Preprint of an article published in Journal of Integrative Neuroscience,
Online Ready, pp. 1-12, 2017.
http://www.worldscientific.com/worldscinet/jin

doi: 10.1142/S0219635216500266

Pupil size behavior during on line processing of sentences

G. Fernández*[1], J. Biondi*[1,2], S. Castro[2], and O. Agamennoni[1,3]

[1]Universidad Nacional del Sur (UNS), Instituto de Investigaciones en Ingeniería Eléctrica (IIIE) (UNS-CONICET), Bahía Blanca, Buenos Aires, Argentina.
[2]Universidad Nacional del Sur (UNS), Departamento de Ciencias e Ingeniería de la Computación, Laboratorio de visualización y computación gráfica (VyGLab), Bahía Blanca, Buenos Aires, Argentina.
[3]Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC), Argentina

Abstract

In the present work we analyzed the pupil size behavior of forty subjects while they read well defined sentences with different contextual predictability (i.e., regular sentences and proverbs). In general, pupil size increased when reading regular sentences, but when readers realized that they were reading proverbs their pupils strongly increase until finishing proverbs' reading. Our results suggest that an increased pupil size is not limited to cognitive load (i.e., relative difficulty in processing) because when participants accurately recognized words during reading proverbs, theirs pupil size increased too. Our results show that pupil size dynamics may be a reliable measure to investigate the cognitive processes involved in sentence processing and memory functioning.

Keywords: Reading; Proverbs; Memory; Pupil size.

1 Introduction

The relationship between cognitive load (i.e. total amount of mental effort) and pupil dilation have attracted the attention of many researchers (See Laeng et al. (2012) for an overview). Although these studies differ in how cognitive load is implemented, they showed that there exist a strong direct correlation between cognitive load and pupil size. In a non-linguistic context, Preuschoff et al. (2011) concluded that pupil size (and therefore, presumably, cognitive load) increases when a stimulus is less expected. Preuschoff et al. (2011) studied pupil dilation while participants performed a simple gambling task. He found that pupil size correlated with unexpectedness, not with the gambling outcome itself (i.e. surprise causes pupil dilation).

Whether unexpectedness of words in sentences also results in pupil dilation is still an open question. When reading a sentence, each current word is integrated with the past words and predictions about upcoming words are generated (e.g., Just et al. (1982); Fernández et al. (2014b, 2015)). The amount of cognitive effort required to process a given word reflects the interplay of word processing and expectancy driven processes (Rayner, 1998; Kliegl et al., 2006). Esterman and Yantis (2009) showed that contextual hints facilitate the information processing. Gazzaley (2013) concluded that working memory performance in improved by contextual hints. The relationship between low predictable upcoming words and cognitive load has been observed in reading studies; the time needed to be read is directly correlated with its surprising values, which is related to different sentence comprehension phenomenas, like garden-path effect (Brouwer et al., 2010) and anti-locality effects (Levy, 2008; Frank and Thompson, 2012). In reading, gaze time duration at each word is correlated with surprising and with low predictable upcoming words (e.g., Boston et al. (2008); Demberg and Keller (2008); Smith and Levy (2008); Fernández et al. (2014b)).

*Gerardo Fernndez and Juan Biondi contributed equally to this work.

Nevertheless, there are only a few works in psycholinguistics analyzing pupil's behavior during reading sentences. Engelhardt et al. (2010) showed that, in comparison with the matching situation, a mismatch between the syntactic and prosodic structure of auditorily presented sentences increase pupil size. Piquado et al. (2010) found a pupil response to syntactic complexity and sentence length in a sentence-listening study.

To the best of our knowledge, there exist only a few published studies in which pupillometry is applied during reading well-defined sentences. Raisig et al. (2012) developed a study with written descriptions of simple events in everyday activities. When the order of presentation was incongruent with the actual temporal order of the described activities they observed an increase in pupil dilation. Just et al. (1982) evaluated reading time and pupil size in a study were they compared object and subject-relative clauses. They found increased reading times and pupil dilatation on the object-relatives that are known to be more difficult to process (Hakes et al., 1976). Moreover, a semantically implausible word increased pupil size compared to a plausible-word. Conversely, Papesh et al. (2012), encounter little perceptual resistance when processing the same words spoken by different speakers, each of whom has a unique vocal structure, speaking rate and pattern of intonation. Further, participants showed that encoding effort of processing spoken words is related to subsequent memory strength. This effect was not limited to encoding, since when participants recognized old items during test, their pupils were again more dilated. In the present study, we examined pupil behavior during on-line sentences processing. We examined whether word properties embedded in sentences with different contextual predictabilities (i.e., regular sentences and proverbs) affect pupillary responses. Using proverbs as reading material allows to examine whether a semantic context facilitates reading processes (Katz and Ferretti, 2001; Fernández et al., 2014b, 2015). Then, it would be possible to counter check whether two kinds of stimulus, an easier one (proverbs) and a less facilitated (regular sentences), affect pupil responses.

2 Methods

2.1 Participants

The group of readers consisted of forty persons (mean age: 58, $SD = 4 : 1$) with professional qualification and no evidence of cognitive decline or impairment in activities of academic and daily living. The mean education trajectory of the readers was 16.1 years ($SD = 1 : 0$). To assess whether subjects comprehended the texts, they were presented with a three alternative multiple-choice question about the sentence in progress on 20% of the sentence trials. Participants answered the questions by moving a mouse and choosing the response with a mouse click. Overall mean accuracy was 96% ($SD = 3 : 1\%$). For a sentence corpus description see Fernández et al. (2014b). For Apparatus technical specifications and for eye movement data analyses see Fernández et al. (2015).

2.2 Analytical strategy for pupil data

We concentrated our analysis on the period that begins with the presentation of the first word of the sentence and ends with the last word. 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 sentence presentation.

2.3 Statistical analysis

Our analyses are based on linear mixed models (LMMs). We used the *lmer* function of the *lme4* package (version 0.999999-2) (Bates and Maechler, 2013) for estimating fixed and random coefficients. This package is supplied in the R system for statistical computing under the GNU General Public License (Version 2, June 1991). The dependent variable was the normalized pupil diameter. Fixed effects in LMM terminology correspond to regression coefficients in standard linear regression models. They can be used to estimate slopes or differences between conditions. The following fixed effects were entered into the model: logit predictability, log frequency, 1/word length. We specified contrasts of sentence type (proverbs vs. regular sentences). In addition, we allowed for varying slopes of pupil diameter with participants and items (sentences) by setting random slopes for participants and items. Instead of estimating differences between conditions, random effects estimate the variance that is associated with the levels of a certain factor. For the LMMs we report regression coefficients (bs), standard errors (SEs), and t-values ($t = b/SE$). Our criterion for referring to an effect as signifficant is $t = b/SE > 1 : 96$.

3 Results

Table 1: Parameter estimates for fixed effects of Linear Mixed Models. Threshold of significance is set at $t = 1.96$.

		Pupil size	
	M	**SE**	**t-value**
Fixed effects			
Normalized Pupil diameter	0.988	0.050	**19.49**
Word Number	0.000	0.001	0.20
Predictabilities (logit)	-0.013	0.003	**-3.62**
Frequencies (log)	-0.005	0.002	**-2.61**
1/Length (characters)	-0.077	0.032	**-2.37**
Sentence type			
Proverbs vs Regular sent.	0.024	0.012	**1.99**
Proverbs vs Regular sent. x Word Number	0.003	0.000	**-2.46**
Proverbs vs Regular sent. x Predictabilities (logit)	0.015	0.005	**3.03**
Proverbs vs Regular sent. x Frequencies (log)	0.002	0.003	1.38
Proverbs vs Regular sent. x Length (characters)	-0.010	0.016	-0.61
Variance components		**Variance**	**SD**
Sentence (n=140)		0.000	0.017
Subject (n=40)		0.093	0.305
Residual (n=11215)		0.006	0.079

In Table 1 we report, (a) main effects when averaging over all predictors i.e., collapsing sentence type (proverbs vs. regular sentences); (b) interactions of word properties, and word number x sentence type. The normalized pupil diameter as a function of sentence type, and word properties are displayed in Figures 1, 2, 3 and 4. The values in Figures are partial effects (for an example of this technique, see Hohenstein and Kliegl (2014)).

As shown in Table 1, the mean pupil diameter significantly increased when reading regular sentences ($t = 1.99$) compared to proverbs, it seems that regular sentences increased readers' cognitive load. When we evaluated whether word properties -averaging over all predictors- affected pupil responses we observed how 1/word length, word frequency and word predictability exerted a significant effect on pupil dilatation ($t = -2.37$, $t = -2.61$, $t = -3.62$, respectively). As shown in Figure 1, longer words increased pupil dilatation because probably longer words are more difficult to process. Quite the contrary, both more frequent and more predictable words decreased the pupil diameter. Word number is not affecting significantly the pupil size when considering averaging collapsing all predictors ($t = 0.20$) (See Table 1 and Figure 4).

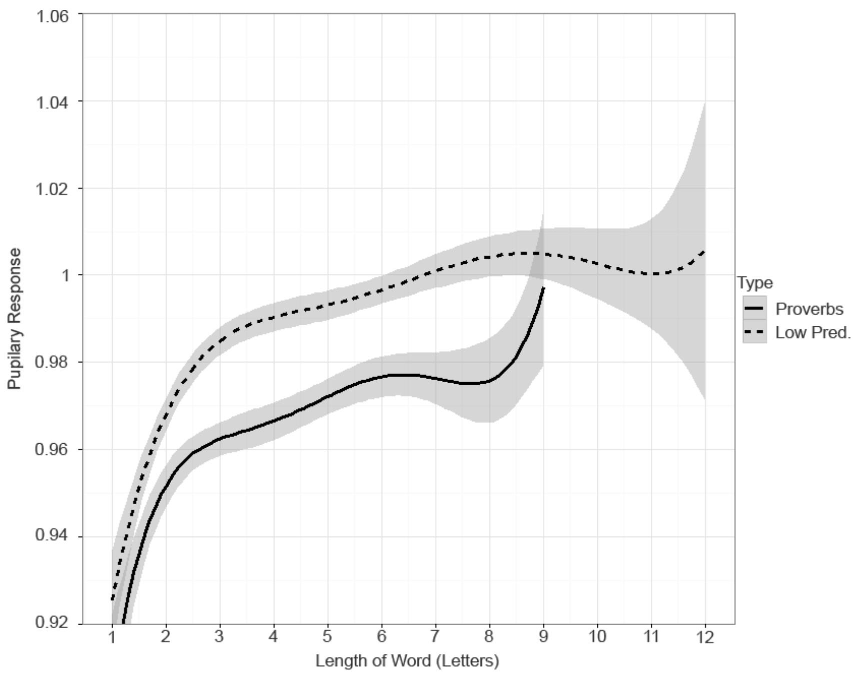

Figure 1: Effect of the length of word on normalized pupil dilatation, broken down by regular sentences and proverbs. Panel reflects regression of normalized pupil dilatation on word on respective length. Shaded areas are 95% confidence intervals.

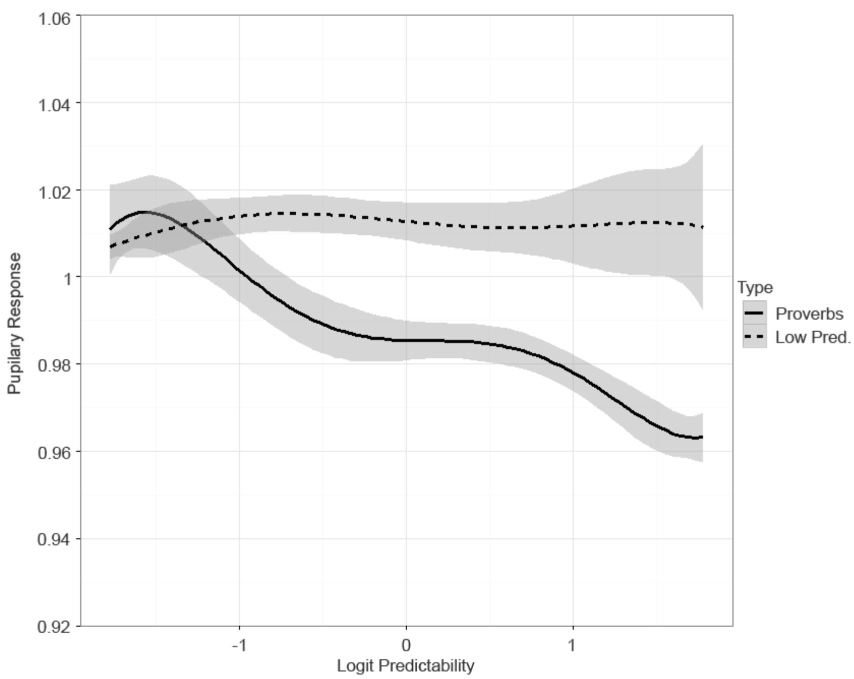

Figure 2: Effect of the predictability of word on normalized pupil dilatation, broken down by regular sentences and proverbs. Panel reflects regression of normalized pupil dilatation on word on respective logits of predictability. Shaded areas are 95% confidence intervals.

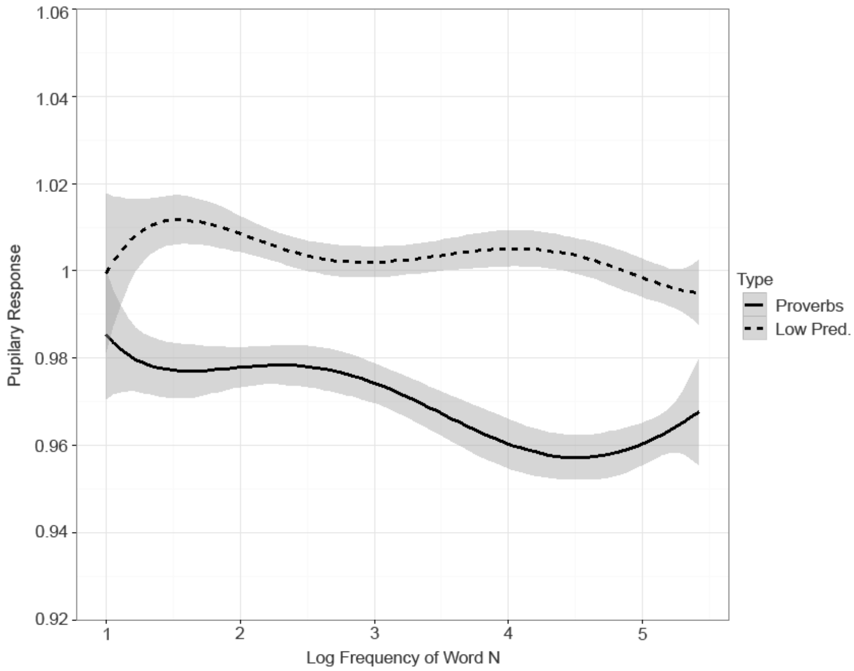

Figure 3: Effect of the frequency of word on normalized pupil dilatation, broken down by regular sentences and proverbs. Panel reflects regression of normalized pupil dilatation on word on respective log of frequency. Shaded areas are 95% confidence intervals.

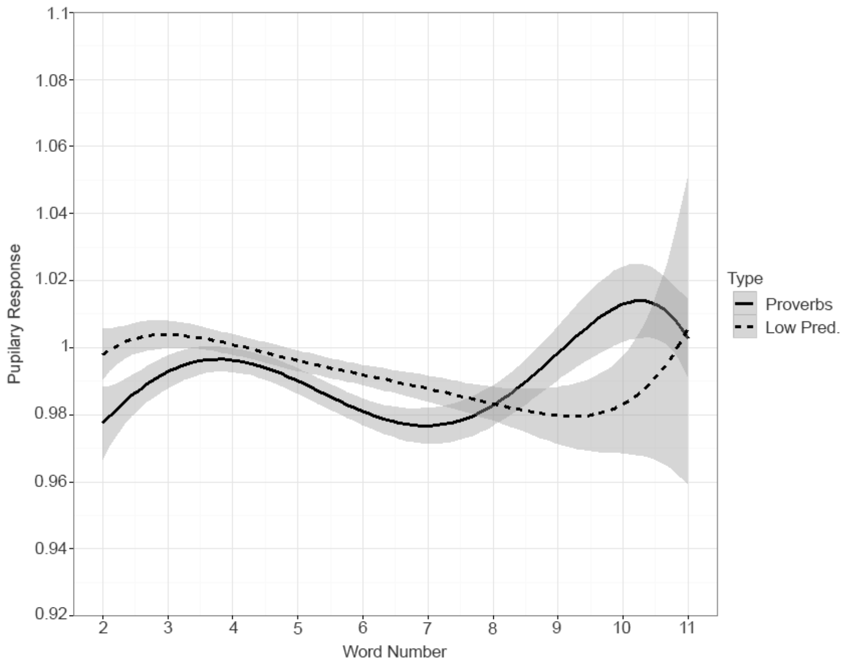

Figure 4: Effect of number of word in sentences on normalized pupil dilatation, broken down by regular sentences and proverbs. Panel reflects regression of normalized pupil dilatation on word on respective number of word. Shaded areas are 95% confidence intervals.

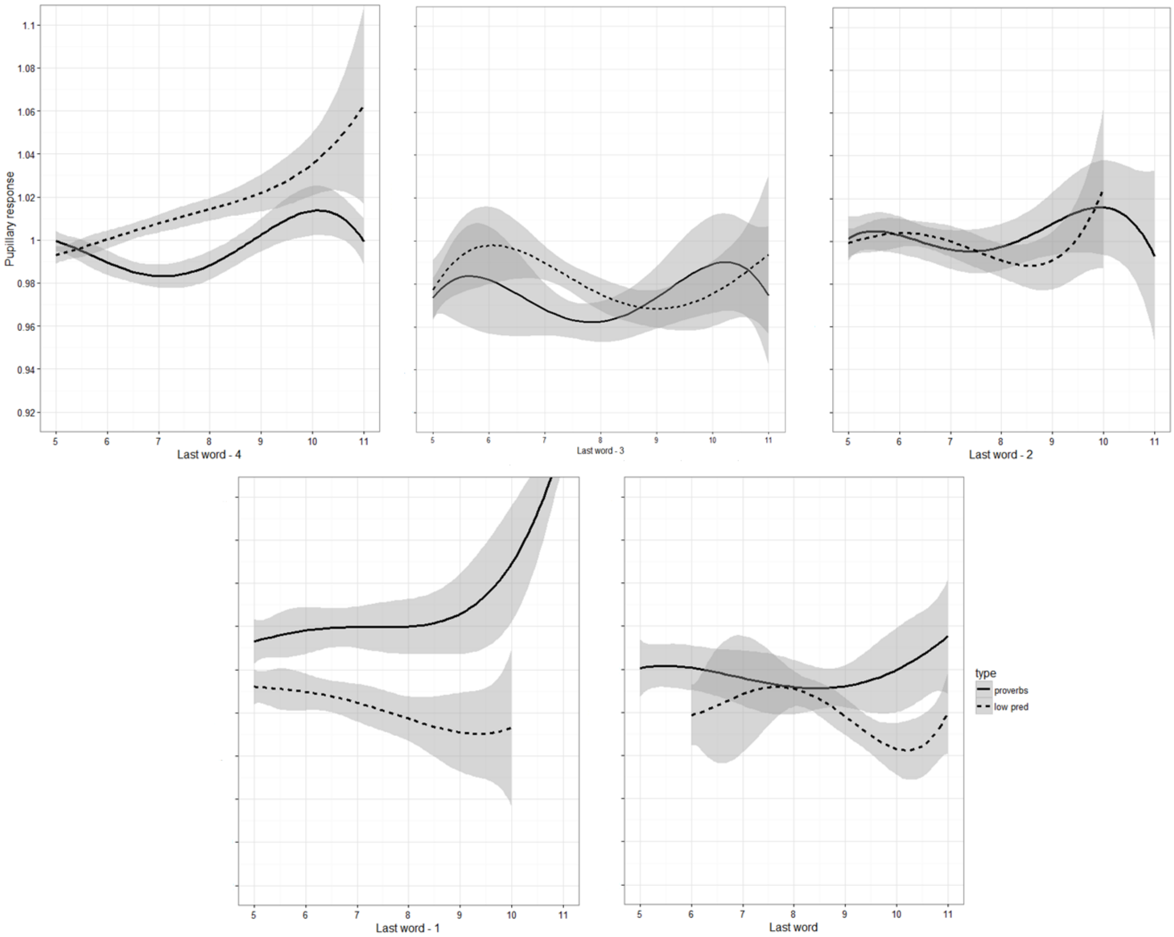

Figure 5: Effect of word position in sentences on normalized pupil dilatation, broken down by regular sentences and proverbs. Panel reflects regression of normalized pupil dilatation on word on respective word position in the sentence. Shaded areas are 95% confidence intervals.

When analyzing predictor per sentence type interaction (i.e., regular sentences vs. proverbs) we noted a significant and interesting effect of word number on pupil size ($t = -2.46$): regular sentences decreased readers' pupil size when comparing with proverbs. Proverbs increased pupil size from the middle to the end of the sentences (See Table 1 and Figure 4). Given that different proverbs have different word lengths and because the pupil increase tends to happen from the middle of the sentence we plotted a Figure that counts word position from the end of the sentence. In this way, the rightmost tick mark on the axis will always refer to the last word in the sentence/proverb, and the tick mark to its left refer to the second-to-last word (i.e., last word -1), and so on. Thus, when analyzing last word processing we appreciated that both sentences produced a similar effect on pupil dilatation although proverbs marginally increased pupil dilatation. Strikingly, the effect is different when analyzing *Last word -1*, where proverbs clearly increased pupil dilatation. It seems that on these words readers recognize proverbs producing a memory cue with an impact on pupil dilatation (see Discussion for an explanation about the "Eureka moment"). *Last word -2* and *word -3* showed similar pupil behavior when processing words, although *Last word -3* in regular sentences evidenced a small increase on pupil dilatation when comparing with proverbs. On the other hand, *Last word -4* showed that regular sentences produced a stronger increase on pupil dilatation. It seems that in *Last word -4* proverbs are not yet identified and words in regular sentences produced a cognitive load increasing readers' pupil size. Finally, only word predictability increased significantly pupil size where regular sentences showed an increased pupil when comparing with proverbs ($t = 3.03$) (See Table 1 and Figure 2).

4 Discussion

This work is, as far as we know, the first one analyzing pupil behavior during reading well defined words embedded in sentences with different contextual predictability. We measure pupil response of participants as they read regular sentences and proverbs. Using sentences with different contextual predictability allowed us to check whether readers devoted greater cognitive effort to process word information and to use contextual hints for improving their reading performance.

In the last 50 years, a large body of studies has confirmed that the pupil behavior changes depending of the task difficulty (Beatty and Kahneman, 1966; Bradshaw, 1968; Hyönä et al., 1995). In general, less effort decreased mean pupil dilatation. This common proxy between a reduced effort and a decreased pupil dilatation seems to be the pattern when analyzing longer words and low-predictable words (see Figures 1, 2). In general, when more effort is required for processing words there is an increase in the mean pupil dilatation. Interestingly, when analyzing regular sentences vs. proverbs emerges a particular pupil behavior. It seems that readers' pupils increase when attentional cues (e.g, knowed words) are provided by proverbs (see Figure 4). As in Papesh et al. (2012), the pupil behavior was sensitive to the content of memory: sentences in which contextual predictability was stronger yielded the largest pupil dilatation. In accordance with Gazzaley (2013), predictive cueing resulted in improved working memory performance and it seems that this phenomenon increase pupil dilatation. Just and Carpenter (1993) proposed that pupillary response as an indicator of how intensely the processing system is operating and interpreted the effect in terms of memory load. Our results suggest that in both, regular sentences and proverbs, the memory load is present when analyzing readers' pupil behavior (see Figures 1, 2 and 3). Additionally, this memory effect seems to be present from the middle of the sentences when analyzing proverbs and on the last words when looking at regular sentences (see Figure 4 and Table 1). Schluroff (1982), considered pupil size changes over time as a reflection of the difficulty of processing resulting from internal properties of the sentences. Our results suggest that both cognitive load in first place and task facilitation when retrieving upcoming words in the second one, increase pupil dilatation. Using regular sentences and proverbs allowed us to explicitly vary the dynamics of memory processing during reading, because highly predictable words in proverbs active top-down processes for predicting upcoming words (See Fernández et al. (2014b) for sentences' description) words.

Alnaes et al. (2014) propose that cognitive factors, such as prior knowledge and expectations, and top-down attentional control interact with incoming sensory signals. They also affirm that this interaction may produce a bias in the competition between objects for access to the working memory (Corbetta and Shulman, 2002; Desimone and Duncan, 1995). The researchers propose that pupillary responses reflect the intensity of mental operations and the allocation of attention across a range of different tasks. Alnaes et al. (2014) observed pupil associated activity in the Locus Coeruleus (LC). The LC is a small brain stem nucleus located in the rostral pons. The noradrenergic projections of the LC are sent to virtually all brain regions with major density to areas known to be important in attentional processing. Further, pupil dilatation associated with cognitive processing are thought to result from an inhibitory effect on parasympathic oculomotor complex generated from the LC (Alnaes et al., 2014). Furthermore, the LC -norepinephrine (NE) based modulatory mechanism can help to establish attentional shifts of either, external or internal stimuli where one event becomes more relevant than each other (Laeng et al., 2012). The LC is also engaged during the process of memory retrieval (Eschenko and Sara, 2008). Posner and Fan (2008) have distinguished between alerting, orienting, and executive networks of the brain. In their model, the alerting network is innervated by the NE system and includes the LC, right frontal cortex, and regions of the parietal cortex. Thus, NE plays a crucial role in energizing the cortical system and promoting adequate levels of activation for cognitive performance. Similar to Papesh et al. (2012), our results suggest that when subjects need to allocate more resources for integrating current words and for predicting upcoming words i.e., when the Eureka moment emerge in proverbs Fernández et al. (2014b) (see Figures 4 and 5) an increase in their pupil sizes happens. As in previous works (Fernández et al., 2014b,a, 2015) it seems that an upcoming predictability effect in proverbs increase pupil diameter while it is retrieved from memory (see Figure 5). As reported (Laeng et al., 2012), the LC response occurs about 100 ms after a relevant event and it takes an additional 60-70 ms for the activity within the LC to reach frontal cortex. Such a delay from the triggering event to NE release at a cortical site is then 150-200 ms. This evidence suggest that pupillary responses could provide a signal of the moment in which the Eureka word emerges. Probing online comprehension processes during both regular sentences and proverbs and tracing their effects on pupil dilatation might give us a tool for the evaluation of cognitive effort and reading facilitation. Our work replicate Hyönä and Pollatsek (2000) results (see Figure 1), where longer words strongly increased pupil size irrespective of which kind of sentences we were analyzing. Finally, word frequency and word predictability influenced pupil responses. Previous researchers have documented that word frequency typically affects pupillary reflexes (Kuchinke et al., 2007; Papesh and S., 2008). As with word frequency, our results show that more predictable words decreased pupil diameter. Interestingly, the predictability effect was stronger for proverbs (see Figure 2 and Table 1).

To conclude, the present findings suggest that both regular sentences and proverbs increases and decreases pupil size. Our work suggests that proverbs are retained in memory, aiding subsequent perception and recognition. As indicated by pupillometry, proverbs strongly influence subjective feelings of memory strength and cognitive demand. Our results show that pupil size dynamics may be a reliable measure to investigate the cognitive processes involved in sentence processing.

Acknowledgments

We thank reviewer's comments that allowed us to improve our paper. This work was partially supported by grants PGI 24/K062, Universidad Nacional del Sur and PICT 2013 - 0403, Agencia Nacional de Ciencia y Tecnologa, Ministerio de Ciencia y Tecnologa, Argentina.

Preprint of an article published in Journal of Integrative Neuroscience, Online Ready, pp. 1-12, 2017.
doi: 10.1142/S0219635216500266

References

Dag Alnaes, Markus Handal Sneve, Thomas Espeseth, Tor Endestad, Steven Harry Pieter van de Pavert, and Bruno Laeng. Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of vision*, 14(4):1–1, 2014.

D Bates and M Maechler. Linear mixed-effects models using s4 classes. r package version 0.999999-2. *R: A Language and Environment for Statistical Computing*, 2013.

Jackson Beatty and Daniel Kahneman. Pupillary changes in two memory tasks. *Psychonomic Science*, 5(10):371–372, 1966.

Marisa Boston, John Hale, Reinhold Kliegl, Umesh Patil, and Shravan Vasishth. Parsing costs as predictors of reading difficulty: An evaluation using the potsdam sentence corpus. *The Mind Research Repository (beta)*, 2(1)(1):1–12, 2008.

JL Bradshaw. Pupil size and problem solving. *The Quarterly journal of experimental psychology*, 20(2): 116–122, 1968.

Harm Brouwer, Hartmut Fitz, and John CJ Hoeks. Modeling the noun phrase versus sentence coordination ambiguity in dutch: evidence from surprisal theory. In *Proceedings of the 2010 Workshop on Cognitive Modeling and Computational Linguistics*, pages 72–80. Association for Computational Linguistics, 2010.

Maurizio Corbetta and Gordon L Shulman. Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, 3(3):201–215, 2002.

Vera Demberg and Frank Keller. Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, 109(2):193–210, 2008.

Robert Desimone and John Duncan. Neural mechanisms of selective visual attention. *Annual review of neuroscience*, 18(1):193–222, 1995.

Paul E Engelhardt, Fernanda Ferreira, and Elena G Patsenko. Pupillometry reveals processing load during spoken language comprehension. *The Quarterly Journal of Experimental Psychology*, 63(4): 639–645, 2010.

Oxana Eschenko and Susan J Sara. Learning-dependent, transient increase of activity in noradrenergic neurons of locus coeruleus during slow wave sleep in the rat: Brain stem–cortex interplay for memory consolidation? *Cerebral Cortex*, 18(11):2596–2603, 2008.

Michael Esterman and Steven Yantis. Perceptual expectation evokes category-selective cortical activity. *Cerebral Cortex*, 5:1245–53, 2009.

Gerardo Fernández, Facundo Manes, Nora P Rotstein, Oscar Colombo, Pablo Mandolesi, Luis E Politi, and Osvaldo Agamennoni. Lack of contextual-word predictability during reading in patients with mild alzheimer disease. *Neuropsychologia*, 62:143–151, 2014a.

Gerardo Fernández, Diego E Shalom, Reinhold Kliegl, and Mariano Sigman. Eye movements during reading proverbs and regular sentences: The incoming word predictability effect. *Language, Cognition and Neuroscience*, 29(3):260–273, 2014b.

Gerardo Fernández, Liliana R Castro, Marcela Schumacher, and Osvaldo E Agamennoni. Diagnosis of mild alzheimer disease through the analysis of eye movements during reading. *Journal of integrative neuroscience*, 14(01):121–133, 2015.

Stefan L Frank and Robin L Thompson. Early effects of word surprisal on pupil size during reading. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, pages 1554–1559, 2012.

A Gazzaley. Top-down modulation deficit in the aging brain: an emerging theory of cognitive aging. *Principles of frontal lobe function*, pages 593–608, 2013.

David T Hakes, Judith S Evans, and Linda L Brannon. Understanding sentences with relative clauses. *Memory & Cognition*, 4(3):283–290, 1976.

Sven Hohenstein and Reinhold Kliegl. Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(1):166, 2014.

Hyönä and Pollatsek. Morphological processing of finnish compound words in reading. *Reading as a Perceptual Process*, pages 65–88, 2000.

Jukka Hyönä, Jorma Tommola, and Anna-Mari Alaja. Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology*, 48(3):598–612, 1995.

Marcel A Just and Patricia A Carpenter. The intensity dimension of thought: pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 47(2):310, 1993.

Marcel A Just, Patricia A Carpenter, and Jacqueline D Woolley. Paradigms and processes in reading comprehension. *Journal of experimental psychology: General*, 111(2):228, 1982.

Albert N Katz and Todd R Ferretti. Moment-by-moment reading of proverbs in literal and nonliteral contexts. *Metaphor and Symbol*, 16(3-4):193–221, 2001.

Reinhold Kliegl, Antje Nuthmann, and Ralf Engbert. Tracking the mind during reading: the influence of past, present, and future words on fixation durations. *Journal of experimental psychology: General*, 135(1):12, 2006.

Lars Kuchinke, Melissa L-H Võ, Markus Hofmann, and Arthur M Jacobs. Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology*, 65(2):132–140, 2007.

Bruno Laeng, Sylvain Sirois, and Gustaf Gredebäck. Pupillometry a window to the preconscious? *Perspectives on psychological science*, 7(1):18–27, 2012.

Roger Levy. Expectation-based syntactic comprehension. *Cognition*, 106(3):1126–1177, 2008.

M. Papesh and Goldinger S. Pupil-blah-metry: Word frequency reflected in cognitive effort. *Annual Meeting of the Psychonomic Society*, 2008.

Megan H Papesh, Stephen D Goldinger, and Michael C Hout. Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology*, 83(1):56–64, 2012.

Tepring Piquado, Derek Isaacowitz, and Arthur Wingfield. Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, 47(3):560–569, 2010.

Michael I Posner and Jin Fan. Attention as an organ system. *Topics in integrative neuroscience*, pages 31–61, 2008.

Kerstin Preuschoff, Bernard Marius t Hart, and Wolfgang Einhäuser. Pupil dilation signals surprise: evidence for noradrenalines role in decision making. *Front Neurosci*, 5:115, 2011.

Susanne Raisig, Herbert Hagendorf, and Elke Van der Meer. The role of temporal properties on the detection of temporal violations: insights from pupillometry. *Cognitive processing*, 13(1):83–91, 2012.

Keith Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3):372, 1998.

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

Nathaniel J Smith and Roger Levy. Optimal processing times in reading: a formal model and empirical investigation. In *Proceedings of the 30th annual conference of the cognitive science society*, pages 595–600, 2008.