UNIVERSITY OF MONTPELLIER

FACULTY OF PHARMACY

Dean LAURENCE VIAN

THESIS M2

MASTER'S DEGREE IN SENSORY AND MOTOR NEUROPROSTHETICS

THE EFFECT OF LUMINANCE ON LISTENING EFFORT ASSESSMENT USING PUPILLOMETRY

By : Florian MALAVAL

Supervisor : **Alexandre LEHMANN**

Centre for Research on Brain, Language and Music – McGill University

Montpellier – September 7th 2018

Director of Education

Jean-Luc PUEL

ACKNOWLEDGMENT

A special thank to

Dr. Alexandre Lehmann, Assistant Professor at McGill University and Adjunct Professor at Université de Montréal, for his friendly welcoming in his researcher team and for supervising and guiding me along this whole internship,

Dr. Mickaël Deroche, Research Associate at McGill University, for his great amiability, for sharing with me his knowledge and his advice and for his precious help in programming,

Dr. Yue Zhang, postdoc researcher, for her kindness, her motivation, her expertise and to help me to analyze data and to write this thesis,

the staff working in the Centre for Research on Brain, Language and Music, for immediately integrated me in their offices,

the staff working in the Centre for Interdisciplinary Research in Music Media and Technology, for lending us an experiment lab, equipment and our study device,

the master degree educational team, to Jerome Bourien and Jean-Charles Ceccato, for contributing in one way or another, to this achievement,

my friends, who always supported me,

my family, whom I owe everything.

SUMMARY

ABSTRACT 4

INTRODUCTION 5

METHODS 12

RESULTS 16

CONCLUSION 21

REFERENCES 23

ABSTRACT

AIM Pupillometry has been used by many studies in recent years to measure listening effort objectively. However, there is no established method yet considering the control on the luminance level during the pupillometry experiments. Past studies have shown that the luminance level might have a significant impact on the task-evoked pupillary responses during detection and memory tasks. But it is unclear whether the luminance level will affect pupillary responses evoked by processing perceived speech stimuli. Although all past pupillometry studies within hearing research controlled for the luminance level in some ways, there is a great inconsistency in the methodology used and this could cause great problems when validating and comparing results from different studies. Therefore, the objective of this work is to quantify the impact of using different levels of light on the pupillary responses evoked by the presence of stationary noise when listening to sentences. As the first study in listening effort research to investigate the interaction between light and task-evoked pupillary responses, the results will highlight the importance of a better experimental control and precise report of the luminance level in future pupillometry studies. In a long term, this will also promote a better experimental framework for reproducible research within the listening effort research. To achieve this objective, we have two specific aims: use different levels of light to check whether this changes the effect size of task-evoked pupillary response; and compare different analysing methods in previous studies to examine whether the light condition affects the parametric test results.

PROTOCOL We used Tobii Glasses Pro2 and set up the experiment in a soundproof room with adjustable room light. Experiment is run on Matlab, with a script synchronising behavioural tasks and pupil recording. A streamline of pupil trace pre-processing and data analysing were designed in Matlab and R.

RESULTS The SNR level, the lighting condition and their interaction are significantly affected tonic pupil diameter and phasic pupil response. We also showed the peak pupil dilation relative to the baseline is significantly greater in medium luminance condition.

INTRODUCTION

CONTEXT

There is no official definition of the listening effort. Hicks and Tharpe (2002), associated it to the cognitive resources necessary for verbal understanding and the Framework for Understanding Effortful Listening (FUEL) as the deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task that involves listening (Pichora-Fuller *et al.*, 2016). More simply, that is the feeling of being tired when you have to work really hard to understand the content of a conversation, as it could be in a crowded restaurant or during a phone call with a crummy signal (Picou, 2013). Speech intelligibility does not reflect this effort necessary to verbal understanding: even if the intelligibility remains high, the listening effort and the speed of sound information processing can be differ according to individuals (Winn, Edwards and Litovsky, 2016). Therefore listening effort assessment methods have emerged, in particular pupillometry. This technique is based on the locus coeruleus (LC) activity. This one of several brainstem neuromodulatory nuclei with diffuse projections throughout the central nervous system (Aston-Jones and Cohen, 2005; Benarroch, 2009). Within the waking state, this system modulates the collection and processing of salient sensory information through a diversity of concentration dependent actions within cortical and subcortical sensory, attention, and memory circuits (Usher, 1999; Berridge and Waterhouse, 2003). On one hand Mather and Harley (2016) noted LC integrity plays a key role in determining late cognitive abilities while, on the other hand Tsukahara, Harrison and Engle (2016) underlined cognitive performances are exactly related to the baseline pupil size. Indeed, these baselines decrease linearly with age (Pfeifer *et al.*, 1983; Winn *et al.*, 1994) and are even lower in patients with Alzheimer's disease (Prettyman, Bitsios and Szabadi, 1997; Benarroch, 2009): these are precisely populations affected by cognitive decline (Rushton and Ankney, 1996; Bäckman *et al.*, 2005; Deary *et al.*, 2009). In addition, a strong relationship between LC activity and pupil diameter have shown in the monkey (Rajkowski, Kubiak and Aston-Jones, 1993; Figure 1) and in the humans (Gilzenrat *et al.*, 2010; Murphy *et al.* 2011, 2014).

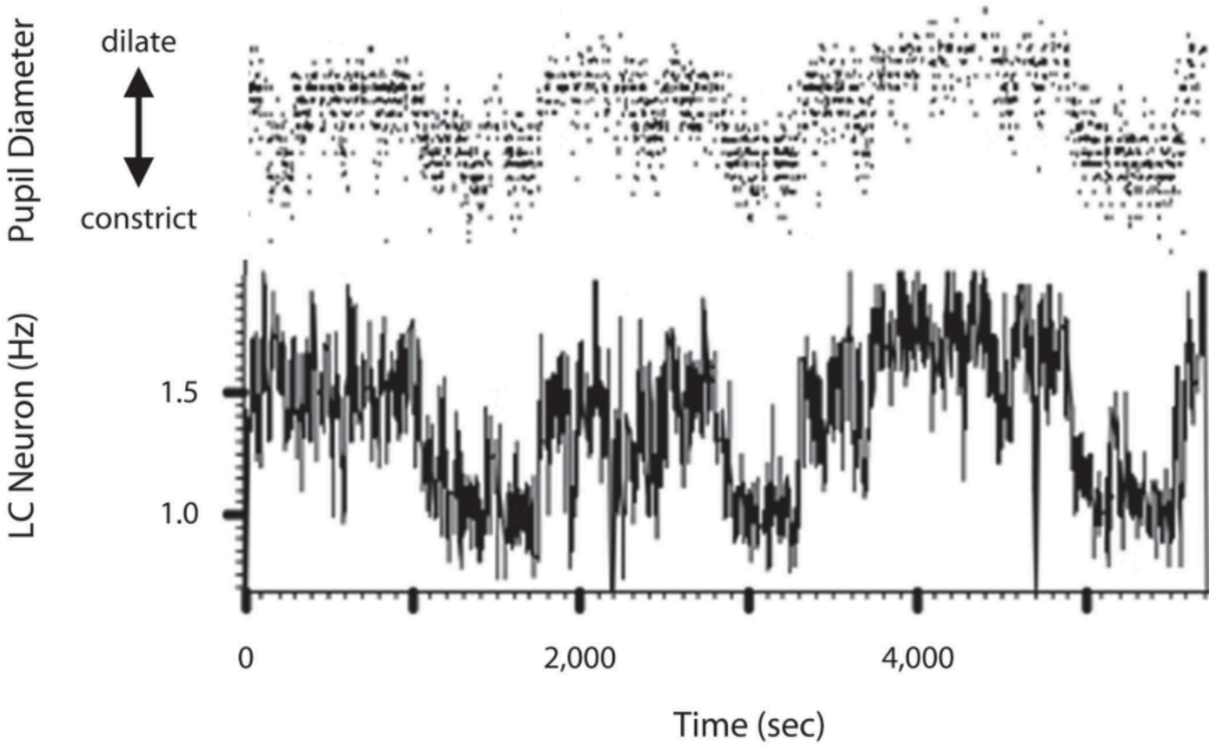


Figure 1. Relationship between tonic pupil diameter and baseline firing rate of a locus coeruleus (LC) neuron in the monkey (Rajkowski, Kubiak, & Aston-Jones, 1993).
Pupil diameter measurements were taken by a remote eyetracking camera at each instant in time at which the monkey achieved fixation of a visual spot during the target detection task.
Note the close positive relationship between pupil diameter and the rate of LC activity.

Handful studies show that listening effort is affected by different listening conditions, even when speech intelligibility is not affected. For instance, Sarampalis *et al.* (2009) showed that noise reduction scheme did not improve normal hearing (NH) listeners' sentence recognition scores at different noise conditions (quiet, +2 dB SNR, -2 dB SNR and -6 dB SNR) but significantly decreased the average reaction time in the simultaneous visual task at -6 dB SNR. Similarly, Wendt, Hietkamp and Lunner (2017) showed that applying noise reduction scheme at ceiling performance reduced listeners' peak pupil dilation (PPD), but did not improve their speech in noise performance. This suggests that these measures on listening effort reveal a different dimension of speech processing that is not well characterised by traditional speech performance measurements and there is therefore considerable interest in assessing the listening effort in clinic. A possible application of this assessment would be to use it as an efficiency indicator for signal processings present in hearing aids or cochlear implants (CI): at first glance those could be not directly improve intelligibility but they could be beneficial for reducing listening effort (Rudner, Rönnberg and Lunner, 2011;

Ng *et al.*, 2013). Within hearing research, pupillometry measuring pupillary responses has been shown to be a valid tool for measuring listening effort in different listening conditions, such as with different masking noise, spectral degradation, speech intelligibility level and syntactic complexity (Beatty, 1982; Granholm *et al.*, 1996; Zekveld and Kramer, 2014; Winn, 2016). Typically, when a speech recognition task gets difficult, for instance with lower signal noise ration (SNR) level or degraded spectral resolution, listeners show a greater task evoked pupillary response (TEPR). This continues until the task becomes so challenging that the listeners « give up ». One of the most investigated factor on listening effort is the type and level of masking noise (Koelewijn *et al.*, 2012, 2014; Zekveld and Kramer, 2014; Ohlenforst *et al.*, 2017, 2018; Wendt, Hietkamp and Lunner, 2017). Thus Ohlenforst *et al.* (2017) investigated PPD across a wide range of SNR in stationary noise masker (-12 dB to +16 dB) for both normal hearing (NH) and hearing-impaired (HI) listeners. They observed an inverse U-shaped relation between PPD and masking noise: listeners showed consistently bigger PPD for lower SNR and smaller PPD for higher SNR. However, when the SNR is approaching quiet (+16 dB), or impossible to obtain any speech information (-12 dB), listeners showed smaller PPD (Figure 2).

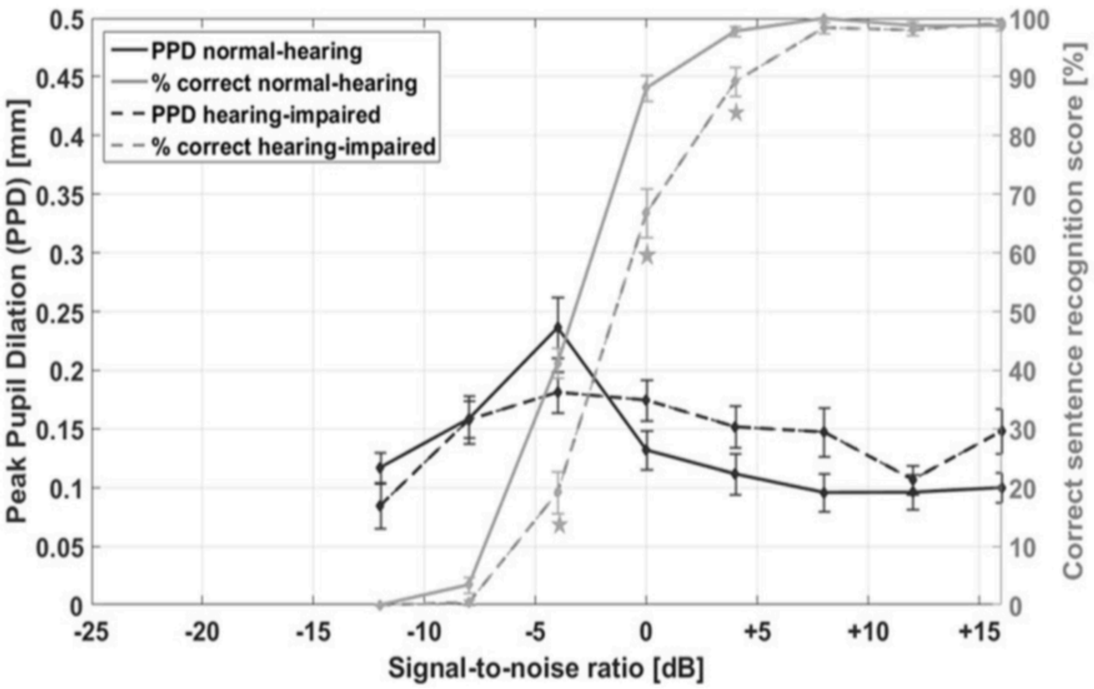


Figure 2. Ohlenforst *et al.* (2017). Peak pupil dilation (PPD) (black color) on the left y-axis and percentage correct sentence recognition scores (gray color) on the right y-axis across signal-to-noise ratios (SNRs) for the stationary masker for normal-hearing (NH) and hearing-impaired (HI) participants. Error bars represent the standard error of the mean. Gray stars indicate significant group differences in sentence recognition performance (NH vs. HI) of $p < 0.006$.

While pupil dilation serves as a reliable « reporter variable » for cognitive effort, it is also sensitive to many other factors, among which luminosity variation is most prominent (Tryon, 1975; Beatty and Lucero-Wagoner, 2000). Two antagonistic muscles control the pupil size: the iris sphincter and the dilator muscle. When light falls on the retina(s), this leads to increased neural activity in the pretectal regions and stimulation of the Edinger-Westphal nucleus, where preganglionic parasympathetic neurons are activated and innervate the ciliary ganglion (Loewenfeld and Lowenstein, 1993; Wang *et al.*, 2016). These in turn command the sphincter muscles of the iris to tighten which leads to pupil constriction. Under the direct control of the autonomic nervous system (ANS), the pupil light reflex reflects the balance between the Sympathetic Nervous System (SNS) and the Parasympathetic Nervous System (PNS) (Figure 3). While the range of pupillary movement in response to luminance levels can vary from less than 1mm to more than 9mm, the largest of cognitively driven movements are about 0.5mm (Beatty, 1982; Winn, Edwards and Litovsky, 2016) suggesting that light reaction is predominant over cognitive pupillary component (Pomplun and Sunkara, 2003; Peysakhovich *et al.*, 2015).

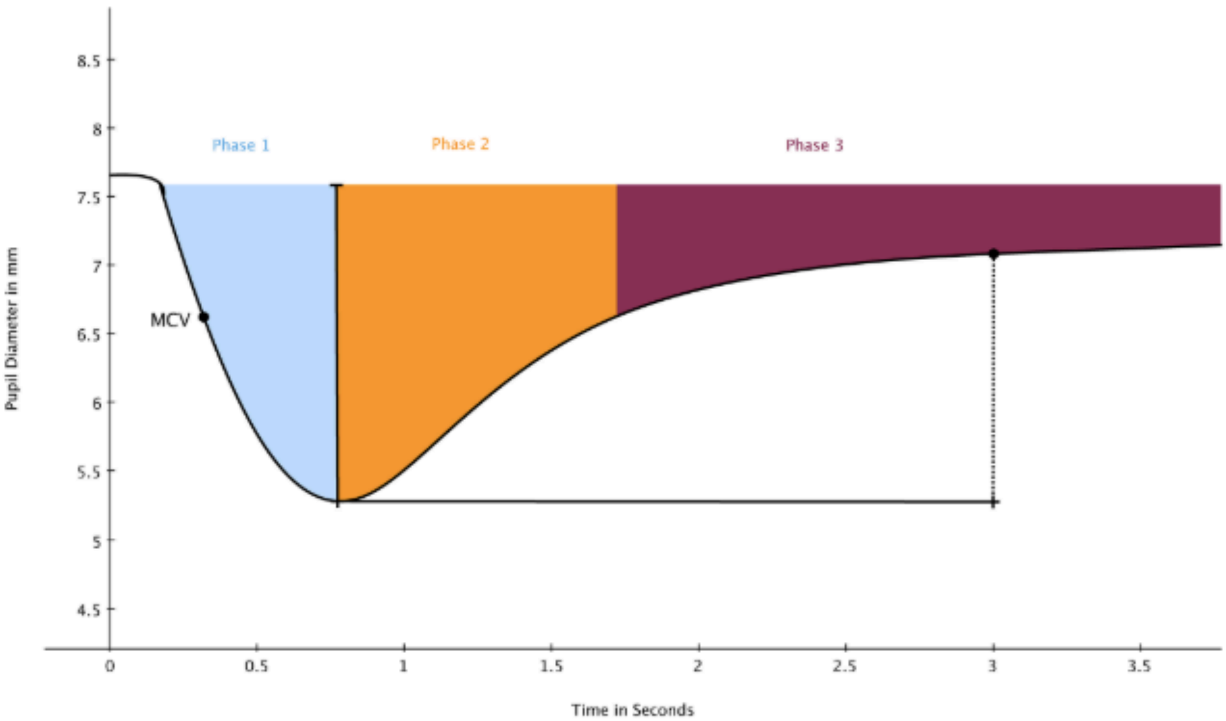


Figure 3. Wang *et al.*, 2016. Demonstration of one pupil light reflex on pupillometry. The light onsets at the '0' point; Phase 1 is a fast constriction mainly controlled by PNS; Phase 2 is a fast redilation under the control of both PNS and SNS; Phase 3 is a slow redilation phase, predominantly controlled by SNS activity.

Yet, only a handful of studies have looked at how light condition could affect TEPRs. Some studies using visual or memory tasks have shown inconclusive, sometimes contradictory, results on whether there is an interaction between cognitive load and luminance condition on pupillary changes, and what this relation might be. Pomplun and Sunkara (2003) performed a visual reaction task where participants had to press a button before a blue circle on a screen expanded to its maximum size. The task was displayed on a screen with either white circles on a black background or black circles on a white background, and of different difficulty levels based on how quick the interval between the tasks was. No interaction was found between the task difficulty and the screen luminance. Xu *et al.* (2011) performed four arithmetic tasks with different levels of difficulty depending on the type of numbers that had to be summed up, combined with four levels of background brightness. There was no report on whether there was an interaction between the task difficulty and the background brightness. Steinhauer *et al.* (2004) conducted two arithmetic tasks, continuously subtracting a random number by 7 (difficult) or adding by 1 (easy), in either dark or moderate room lighting. During the pre-task period (i.e. baseline pupil size measurement period), no interaction was found between the task difficulty and the light condition. In the response period, on the other hand, there was a significant interaction: in the dark lighting setting, there was no difference in mean pupil size compared to the baseline pupil size between the two tasks but a significantly greater mean pupil dilation for the Subtract 7 than the Add 1 task. This suggests that the pupillary changes observed during a cognitive task partially depend on the physical light condition. Peysakhovich *et al.* (2015) conducted a short-term memory task where participants were asked to either recall (with load on memory) or not recall (without load on memory) a series of auditory presented digits. The task difficulty was also controlled by the number of digits required to be recalled (5, 7 or 9 digits). Screen luminance was changing from trial to trial amongst black, gray or white. Similarly, no difference was found in the pre-task baseline period, but in darker lighting condition, the same amount of load on memory induced higher increase in pupil diameter. Peysakhovich, Vachon and Dehais (2017) required participants to perform a N-back recall Task coupled with an arithmetic task. Participants either added or subtracted two numbers displayed on the screen and were asked to respond whether it matched the result from one block back (1-back) or two blocks back (2-back). The screen was either grey (low light) or white (high light). Results showed that differences in baseline pupil diameter between the 1-back (easy) and the 2-back (difficult) tasks were larger under the gray background condition compared to the white background condition. No effect of light and no interaction were found for the TEPR.

It is important to note that tasks used in those studies vary greatly in cognitive functions and difficulty levels. While tasks like visual reaction and immediate arithmetic assignments require transient investment in cognitive resources, tasks like digits recall require constant staining of mental effort. This adds to the complexity already caused by the limited number of studies and little replication on this research topic. However, it is clear that extra care needs to be taken when controlling the luminance level during a pupillometry experiment, as it might influence the effect size of the main experimental factors. Although all past pupillometry studies within hearing research controlled the luminance level in some ways, there is no consensus yet on how to conduct and report this experimental control due to the lack of research on how the light conditions could affect TEPR during listening tasks. Studies typically measure and control room ambience lighting levels, instead of the magnitude of light hitting the participants' eyes (Zekveld, Kramer and Festen, 2011; Koelewijn *et al.*, 2012, 2014, Kuchinsky *et al.*, 2013, 2014; Zekveld and Kramer, 2014; Winn, 2016; Winn, Edwards and Litovsky, 2016). During the listening tasks, although ambience light was fixed, participants were required to focus on an object either on a distant white wall (Zekveld, Kramer and Festen, 2011; Koelewijn *et al.*, 2012; Zekveld and Kramer, 2014) or on an illuminated screen (Kuchinsky *et al.*, 2013, 2014; Zekveld and Kramer, 2014; Winn, 2016; Winn, Edwards and Litovsky, 2016). Benedetto *et al.* (2014) have shown that the screen luminance, which was closers, attracted most of participants' attention and had a more significant effect on the pupil size compared to ambient lighting during a digital reading task. Therefore, without a more precise measure and control on the light condition during listening tasks, it is unclear whether results in different listening effort studies are directly comparable. Furthermore, the magnitude of the pupil light reflex is also influenced by the colour and the duration of the light stimulus (Bremner, 2012; Ishikawa *et al.*, 2012).

Considering that many listening effort studies used a variety of visual cues on an illuminated screen to help listeners fix an object and remind them to respond within a trial (Kuchinsky *et al.*, 2013, 2014; Winn, 2016; Winn, Edwards and Litovsky, 2016), it is unclear how independent the observed TEPRs during speech recognition tasks are from the pupil light reflex. The lack of research on this topic also leads to complications in data analysis for pupillometry studies. Past studies typically extract PPD by subtracting (or dividing) the peak by the baseline pupil diameter, assuming that they are two independent components during a cognitive task (Beatty, 1982). However, recent study found that light condition seems to affect the tonic pupil diameter (i.e. the baseline) and phasic pupil response (i.e. peak pupil dilation PPD) differently (Steinhauer *et al.*, 2004; Peysakhovich, Vachon and Dehais, 2017). A bad control of the luminance level could then mean that the measures of the two components are confounded, hence making the subtraction or the division between the two not a reliable index for cognitive effort. Furthermore, with a luminance level currently applied in each study not controlled and not precisely reported, it is likely that the baseline pupil diameter starts at different levels for each participant. Considering that the range of pupillary movement depends on the baseline pupil size, larger baseline pupil diameters will allow more room for constriction (Newsome and Loewenfeld, 1971). In this case, participants starting with a larger baseline pupil size might show a different task evoked pupillary response than participants starting with a smaller baseline pupil size. This is due to biased calculation rather than on the difference in listening effort.

PURPOSE

To investigate whether the luminance level affects the pupillary responses related to listening effort and parametric test results, we ran a well-studied listening effort test paradigm in different lighting conditions. This is the first study in listening effort research to compare TEPRs in different lighting conditions. By showing how physical luminance levels affect TEPRs, this study highlights the importance to establish a better experimental control in future pupillometry studies to avoid light-evoked pupillary responses affecting pupillary responses related to listening effort.

METHODS

PARTICIPANTS

18 normal-hearing adults have been recruited (aged between 20 and 55 years old). All participants were either English (8 adults) or French (10 adults) native speakers. In order to ensure that participants could clearly see the visual cues displayed on the screen, we verified their binocular vision at 10ft, without the use of any vision correction device, by using a Snellen Eye chart. A visual acuity of 10/10 or better was required. We considered a normal hearing when binaural pure tone average (0.25, 0.5, 1, 2, 4, 8 kHz), assessed with headphones, was inferior of 30 dB HL.

MATERIAL

Lab. Test administration took place in a soundproof room furnished with light dimmer devices, located in the CIRMMT (Centre for Interdisciplinary Research in Music Media and Technology, housed at the Schulich School of Music at McGill University, Montreal, Canada).

Hardware & Software. The pupil size was measured with the infrared binocular eyetracker Tobii Glasses Pro2 with a 50 Hz sampling rate per camera (two cameras for each eye). Auditory stimuli were presented by the headphones Beyer Dynamics DT 990 Pro (calibrated at 1 kHz at 65 dB SPL) linked to the external soundcard Focusrite Scarlett 2i4. The luminance level was picked up with the light meter TES-1335. The visual cues were displayed on the Panasonic Viera TH-65PHD7 plasma monitor. Experiments were run from the laptop Dell Precision3 (OS: Windows 10 Pro, v1709) in Matlab 2016b, using Psychtoolbox and custom software and data were analysis thanks to Matlab 2016b and R.

Stimuli. We used IEEE sentences recorded from a male native American English speaker and HINT sentences recorded from a male Quebecois French speaker.
Sentences were masked by speech-shaped noise (filtered on the long-term excitation pattern of the entire material, respectively in English or French). The SNR levels were

chosen due to the consistent results in sentence recognition scores and PPD in the past literature using stationary noise (Ohlenforst *et al.*, 2017, 2018). In those experiments, sentence recognition scores were significantly worse with a SNR of 0 (80%), but reached ceiling above with a SNR of +8 dB. Meanwhile, PPD is significantly bigger with a SNR of 0 dB than +8 dB, and no significant difference above +8 dB. Therefore, a similar pattern should be observed in this study.

DESIGN & PROCEDURE

Considering that our lab only adopted the pupillometry technique recently, the first step was to replicate experiment designed to validate that the new eyetracker and lab set-ups were robust and precise enough to capture event-evoked pupillary responses. 12 participants (6 English native speakers and 6 French native speakers) participated in this first experiment.

Participants comfortably sat on a rigid chair in the soundproof room, 2m in front of the screen monitor. They wore infrared binocular eyetracker as well as headphones directly connected to the external soundcard (all audio stimuli were presented through them). The room and the screen luminance levels were adjusted to reach around 80 lx (measured using a light meter with the sensor positioned at the same height as the participants' left eye and facing the screen). The luminance levels were fixed throughout the experiments, to avoid changes in light level inducing task-unrelated pupillary response.

After listening instructions and being presented with a demo of the experiment procedure and its instructions, participants firstly listened to five sentences (excluded from the test) at +14 dB SNR to familiarise themselves with the test and typical sentences of the speech material.

In each trial, the presentation of the speech-shaped noise masker (or quiet in the quiet condition) started 3s before the onset of the sentences and was always fixed at 65 dB. This was set to provide time for the pupils to recover from the previous trial to avoid carry-over effect (1s) and to measure pre-task baseline pupil diameter (2s). Participants were instructed to focus on a black cross displayed at the center of the screen. After 3s, the sentence was played along with the continuous noise (or without noise in the quiet condition), and the presentation of the masker noise (or quiet in the quiet condition) was turned off 2s after the offset of the sentence, to allow the pupil to reach its peak. Upon the masker offset, participants were prompted by the black cross turning into a circle displayed at the screen center to repeat the sentence. This delayed verbal response ensured that the participants' speech motor commands did not affect the pupillary response corresponding to processing of the sentence perceived.

Their verbal responses were scored by the experimenter based on the number of key words correctly repeated. Then the program proceeded to the next trial (i.e. the next sentence). Twenty sentences were tested for each condition (0 dB, +7 dB, +14 dB and quiet), and the sequence of the four conditions was randomised for each participant.

Finally, in order to investigate the interaction between the light and the task-evoked pupillary responses, we renewed exactly this same design (Figure 4) with three lighting levels (around 0, 80 and 220 lx) and two SNR levels (0 dB and +14 dB) totalling (3x2) 6 conditions. These conditions were also randomised for each participant.

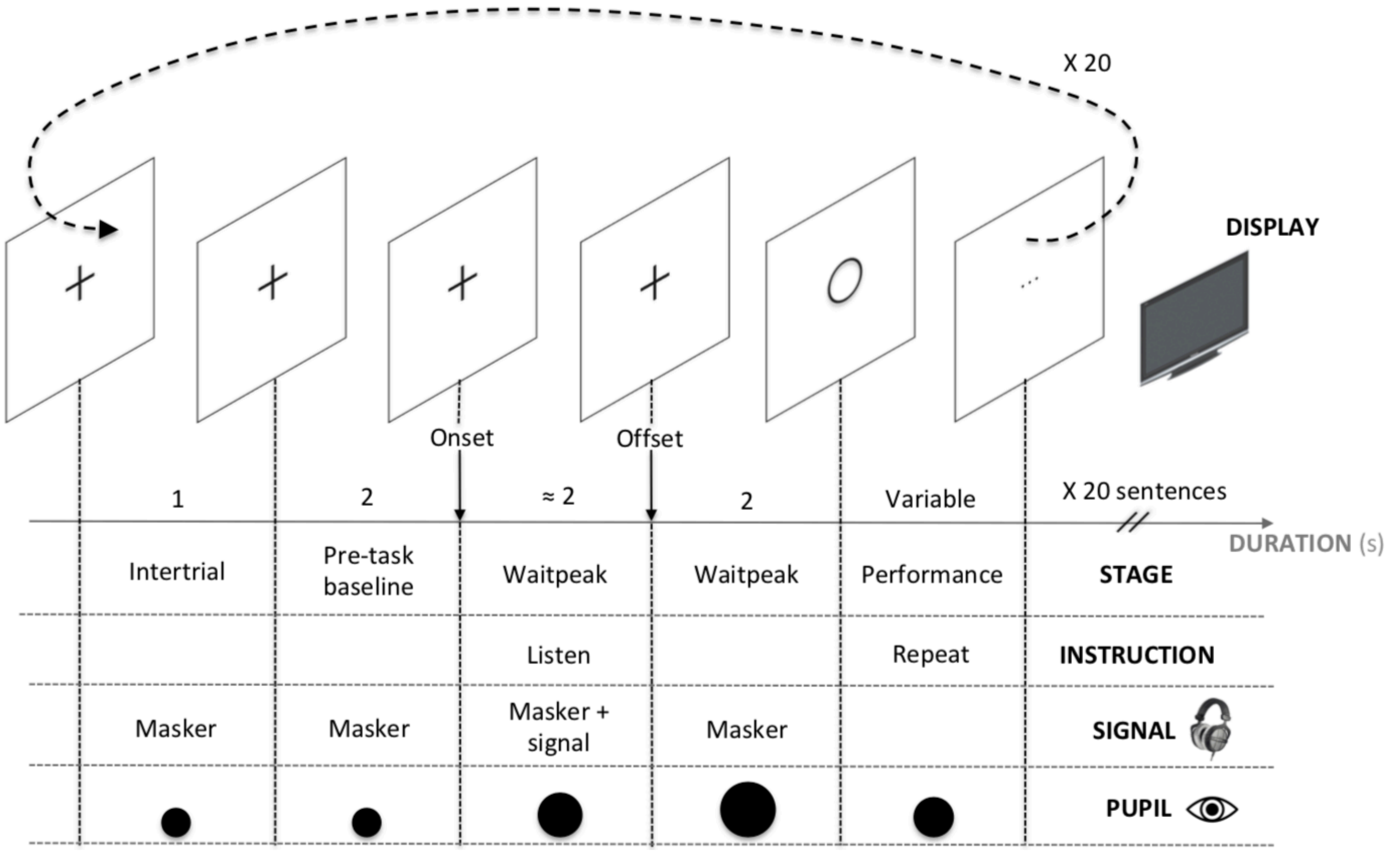


Figure 4. Procedure illustration for one trial. Pupil sizes drawn only represent an indication of the main trend; those may be different during real trials.

DATA ANALYSIS

To examine the effect of SNR and the lighting condition on sentence recognition, a logistic regression model is fitted to the listeners' sentence recognition performance using SNR condition and lighting condition as independent variables. The model is compared to the base model with only the intercept using the Chi-squared test based on the changes in the model deviance. Variables are only retained if they significantly improve the model's fit ($p < 0.05$). Relation amongst each level of the two variables and their interaction are examined with the post hoc Wald test using the coefficient and its standard error ($p < 0.05$). Pupil traces are cleaned using procedures similar to previous studies prior to the analysis (Zekveld, Kramer and Festen, 2011; Zekveld and Kramer, 2014). Baseline pupil diameter in each trial is calculated as averaged pupil traces 2s before the start of the sentence. Usually, the rest of the pupil size measurements are subtracted by the baseline pupil diameter in order to obtain the TEPR. Pupil diameter values below 3 standard deviations (SD) of the whole recording's mean are coded as blinks. Traces within 25 data points, i.e. 500 ms before the start and after the end of the blink are cubically interpolated in Matlab, to decrease the impact of the obscured pupil from blinks. Trials with over 20% of the data points interpolated are excluded. All valid traces are then low-pass filtered at 10 Hz with a first order Butterworth filter to preserve only cognitively related pupil size modulation (Klingner, Kumar and Hanrahan, 2008). Two indices of the TEPR (mean pupil dilation and peak pupil dilation) are obtained from the processed traces within the time window between the onset of sentence and offset of the masker. To examine the effect of SNR and lighting conditions on these indices, two logistic regression models are fitted to each pupil response index using SNR and lighting conditions as independent variables. The model is compared to the base model with only the intercept using the Chi-squared test to examine whether SNR is a significant main effect ($p < 0.05$). The interaction between two independent variables is examined with the post hoc Wald test ($p < 0.05$).

Another two linear regression models are built with the baseline pupil diameter and PPD as dependent variables using SNR and lighting condition as independent variables to examine whether there is any different lighting impact on the tonic and on the phasic components of the pupil responses.

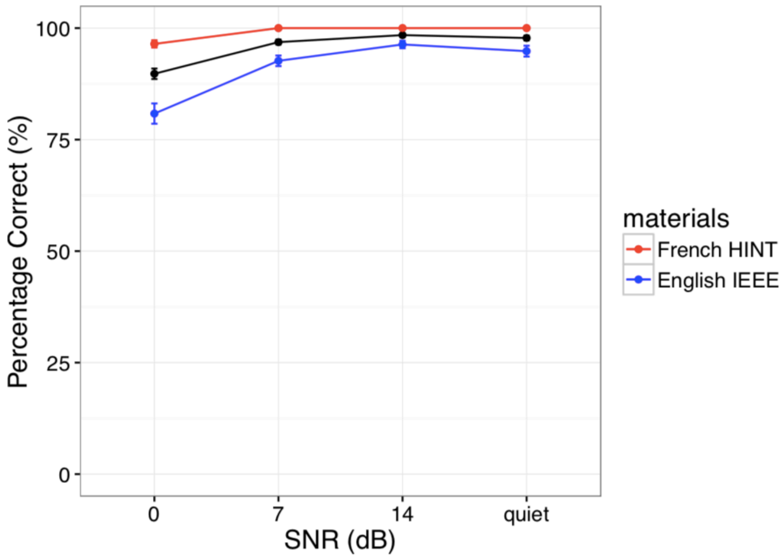
RESULTS

EXPERIMENT OF REPLICATION (n = 12)

Behaviour. There is a significant main effect of SNR ($p < 0.001$), language ($p < 0.001$) and a significant interaction ($p < 0.001$) on the sentences recognition.

At 0 dB SNR, French speech material have higher sentence recognition than English speech material ($p < 0.001$). For the English sentences, 0 dB SNR has lower sentence recognition score than 7 dB SNR ($p < 0.001$), 14 dB SNR ($p < 0.001$) and quiet ($p < 0.001$) whereas 7 dB SNR has lower sentence recognition than 14 dB SNR ($p < 0.05$) but no significant difference with quiet ($p = 0.12$). For French sentences, there is no significant difference between all the SNRs level.

Figure 5. Performance scores in function of SNR level and speech materials. Red curve represents results for French native speakers and blue curve represents results for English native speakers. Error bars represent the standard error of the mean.



Pupillary activity.

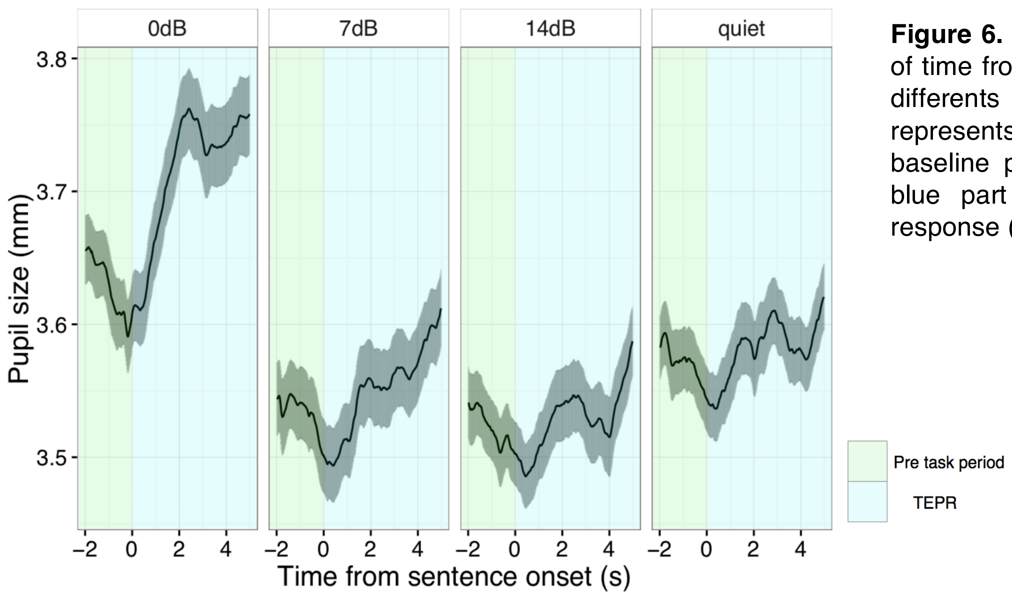
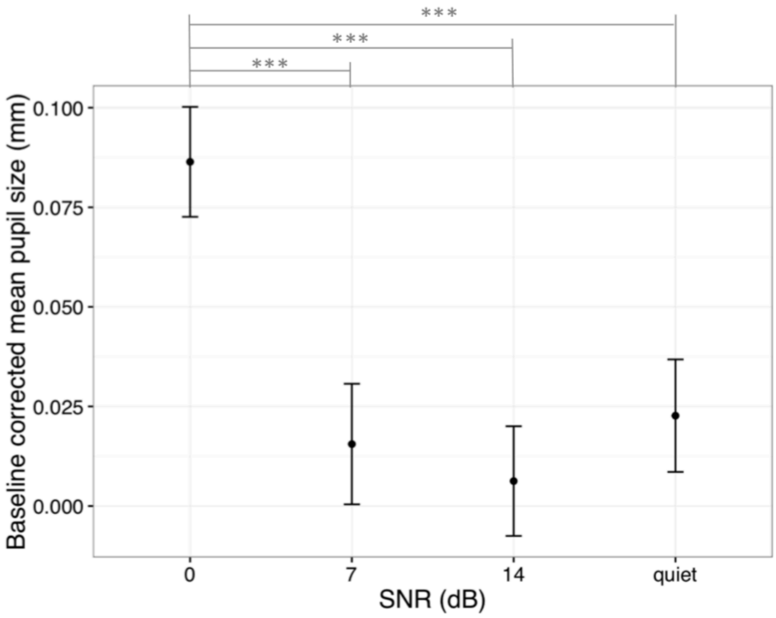


Figure 6. Pupil size (y-axis) in function of time from stimuli onset (x-axis) for 4 differents SNR levels. Green part represents the pre task period (where baseline pupil size is calculated) and blue part is the task evoked pupil response (TEPR).

For baseline corrected mean pupillary responses (Figure 7), only SNR is a significant main effect ($p < 0.001$), no effect of language or the interaction between the two. 0 dB SNR evokes bigger responses than 7 dB SNR ($p < 0.001$), 14 dB SNR ($p < 0.001$) and quiet ($p < 0.001$). 7dB SNR has no difference with 14 dB SNR ($p = 0.94$) and quiet ($p = 0.64$). 14 dB SNR has no difference with quiet ($p = 0.69$).

Figure 7. Baseline corrected mean pupil size (y-axis) in function of SNR levels (x-axis). Error bars represent the standard error of the mean.



For event-evoked peak pupillary response (Figure 8), only SNR is a significant main effect ($p < 0.001$), no effect of language or the interaction between the two. 0 dB SNR evokes bigger baseline corrected peak pupil size than 7 dB SNR ($p < 0.001$), 14 dB SNR ($p < 0.001$) and quiet ($p < 0.001$). 7 dB SNR evokes similar phasic pupil response as 14 dB SNR ($p = 0.66$), and quiet ($p = 0.29$). No difference between 14 dB SNR and quiet ($p=0.13$).

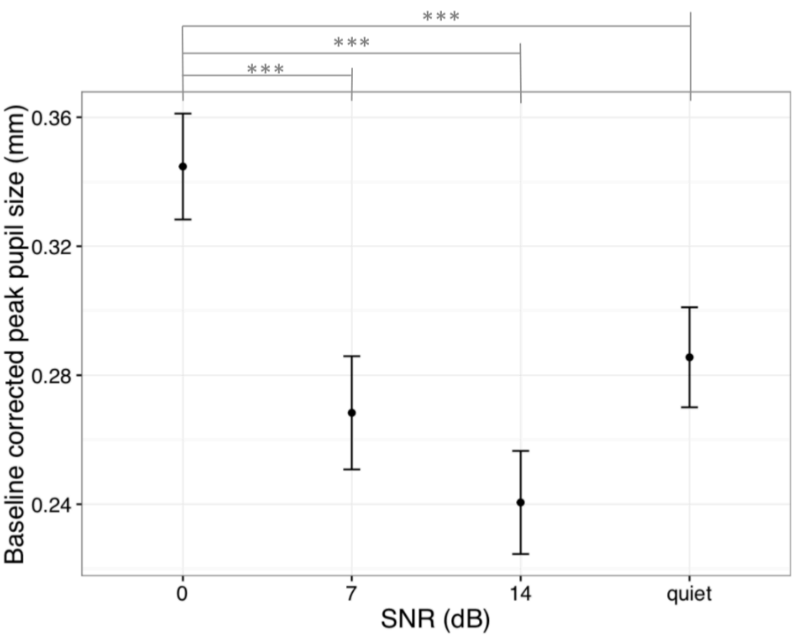


Figure 8. Baseline corrected peak pupil size (y-axis) in function of SNR levels (x-axis). Error bars represent the standard error of the mean.

LIGHTING EXPERIMENT (n = 18)

Behaviour. SNR has a significant main effect ($p < 0.001$), but luminance level ($p = 1$) and the interaction between both ($p = 0.94$) have no significant effect on sentences recognition (Figure 9). 14 dB SNR has higher sentence recognition scores than 0 dB SNR ($p < 0.001$).

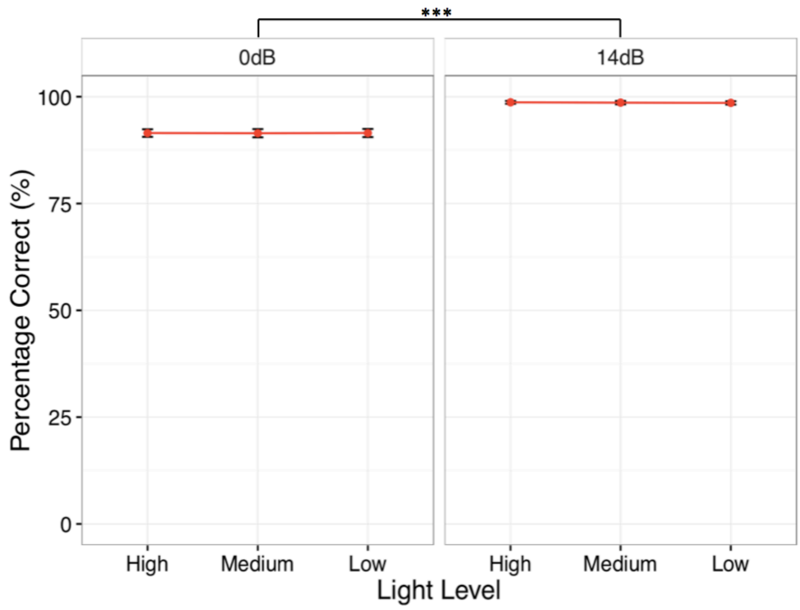


Figure 9. Intelligibility performance (y-axis) in function of lighting condition (x-axis) and SNR level. Error bars represent the standard error of the mean.

Pupillary activity.

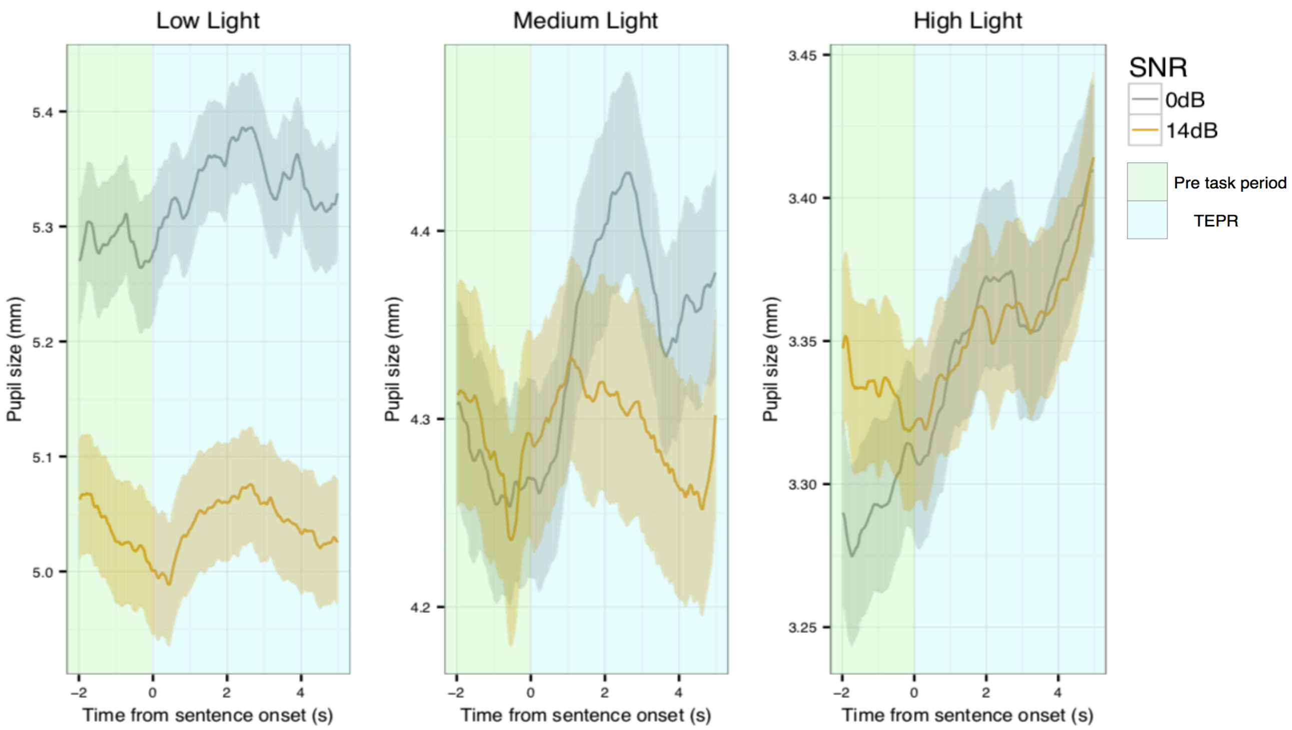


Figure 10. Pupil size (y-axis) in function of time from stimuli onset (x-axis) for 3 differents lighting conditions. Note the y-scale is different. Green part represents the pre task period (where baseline pupil size is calculated) and blue part is the task evoked pupil response (TEPR). Silver line represents 0 dB SNR and gold line represents 14 dB SNR.

Baseline pupil size. Luminance level ($p < 0.001$), SNR ($p < 0.001$) and their interaction ($p < 0.05$) affect significantly the tonic pupil diameter (Figure 11). Low light level gives bigger baseline than high light level ($p < 0.001$) and than medium light level ($p < 0.001$) whereas medium light level gives bigger baseline than high light level ($p < 0.001$). 14 dB SNR has smaller baseline than 0 dB SNR (< 0.001). In low light, 14 dB SNR has smaller baseline than 0 dB SNR ($p < 0.001$); no significant difference in medium light ($p = 0.76$) and in high light ($p=0.15$).

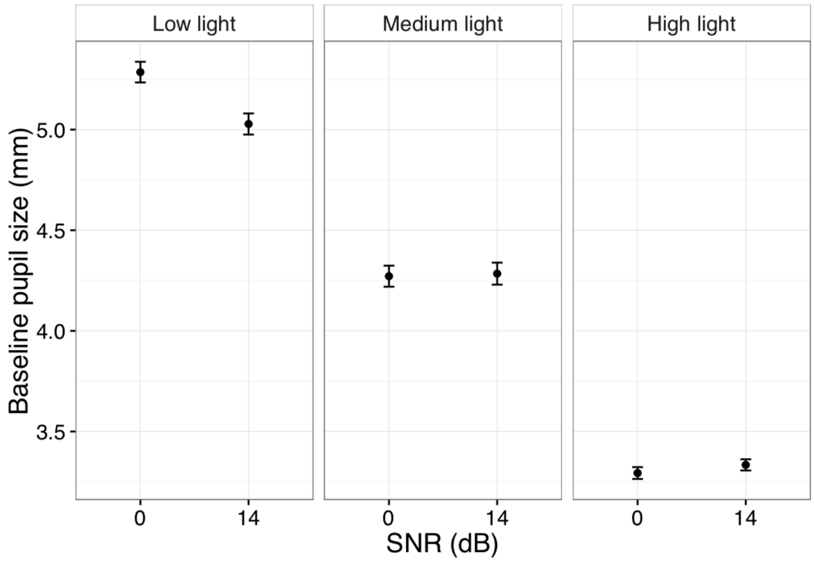


Figure 11. Baseline pupil size in function of SNR levels and lighting conditions. Error bars represent the standard error of the mean.

Peak pupil size. Luminance level ($p < 0.001$), SNR ($p < 0.001$) and their interaction ($p < 0.001$) affect significantly the peak pupil size (Figure 12). The low light condition gives bigger PPD than high light level ($p < 0.001$) and than medium light level ($p < 0.001$) whereas the medium light condition gives bigger peak size than high light level ($p < 0.001$). 14 dB SNR gives smaller peak size than 0 dB SNR ($p < 0.001$). In low light ($p < 0.001$), medium light ($p < 0.001$) and high light condition ($p < 0.05$), 0 dB SNR has bigger absolute peak size than 14 dB SNR.

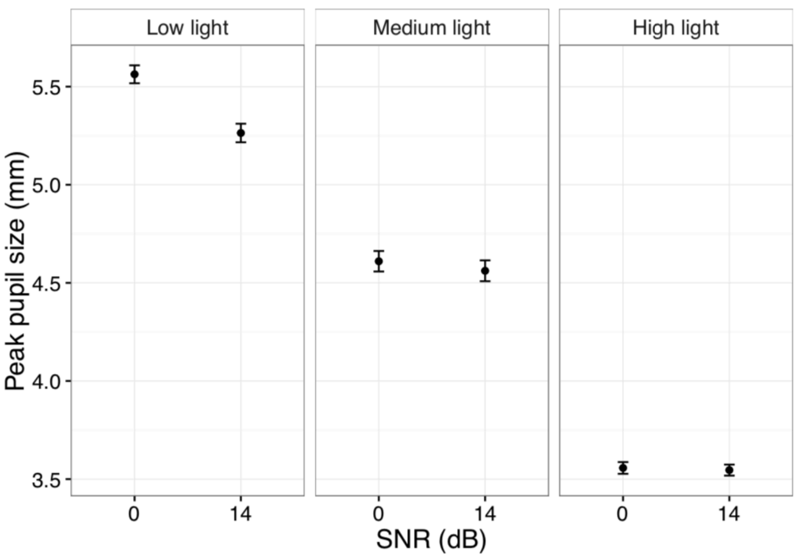


Figure 12. Peak pupil dilation (y-axis) in function of SNR levels (x-axis) for 3 differents lighting conditions. Error bars represent the standard error of the mean.

Using absolute peak and baseline pupil size difference, SNR ($p < 0.001$) and luminance level ($p < 0.001$) are significant main factors, but not their interaction (Figure 13). In all luminance levels, the phasic pupil response relative to the baseline is bigger in 0 dB SNR than 14 dB SNR ($p < 0.001$). Across two SNR levels, medium luminance level evokes bigger PPD relative to the baseline than low ($p < 0.05$) and than high level ($p < 0.001$) whereas PPD relative to the baseline is similar between low and high level ($p = 0.08$).

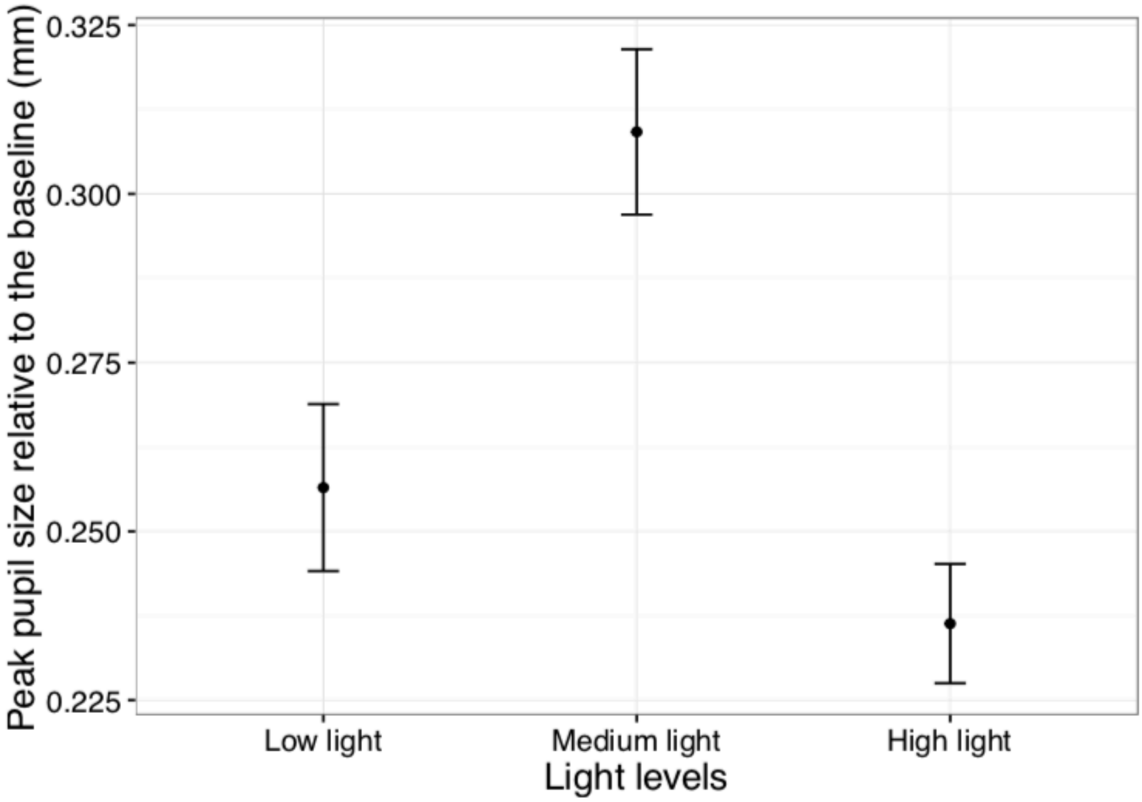


Figure 13. PPD (y-axis) in function of lighting levels (x-axis). These absolute values of PPD are obtained in substracting peak found in the blue area by baseline calculated in the green area (Figure 10). SNR levels (0 dB and +14 dB) are here averaged (due to the no significant interaction). Error bars represent the standard error of the mean.

CONCLUSION

DISCUSSION

Experiment of replication. The first step was to validate our devices and experimental procedure allowing the collection of reliable and robust pupil measurements. In this aim, we chose to used a similar design as in the Ohlenforst *et al.* (2017, 2018) studies. The only difference concerned our participants: they spoke two different mother tongues. Unfortunately, the listeners' sentence recognition performance was significantly better with the French sentences than the English ones. On the other hand, though there was this significant difference in sentence recognition performance, we did not observe any significant interaction of speech material on the pupil size probably due to the small sample size.

However, we successfully replicated these previous studies because we observed the same trend: a significant main effect of SNR on sentence recognition (at 0 dB SNR, performance score lower than all SNR levels; Figure 5) and a significant main effect of SNR on the TEPR (at 0 dB SNR, PPD and mean pupillary responses are bigger compare to all SNR levels; Figures 7 et 8).

Lighting experiment. As previously found in the experiment of replication, the SNR level had a significant main effect on the sentence's recognition and, as expected, the luminance level did not interfere on this behavioural task (Figure 9). On the other hand, the SNR level, the lighting condition and their interaction are significantly affected the tonic pupil diameter and the phasic pupil response (Figures 11 & 12), while the Peysakhovich, Vachon and Dehais (2017) results were significant only for the baseline and those of Steinhauer *et al.* (2004) only for the TEPR. That means that the luminance level affect the potential of observing a listening effort change and, although the tasks evoked were not auditory, we join these researchers on the importance of systematically reporting the luminance level when using pupillometry technique.

Otherwise the absolute difference between peak pupil size and baseline pupil size is the reference to present the PPD. In this way we showed that difference value was significantly affected by the luminance level. Moreover across two SNR levels, it was the medium light that evoked bigger phasic pupil response than other light conditions (Figure 13).

CONCLUSION

The listening effort is a new dimension within auditory research that should not be overlooked and pupillometry appears as a promising and reliable method to evaluate it.

One of our aims was to show a possible interaction between the PNS system and the pupillometry measurements. Thus, after validating our materials and methods of using pupillometry by observing results from previous studies, we examined the effect of luminance on listening effort via pupillometry. We noted the importance of controlling the amount of light coming in the participant's eyes in showing the significant luminance impact on the pupil size when the difficulty of the listening tasks varied.

Another aim was to examine whether the light condition affected the parametric test results. In response to this question, we propose to use a medium light environment (around 80 lux) to show greater phasic pupil responses during the listening tasks.

For further, we only studied the impact of three luminance conditions; a future study could analyze them more thoroughly.

Meanwhile, there is evidence that the HI listeners might have different PNS system as seen with NH listeners (Wang *et al.*, 2016). This suggests that the interaction between light and task-evoked pupillary responses might be different in NH and HI or CI users. Therefore, this study also serves as a reference when comparing pupillometry results between different hearing populations.

REFERENCES

Aston-Jones, G. and Cohen, J. D. (2005) ‘An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance’, *Annual Review of Neuroscience*, 28(1), pp. 403–450. doi: 10.1146/annurev.neuro.28.061604.135709.

Bäckman, L. *et al.* (2005) ‘Cognitive impairment in preclinical Alzheimer’s disease: A meta-analysis.’, *Neuropsychology*. doi: 10.1037/0894-4105.19.4.520.

Beatty, J. (1982) ‘Task-evoked pupillary responses, processing load, and the structure of processing resources.’, *Psychological bulletin*, 91(2), pp. 276–92. doi: 10.1037/0033-2909.91.2.276.

Beatty, J. and Lucero-Wagoner, B. (2000) ‘The pupillary system’, in *Handbook of psychophysiology (2nd ed.).*

Benarroch, E. E. (2009) ‘The locus ceruleus norepinephrine system: Functional organization and potential clinical significance’, *Neurology*, 73(20), pp. 1699–1704. doi: 10.1212/WNL.0b013e3181c2937c.

Benedetto, S. *et al.* (2014) ‘Effects of luminance and illuminance on visual fatigue and arousal during digital reading’, *Computers in Human Behavior*. Elsevier Ltd, 41, pp. 112–119. doi: 10.1016/j.chb.2014.09.023.

Berridge, C. W. and Waterhouse, B. D. (2003) ‘The locus coeruleus–noradrenergic system: modulation of behavioral state and state-dependent cognitive processes’, *Brain Research Reviews*, 42(1), pp. 33–84. doi: 10.1016/S0165-0173(03)00143-7.

Bremner, F. D. (2012) ‘Pupillometric Evaluation of the Dynamics of the Pupillary Response to a Brief Light Stimulus in Healthy Subjects’, *Investigative Opthalmology & Visual Science*, 53(11), p. 7343. doi: 10.1167/iovs.12-10881.

Deary, I. J. *et al.* (2009) ‘Age-associated cognitive decline’, *British Medical Bulletin*, 92(1), pp. 135–152. doi: 10.1093/bmb/ldp033.

Gilzenrat, M. S. *et al.* (2010) ‘Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function’, *Cognitive, Affective, & Behavioral Neuroscience*, 10(2), pp. 252–269. doi: 10.3758/CABN.10.2.252.

GRANHOLM, E. *et al.* (1996) ‘Pupillary responses index cognitive resource limitations’, *Psychophysiology*, 33(4), pp. 457–461. doi: 10.1111/j.1469-8986.1996.tb01071.x.

Hicks, C. B. and Tharpe, A. M. (2002) ‘Listening Effort and Fatigue in School-Age Children With and Without Hearing Loss’, *Journal of Speech Language and Hearing Research*, 45(3), p. 573. doi: 10.1044/1092-4388(2002/046).

Ishikawa, H. *et al.* (2012) ‘Effects of selective-wavelength block filters on pupillary light reflex under red and blue light stimuli’, *Japanese Journal of Ophthalmology*. doi: 10.1007/s10384-011-0116-1.

Klingner, J., Kumar, R. and Hanrahan, P. (2008) ‘Measuring the task-evoked pupillary response with a remote eye tracker’, in *Proceedings of the 2008 symposium on Eye tracking research & applications - ETRA ’08*. New York, New York, USA: ACM Press, p. 69. doi: 10.1145/1344471.1344489.

Koelewijn, T. *et al.* (2012) ‘Pupil dilation uncovers extra listening effort in the presence of a single-talker masker . PubMed Commons’, 33(2), pp. 291–300. doi: 10.1097/AUD.0b013e3182310019.Pupil.

Koelewijn, T. *et al.* (2014) ‘The pupil response is sensitive to divided attention during speech processing’, *Hearing Research*. The Authors, 312, pp. 114–120. doi: 10.1016/j.heares.2014.03.010.

Kuchinsky, S. E. *et al.* (2013) ‘Pupil size varies with word listening and response selection difficulty in older adults with hearing loss’, *Psychophysiology*, 50(1), pp. 23–34. doi: 10.1111/j.1469-8986.2012.01477.x.

Kuchinsky, S. E. *et al.* (2014) ‘Speech-perception training for older adults with hearing loss impacts word recognition and effort.’, *Psychophysiology*, 51(10), pp. 1046–57. doi: 10.1111/psyp.12242.

Loewenfeld, I. E. and Lowenstein, O. (1993) ‘The pupil : Anatomy, physiology, and clinical applications’, 2(9). Available at: https://academic.oup.com/brain/article-lookup/doi/10.1093/brain/124.9.1881.

Mather, M. and Harley, C. W. (2016) ‘The Locus Coeruleus: Essential for Maintaining Cognitive Function and the Aging Brain’, *Trends in Cognitive Sciences*. Elsevier Ltd, 20(3), pp. 214–226. doi: 10.1016/j.tics.2016.01.001.

Murphy, P. R. *et al.* (2011) ‘Pupillometry and P3 index the locus coeruleus-noradrenergic arousal

function in humans', *Psychophysiology*, 48(11), pp. 1532–1543. doi: 10.1111/j.1469-8986.2011.01226.x.
Murphy, P. R. *et al.* (2014) 'Pupil diameter covaries with BOLD activity in human locus coeruleus', *Human Brain Mapping*, 35(8), pp. 4140–4154. doi: 10.1002/hbm.22466.
Newsome, D. A. and Loewenfeld, I. E. (1971) 'Iris mechanics II. Influence of pupil size on details of iris structure', *American Journal of Ophthalmology*. doi: 10.1016/0002-9394(71)90133-4.
Ng, E. H. N. *et al.* (2013) 'Effects of noise and working memory capacity on memory processing of speech for hearing-aid users', *International Journal of Audiology*, 52(7), pp. 433–441. doi: 10.3109/14992027.2013.776181.
Ohlenforst, B. *et al.* (2017) 'Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation', *Hearing Research*. Elsevier B.V, 351, pp. 68–79. doi: 10.1016/j.heares.2017.05.012.
Ohlenforst, B. *et al.* (2018) 'Impact of SNR, masker type and noise reduction processing on sentence recognition performance and listening effort as indicated by the pupil dilation response', *Hearing Research*. Elsevier B.V, 365, pp. 90–99. doi: 10.1016/j.heares.2018.05.003.
Peysakhovich, V. *et al.* (2015) 'Frequency analysis of a task-evoked pupillary response: Luminance-independent measure of mental effort', *International Journal of Psychophysiology*. Elsevier B.V., 97(1), pp. 30–37. doi: 10.1016/j.ijpsycho.2015.04.019.
Peysakhovich, V., Vachon, F. and Dehais, F. (2017) 'The impact of luminance on tonic and phasic pupillary responses to sustained cognitive load', *International Journal of Psychophysiology*. Elsevier B.V., 112, pp. 40–45. doi: 10.1016/j.ijpsycho.2016.12.003.
Pfeifer, M. A. *et al.* (1983) 'Differential changes of autonomic nervous system function with age in man.', *The American journal of medicine*, 75(2), pp. 249–58. doi: 0002-9343(83)91201-9.
Pichora-Fuller, M. K. *et al.* (2016) 'Hearing Impairment and Cognitive Energy', *Ear and Hearing*, 37, p. 5S–27S. doi: 10.1097/AUD.0000000000000312.
Picou, E. M. (2013) '20Q: Listening effort - we know it's a problem but how do you measure it?', *AudiologyOnline*. Available at: https://www.audiologyonline.com/articles/listening-effort-we-know-it-12032.
Pomplun, M. and Sunkara, S. (2003) 'Pupil dilation as an indicator of cognitive workload in human-computer interaction', in *Proceedings of the International Conference on HCI*. New York, New York, USA: ACM Press. Available at: http://www.academia.edu/download/2928509/1xb5jw1d6oycvux.pdf.
Prettyman, R., Bitsios, P. and Szabadi, E. (1997) 'Altered pupillary size and darkness and light reflexes in Alzheimer's disease.', *Journal of neurology, neurosurgery, and psychiatry*, 62(6), pp. 665–8. doi: 10.1136/jnnp.62.6.665.
Rajkowski, J., Kubiak, P. and Aston-Jones, G. (1993) 'Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention', *Society for Neuroscience Abstracts*, 19, p. 974.
Rudner, M., Rönnberg, J. and Lunner, T. (2011) 'Working Memory Supports Listening in Noise for Persons with Hearing Impairment', *Journal of the American Academy of Audiology*, 22(3), pp. 156–167. doi: 10.3766/jaaa.22.3.4.
Rushton, J. P. and Ankney, C. D. (1996) 'Brain size and cognitive ability: Correlations with age, sex, social class, and race', *Psychonomic Bulletin & Review*. doi: 10.3758/BF03210739.
Sarampalis, A. *et al.* (2009) 'Objective measures of listening effort: effects of background noise and noise reduction.', *Journal of speech, language, and hearing research : JSLHR*, 52(5), pp. 1230–40. doi: 10.1044/1092-4388(2009/08-0111).
Steinhauer, S. R. *et al.* (2004) 'Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing', *International Journal of Psychophysiology*, 52(1), pp. 77–86. doi: 10.1016/j.ijpsycho.2003.12.005.
Tryon, W. W. (1975) 'Pupillometry: A Survey of Sources of Variation', *Psychophysiology*, 12(1), pp. 90–93. doi: 10.1111/j.1469-8986.1975.tb03068.x.
Tsukahara, J. S., Harrison, T. L. and Engle, R. W. (2016) 'The relationship between baseline pupil size and intelligence', *Cognitive Psychology*. Elsevier Inc., 91, pp. 109–123. doi: 10.1016/j.cogpsych.2016.10.001.
Usher, M. (1999) 'The Role of Locus Coeruleus in the Regulation of Cognitive Performance', *Science*, 283(5401), pp. 549–554. doi: 10.1126/science.283.5401.549.
Wang, Y. *et al.* (2016) 'Parasympathetic Nervous System Dysfunction, as Identified by Pupil Light Reflex, and Its Possible Connection to Hearing Impairment', *PLOS ONE*. Edited by Z. Kapoula, 11(4), p. e0153566. doi: 10.1371/journal.pone.0153566.

Wendt, D., Hietkamp, R. K. and Lunner, T. (2017) 'Impact of Noise and Noise Reduction on Processing Effort', *Ear and Hearing*, 38(6), pp. 690–700. doi: 10.1097/AUD.0000000000000454.
Winn, B. *et al.* (1994) 'Factors affecting light-adapted pupil size in normal human subjects.', *Investigative ophthalmology & visual science*, 35(3), pp. 1132–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8125724.
Winn, M. B. (2016) 'Rapid Release From Listening Effort Resulting From Semantic Context, and Effects of Spectral Degradation and Cochlear Implants', *Trends in Hearing*, 20, p. 233121651666972. doi: 10.1177/2331216516669723.
Winn, M. B., Edwards, J. R. and Litovsky, R. Y. (2016) 'The Impact of Auditory Spectral Resolution on Listening Effort Revealed by Pupil Dilation.', *Ear and hearing*, 36(4), pp. e153-65. doi: 10.1097/AUD.0000000000000145.
Xu, J. *et al.* (2011) 'Pupillary Response Based Cognitive Workload Measurement under Luminance Changes', in, pp. 178–185. doi: 10.1007/978-3-642-23771-3_14.
Zekveld, A. A. and Kramer, S. E. (2014) 'Cognitive processing load across a wide range of listening conditions: Insights from pupillometry', *Psychophysiology*, 51(3), pp. 277–284. doi: 10.1111/psyp.12151.
Zekveld, A. A., Kramer, S. E. and Festen, J. M. (2011) 'Cognitive load during speech perception in noise: the influence of age, hearing loss, and cognition on the pupil response.', *Ear and hearing*, 32(4), pp. 498–510. doi: 10.1097/AUD.0b013e31820512bb.