interaction between verb type and word order, showing that participants took longer to respond questions when the sentence arguments did not follow the more-to-less prominent order. Both the main effect of Verb Type and the interaction are consistent with the effects of Verb Type and the interaction between Verb Type and Word Order found at Region 6 (Second Proper Name), this is, when participants have all the information needed to accurately understand the event. The latter suggests that the effects of wrongly computing the prominence status of an argument are so robust that they may persist several seconds afterwards, when participants need to recall the sentence that have just read in order to retrieve the correct argument structure of the sentence and successfully respond the question.

On the other hand, total reading times are informative of the resources that participants used in order to respond the questions accurately and as fast as possible, as they were instructed before the beginning of the experiment. One of the objectives of the current study was to analyse whether prominence computation would still have an impact on reading times in a more naturalistic reading setting, as compared to the reading paradigm used in a self-paced reading experiment. The outcome of the accuracy measure along with the total reading time measure shows that participants are able to accurately solve the problems posed by sentences with arguments that did not follow the canonical prominence hierarchy. We believe that this was possible because the task itself enables readers to regress to previous regions, and to skip words that are not so relevant for understanding "who did what to whom." The cost of doing so is time, reflected in the increase of reading time in both problematic sentences (Act OVS and ObjExp psych SVO).

The estimation of the total reading time that participants spent in each condition has also shown that even in natural reading conditions there is a difference between those sentences that required a reversal of arguments thematic hierarchy. Although the difference in accuracy between these two conditions was not significant, participants spent longer time reading and rereading those sentences with SVO word order and ObjExp psych verbs than sentences with OVS word order and activity verbs. This means that in both conditions participants reached the same accuracy level, but the cost of achieving the same comprehension level was higher for sentences with ObjExp psych verbs and SVO word order than for the other case. These results partly replicate what Gattei, Dickey, et al. (2015) found for accuracy between these sentences in the self-paced reading task. Questions about sentences with ObjExp psych verbs and SVO word order only reached 60% accuracy. The authors hypothesised that reassigning thematic roles might be more difficult when an "Actor" has to be downgraded to the macrorole of "Undergoer" (as it would occur in SVO sentences with ObjExp psych verbs) than when the reinterpretation goes from "Undergoer" to "Actor."

A final issue that needs to be considered on the discussion of the off-line task is the role of sentence acceptability. Prior literature has shown that sentence acceptability plays an important role in sentence comprehension, modulating metrics such as response times and response accuracy (Ditman, Holcomb, & Kuperberg, 2007; Wagers, Lau, & Phillips, 2009). In order to disentangle whether global sentence acceptability played any role in the results of the offline comprehension task, we decided to run an acceptability judgement task of all critical sentences and to reanalyse comprehension response times, total reading time and accuracy rates by only considering sets of sentences matched in acceptability. Following a similar procedure performed in Gattei, Dickey, et al. (2015), we expected that, if results of Experiment 1 were mainly yielded by sentences' mean acceptability, there should be no thematic reanalysis effects when analysing a subset of sentences with matched acceptability ratings. On the contrary, the persistence of such effects in the subset analysis would mean that it was arguments' misinterpretation that had led to higher response and reading times.

Experiment 2: acceptability judgement task

All stimuli used in the eye-tracking experiment and 200 filler sentences were used in an acceptability judgement task in order to disentangle whether difficulty for understanding SVO sentences with ObjExp psych verbs and OVS sentences with activity verbs was related to sentence acceptability. Evidence has shown that reading less acceptable sentences trigger higher response times and response accuracy in off-line comprehension tasks (Ditman et al., 2007; Wagers et al., 2009). This issue has been discussed at Gattei, Dickey, et al. (2015). Twenty of the filler sentences were modified so that they would become unacceptable sentences in the language, and the syntactic constituents of 20 other filler sentences were scrambled so that they would become acceptable but not very common sentences of the language, the rationale being that acceptability differed among sentences and participants were obliged to think about their response.

Methods

Participants

One hundred and sixty speakers of Spanish from Argentina (94 females) participated in this experiment voluntarily. Age ranged from 18 to 63 years old (*M*=32.4 years old). None of the subjects had a history of prior neurological disease, drug or alcohol abuse, psychiatric disorders, developmental speech/language disorders, or learning disabilities. All participants had normal or corrected-to-normal vision. All of them provided written consent prior to the study.

Procedure

Participants were tested individually through the Ibex Farm application (developed by Alex Drummond, McGill University, http://spellout.net/ibexfarm). Trials were presented randomly according to a Latin square design. Participants were assigned to one out of two lists, comprehending sentences from the different sets (I and II in Appendix A). Thus, each participant saw 24 critical sentences, six sentences per condition. Participants were asked to judge on a 5- point scale how the sentence they had just read sounded to them. They were reminded that their answer should be driven according to whether they found the sentence acceptable (5 points), fairly acceptable (4), not that acceptable (3 points), doubtfully acceptable or unlikely to be acceptable (2 points), or not an acceptable sentence of Spanish (1 point). Although no time limit was set for each of the trials, participants were encouraged to rate sentences as fast as possible and to base their answers on their own intuition. Apart from this, half of the filler sentences included a comprehension question that participants had to respond with "yes" or "no" by clicking on the right answer. In this way, we forced participants to pay attention to what they were reading.

Data analysis

Mean acceptability was computed for each critical item. A one-way ANOVA between conditions was used in order to test whether there were significant differences among them. In order to clarify whether computation of prominence status played a role in off-line comprehension of Experiment 1, nine items from each condition that were matched on mean acceptability were chosen, and all off-line measures were reanalysed following the same procedure as in Experiment 1.

Results

Acceptability judgement task. A one-way ANOVA test revealed a significant difference among conditions, $F(3, 1759) = 161$, $p<0.01$. A multiple comparisons test showed that mean acceptability of subject-initial sentences was significantly lower when the sentence included an ObjExp psych verb, $\beta = -1.188$, SE=0.081, z=−14.662, p<0.0001. On the other hand, mean acceptability of object initial sentences was significantly lower when the sentence included an activity verb, $\beta = -1.309$, SE=0.080, z=−16.321, p<0.0001. This means that sentences that include a non-canonical arguments order are overall less acceptable than sentences whose first argument is the most prominent one.

Sentences acceptability matching. We chose nine items from each condition that were matched on mean acceptability in order to rule out that prominence miscomputation effects were caused by differences on sentences' global acceptability. A one-way ANOVA revealed no difference between the four groups of sentences, $F(3, 32) = 1.459$, p=0.244. Table 4 shows mean acceptability and standard error for the sentences selected from each condition. The outcome of the comprehension task for these thirty-six sentences was then re-analysed. Results of this reanalysis are described below.

Question accuracy. Figure 3(a) shows mean accuracy for the subset of 36 sentences according to condition. Results show that mean accuracy for this subset of 36 critical items was 89.83% (SE=1.39). Analyses of differences in accuracy according to verb type and word order revealed main effects of word order and verb type. On average, participants responded better to questions about SVO sentences ($M = 91.98\%$, SE=1.77) than about OVS sentences ($M = 87.66\%$, SE=2.5), $\beta = -0.351$, SE=0.178, z=−1.965, p=0.049. They also performed better for questions about sentences with activity verbs ($M = 91.98\%$, SE=1.77) than for questions about sentences with ObjExp psych verbs ($M = 87.66\%$, SE=1.77), $\beta = -0.351$, SE=0.179, z=−1.965, p=0.049. An interaction between both factors was also found, $\beta = 0.389$, SE=0.179, z=2.177, p=0.029. Tukey's multiple comparisons test revealed that there was a tendency towards significance in the difference between mean accuracy of SVO sentences with activity verbs ($M = 96.67\%$, SE=1.64) and the mean accuracy of the other three conditions (ActOVS: $\beta = -1.450$, SE=0.578, z=−2.506, p=0.05; ObjExpSVO: $\beta = -1.450$, SE=0.578, z=−2.506, p=0.05; ObjExpOVS: $\beta = -1.362$, SE=0.583, z=−2.337, p=0.08).

Table 4. Mean acceptability and standard error of the sentences selected for the reanalysis of the comprehension task of the current eye-tracking study.

Verb type	Word order	Verb	M	*SE*
Activity	OVS	agradecer	3.00	.24
		cocinar	2.82	.27
		enseñar	3.63	.29
		hablar	2.96	.26
		implorar	2.75	.35
		insistir	2.77	.17
		mentir	2.93	.32
		sonreir	3.40	.34
		suplicar	3.30	.37
Activity	SVO	aconsejar	2.78	.27
		apuntar	3.63	.22
		cobrar	3.28	.24
		cocinar	3.48	.24
		contestar	3.49	.23
		llorar	2.94	.30
		protestar	3.19	.25
		robar	3.64	.27
		rogar	3.44	.26
ObjExp	OVS	asombrar	3.45	.34
		cansar	3.25	.28
		conmover	3.14	.35
		deprimir	2.76	.25
		disgustar	3.71	.28
		horrorizar	3.52	.25
		importar	2.94	.31
		impresionar	3.50	.32
		interesar	3.61	.29
ObjExp	SVO	agradar	3.43	.33
		atraer	3.06	.31
		desagradar	3.79	.35
		divertir	3.06	.39
		encantar	3.00	.33
		fascinar	3.38	.46
		gustar	3.25	.28
		importar	2.81	.32
		repugnar	3.00	.38

Notes: ObjExp, object experiencer psych verbs; OVS, object-verb-subject; SVO, subject-verb-object.

Response times. Figure 3(b) shows mean response times (RTs) for the subset of 36 sentences chosen in the acceptability judgement task according to condition. Analyses of differences in response times revealed a main effect of verb type, $\beta = 0.092$, *SE*=0.023, *t*=3.960. On average, participants took longer to respond questions of sentences that included ObjExp psych verbs (*M*=4500 ms., *SE*=205), than questions about sentences that included activity verbs (*M*=3595 ms., *SE*=143). The interaction between verb type and word order was marginally significant: $\beta = -0.0449$, *SE*=0.0234, *t*=−1.918. The Tukey post-hoc test showed that participants responded questions about SVO sentences with activity verbs significantly faster (*M*=3257 ms., *SE*=165) than questions about sentences of the other three conditions (ActOVS: $\beta = -0.171$, *SE*=0.066, *z*=−2.588, *p*=0.04; ObjExpSVO: $\beta = 0.274$, *SE*=0.066, *z*=4.156, *p*<0.001; ObjExpOVS: $\beta = -0.266$, *SE*=0.065, *z*=4.040, *p*<0.001).

Total reading time. Figure 3(c) shows the average total reading time for the subset of 36 sentences according to condition. Analyses of differences in total reading times of the subset of 36 critical items revealed a significant interaction between word order and verb type, $\beta = -0.066$, *SE*=0.019, *t*=−3.507. The multiple comparisons test showed that participants took longer time to read SVO sentences when the sentence included an ObjExp psych verb than when it included an activity verb, $\beta = 0.140$, *SE*=0.053, *z*=2.637, *p*=0.04. Difference in total reading time between OVS conditions showed that participants read sentence slower when the sentence included an activity verb than when it included an ObjExp psych verb. This difference was marginally significant, $\beta = 0.125$, *SE*=0.053, *z*=2.336, *p*=0.089.

Discussion

Experiment 2 was included in order to test whether there were significant differences in the acceptability of sentences that required a reversal of thematic arguments,

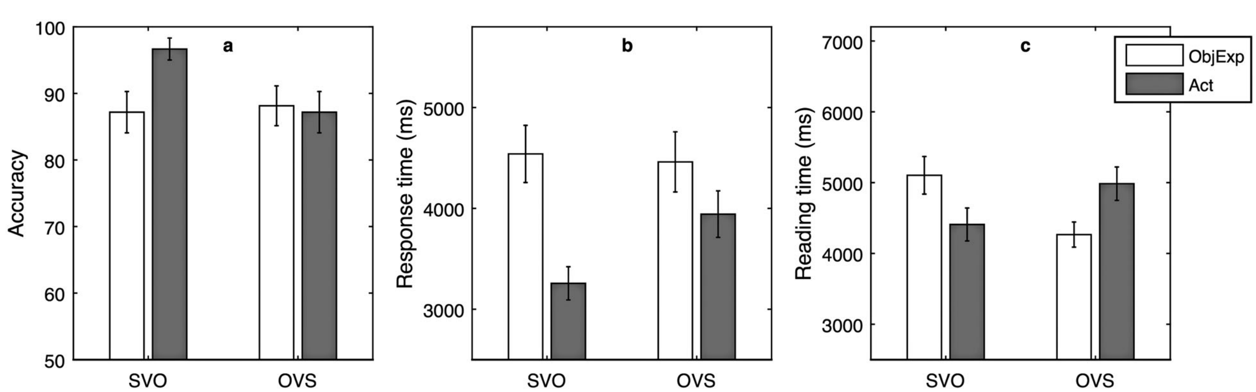

Figure 3. Percentage of accurate answers (panel a), mean response times for the comprehension question (panel b), and mean total reading time (panel c) for the subset of 36 critical sentences chosen from the acceptability judgement task according to verb type (ObjExp psych verb vs. Act) and word order (SVO vs. OVS). Error bars correspond to Standard Error of the Mean. ObjExp psych verb, object experiencer psych verb; Act, activity verb; SVO, subject-verb-object; OVS, object-verb-subject.

and whether the results from the comprehension task of Experiment 1 mainly depended on sentence acceptability instead of being a correlate of the cognitive cost imposed by the wrong computation of the arguments prominence status. Results of the acceptability judgement task showed that acceptability was significantly different according to their Verb Type and Word Order. Thus, a subset of sentences with similar acceptability was chosen and data from the comprehension task was reanalysed. Results of this reanalysis task confirm that main effects and interactions between verb type and word order persist for all accuracy, response time and total reading time, with a small decrease in the statistical significance value, mostly in the interactions' multiple comparisons tests. The pattern of differences in questions response time also shows facilitation for SVO sentences with activity verbs when compared to the other three conditions. This pattern also takes place in accuracy rate. Once again, this outcome was expected, as this condition is the one that shows no modification neither from the canonical word order of the language (SVO: Contreras, 1976; Hernanz & Brucart, 1987) nor from the arguments prominence hierarchy. Results of the reanalysis suggest that arguments prominence hierarchy is thus a robust predictor for accuracy, response times and mainly for reading time of grammatical sentences.

Conclusions

Results of the current experiments have shown that the parser processes arguments incrementally, and generates predictions about the class of the upcoming verb very early; in fact, as soon as the first argument is read. In the sentences tested in this study, such predictions were based on morphosyntactic information like word order and morphological case marking. When there was a mismatch between the verb class predicted and the input, readers spent longer time fixating on the verb and the following regions. Furthermore, having the freedom to read naturally, readers regressed to previous regions more and for longer time when they faced this mismatch. In order to respond to the comprehension question correctly, whenever they had to reverse the argument structure of the event, participants specifically regressed to the region that was more salient in terms of the syntactic information it provided, this is, the object. Results of this task also showed that participants performed better in this study than in a self-paced reading task with similar materials. The latter two findings show that paradigm-specific demands may modulate accuracy rates related to sentence interpretation and strategies for solving off-line tasks. However, difficulties in processing may come to light if other types of cognitive measures are considered. In this study, participants improved performance in non-canonical sentences at the expense of spending longer time reading them.

Note

1. The dative clitic can be understood as the incorporation of a new argument to the logic structure "responder (*x*,*y*)," where *x* is the Actor and *y* is a linguistic object that acts as a stimulus (e.g. a letter, a question, etc.). The clitic "le" corresponds to argument *z*, the recipient of that letter, question. Once it is added to the structure, it ranks higher for Undergoer than *y* just like a ditransitive structure in English of the kind "Mary sends John a letter." This is an extremely productive process in Spanish (Cuervo, 2010).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Consejo Nacional de Investigaciones Científicas y Técnicas [National Scientific and Technical Research Council (CONICET)] as part of the research project entitled "Relaciones entre procesos léxicos y procesos sintácticos en la producción y la comprensión de lenguaje" under [grant number PIP 112 201101 00994].

ORCID

Carolina A. Gattei http://orcid.org/0000-0002-7490-3363

References

Arnaiz, A. R. (1998). An overview of the main word order characteristics of Romance. In A. Siewierska (Ed.), *Constituent order in the languages of Europe* (Vol. 20, pp. 47–74). Berlin: Mouton de Gruyter. doi:10.1515/9783110812206.47

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A*, *58*(6), 1065–1086. doi:10.1080/02724980443000476

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using r*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511801686

Bader, M., & Bayer, J. (2006). *Case and linking in language comprehension: Evidence from German* (Vol. 34). Dordrecht: Springer Science & Business Media. doi:10.1007/1-4020-4344-9

Bakovic, E. (1998). Optimality and inversion in Spanish. In P. Barbosa (Ed.), *Is the best good enough?: Optimality and competition in syntax* (pp. 35–58). Cambridge, MA: MIT Press.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*(3), 364–390. doi:10.1016/0010-0285(85)90013-1

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*(3), 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014, June). Fitting linear mixed-effects models using lme4. ArXiv e-prints.

Belletti, A., & Rizzi, L. (1988). Psych-verbs and θ-theory. *Natural Language & Linguistic Theory*, *6*(3), 291–352. doi:10.1007/BF00133902

Beretta, A., Schmitt, C., Halliwell, J., Munn, A., Cuetos, F., & Kim, S. (2001). The effects of scrambling on Spanish and Korean agrammatic interpretation: Why linear models fail and structural models survive. *Brain and Language*, *79*(3), 407–425. doi:10.1006/brln.2001.2495

Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, *113*(4), 787–821. doi:10.1037/0033-295X.113.4.787

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2002). Beyond syntax: Language-related positivities reflect the revision of hierarchies. *NeuroReport*, *13*(3), 361–364.

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2003). Eliciting thematic reanalysis effects: The role of syntax-independent information during parsing. *Language and Cognitive Processes*, *18*(3), 269–298. doi:10.1080/01690960244000018

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage*, *26*(1), 221–233. doi:10.1016/j.neuroimage.2005.01.032

Bornkessel-Schlesewsky, I., & Schlesewsky, M (2009). *Processing syntax and morphology: A neurocognitive perspective*. Oxford: Oxford University Press. doi:10.1075/fol.18.2.07mac

Brennan, J., & Pylkkänen, L. (2010). Processing psych verbs: Behavioural and MEG measures of two different types of semantic complexity. *Language and Cognitive Processes*, *25*(6), 777–807. doi:10.1080/01690961003616840

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 53–77). Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198566816.003.0003

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. van Gompel (Ed.), *Eye movements: A window on mind and brain* (pp. 341–372). Amsterdam: Elsevier Science. doi:10.1016/B978-008044980-7/50017-3

Contreras, H. (1976). *A theory of word order with special reference to Spanish* (No. 9). North-Holland.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: Eye tracking with MATLAB and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617. doi:10.3758/BF03195489

Crocker, M. W. (2012). *Computational psycholinguistics: An interdisciplinary approach to the study of language* (Vol. 20). Dordrecht: Springer Science & Business Media. doi:10.2307/417577

Cuervo, M. C. (2010). Some dative subjects are born, some are made. In *Selected proceedings of the 12th hispanic linguistics symposium* (pp. 26–37). Somerville, MA: Cascadilla Proceedings Project. Retrieved from http://www.lingref.com/cpp/hls/12/paper2403.pdf.

Cupples, L. (2002). The structural characteristics and on-line comprehension of experiencer-verb sentences. *Language and Cognitive Processes*, *17*(2), 125–162. doi:10.1080/01690960143000001

Davis, C. J., & Perea, M. (2005). Buscapalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods*, *37*(4), 665–671. doi:10.3758/BF03192738

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology*, *44*(6), 927–935. doi:10.1111/j.1469-8986.2007.00593.x

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*, 547–619. doi:10.2307/415037

Dröge, A., Maffongelli, L., & Bornkessel-Schlesewsky, I. (2014). Luigi piace a Laura? In A. Bachrach, I. Roy, & L. Stockall (Eds.), *Structuring the argument: Multidisciplinary research on verb argument structure* (Vol. 10, pp. 83–118). Amsterdam: John Benjamins Publishing Company. doi:10.1075/lfab.10.05

Franco, J. (2013). Towards a typology of psych verbs: Evidence from Spanish. Anuario del Seminario de Filología Vasca "Julio de Urquijo" 0, 119–134.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*(2), 178–210. doi:10.1016/0010-0285(82)90008-1

Frisch, S., & Schlesewsky, M. (2001). The N400 reflects problems of thematic hierarchizing. *NeuroReport*, *12*(15), 3391–3394. doi:10.1097/00001756-200110290-00048

Gattei, C. A., Dickey, M. W., Wainselboim, A. J., & París, L. (2015). The thematic hierarchy in sentence comprehension: A study on the interaction between verb class and word order in Spanish. *The Quarterly Journal of Experimental Psychology*, *68*(10), 1981–2007. doi:10.1080/17470218.2014.1000345

Gattei, C. A., Tabullo, Á., París, L., & Wainselboim, A. J. (2015). The role of prominence in Spanish sentence comprehension: An ERP study. *Brain and Language*, *150*, 22–35. doi:10.1016/j.bandl.2015.08.001

Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, *111*(1), 1–23. doi:10.1016/j.cognition.2008.12.006

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., von Cramon, D. Y., & Schlesewsky, M. (2006). Linguistic prominence and broca's area: The influence of animacy as a linearization principle. *NeuroImage*, *32*(3), 1395–1402. doi:10.1016/j.neuroimage.2006.04.213

Gutiérrez-Bravo, R. (2007). Prominence scales and unmarked word order in Spanish. *Natural Language & Linguistic Theory*, *25*(2), 235–271. doi:10.1007/s11049-006-9012-7

Haupt, F. S., Schlesewsky, M., Roehm, D., Friederici, A. D., & Bornkessel-Schlesewsky, I. (2008). The status of subject–object reanalyses in the language comprehension architecture. *Journal of Memory and Language*, *59*(1), 54–96. doi:10.1016/j.jml.2008.02.003

Hernanz, M. L., & Brucart, J. M. (1987). *La sintaxis: Principios teóricos*. Barcelona: Ed. Crítica.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. doi:10.1037/0033-295X.87.4.329

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1-2), 262–284. doi:10.1080/09541440340000213

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*(1), 12–35. doi:10.1037/0096-3445.135.1.12

Kretzschmar, F., Bornkessel-Schlesewsky, I., Staub, A., Roehm, D., & Schlesewsky, M. (2012). Prominence facilitates ambiguity resolution: On the interaction between referentiality, thematic roles and word order in syntactic reanalysis. In M. Lamers, & P. de Swart (Eds.), *Case, word order and prominence* (Vol. 40, pp. 239–271). Dordrecht: Springer. doi:10.1007/978-94-007-1463-2_11

Lamers, M., & De Swart, P. (2012). *Case, word order and prominence: Interacting cues in language production and comprehension*. Dordrecht: Springer Science & Business Media. doi:10.1007/978-94-007-1463-2

Logacev, P., & Vasishth, S. (2013). em2: A package for computing reading time measures for psycholinguistics. [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=em2 (R package version 0.9).

Manouilidou, C., de Almeida, R. G., Schwartz, G., & Nair, N. P. V. (2009). Thematic roles in Alzheimer's disease: Hierarchy violations in psychological predicates. *Journal of Neurolinguistics*, *22*(2), 167–186. doi:10.1016/j.jneuroling.2008.10.002

Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition*, *30*(4), 551–561. doi:10.3758/BF03194956

Ocampo, F (1995). The word order of two-constituent constructions in spoken Spanish. In P. A. Downing & M. Noonan (Eds.), *Word order in discourse* (Vol. 30, pp. 425–447). Amsterdam: John Benjamins. doi:10.1075/tsl.30.15oca

Piñango, M. M. (2006). Thematic roles as event structure relations. In I. Bornkessel, M. Schlesewsky, B. Comrie, & A. D. Friederici (Eds.), *Semantic role universals and argument linking: Theoretical, typological, and psycholinguistic perspectives* (Vol. 165, pp. 303–326). Berlin: Walter de Gruyter. doi:10.1515/9783110219272.303

Pinheiro, J. C., & Bates, D. M. (2000). Linear mixed-effects models: Basic concepts and examples. In J. C. Pinheiro, & D. M. Bates (Eds.), *Mixed-effects models in s and s-plus* (pp. 3–56). New York: Springer. doi:10.1007/0-387-22747-4_1

Primus, B. (2012). Animacy, generalized semantic roles, and differential object marking. In M. Lamers, & P. de Swart (Eds.), *Case, word order and prominence* (Vol. 40, pp. 65–90). Dordrecht: Springer. doi:10.1007/978-94-007-1463-2_4

Pulvermüller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language*, *110*(2), 81–94. doi:10.1016/j.bandl.2008.12.001

R Core Team. (2013). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from http://www.R-project.org/.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the EZ reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 720–732. doi:10.1037/0096-1523.30.4.720

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research*, *16*(8), 829–837. doi:10.1016/0042-6989(76)90143-7

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, *3*(2), 245–248. doi:10.3758/BF03212426

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(5), 1188–1200. doi:10.1037/0096-1523.22.5.1188

Roehm, D., Schlesewsky, M., Bornkessel, I., Frisch, S., & Haider, H. (2004). Fractionating language comprehension via frequency characteristics of the human EEG. *NeuroReport*, *15*(3), 409–412. doi:10.1097/00001756-200403010-00005

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74*(1), 5–35. doi:10.3758/s13414-011-0219-2

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences*, *7*(11), 489–493. doi:10.1016/j.tics.2003.09.010

Staub, A., & Clifton Jr., C. (2006). Syntactic prediction in language comprehension: Evidence from either… or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425–436. doi:10.1037/0278-7393.32.2.425

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 327–342). Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780198568971.013.0019

Suñer, M. (1982). *Syntax and semantics of Spanish presentational sentence-types*. Washington, DC: Georgetown University School of Language. doi:10.1075/sl.8.3.08put

Thompson, C. K., & Lee, M. (2009). Psych verb production and comprehension in agrammatic Broca's aphasia. *Journal of Neurolinguistics*, *22*(4), 354–369. doi:10.1016/j.jneuroling.2008.11.003

Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139166799

Van Valin Jr., R. D. (2005). *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511610578

Vasishth, S., von der Malsburg, T., & Engelmann, F. (2013). What eye movements can tell us about sentence comprehension. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(2), 125–134. doi:10.1002/wcs.1209

von der Malsburg, T., & Vasishth, S. (2013). Scanpaths reveal syntactic underspecification and reanalysis strategies. *Language and Cognitive Processes*, *28*(10), 1545–1578. doi:10.1080/01690965.2012.728232

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*(2), 206–237. doi:10.1016/j.jml.2009.04.002

Wang, L., Schlesewsky, M., Bickel, B., & Bornkessel-Schlesewsky, I. (2009). Exploring the nature of the subject-preference: Evidence from the online comprehension of simple sentences in Mandarin Chinese. *Language and Cognitive Processes*, *24*(7-8), 1180–1226. doi:10.1080/01690960802159937

Wang, L., Schlesewsky, M., Philipp, M., & Bornkessel-Schlesewsky, I. (2012). The role of animacy in online argument interpretation in Mandarin Chinese. In M. Lamers, & P. de Swart (Eds.), *Case, word order and prominence* (pp. 91–119). Dordrecht: Springer. doi:10.1007/978-94-007-1463-2_5

Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology*, *36*(5), 559–570. doi:10.1111/1469-8986.3650559

Wolff, S., Schlesewsky, M., Hirotani, M., & Bornkessel-Schlesewsky, I. (2008). The neural mechanisms of word order processing revisited: Electrophysiological evidence from Japanese. *Brain and language*, *107*(2), 133–157. doi:10.1016/j.bandl.2008.06.003