The Index of Cognitive Activity as a Measure of Linguistic Processing

Vera Demberg (vera@coli.uni-saarland.de) and
Evangelia Kiagia (ekiagia@coli.uni-saarland.de) and
Asad Sayeed (asayeed@coli.uni-saarland.de)
Cluster of Excellence, Saarland University
Campus C7.4, 66123 Saarbrücken, Germany

Abstract

This paper reports experimental results on the Index of Cognitive Activity (ICA), a recent micro-level measure in pupillometry which relates processing load to the frequency of rapid small variations in pupil diameter. We collected pupil size data during three language processing tasks (German subject vs. object relative clauses, a grammatical gender violation, and semantic anomalies). We report the first results on the Index of Cognitive Activity for language processing and show that the ICA is responsive to all of our manipulations. We also compare the ICA results to overall pupil dilations and self-paced reading as a measure of processing difficulty. Overall, the use of the ICA as opposed to traditional pupillometry seems promising, as our data provide initial evidence that the ICA may be a faster and more fine-grained measure of cognitive load than overall pupil dilation.

Keywords: Pupillometry, Index of Cognitive Activity, Relative Clause, Self-paced reading, Language, Semantic Anomaly, German

Introduction

The size of the pupil has long been known to reflect arousal (Hess & Polt, 1960) and cognitive load in a variety of different contexts such as arithmetic problems (Hess & Polt, 1964), digit recall (Kahneman & Beatty, 1966), attention (Beatty, 1982) as well as language complexity (Schluroff, 1982; Just & Carpenter, 1993; Hyönä, Tommola, & Alaja, 1995), grammatical violations (Gutiérrez & Shapiro, 2010), and context integration effects (Engelhardt, Ferreira, & Patsenko, 2010). All of these studies have looked at the macro-level effect of the overall dilation of the pupil as response to a stimulus. Recently, a novel micro-level measure of pupil dilation has been proposed, called the "Index of Cognitive Activity" or ICA (Marshall, 2000, 2002, 2007), which does not relate processing load to the overall changes in size of the pupil, but instead counts the frequency of rapid small pupil size variations, which are usually discarded as pupillary hippus (Beatty & Lucero-Wagoner, 2000). The ICA has been argued to be robust to changes in ambient light and eye movements and can therefore be hoped to be more reliable and robust than traditional pupil dilation. Furthermore, as it does not use the overall dilation of the pupil, which can vary as a function of lighting and individual, the frequency of the rapid pupil size changes is argued to be comparable across tasks and subjects.

If it reliably reflects processing load, the ICA would be a convenient method to assess processing load using an eye-tracker in environments where lighting cannot be held constant or where eye or head movements cannot be avoided. The Index of Cognitive Activity could thus complement the range of experimental paradigms currently in use. For example, the ICA might be an interesting measure to use within the visual world paradigm in order to assess visual attention and cognitive load simultaneously without confounding the cognitive load measure due to the eye-movements.

To the best of our knowledge, the present paper is the first to evaluate the ICA on single-task language processing. In the following, we report three bench-mark experiments: a relative clause experiment, a semantic anomaly experiment, and a gender mismatch experiment. These experiments were chosen because reliable effects of these manipulations have previously been shown in the literature such that we would expect large effects; see e.g., (Bader & Meng, 1999) for the German relative clause asymmetry, (Friederici, Pfeifer, & Hahne, 1993) for semantic anomaly processing and (Hagoort & Brown, 1999) for grammatical gender mismatch. Failure of the ICA to reflect the manipulations would allow us to conclude that the ICA is not a suitable measure for capturing effects of linguistic processing difficulty.

The Index of Cognitive Activity

The Index of Cognitive Activity is a patented measure of cognitive load which has only been evaluated (in published work) on a small range of tasks (Marshall, 2000, 2002, 2007; Schwalm, Keinath, & Zimmer, 2008), including digit span tasks and a simulated driving task. The underlying assumption behind using the ICA as a measure of cognitive processing is the observation that pupil size can be affected by two different processes: lighting conditions and cognitive activity. For overall pupil dilation, these two effects are confounded because there is a so-called light reflex even in steady light, meaning that the pupil oscillates irregularly and continually. The pupil dilation is controlled by two groups of muscles: the circular ones, which make the pupil contract, and radial muscles, which make the pupil dilate. Because the activation and inhibition processes are different for reactions due to light and reactions due to cognitive activity, there is a difference in the patterns observed: dilations due to cognitive activity are very short and abrupt. The idea behind the ICA is therefore to perform a wavelet analysis on the pupil dilation record to remove all large oscillations and retain only the very short and rapid events larger than a specified threshold, which can be attributed to the effect of cognitive activity.

To obtain a continuous measure, blinks are factored out by linear interpolation of adjacent events. The ICA events per second are counted, divided by the number of expected ICA events per second (30), and the resulting number is then transformed using the hyperbolic tangent function, in order to obtain a number between zero and one[1]. When using the EyeTracking.Inc software, an ICA value per second is produced. To obtain finer granularity, we also calculated a ICA value per 100msec from the ICA events (i.e. the rapid dilation events). Due to the short time span, we could not interpolate for blinks (which take about 100msec) and therefore simply excluded time frames during which a blink or partial blink occurred from our analysis.

Experimental Setup

We conducted our experiments with 24 participants (18 female, 6 male; average age 23.8), who received course credit for their participation.

All materials were presented in a self-paced word-by-word presentation mode in the middle of the screen in order to minimize eye-movements. Each sentence was followed by a question asking whether the sentence had been grammatical and made sense. Participants responded yes or no using a response pad. Answers were balanced so that "yes" was the correct answer half of the time. Experiment duration was 20-30 minutes. We created randomized lists such that each participant saw only one condition of each item and saw the items in random order. Each of the lists included 24 stimuli per experiment as well as 72 fillers.

We recorded pupil dilations on both eyes using the head-mounted SRI EyeLink II eyetracker at 250 Hz.

Relative Clause Experiment

Experimental Materials The linguistic stimuli for our first experiment consisted of 24 German subject and object relative clauses which were locally ambiguous. This means that the relative clause starting at *die* would be initially interpreted as a subject RC, and only disambiguated to an object RC in one of the conditions at the verb *haben*. These ORC sentences are thus very hard to process. The items were loosely based on (Bader & Meng, 1999), see Example (1).

(1) *Die Nachbarin, [**die**$_{sg,\ nom/acc}$ einige$_{pl,\ nom/acc}$ der Mieter auf Schadensersatz verklagt **hat**$_{sg}$/ **haben**$_{pl}$]$_{relative\ clause}$, traf sich gestern mit Angelika.*
"The neighbor, [whom some of the tenants sued for damages / who sued some of the tenants for damages]$_{relative\ clause}$, met Angelika yesterday."

Data Analysis and Results

Data was analysed using the R lme4 (Baayen, Davidson, & Bates, 2008) and mgcv (Wood, 2001) packages. After removing blinks and partial blinks, we centered and normalized pupil area and removed outliers. We aligned the data for the onsets of critical regions.

Self-paced reading Self-paced reading is a well-established method for measuring linguistic processing difficulty. As Figure 1 shows, we find the expected significantly longer reading times in the disambiguating region as well as in the following word for the object relative clause

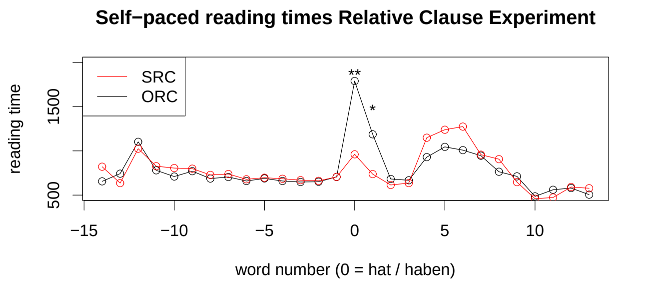

Figure 1: Self-paced reading results for the RC Experiment.

as opposed to the subject relative clause. The effect on the following word is likely a spill-over effect. No other significant effects were found.

Question Answer Accuracy After each sentence, we asked participants to judge whether the sentence had made sense. All of the relative clauses were grammatically correct and semantically acceptable. However, participants responded with "no" significantly more often after an object relative clause than following a subject relative clause (answer accuracy ORC: 63.8%, answer accuracy SRC: 92.3%, difference significant at $p < 0.0001$ in a two-sided binomial test).

Index of Cognitive Activity For the ICA regression analyses, we chose a window from 250msec till 1150msec after the word of the critical region first appears on the screen. The rationale behind this time window is that it takes a short moment for the word to be perceived and processed before we can expect to see a first reaction in the pupil. As we were using self-paced reading, we cannot control when the next word appears on the screen, so we decided for a fixed time window independent of whether participants have already pressed the button to go to the next word or not. Note that the average reading time of the disambiguating region in the SRC case is 960msec, while for the ORC case it is 1789msec. For the SRC data, the next word may thus just have appeared for some subjects during the first second after the critical region, while for ORC sentences, subjects would typically not yet have pressed the space bar during the time period for which we evaluate the ICA.

Pupil dilation was recorded for both eyes, and the ICA is by definition calculated for both eyes independently.

Left Eye We ran a linear mixed effects regression model with the left-eye ICA data aggregated per 100msec as a response variable. We included the relative clause type as a predictor, random intercepts for subjects and items, and a random slope for relative clause type by both subject and item. These random slopes did not significantly improve model fit, and we found that `relative clause type=subject RC` was a significant negative predictor in models with and without random slopes, Table 1.

To get a better intuition of what the ICA data looks like, we also plotted a spline plot (smoothed fitting showing how

[1]The method is patented, and the analysis program has to be licensed from EyeTracking, Inc., San Diego, CA.

Table 1: Linear mixed effects model with the ICA as a response variable and item type as a predictor; response variable is ICA for the left eye. Model includes random intercepts and random slopes for item type under both subject and item.

left eye ICA	Estimate	Std. Error	t value	Signif
(Intercept)	0.831987	0.008271	100.60	***
Subject RC	-0.013866	0.006414	-2.16	*

right eye ICA	Estimate	Std. Error	t value	Signif
(Intercept)	0.821605	0.009070	90.58	***
Subject RC	-0.008202	0.005510	-1.49	

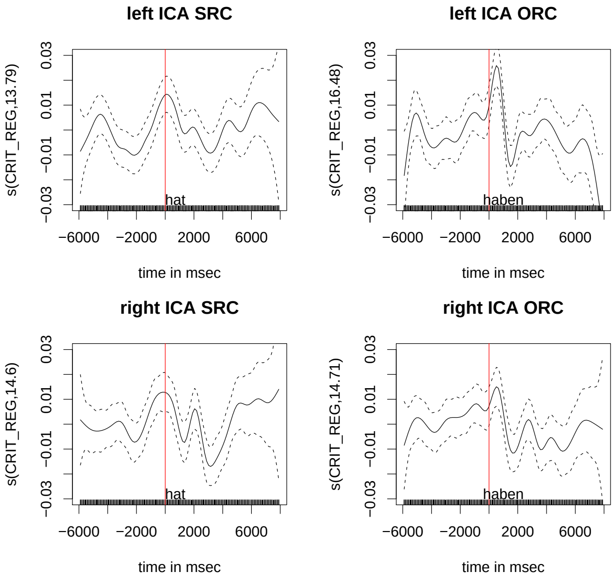

Figure 2: Spline plot k=30 for subject relative clause data (left) and object relative clause data (right) for the ICA for each eye, time point 0 shows the onset of the disambiguating region ”hat” / ”haben”.

the ICA evolves across an item), time-locked for the onset of the disambiguating region, see top plots in Figure 2. We can also clearly see the difference between the ICA in the subject vs. object relative clause conditions in this plot: the ICA values rise sharply during the first second after the disambiguating region in the ORC case, which is not observable in the SRC case. This plot also is confirming evidence that we should expect an effect of the ICA during the first second of the critical region. In the subject relative clause case, we also see slightly elevated ICA during the ambiguous region, but it falls again right after the ambiguity is resolved.

Right Eye We ran the same regression analysis for the ICA of the right eye. We found the same tendency, but no significant effect, see Table 1 (again, this result is independent of whether we include random slopes for relative clause type by subject or not). This result is consistent with the spline plot of the right eye ICA shown in bottom plots of Figure 2: while there is a small tendency to an effect similar to the one for the left eye, it is much smaller for the right eye.

Pupil Dilation Next, we test whether the more traditional measure of overall pupil dilation (centered and normalized for each subject) is in line with our findings for the ICA. We know that pupil dilation can be expected to peak approximately 1.2 seconds after the stimulus. We therefore include in our regressions a 2 second time window starting at 250 msec after the disambiguating word first appears on the screen. We here focus on the rate of dilation as opposed to comparing pupil sizes directly (Engelhardt et al., 2010). Our regression models again include random intercepts, as well as random slopes of relative clause type by subject and item. (Models including random slopes for the interaction of relative clause type and time did not converge.)

Left Eye Figure 3 shows that the overall pupil dilation rises after the disambiguating region in both conditions, but that the rise is much faster in the ORC condition than in the SRC condition. We calculated linear mixed effects models with normalized left eye pupil area as a response variable and the relative clause type, time, and the interaction between them as explanatory variables. The `time` variable measures how much time has gone by since the disambiguating word first appeared on the screen. Table 2 shows no main effect of relative clause type, but a significant positive effect of time (i.e., the pupil dilates following the disambiguating region in both conditions) and a significant negative interaction between relative clause type and time. This interaction reveals that the pupil dilates significantly less quickly in the subject relative clause condition than in the object relative clause condition, as observed in top plots of Figure 3.

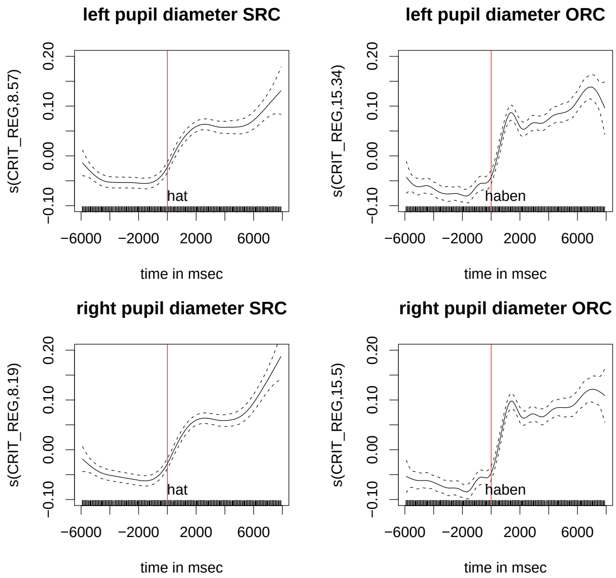

Figure 3: Overall pupil dilation for the left eye. Time 0 = onset of disambiguating region.

Table 2: A linear mixed effects model shows that pupil dilates significantly more slowly in the subject relative clause condition than in the object relative clause.

left pupil area	Estimate	Std. Error	t value	Sign
(Intercept)	4.407e-03	2.584e-02	0.171	
Subject RC	-8.353e-03	1.770e-02	-0.472	
Time	7.397e-05	8.398e-06	8.809	***
SRC:Time	-3.136e-05	1.181e-05	-2.654	**

Table 3: Linear mixed effects model for right pupil area.

right pupil area	Estimate	Std. Error	t value	Sign
(Intercept)	3.309e-03	2.637e-02	0.125	
Subject RC	-1.565e-02	1.812e-02	-0.863	
Time	8.285e-05	8.370e-06	9.898	***
SRC:Time	-3.159e-05	1.178e-05	-2.683	**

Right Eye For the right eye pupil area, we observe almost identical effects as for the left eye (see Fig. 3 and Table 3).

Discussion

All of our measures (self-paced reading times, question response accuracy, Index of Cognitive Activity, and overall pupil dilation) were consistent in that we found statistically significant evidence that the object relative clause was harder to process than the subject relative clause. Interestingly, we found a significant effect of the ICA only in the left eye, but not in the right eye, while both eyes showed a significant effect of overall dilation. This is particularly interesting in conjunction with results from a dual-task experiment which used the same stimuli (but presented aurally) and came to the same result (Engonopoulos, Sayeed, & Demberg, 2013). Additionally, in that dual task experiment, we found that the right eye ICA reflected the primary task (driving) better than the left eye's ICA. Taken together, these findings may point to a brain-hemispheric difference which is reflected in the ICA.

Semantic Violation Experiment

The second experiment tests whether the results concerning the sensitivity of the ICA to linguistic load can be replicated in an experiment with semantic violations.

Materials

We created 24 items that contained a semantic violation such that the direct object argument of the verb did not match the selectional restriction of the verb, as shown in (2). We contrasted this with a version of the sentence where the verb was chosen to fit the direct object and measured the semantic mismatch effect on the direct object (marked in bold in our example item). Once again, we mixed these items with 72 fillers, and each person only saw one version of each item (12 semantically anomalous ones and 12 good ones). We made sure that the critical region was not sentence final to avoid confusion with sentence wrap-up effects.

(2) *Max singt / arbeitet als* ***Rechtsanwalt*** *bei einer großen Firma.*
"Max is singing / working as a lawyer for a large company."

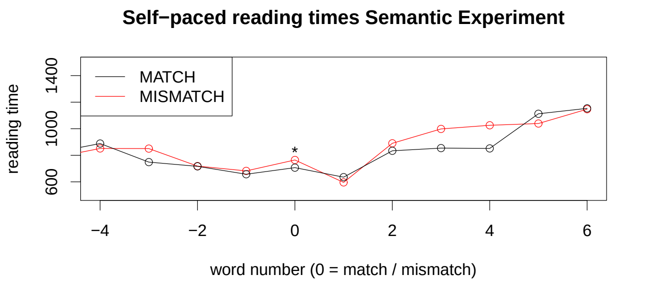

Figure 4: Self-paced reading for the semantic anomaly materials. Semantic Anomaly is located at word number 0.

Table 4: ICA mixed effects regression results for semantic fit study, for period of 250msec – 1150msec after onset of critical region.

left eye ICA	Estimate	Std. Error	t value	Sign
(Intercept)	0.817823	0.008072	101.32	***
semantic violation	0.014805	0.005985	2.47	*
right eye ICA	Estimate	Std. Error	t value	Sign
(Intercept)	0.819334	0.007654	107.04	***
semantic violation	0.004450	0.005691	0.78	

Data Analysis and Results

We use the same regression analysis methods as in the relative clause experiment.

Self-paced reading We used linear mixed effects models to test for significance at each word (items were aligned with respect to the critical region), and found a significant difference between conditions directly on the critical region: reading times were significantly longer for semantically anomalous direct objects than for semantically fitting ones. The differences in average reading times in later parts of the sentence (see words number 3 and 4 in Figure 4) were not significant. In this experiment, we did not observe the spill-over to the word after the critical region which we observed in the relative clause experiment.

Question Answer Accuracy We once again asked participants to judge whether the sentence had made sense. Question answer accuracy was 99% correct for the sentences without semantic violations and 96% correct for sentences with semantic anomalies. This indicates that the items from our semantic experiment were easier to process than the relative clauses from the previous experiment, which may also explain the lack of spill-over in self-paced reading.

Index of Cognitive Activity For the left eye's ICA, we find a significant effect of the semantic anomaly, showing significantly more ICA events compared to the semantically normal condition, see Table 4. As in the first experiment, the linear model compares ICA values for the time span from 250msec till 1150msec after the critical word first appeared on the screen. Again, the main effect is stable irrespective of the inclusion of random slopes for condition by subject and

Table 5: Pupil area mixed effects regression results for semantic violation study, for period of 350msec – 1s after onset of critical region.

left pupil area	Estimate	Std. Error	t value	Sign
(Intercept)	-4.775e-02	2.477e-02	-1.928	
semantic violation	-2.709e-02	3.316e-02	-0.817	
time	3.105e-05	8.251e-06	3.766	***
sem violation:time	4.588e-05	1.152e-05	3.982	***

right pupil area	Estimate	Std. Error	t value	Sign
(Intercept)	-4.199e-02	2.518e-02	-1.668	
semantic violation	-1.097e-02	3.384e-02	-0.324	
time	4.154e-05	8.224e-06	5.051	***
sem violation:time	3.800e-05	1.148e-05	3.309	***

item. Very interestingly, the lack of effect (but tendency in the expected direction) of our linguistic manipulation on the right eye's ICA is also replicated, see bottom of Table 4. The results of our semantic fit study are hence very consistent with the results of the relative clause study.

Pupil Dilation In our linear mixed effects model analysis of the overall pupil area in the two seconds following the onset of the critical region, we find a main effect of pupil dilation in both eyes as well as a significant interaction of time and our semantic condition. The pupil dilates more quickly in the semantic violation condition than in the semantically normal sentences, see Table 5. These findings are consistent with the ICA analysis of the semantic violation experiment, with the data from the relative clause experiment, and with earlier reports on pupil dilation as a measure of cognitive load.

Gender Mismatch Experiment

Our third experiment aims to test whether the ICA is sensitive to grammatical gender mismatch. Again, we used German materials. All analysis methods are identical to the previous experiments.

Materials

The materials for the gender mismatch experiment included 24 items, each participant saw 12 grammatically correct items and 12 grammatically incorrect items, where the gender of the determiner and adjective did not match the grammatical gender of the noun (see Example (3), with the noun in bold face). We again made sure that the critical region was not sentence-final to avoid confusion with sentence wrap-up effects.

(3) Simone hatte eine(n) schreckliche(n) **Traum** und keine Lust zum Weiterschlafen.
"Simone had a$_{[masc/fem]}$ horrible$_{[masc/fem]}$ dream and didn't feel like sleeping any longer."

Data Analysis and Results

We expected to find an effect at the critical word. Note that this is the only experiment of the ones reported in this study which uses an ungrammatical condition.

Self-paced reading For self-paced reading, we found significantly longer reading times on the critical region where

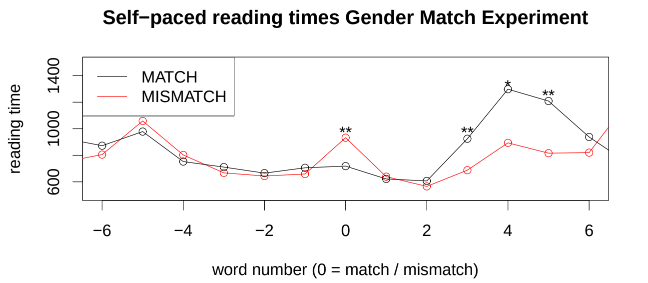

Figure 5: Self-paced reading results for grammatical gender violation experiment. Critical region at word number 0.

the gender mismatch happens. The effect is more local than for the relative clause experiment: there is no spill-over effect to the following word. We do however see a large facilitation effect for the mismatch condition as compared to the grammatically correct condition towards the end of the sentence. We hypothesize that this effect is due to our questions on whether the sentence is grammatical and makes sense: After the gender mismatch, subjects know they will be able to answer the question with "no" and stop paying attention to the end of the sentence.

Question Answer Accuracy Once again, participants were asked whether the given sentence had made sense. Answer accuracy was identical (96% correct) for both conditions.

Index of Cognitive Activity Similar to the relative clause experiment, we find a significant effect of condition on the ICA of the left eye: during the second following the critical word, the ICA is significantly higher in the gender mismatch condition than in the gender-correct condition, see Table 6. We find, however, that the effect seems to start slightly later: when we include random slopes under subject and items, the main effect for condition only reaches significance starting from 350msec after the critical word first appeared on the screen. A spline plot (not included here for space reasons) confirms the impression that the effect seems to start 100msec later. Note though that this might be due to the different nature of the experiment: the gender mismatch is not possible to expect (as the disambiguating region in the relative clause) and is the only ungrammatical material (except fillers) in the experiment.

Another difference with the relative clause experiment is that we do find a significant effect of gender mismatch on the right eye's ICA as well (see Table 6).

Pupil Dilation For pupil dilations, we again find that the pupil dilates during the critical region and that it dilates significantly more quickly in the gender mismatch condition than in the gender matching sentences (Table 7).

Overall Discussion and Conclusions

These three first language experiments using the Index of Cognitive Activity yielded remarkably consistent results: for

Table 6: Left eye ICA mixed effects regression results for gender mismatch study, for period of 350msec – 1s after onset of critical region.

left eye ICA	Estimate	Std. Error	t value	Sign
(Intercept)	0.835114	0.010077	82.88	***
gender mismatch	0.018324	0.008472	2.16	*

right eye ICA	Estimate	Std. Error	t value	Sign
(Intercept)	0.832818	0.008763	95.04	***
gender mismatch	0.015909	0.007612	2.09	*

Table 7: Pupil area for the time span from 250msec till 2s after the onset of the critical region. Reported models include random intercepts under item and subject, as well as random slopes for condition.

left pupil area	Estimate	Std. Error	t value	Sign
(Intercept)	-2.504e-02	2.647e-02	-0.946	
gender mismatch	1.383e-02	3.530e-02	0.392	
time	4.331e-05	8.145e-06	5.317	***
gend mism:time	5.151e-05	1.171e-05	4.398	***

right pupil area	Estimate	Std. Error	t value	Sign
(Intercept)	-2.502e-02	2.616e-02	-0.956	
gender mismatch	1.847e-02	3.630e-02	0.509	
time	4.840e-05	8.283e-06	5.843	***
gend mism:time	5.214e-05	1.191e-05	4.378	***

the subject vs. object relative clause experiment and the semantic violation experiment, we found significant effects in the left eye's ICA measure, showing that significantly more rapid dilations were recorded during the period of 250msec – 1150msec after the onset of the critical word in the condition that causes more processing load (object relative clause and semantic violation, respectively). In both of these experiments, the right eye's ICA showed the same tendency but did not reach significance. This observation is consistent with findings from a recent dual-task experiment which measured the ICA during processing of the same relative clauses as in this experiment, but with a simultaneous steering task (Engonopoulos et al., 2013), where we found that the left eye was sensitive to our language manipulation, while the ICA of the right eye showed stronger effects of the steering task.

The grammatical gender experiment reported in this paper also showed similar results, with the grammatical gender violation being a significant positive predictor of the ICA of both eyes in the time span starting 350msec up to 1 sec after the critical word first appeared on the screen. Further experiments will be necessary to determine whether the slightly longer latency and effect in both eyes' ICA is due to the type of experimental manipulation, or whether it is coincidental.

All expected effects were found not only in the ICA, but also in the self-paced reading measure and the speed of the overall pupil dilation. We note, however, that the ICA effect can be measured earlier than the macro-scale pupil dilation effect.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*.

Bader, M., & Meng, M. (1999). Subject-object ambiguities in german embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research*, *28*(2).

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, *91*(2), 276.

Beatty, J., & Lucero-Wagoner, B. (2000). *The pupillary system*. Cambridge University Press.

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly journal of experimental psychology*, *63*, 639—645.

Engonopoulos, N., Sayeed, A., & Demberg, V. (2013). Language and cognitive load in a dual task. In *Proc. of CogSci*.

Friederici, A., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183–192.

Gutiérrez, R. S., & Shapiro, L. P. (2010). Measuring the time-course of sentence processing with pupillometry. In *Cuny conference on human sentence processing*.

Hagoort, P., & Brown, C. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research*, *28*(6), 715–728.

Hess, E., & Polt, J. (1960). Pupil size as related to interest value of visual stimuli. *Science*.

Hess, E., & Polt, J. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*.

Hyönä, J., Tommola, J., & Alaja, A. (1995). 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.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: pupillometric indices of sentence processing. *Canadian journal of experimental psychology*, *47*(2).

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*.

Marshall, S. (2000). *U.s. patent no. 6,090,051.*

Marshall, S. (2002). The index of cognitive activity: Measuring cognitive workload. In *proc. 7th conference on human factors and power plants* (pp. 7–5).

Marshall, S. (2007). Identifying cognitive state from eye metrics. *Aviation, space, and environmental medicine*, *78*.

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

Schwalm, M., Keinath, A., & Zimmer, H. (2008). Pupillometry as a method for measuring mental workload within a simulated driving task. *Human Factors for assistance and automation*, 1–13.

Wood, S. (2001). mgcv: GAMs and generalized ridge regression for R. *R news*, *1*(2), 20–25.