**Online supplementary material for the article *Pupillometry in linguistic research – an introduction and review for second language researchers***

The main text described the most common ways in which pupil data have been analyzed to date. However, much data is lost by averaging over a trial or by taking only the peak dilation in a trial. Another problem arises when the largest dilation coincides with the end of a trial across conditions, in which case the analysis of peaks would be meaningless. One way around these drawbacks is to use time-series analysis, or growth-curve analysis, where change is modeled over time during the course of a trial. Describing this statistical technique here in detail is beyond the scope of the article and the reader is referred to pertinent introductions to growth-curve analysis for psycholinguists (Mirman, 2014; Mirman, Dixon, & Magnuson, 2008) as well as more general introductions to mixed-effects modeling (e.g., Baayen, 2008; Baayen, Davidson, & Bates, 2008; Gelman & Hill, 2007). In a nutshell, a regression model is fit to the data that includes not only a linear term for time but also the polynomials (time)2, (time)3 and so forth (e.g., *f*(*x*) = *ax* + *bx*2 + *cx*3 + *d*, where *x* is time) so as to approximate the curvature of the pupil dilation and constriction over the course of a trial. The number of polynomials included in the model depends on the shape of the pupil response. For example, Kuchinsky et al. (2013) used a polynomial function of fifth degree in their model to account for a second peak in dilation towards the end of a trial. The effect of an experimental manipulation can then be investigated by including interaction terms of experimental condition with the polynomials. Because this analysis is performed on the whole data set, it can be more sensitive to effects than an analysis of PA and PL (see Kuchinsky et al., 2013), but at the same time, it is computationally very intensive and computational resources needed for complex models may exceed the capacity of regular computers. Furthermore, the interpretation of interactions may not always be straightforward and researchers should be careful to relate any significant effects to their theoretical importance (for an example see Kalénine, Mirman, Middleton, & Buxbaum, 2012).

Yet another way to analyze pupil data was employed by Engelhardt et al. (2010). As described above, participants in this experiment heard sentences for which sentence prosody and syntactic structure were either congruent or conflicting. In their analysis, the authors focused on a time window of 1.2 s following the disambiguating word (The length of the time window was based on Just and Carpenter (1993). However, it should be noted that a later study (Gagl, Hawelka, & Hutzler, 2011) found that the pupil data reported in Just and Carpenter may have been confounded by gaze position. Therefore, it is recommended to visually inspect the data before the analysis to ensure that a 1.2 s time window is indeed appropriate). Engelhardt et al. then fit separate simple linear regression models for each participant in each condition and performed their subsequent analyses on the slope estimates of these regression models. Their reasoning was that a steeper slope would indicate greater processing effort. However, when researchers use a less common method of analysis, it may be advisable to also report PL and PA so that results across different studies can be compared. In addition, a visual representation of the data should always be included to help interpret the results of statistical tests. Another method that also preserves the time-course information of the pupil response is functional data analysis and the reader is referred to Jackson and Sirois (2009) and Sirois and Brisson (2014) for more details about this type of analysis. Further ways of analyzing pupil data are described in Demberg and Sayeed (2016) and McCloy, Larson, Lau, and Lee (2016). Interested readers are also referred to the *R* package eyetrackingR, which is described on the accompanying website (eyetracking-r.com).

References

Baayen, R. (2008). *Analyzing linguistic data*. Cambridge University Press.

Baayen, R., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. doi:10.1016/j.jml.2007.12.005

Demberg, V., & Sayeed, A. (2016). The frequency of rapid pupil dilations as a measure of linguistic processing difficulty. *PLoS ONE*, *11*(1), 1–29. doi:10.1371/journal.pone.0146194

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly Journal of Experimental Psychology*, *63*(4), 639–45. doi:10.1080/17470210903469864

Gagl, B., Hawelka, S., & Hutzler, F. (2011). Systematic influence of gaze position on pupil size measurement: analysis and correction. *Behavior Research Methods*, *43*(4), 1171–81. doi:10.3758/s13428-011-0109-5

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press.

Jackson, I., & Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Developmental Science*, *12*(4), 670–9. doi:10.1111/j.1467-7687.2008.00805.x

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, *47*(2), 310–39.

Kalénine, S., Mirman, D., Middleton, E. L., & Buxbaum, L. J. (2012). Temporal dynamics of activation of thematic and functional knowledge during conceptual processing of manipulable artifacts. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(5), 1274–95. doi:10.1037/a0027626

Kuchinsky, S. E., Ahlstrom, J. B., Vaden, K. I., Cute, S. L., Humes, L. E., Dubno, J. R., & Eckert, M. a. (2013). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology*, *50*(1), 23–34. doi:10.1111/j.1469-8986.2012.01477.x

McCloy, D. R., Larson, E. D., Lau, B., & Lee, A. K. C. (2016). Temporal alignment of pupillary response with stimulus events via deconvolution. *The Journal of the Acoustical Society of America*, *139*(3), EL57-EL62. doi:10.1121/1.4943787

Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: Chapman and Hall / CRC Press.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*(4), 475–494. doi:10.1016/j.jml.2007.11.006

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, (Ea 4700), n/a-n/a. doi:10.1002/wcs.1323