Pupillometry in language acquisition research?

Tom Fritzsche
tomfri@uni-potsdam.de

CLEAR workshop Groningen September 10, 2010
Abstract

Pupil size is not only influenced by exogenous factors (e.g. luminance) but by endogenous ones as well, especially emotion, arousal and cognitive load. Automatic eyetrackers log pupil size along with gaze data - so why not make use this information?

The technique of pupillometry/TEPRs is introduced and findings from the literature are presented before discussing methodological issues and the potential application in language acquisition research.
Outline

1. Why pupillometry
2. Basics
3. Classics
   Hess & Polt 1960, Kahneman & Beatty 1966
4. Infant studies
   Gredebäck & Melinder 2010, Jackson & Sirois (2009)
5. Pupil size & language
   Just & Carpenter (1993); Engelhardt et al. (2010)
6. Methodology
7. Example
8. Discussion
1. Why pupillometry?

- "Pupil workshop" at the University of Potsdam in March 2010
- Increasing no. of papers in the last years
1. Why pupillometry?

Pupil size is

• A sensitive, reliable, consistent measure of cognitive load
• A dynamic (online) measure of processing
• Ideal for infant research
• Provided by most automatic eyetrackers e.g. Tobii, SMI, EyeLink
• Compatible with existing VoE paradigms
• Available in addition to other measures
2. Basics: History

- Connection between cognitive processing and pupillary dilation at the end of the 19th century

- "Every active intellectual process, every psychological effort, every exertion of attention, every active mental image, regardless of content, particularly every affect just as truly produces pupil enlargement as does every sensory stimulus."
  Oswald Bumke in 1911, (cited from Hess 1975 p. 23-24)

- Pupillometry pioneer: Otto Lowenstein
  Thompson 2005

  - book comprises the work of Lowenstein and Loewenfeld
  - standard reference on pupil research

- Since 1960s: Jackson Beatty, Eckhard Hess, Daniel Kahneman
2. Basics: Parameters

- Pupil size controlled by two opposing muscles
- Typical size: 3-5 mm (range: 1-9 mm)
  gets smaller by .04 per year
- Psychosensory reflex
  - After any sensory occurrence, 
    external (visual, auditory, tactile etc.) or 
    internal (emotions, mental processes, intentional effort), > dilation
    Exception: light -> constriction
- Hippus (pupillary unrest)
- Movements:
  - Begin after 200 ms
  - Peak .5 to 2 s after stimulus
- Pupil diameter covaries with
  Kahneman 1973, Bradley et al. 2008
  - Skin conduction
  - Heart rate (less reliably)

from: Beatty & Lucero-Wagoner 2000
2. Basics: Influencing factors

- Luminance
- Arousal
- Emotion
- Attention
- Cognitive load/intensity, resource allocation, mental effort
- Task complexity
- Intelligence  
  Ahern & Beatty 1977, van der Meer et al. 2010
2. Basics: TEPR

• TEPR = task-evoked pupillary responses
  – Non-reflexive phasic pupillary movement
  – No causal link to but reporter variable for brain functions

• Changes of about .1-.5 mm
  – Independent of baseline dilation Bradshaw 1969
  – Comparable between participants and different labs (Normalization unnecessary Beatty & Lucero-Wagoner 2000)

• Response delay 200-300 ms
• Peak at about 1200 ms
2. Basics: TEPR – index of cognitive load

- Arithmetic problems Hess & Polt 1964
- Memory
  - STM load Kahneman & Beatty 1966
  - LTM retrieval Beatty & Kahneman 1966
- Perception (e.g. pitch discrimination)
- Language
  - Syntactic complexity Schluroff 1982, Just & Carpenter 1993
  - Grammaticality violations Gutiérrez & Shapiro 2010
  - Sentence comprehension Wright & Kahneman 1971
  - Context integration Engelhardt et al. 2010
- Attention Beatty 1982, 1988
- Responding (similar to ERPs - CNV or RP)
- Standard tests of "concentration"

- Pictures shown for 10 s after a 10 s baseline
- Participants: 6 (2, 4)
- Data points every 500 ms

- Highly reliable results (1 day retest)
- Interpretation: Pupil size reflects interest value
- Bidirectional TEPR (later: no confirmation)
  - pleasant > dilation
  - unpleasant > constriction
3. Classics: Kahneman & Beatty 1966

- Pupil size reflects memory load
- Participants: 5
- digit (3-7) memory task

- Memorization: Dilation
- Recall: Constriction
- Practice effects: reduction in repetition -> processing load

Replication by Klingner et al. 2008 with a Tobii 1750
4. Infant Studies: Gredebäck & Melinder 2010

• Action understanding in infants

• Participants
  – 28 6-month olds
  – 28 12-month olds

• Pupil size varies with rationality
  – Larger dilation for non-rational vs. rational goal

• Replication with familiarity bias removed
4. Infant Studies: Jackson & Sirois 2009

- Event understanding: Violation of Expectation (VoE) paradigm
- Participants: 24 8-month olds
- Pupil size bigger for impossible vs. possible novel events
5. Pupil size & language: Just & Carpenter 1993

- Reading study with 35 adults
- Processing in subject vs. object relative clauses
  \[\text{The reporter that ... attacked the senator} \]
  \[\text{the senator attacked... admitted the error publicly after the hearing.}\]
- Measurement after first fixation of the main verb for max. 3 s
- Results
  - Larger pupil size in object vs. subject RCs
  - Later peak in ORC vs. SRC (116 ms)
  - Pupil size is proportional to
    - reading time
    - error rate (of probe questions)
  - TEPR reflects intensity of processing
Pupil size & language: Engelhardt et al. 2010

- Listening study with 2x18 adults
- Integration of prosody and visual context in garden-path sentences
  
  \[While\ the\ woman\ cleaned\ (#)\ the\ dog\ that\ was\ big\ and\ brown\ stood\ in\ the\ yard.\]

- Measurement after 200 ms of onset disambiguating word for 1200 ms

- Results
  - Larger pupil size for conflicting vs. congruent prosody
  - If additional visual context was
    - consistent (\textit{woman cleans not a dog}): no effect of prosody
    - inconsistent (\textit{woman cleans a dog}): effect of prosody
  - TEPR reflects processing effort of spoken language comprehension
6. Methodology: Luminance

- Pupillary light reflex (initial constriction) is much bigger and masks TEPRs
- CONTROL for
  - Luminance of stimuli
  - Lighting of room

from: Bradley et al. 2008
6. Methodology: Steps in pupil analysis

1. Data inspection & artifact rejection
2. Noise filtering
   • Low pass with 10 or 15 Hz
3. Replacement of missing data
   • Use of other eye (if available)
   • Linear interpolation
4. Averaging both pupil size values
5. Baseline creation
6. Calculating baseline-adjusted measures
7. (Normalizing)
6. Methodology: Pupillometry parameters

- **Mean size**
  - Average in an interval of interest
- **Peak size**
  - Maximum dilation in an interval of interest
- **Peak latency**
  - Amount of time before peak size is reached
- **Baseline**
  - Often 500 ms before critical stimulus
  - Neutral stimulus
6. Methodology: Pupil size units

• Diameter units
  – \textit{mm}
    Tobii: mean/real (difference: T/X series vs. 1750)
    SMI: horizontal and vertical
  – \textit{pixels} (SMI)
  – \textit{arbitrary unit} (EyeLink)

• Pupil area (also different units)
  
  
  
  Not directly comparable between different systems
  Not comparable between different participants (in raw form - like EEG data)
6. Methodology: Normalization

- Really needed?
- Normalized data, relative to
  - Baseline
  - Participant's pupil mean
  - Condition's pupil mean
6. Methodology: Analysis techniques

- Means (of time windows)
  Laeng et al. 2007; Falk-Ytter 2008
- Slope in a time window
  Engelhardt et al. 2010
- B-splines
  Jackson & Sirois 2009
- Like EEG/ERP data
- PCA
- Fixation-based pupil measures
6. Methodology: Pupil vs. Gaze

<table>
<thead>
<tr>
<th>Property</th>
<th>Pupil</th>
<th>Gaze</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay</td>
<td>≈ 200-300 ms peak at ≈ 1200 ms</td>
<td>≈ 200 ms</td>
</tr>
<tr>
<td>Repetition effects</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>AoI-based analysis</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Light sensitivity</td>
<td>High</td>
<td>None</td>
</tr>
<tr>
<td>Missing Data</td>
<td>Easy to interpolate</td>
<td>Hard to recover</td>
</tr>
</tbody>
</table>

- Design should be optimized for pupil or gaze
7. Ex: Visual Categorization Study

Fritzsche & Höhle

- Can infants categorize different nonsense syllable strings by mapping them onto two characters?
- Familiarization phase with 2 conditions: FSG, PSG
- Participants
  - 17 five-year-old children (mean 5.7 yrs)
  - 24 adults (mean 23.5 yrs)
7. Ex: Stimuli

Familiarization Trial

Test Trial
7. Ex: Gaze data
7. Ex: Pupil data management

1. Linear interpolation of missing data

<table>
<thead>
<tr>
<th>data points</th>
<th>chunk length</th>
</tr>
</thead>
<tbody>
<tr>
<td>5yo</td>
<td>35%</td>
</tr>
<tr>
<td>adults</td>
<td>21%</td>
</tr>
</tbody>
</table>

2. Combining both pupil size values into one

3. Baseline 500 ms pre-stimulus

4. Normalization by mean participant pupil diameter (just to see what changes)
7. Ex: Pupil data

<table>
<thead>
<tr>
<th>Pupil size change in Familiarization</th>
</tr>
</thead>
<tbody>
<tr>
<td>5yo</td>
</tr>
<tr>
<td>Auditory Presentation</td>
</tr>
<tr>
<td>adults</td>
</tr>
<tr>
<td>Auditory Presentation</td>
</tr>
</tbody>
</table>

Baseline data

![Graph showing pupil size change over time](image-url)
7. Ex: Summary

- TEPR gives additional information
  - Greater dilation in PSG vs. FSG condition for the familiarization in adults (although equal gaze patterns)
  - This condition has more correct looks in the test phase
  - TEPR seems to index the processing load of processing stimuli of this condition
So ...

- Pupillometry in language acquisition research?

It's worth a try!

*FIG. 6. Recording pupil movements (late 1920s). This is one of Lowenstein’s early arrangements. The lights were still very hot and had to be cooled by circulating water.*
8. Discussion
References


