

This article was downloaded by: [Yamila Sevilla]

On: 08 May 2014, At: 06:30

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## The Quarterly Journal of Experimental Psychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pqje20>

### Pupillary dynamics reveal computational cost in sentence planning

Yamila Sevilla<sup>ab</sup>, Mora Maldonado<sup>a</sup> & Diego E. Shalóm<sup>bc</sup>

<sup>a</sup> Instituto de Lingüística, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina

<sup>b</sup> Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

<sup>c</sup> Laboratorio de Neurociencia Integrativa, Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina

Accepted author version posted online: 08 Apr 2014. Published online: 06 May 2014.

To cite this article: Yamila Sevilla, Mora Maldonado & Diego E. Shalóm (2014): Pupillary dynamics reveal computational cost in sentence planning, *The Quarterly Journal of Experimental Psychology*, DOI: [10.1080/17470218.2014.911925](https://doi.org/10.1080/17470218.2014.911925)

To link to this article: <http://dx.doi.org/10.1080/17470218.2014.911925>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or

distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

# Rapid communication

## Pupillary dynamics reveal computational cost in sentence planning

Yamila Sevilla<sup>1,2</sup>, Mora Maldonado<sup>1</sup>, and Diego E. Shalóm<sup>2,3</sup>

<sup>1</sup>Instituto de Lingüística, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina

<sup>2</sup>Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

<sup>3</sup>Laboratorio de Neurociencia Integrativa, Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina

This study investigated the computational cost associated with grammatical planning in sentence production. We measured people's pupillary responses as they produced spoken descriptions of depicted events. We manipulated the syntactic structure of the target by training subjects to use different types of sentences following a colour cue. The results showed higher increase in pupil size for the production of passive and object dislocated sentences than for active canonical subject-verb-object sentences, indicating that more cognitive effort is associated with more complex noncanonical thematic order. We also manipulated the time at which the cue that triggered structure-building processes was presented. Differential increase in pupil diameter for more complex sentences was shown to rise earlier as the colour cue was presented earlier, suggesting that the observed pupillary changes are due to differential demands in relatively independent structure-building processes during grammatical planning. Task-evoked pupillary responses provide a reliable measure to study the cognitive processes involved in sentence production.

*Keywords:* Language production; Pupillometry; Grammatical planning; Cognitive effort; Complexity.

According to a speaker's intentions, sentences may differ in many ways. Similar meanings can be expressed using different words, and the same words can be arranged in different ways to convey slightly distinct messages. To account for this property, most models of sentence production (Bock & Levelt, 1994) assume the existence of two separate sets of mechanisms: lexical retrieval processes and structure-building processes (V. S. Ferreira & Slevc, 2007, for a review). While the former recruit the appropriate words for representing the elements of the intended message, the latter deal

with preparing the syntactic framework that expresses the relation between them. The interplay between these two sorts of processes and the extent to which they are independent has been the focus of an intense debate. Approaches range from a strong lexicalist account that makes lexical selection a pre-requisite for structure building (Levelt, Roelofs, & Meyer, 1999) to a strong syntactic account that posits that independent abstract processes are responsible for structural configuration and word ordering in a way that is not mediated by lexical retrieval (Bock, 1990). Regardless of the adopted

---

Correspondence should be addressed to Yamila Sevilla, Instituto de Lingüística, Facultad de Filosofía y Letras, Universidad de Buenos Aires, 25 de Mayo 217/221, 1°, (C1002ABE) Ciudad Autónoma de Buenos Aires, Argentina. E-mail: [ysevilla@filo.uba.ar](mailto:ysevilla@filo.uba.ar); [yamilasevilla@gmail.com](mailto:yamilasevilla@gmail.com)

This work was partially supported by National Research Council (CONICET) [grant number PIP 112 201101 00994].

view, theories must assume the existence of specific syntactic operations that would ultimately determine the linear form of a sentence. The kind and number of operations of this sort involved during the encoding characterize sentence complexity.

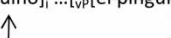
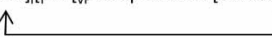

It is widely accepted that some syntactic structures are more difficult to process than others, based on various factors, including their internal properties. Thus, for instance, the presence of constituent displacement during sentence derivation is supposed to correlate with cognitive demand during planning. However, there are some movement operations for which no computational effort was empirically evinced (Phillips, 1996). To account for this issue, some proposals regarding online linguistic computation (Corrêa, 2011; Phillips & Lewis, 2013) suggest that two sorts of movement operations are involved in a linguistic derivation: those required for the linear positioning in the canonical order of the language, and those that displace constituents from the position where they are semantically interpreted due to discourse requirements. According to Corrêa (2011), when producing a sentence, the movement operations required for the linear positioning of constituents in canonical thematic order do not need to be carried out on line and are therefore costless, whereas the displacement of constituents due to discourse constraints—which causes noncanonical thematic order—has additional cost. Hence, greater processing load would be predicted for passive sentences, focused structures, “WH” questions, and relative clauses in contrast with structures in the canonical word order. Based on different assumptions about what causes complexity, and therefore processing difficulty, other accounts (see F. Ferreira, 1994; Gennari, Mirkovic, & Macdonald, 2012; Just & Carpenter, 1993; Macdonald, 2013) make similar sort of predictions. Using speech latency measures, a few studies focused specifically on the cognitive demand of producing active versus passive sentences and documented that actives are easier to formulate and articulate (F. Ferreira, 1994; Tannenbaum & Williams, 1968).

In this study we investigate the cognitive demand associated with syntactic complexity,

studying pupil responses related to the production of sentences with different structures. The task-evoked pupillary responses (TEPR) have been used as reliable neurophysiological index of cognitive effort in different domains (Beatty, 1982; Beatty & Lucero-Wagoner, 2000), including perception, mental arithmetic, working memory load, and attention. Regarding language, in studies using lexical decision tasks and naming, frequency effects correlated with pupil diameter, suggesting a relation between cognitive effort and word processing (Kuchinke, Vo, Hofmann, & Jacobs, 2007; Papesh & Goldinger, 2012).

More relevant to our purposes, in the sentence comprehension domain, TEPR have provided a measure of differential processing cost according to the type of structure and its syntactic complexity. In a recent study investigating the role of different sources of information during spoken sentence comprehension, Engelhardt and his colleagues (Engelhardt, Ferreira, & Patsenko, 2010) examined the influence of prosody and visual context on processing effort associated with the resolution of temporary syntactic ambiguities, and they demonstrated that pupillometry is a sensitive measure to quantify differential cognitive cost in syntactic processing. Pupil size changes over time were also used to investigate constituent order priming during online sentence comprehension (Scheepers & Crocker, 2004). This work considered pupil dilation as a proxy for processing difficulty in disambiguation. Older studies have straightforwardly examined the effects of syntactic complexity on processing effort. In a seminal work, Schluroff (1982) presented strong evidence relating TEPR amplitude to syntactic complexity of English sentences. Schluroff showed that mean pupillary dilation significantly correlated with syntactic complexity, suggesting a close and fine-grained relation between structural properties of utterances and computational demands in online processing. Just and Carpenter (1993) also used pupillary dilation responses as an indicator of cognitive load imposed by syntactic structure in sentence comprehension during reading. They contrasted the syntactic processing of simpler versus more complex sentences. The more

Table 1. Target structures

(1)	a. [ <sub>IP</sub> [El pingüino] <sub>i</sub> ... [ <sub>VP</sub> [el pingüino] <sub>i</sub> aplaude [al ratón]]]. 	<b>ACTIVE</b> CANONICAL THEMATIC ORDER PER MOVEMENT
	b. El pingüino aplaude al ratón. <i>The penguin applauds the mouse.</i>	
(2)	a. [ <sub>IP</sub> [El ratón] <sub>i</sub> ] [ <sub>i</sub> ... [ <sub>VP</sub> es aplaudido [el ratón] <sub>i</sub> ] [por el pingüino]]. 	<b>PASSIVE</b> NON-CANONICAL THEMATIC ORDER PER MOVEMENT
	b. El ratón es aplaudido por el pingüino. <i>The mouse is applauded by the penguin.</i>	
(3)	a. [ <sub>TOPP</sub> [Al ratón] <sub>i</sub> ] ... [ <sub>IP</sub> [ <sub>VP</sub> lo <sub>i</sub> aplaude ... [el pingüino] [al ratón] <sub>i</sub> ]]. 	<b>CLLD</b> CANONICAL THEMATIC ORDER PER MOVEMENT
	b. Al ratón lo aplaude el pingüino. <i>The mouse [clitic: him] applauds the penguin.</i>	

complex sentences induced not only increased latency but also larger pupillary responses than their simple counterparts.

Despite these robust and reliable results in comparable domains, pupillometry has been scarcely used as investigation tool for language production, and no studies exist, to our knowledge, examining pupil dynamics during sentence planning.

In order to investigate whether different cognitive effort is associated with different syntactic structures in sentence production we analysed the pupil size changes during the oral description of events. We manipulated the syntactic structure that should be used to describe the pictures. Target structures were Spanish canonical subject–verb–object active sentences and two different non-canonical structures: passive sentences and clitic left dislocated sentences (CLLD), which are active topicalized structures where the object is placed at the beginning of the sentence (Table 1).

To further investigate whether lexical retrieval processes can be separated from syntactic processes, we studied the timing of cognitive effort during planning by manipulating the time of the presentation of the structure-building cue. Thus, the colour frame was presented simultaneously (*unplanned condition*)

or before the picture ( $-1$  s; *planned condition*). An initial *training* task, where the structures were presented in blocks, was implemented.

According to our predictions, TEPR should reflect differential processing cost related to complexity, showing a higher pupil diameter increase for noncanonical thematic order (passive and CLLD sentences) than for canonical sentences. Moreover, if TEPR reflect the cost of structure-building processes, then the differential TEPR should be observed earlier as the structure cue is presented earlier.

## Method

### Participants

Twenty subjects took part in the experiment. All participants were native speakers of Rioplatense Spanish, aged 21–35 years, and reported normal or corrected-to-normal vision. Subjects were recruited from University of Buenos Aires general population and were paid for their participation.

### Materials and apparatus

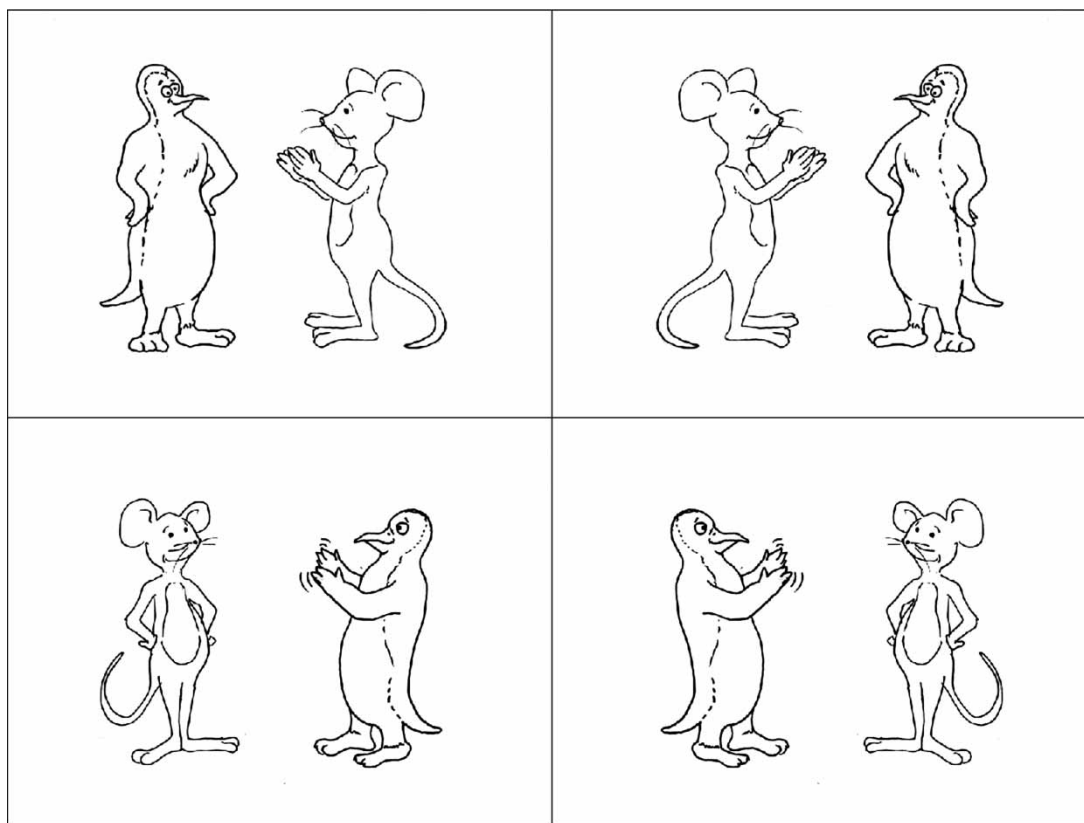
The experimental pictures were black-on-white line drawings of 27 simple transitive events. Two

role-traded versions of each event were created, switching the event participants (agent and patient), in order to control perceptual and lexical properties of the images. We used two mirror-imaged versions of each of these 54 images (see Figure 1), giving a total of 108 stimuli. The complete list of events can be found in the Appendix.

Stimuli were presented on a 19" CRT monitor (1024 × 768 pixels resolution; frame rate 60 Hz). Participants were seated in front of the monitor with the head positioned on a chin and forehead rest at a distance of 70 cm from the monitor. Participants were instructed to rely on the forehead rest, in order to reduce head movements during verbal responses. The illumination of the room was kept constant throughout testing sessions.

Speech was recorded during each trial. Vocalization onsets were extracted offline using a semiautomated procedure (Protopapas, 2007).

Pupil diameter was monitored using a desktop-mounted, video-based eye tracker (EyeLink 1000, SR Research Ltd., Ontario, Canada) at a sampling rate of 1000 Hz. Participant's gaze was calibrated with a standard 13-point grid. Only left-eye data were used for the analysis. The experiment was implemented in Matlab (Mathworks, Natick, MA) using Psychophysics Toolbox (Brainard, 1997). Pupil size during blinks and saccades was replaced by a linear interpolation from the last valid sample before the beginning of blink or saccade to the first valid sample after the end of each event.



**Figure 1.** Example of the stimuli. For each event, four images were created: two role-traded and two mirrored versions. Both agent and patient of all events were animals, in order to control the animacy of the participants. All pictures had the same luminosity and a similar disposition in space.

### Task and procedure

Participants were asked to describe a picture in one sentence, using the verb presented at the beginning of the trial. Although there was not a time limit, subjects were required to respond as fast as possible.

In order to force the use of the three different types of structures (active, passive and CLLD sentences), participants were trained to associate each structure to one of three different colour cues (green, violet, or orange). During the *training*, the participants saw 54 images (all 27 events in both role-traded versions), separated in three blocks of 18 trials. Before the beginning of each block, the subjects were instructed to use one particular structure. A colour frame (0.8° of visual angle in width) was drawn along the edges of the screen, surrounding the image (see Figure 2, left). Three different colours were used for the three blocks. To avoid any influence of the particular order of trials or of the colour associated with each structure, the experimental pictures were distributed across 12 lists with balanced combinations of three-colour frames. Each block was preceded by four practice trials, taken from a different set of six transitive events (with their four versions).

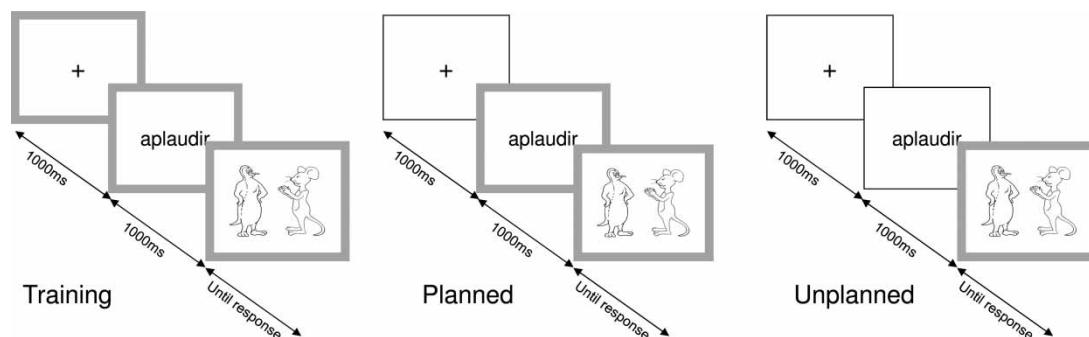
In the experimental task, structural conditions (colour frames) were presented pseudorandomly, so participants were unable to predict the condition of the following trial. In each trial, participants had to produce a particular sentence type according to the colour associations learned during the training.

We also manipulated the time at which the frame appeared on each trial. In the planned condition the colour frame appeared 1000 ms before the picture, simultaneously with the verb (Figure 2, centre). In the unplanned condition the colour frame appeared simultaneously with the picture (Figure 2, right).

During the experimental task, participants saw the complete set of 108 images, separated in two groups. In the first group we presented the mirrored versions of the 54 images used in the training task. In the last group, the 54 images of the training task were used. Moreover, no structure–image pair was repeated along the experiment. In other words, the two times a verb was used in a certain structure, the picture representing the event was different (either its role-trade version or its mirrored version).

### Analytical strategy for pupil data

We concentrated our analysis on the period that begins with the presentation of the image and ends with the vocalization onset. However, the length of this period is different for each trial, not permitting direct event-locked analyses. To deal with this issue, we rescaled the time of each individual trial and converted it to a normalized period, dividing it by the time elapsed between picture presentation and the onset of response vocalization. In normalized time,  $t = 0$  corresponds to the image



**Figure 2.** Schematic of task sequences. Each trial began with a centred fixation cross (1000 ms), followed by the infinitive form of the required verb; 1000 ms later the picture appeared. The picture was present until participants finished talking and pressed a key indicating the end of the trial. Voice recordings began with image presentation and continued until 1000 ms after the participants pressed the key. The timing conditions (training, planned, unplanned) differed in the moment at which the coloured frame was presented.

appearance, and  $t = 1$  corresponds to vocalization onset of the first named participant.

Also, a pupil size normalization procedure was applied on each individual trial, dividing pupil size data by the mean baseline value, defined as a period equivalent to 25% of the normalized time before picture presentation. Pupil size was then sampled in bins of 1/100 of the normalized time. We submitted each bin to one-way analyses of variance (ANOVAs) and  $t$  tests to compare structural conditions across all subjects. This implies 100 comparisons for each test, only before the onset of vocalization. We used a window-based correction (Dehaene et al., 2001; Kamienkowski, Ison, Quiroga, & Sigman, 2012; Luck, 2005), filtering these multiple comparisons across time samples with the following criteria: (a) We kept only samples with  $p < .01$ ; (b) a given time sample was considered significant if it was part of a cluster of 10 or more consecutive significant time samples (10% of the vocalization reaction time).

## Results

### *Analysis of vocal responses*

In order to test the effectiveness of the colour-structure association established in the training task, we measured the fraction of incorrectly used structures among the responses. From a grand total of 2160 trials, only 47 (2.18%) were rejected due to wrong structure. We also rejected 96 trials due to no answer (7 trials), nonfluent utterances (55 trials), inverted roles (16 trials), or response time outliers (3 standard deviations, 18 trials).

For each trial, we measured the response time (RT) as the time elapsed from the presentation of the picture to the vocalization of the first named participant (Figure 3). We submitted the data to an ANOVA analysis with timing (planned and unplanned conditions) and structure (active, passive, CLLD) as independent factors. We observed a significant effect of the timing, significant effect of structure only in  $F_2$  analysis, and no significant interaction [ $F_1$ : timing,  $F(1, 114) = 19.75$ ,  $p < .0001$ ; structure,  $F(2, 114) = 1.67$ ,  $p = .19$ ; interaction,  $F(2, 114) = 0.20$ ,  $p = .82$ ;

$F_2$ : timing,  $F(1, 156) = 119.33$ ,  $p < .0001$ ; structure,  $F(2, 156) = 8.85$ ,  $p = .0002$ ; interaction,  $F(2, 156) = 1.41$ ,  $p = .25$ ]. The disparity of  $F_1$  and  $F_2$  in the effect of structure could be explained by the difference in the variabilities when grouping by items or subjects (see standard errors in Table 2).

### *Analysis of pupil diameter*

We concentrate our analysis on the period that begins with the presentation of the image and ends with the vocalization onset (RT), since we expect most of the planning to occur in this time. Pupil diameter in normalized time ramps up after the presentation of the image, for all time conditions and all structure types (TEPR, Figure 4, left panels). Furthermore, some differences arise among conditions. In the planned condition, the increase in pupil size begins earlier, at image presentation. In the unplanned condition the increase begins later, at about the midpoint between image presentation and vocalization onset.

More interestingly, we found consistent differences between sentence types. Pupil growth was larger for the more complex structures (passive and CLLD) than for the simpler (active) condition. Moreover, structure differences were observed earlier for the planned than for the unplanned condition. These differences are more evident when subtracting the active from the other two structures ( $\Delta$ TEPR, Figure 4, right panels). However, since the temporal course of sentence planning differs between planned and unplanned conditions, we considered it inadequate to directly compare the time course of these two conditions. Hence we submitted TEPR of each timing condition (planned and unplanned) to separate one-way ANOVAs with the described window-based correction, for the three levels of structure (active vs. passive vs. CLLD). Significant effects of structure were seen for both planned and unplanned conditions, but these effects rose later for the unplanned condition than for the planned condition ( $F_1$ : planned, significant in the period 0.18–0.34 in normalized time; unplanned, significant in the period 0.51–0.65;  $F_2$ : planned, significant in the period



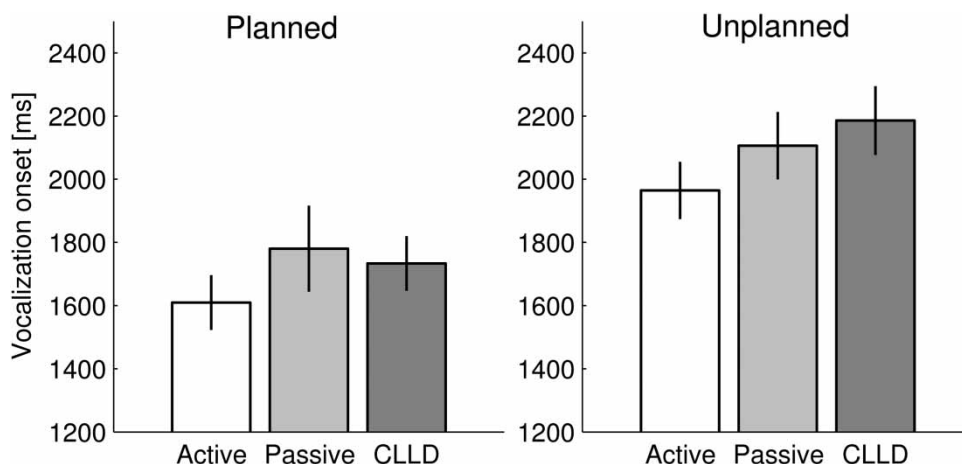


Figure 3. Response times for all timing and structure conditions. Error bars are standard errors of subjects' means. CLLD = clitic left dislocated sentences.

0.14–0.42; unplanned, significant in the period 1.16–1.24).

To further dissect these effects, we performed direct comparisons of all three pairs of structures using *t*-tests using the same window-based correction. In the planned condition, the passive–active difference was significant in the period 0.15–0.30, and the CLLD–active differences were significant in the period 0.15–0.42 in normalized time (see top bars in right panels of Figure 4). In the unplanned condition, the passive–active difference was significant in the period 0.78–1.15, while the CLLD–active difference was significant in the period 0.48–1.83. A similar pattern of results was obtained under  $F_2$  analysis (planned: passive–active significant at 0.14–0.30, CLLD–active at 0.12–0.99; unplanned: passive–active significant at 0.82–1.04 if significance is changed to .05, CLLD–active at 0.84–1.38). The CLLD–passive difference did not reach significance in any timing condition.

We also analysed the pupillary responses obtained during the training phase, in order to assess whether differences among structures are also present in a blocked presentation. TEPR results obtained during training phase were then submitted to the same analyses of the other timing conditions. In this case, no significant differences emerged for any pair of structures.

## Discussion

This work is, as far as we know, the first study applying pupillometry to sentence production. We measured TEPR of participants as they orally described depicted scenes of simple transitive events. A consistent increase in pupil size related to the task was observed as people planned and produced spoken picture descriptions using different sentence structures, showing that pupil size dynamics may be a reliable measure to investigate the cognitive processes involved in sentence production.

The first aim of our study was to investigate the processing load connected to syntactic operations in sentence production. In particular, we studied pupillary dynamics as a proxy for cognitive effort associated with the particular kind of movement operations that eventually give rise to noncanonical thematic order, as they are supposed to be computationally more demanding. Following a colour cue presented simultaneously (unplanned condition) or before the picture (planned condition), participants produced active canonical sentences, and noncanonical passive or CLLD sentences. As predicted, higher increase in pupil size was observed when participants produced passive or CLLD sentences than when they produced active sentences. Arguably, there is a cost involved in switching

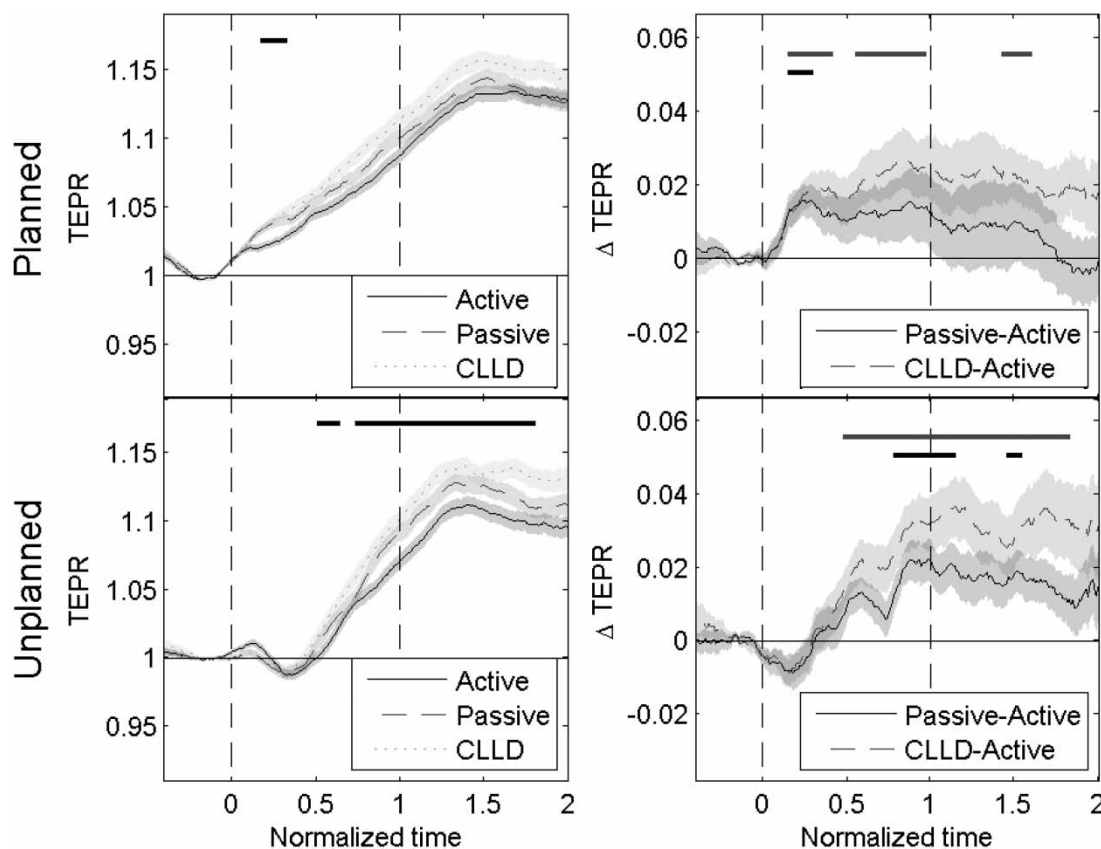
**Table 2.** Mean vocal responses

Type of analysis	Timing condition	Active	Passive	CLLD
$F_1$ by subjects	Planned	1610 (87)	1780 (136)	1733 (86)
	Unplanned	1964 (91)	2106 (107)	2186 (109)
$F_2$ by items	Planned	1606 (42)	1763 (41)	1714 (47)
	Unplanned	1955 (35)	2086 (41)	2171 (45)

Note. Standard errors in parentheses. Response time expressed in milliseconds for each condition. CLLD = clitic left dislocated sentences.

conditions and making the proper association that could be reflected in pupil size changes. This cost is, however, equivalent for all conditions. Thus,

the differential TEPR increase suggests that producing a sentence that involves noncanonical order is harder than producing structures in the subject-



**Figure 4.** Pupillary responses, for all time and structure conditions. Normalized time = 0 corresponds to picture presentation, and normalized time = 1 corresponds to vocalization onset of the first named participant. Left panels: task evoked pupillary responses (TEPR). The bars at the top correspond to periods of significance of each difference (analysis of variance, ANOVA, with window-based correction,  $p < .01$ ). Right panels:  $\Delta$ TEPR, subtracting active condition. The bars at the top correspond to periods of significance of each difference ( $t$ -test with window-based correction,  $p < .01$ ). Black bar corresponds to passive-active difference, and grey bar corresponds to CLLD-active difference (CLLD = clitic left dislocated sentences). Shaded areas are standard errors of subjects' means.

verb–object canonical thematic word order in Spanish (Figure 4).

A series of studies seems to confirm parallel results in the sentence comprehension domain (Engelhardt et al., 2010; Just & Carpenter, 1993; Scheepers & Crocker, 2004; Schlurhoff, 1982). While these works reveal uniformly that some structures are more difficult to process than others, no one single interpretation emerges from the studies. Just and Carpenter (1993) proposed that pupillary response is an indicator of how intensely the processing system is operating and interpreted the effect in terms of memory load. Schlurhoff (1982), in his paper, considered pupil size changes over time as a reflection of the difficulty of processing resulting from internal properties of the sentences. In his study, syntactic complexity (independently measured) was shown to be a better predictor of cognitive effort than sentence length. As in the results of these studies, our data showed that structural configurations differentiated each other in terms of the pupillary changes that they elicited, suggesting increased processing cost in sentences with noncanonical thematic order, as indicated by larger TEPR for passive and CLLD sentences. Whether the effect reflects memory load demands or more specific cognitive operations remain unclear from our results.

Also, these findings seem to fit well with previous works on the cost of producing active versus passive sentences (F. Ferreira, 1994; Tannenbaum & Williams, 1968). Together, these studies showed faster latencies for actives than for passive sentences and proposed differential processing effort for these structures. However, in our study, we found no significant differences between structural conditions in speech latencies, most likely due to large variability of the subjects' data. Overall, speech onset times in sentence production tasks have proven to be a quite elusive measure (Allum & Wheeldon, 2007; Griffin & Spieler, 2006). As it has been extensively discussed in the literature, response times in sentence production show an enormous variability, in part because speakers could adjust the scope of their planning according to different factors, both external and internal, like time pressure, cognitive load, or the

particular strategies used to solve a task efficiently (F. Ferreira & Swets, 2002; Konopka, 2012; Wagner, Jescheniak, & Schriefers, 2010). On the contrary, pupil dilation occurs automatically and with relative independence of the task strategy. This suggests that TEPR may be a more sensitive measure of cognitive cost than RT for studying sentence production and therefore the processes involved in grammatical planning.

We also measured pupil size changes during the training, where subjects were presented with blocks of pictures to be described using only one sentence type per block. No differences emerged in pupil dynamics between structural conditions, indicating that none of the three sentential structures imposed more processing demands on the participants than the others during training. There are two possible explanations for this lack of structural effect. It is possible that, as a result of the blocked design, no online planning is required to perform the task. According to this interpretation, the cognitive effort observed in the experimental trials is due to differential demands during planning and does not emerge if structure can be preplanned. These results go in line with Corrêa's finding regarding processing load in the production of Portuguese subject versus object relative clauses in planned and unplanned conditions (Corrêa, Augusto, & Marcilese, 2009).

Under an alternative explanation, the blocked presentation of the trials in the training phase might reduce the competition between structures, hence reducing the cognitive effort associated with deciding what structure to use to describe the depicted event. According to this competition view, when available, different syntactic frames become simultaneously preactivated and compete with one another for selection (Myachykov, Scheepers, Garrod, Thompson, & Fedorova, 2013). Thus, the speaker does not only need to select the preferred structure among competitors, but also needs to inhibit the nonpreferred. As no alternatives are available during training, no extra processing cost is observed. This kind of explanation would account for the structural effect of the experimental trials in terms of the differential effort addressed to suppress the nonchosen

structure. Under this view, active structure might be harder to inhibit, leading to bigger effort and larger pupil growth for passive and CLLD sentences.

To further investigate whether the syntactic processes that lead to these different structural configurations can be dissociated from lexical retrieval, we manipulated the time at which the colour cue that triggered the structure-building processes was presented. In the unplanned condition, where both picture and colour frame were presented simultaneously, vocal responses were consistently delayed, as compared to the planned condition, where the frame was presented 1000 ms before the picture (Figure 3). In the planned condition, the pupil begins to dilate earlier than in the unplanned condition, and the difference between structural conditions (active versus passive and CLLD sentences) also rises earlier. Moreover, in the planned condition, the pupil begins to dilate as soon as the image appears (Figure 4). As observed by Scheepers and Crocker (2004), pupillary responses are relatively slow (with a latency of about 200–400 ms), but still fast enough to enable the identification of their triggering events (see also Beatty & Lucero-Wagoner, 2000). Thus, it can be inferred that the reflected cognitive effort has started before the presentation of the picture, namely before the information relevant for lexical access of the first word of the sentence is available. More specifically, in the planned condition, the difference between structural conditions rises early (0.18 in normalized time, or about 250 ms). Since this time is too short to be attributable to a pupillary response to the image, it seems that participants can make use of the specific information about the type of structure before the image is processed. Our results suggest that, if necessary, structure processes can start working even before the lexical access corresponding to the first word of the sentence has been initiated. These results go in line with previous findings both in comprehension (Frazier, 1995) and in production (Bock & Loebell, 1990; Konopka & Bock, 2009) and support the idea that speakers must have autonomous syntactic mechanisms that can generate abstract sentence representations for utterances.

Even though syntactic choice has not been the focus of our analysis, it is obvious that in normal language, speakers produce different word orders, including noncanonical sentences, for a variety of reasons, most of them context dependent. This fact underlines the need to expand these findings to less constrained situations, essentially to tasks that do not force speakers to produce particular structures. More naturalistic situations, in which the different syntactic structures are motivated by discourse context, may serve to disentangle the cognitive effort caused by structure building from other factors affecting cognitive effort in sentence production. Moreover, it is possible that multiple factors—including computational complexity, frequency of use, lexical bias, and context appropriateness, as well as individual differences—cooperate in the observed processing difficulty. Hence, integrative approaches are needed to establish specific contributions to cognitive effort in sentence production.

Original manuscript received 23 August 2013

Accepted revision received 18 February 2014

First published online 6 May 2014

## REFERENCES

- Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 33(4), 791–810.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91(2), 276–292.
- Beatty, J., & Lucero-Wagoner, B. (2000). Pupillary system. Chapter 6. In J. T. Cacioppo, L. G. Tassinary & G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 142–161). Cambridge: Cambridge University Press.
- Bock, K. (1990). Structure in language. Creating form in talk. *American Psychologist*, 45(11), 1221–1236.
- Bock, K., & Levelt, W. L. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). London: Academic Press.
- Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition*, 35(1), 1–39. doi:0010-0277(90)90035-I[pii]

- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Corrêa, L. M. S. (2011). Minimalist perspectives for psycholinguistic research. *Avances en Psicología Latinoamericana*, 29, 214–230.
- Corrêa, L. M. S., Augusto, M., & Marcilese, M. (2009). Resumptive pronouns and passives in the production of object relative clauses: Circumventing computational cost. Paper presented at the CUNY: Conference on Human Sentence Processing, University of California.
- Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J. F., & Poline, J. B. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neurosciences*, 4(7), 752–758.
- Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly Journal of Experimental Psychology*, 63(4), 639–645. doi:917841132[pii]10.1080/17470210903469864
- Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33(6), 715–736.
- Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46(1), 57–84.
- Ferreira, V. S., & Slevc, L. R. (2007). Grammatical encoding. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 453–469). Oxford: Oxford University Press.
- Frazier, L. (1995). Constraint satisfaction as a theory of sentence processing. *Journal of Psycholinguistic Research*, 24(6), 437–468.
- Gennari, S. P., Mirkovic, J., & Macdonald, M. C. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology*, 65(2), 141–176.
- Griffin, Z. M., & Spieler, D. H. (2006). Observing the what and when of language production for different age groups by monitoring speakers' eye movements. *Brain and Language*, 99(3), 272–288. doi:http://dx.doi.org/10.1016/j.bandl.2005.08.003
- Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, 47(2), 310–339.
- Kamienkowski, J. E., Ison, M. J., Quiroga, R. Q., & Sigman, M. (2012). Fixation-related potentials in visual search: A combined EEG and eye tracking study. *Journal of Vision*, 12(7), 4. Retrieved from <http://www.journalofvision.org/content/12/7/4.full>
- Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, 66(1), 143–162.
- Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, 58(1), 68–101. doi:S0010-0285(08)00037-6[pii]10.1016/j.cogpsych.2008.05.002
- Kuchinke, L., Vo, M. L., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, 65(2), 132–140. doi:10.1016/j.ijpsycho.2007.04.004
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral Brain Science*, 22(1), 1–380; discussion 38–75.
- Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- Macdonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4, 226. Retrieved from <http://journal.frontiersin.org/Journal/10.3389/fpsyg.2013.00226/full>
- Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: The case of English and Russian. *Quarterly Journal of Experimental Psychology*, 66(8), 1601–1619.
- Papesh, M. H., & Goldinger, S. D. (2012). Pupil-BLAHmetry: Cognitive effort in speech planning reflected by pupil dilation. *Attention, Perception, & Psychophysics*, 74(4), 754–765. doi:10.3758/s13414-011-0263-y
- Phillips, C. (1996). *Order and structure*. (Ph D), Massachusetts Institute of Technology. Retrieved from <http://hdl.handle.net/1721.1/10666>
- Phillips, C., & Lewis, S. (2013). Derivational order in syntax: Evidence and architectural consequences. *Studies in Linguistics*, 6, 11–47.
- Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, 39(4), 859–862. doi:10.3758/BF03192979
- Scheepers, C., & Crocker, M. W. (2004). Constituent order priming from reading to listening: A visual-world study. In M. Carreiras & C. J. Clifton (Eds.), *The on-line study of sentence comprehension: eyetracking, ERP, and beyond* (pp. 167–185). New York: Psychology Press.

Schluroff, M. (1982). Pupil responses to grammatical complexity of sentences. *Brain and Language*, 17(1), 133–145.

Tannenbaum, P. H., & Williams, F. (1968). Generation of active and passive sentences as a function of subject or object focus. *Journal of Verbal Learning and Verbal Behavior*, 7(1), 246–250.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36(2), 423–440. doi:10.1037/a0018619

## APPENDIX

### List of events

#### *Experimental items*

Acariciar (to caress)  
 Afeitarse (to shave)  
 Apedrear (to throw stones)  
 Aplaudir (to applaud)  
 Arañar (to scratch)  
 Arrastrar (to pull)  
 Atar (to tie)  
 Atrapar (to catch)  
 Castigar (to punish)  
 Empujar (to push)  
 Escupir (to spit out)  
 Espiar (to spy)  
 Filmar (to film)  
 Fotografar (to photograph)  
 Golpear (to punch)  
 Hamacar (to swing)

Manchar (to stain)  
 Maquillar (to make up)  
 Mojar (to wet)  
 Olfatear (to sniff)  
 Patear (to kick)  
 Peinar (to comb)  
 Pinchar (to prick)  
 Pisar (to step on)  
 Rescatar (to rescue)  
 Salpicar (to splash)  
 Saludar (to greet)

#### *Practice items*

Abrazar (to hug)  
 Bañar (to bath)  
 Besar (to kiss)  
 Enterrar (to bury)  
 Morder (to bite)  
 Perseguir (to chase)