Applied Linguistics 2020: 1–7
doi:10.1093/applin/amaa041

REVIEW

Using Eye-Tracking Technology in Applied Linguistics and SLA: A Review of Conklin *et al.* (2018) and Godfroid (2020).

K. Conklin, A. Pellicer-Sánchez, and G. Carrol: EYE-TRACKING: A GUIDE FOR APPLIED LINGUISTICS RESEARCH. Cambridge University Press, 2018.
A. Godfroid: EYE TRACKING IN SECOND LANGUAGE ACQUISITION AND BILINGUALISM: A RESEARCH SYNTHESIS AND METHODOLOGICAL GUIDE. Routledge, 2020.

INTRODUCTION

As the authors of both of these texts note, online measurement of cognitive processing during language tasks is still a remarkably new development in the fields of Applied Linguistics (AL), Bilingualism, and Second-Language Acquisition (SLA). However, interest has been growing rapidly. One particular sub-area of interest in online processing is the study of eye movements, and these two books provide a helpful introduction to several key components: (i) an overview of eye tracking as a measure of attention; (ii) methodology for applied eye-tracking research; and (iii) how to collect and analyze eye-tracking data. The mechanics of study design and implementation, and challenges with data analysis and reporting, frame each text.

We found it best to structure this review as a conversation grounded in where we each found the books useful, and where supplemental information would be beneficial. Jonathan Malone provides an initial sketch of the theoretical issues surrounding eye-tracking methodology in AL/SLA ('Theoretical considerations'). Wei Yi contributes a number of methodological considerations from the perspective of an experienced applied researcher ('Methodological considerations'). Finally, Kaiwen Man addresses data standardization, psychometric issues, and the types of data modeling that could account for variability across studies ('Analytical considerations').

THEORETICAL CONSIDERATIONS

Employing terms such as 'richer', 'concurrent', 'online', 'ongoing', and 'time-sensitive' to describe the methodology, the authors of both books spend considerable time building the argument that online measurement is essential to language research. Each book traces its roots in cognitive psychology to the way eye-movement data serve as a measure of attention, often referencing Rayner's (2009) seminal account of eye-movement data and attention. Both also discuss assumptions behind the *eye-mind hypothesis* originally from Just and Carpenter (1980), the theoretical link between physical eye movements

and cognitive processing. Godfroid also provides a helpful comparison of how eye tracking has been modeled on page 52, including a discussion of cognitive modeling that assumes greater or lesser oculomotor (biological, low-level) control compared with more direct cognitive (higher-level) control. Godfroid also helpfully compares eye tracking with other online methodologies, for example, think-aloud protocols, self-paced reading, and event-related potentials (ERPs). Eye tracking and ERPs benefit from not requiring a secondary task from the participant to retrieve meaningful data, which exempt them from reactivity effects to which think-aloud and self-paced reading data are susceptible (Leow and Morgan-Short 2004).

Both books do well in summarizing primary areas of work to this point in AL/SLA/Bilingualism, but in different ways. Each spends substantial time reporting contemporary eye-tracking research on reading text. Audio–visual input, whether through subtitles or the visual world paradigm, is explored in some depth, with summaries of important studies in L1 research. Given Godfroid's scope, her book provides a much more detailed synthesis of the studies that have been conducted in L2 research fields, whether Bilingualism or SLA, whereas Conklin *et al.*'s broader scope of AL precludes such a sweeping review, leaving their report as a selected summary of relevant studies. Both books make proposals, with study designs embedded into Conklin *et al.*'s literature review chapters, whereas previous strands of work are expanded in Godfroid into an entire chapter of proposed research designs. As an SLA-focused researcher, I found Godfroid's sections on both areas to be more useful in understanding the breadth of previous and future work in the field, but the broader view of AL research in Conklin *et al.* may well be more useful to others, depending on context.

Ultimately, both books reveal basic theoretical assumptions behind online processing measurement, none of which are terribly complex or difficult to test. I was left wondering why, given the explosive advancement of technology in the 40 years since Just and Carpenter (1980), there has not been more and better work done in recent years. It seemed that many of Conklin *et al.*'s sample studies across their three theoretical chapters were fraught with methodological issues (which the authors note), while Godfroid's synthesis found only 52 *total* studies examining eye tracking with reading (including multimodal studies), and even fewer ($k = 32$) visual world studies. This could be indicative of the relative youth of the AL/SLA/Bilingualism fields more broadly, but it was surprising.

Given the value of online methodology in addressing processing in real time, one primary takeaway from both books is that there is a wide-open opportunity for those who are willing and have access to eye-tracking equipment. The fields of AL/SLA/Bilingualism are clearly in need of a wider base of rigorous studies building individual research questions into testable hypotheses and theoretical modeling for the interrelationships between eye movements, attention, language processing, and language learning. The work to this point as summarized in both books has been underwhelming at best,

whether from the design issues mentioned by Conklin *et al.* or the lack of a substantial body of L2 work evidenced in Godfroid. I am hopeful that these two books and the growing interest in online measurement will result in data that draw issues in language processing and acquisition into sharper relief, allowing for researchers and teachers to gain insights into language processing, as well as potentially inform teaching and learning.

METHODOLOGICAL CONSIDERATIONS

For anyone aspiring to design and carry out an eye-tracking experiment, it is necessary to gain a good understanding at conceptual and technical levels. Conklin *et al.* and Godfroid discuss the advantages of eye tracking in the beginning chapter of each book, laying a solid foundation for the reader to reflect on what eye tracking can do and whether the technique is useful in their context. Following the introduction section, Conklin *et al.* explain the function of eye-tracking hardware and software, available eye-tracking systems, and how to assess the quality of data based on sampling rate, accuracy, precision, and latency, closing with guidelines for selecting eye trackers for research purposes (Chapter 2). Godfroid also addresses technical specifications, but they occupy a less prominent role in the final chapter of her book. With respect to the design and administration of eye-tracking experiments, both books cover types of experimental design, the creation of counterbalanced stimuli lists with well-matched materials, sample size and statistical power, defining regions of interest, the selection of a combination of early and late eye-tracking measures, and the general procedure of building and running an experiment. Godfroid's book complements Conklin *et al.* by pointing out possible solutions to reduce Type I error due to the dependency of eye-tracking measures (Chapter 5), as well as by providing a possible rule of thumb regarding the number of items and participants needed to achieve a medium effect size for eye-tracking SLA studies (p. 156). Furthermore, she elaborates on the management of eye-tracking labs and offers bulleted lists of tips for beginners.

Both books carefully walk readers through essential conceptual and technical aspects of eye tracking, with hands-on research experience integrated. Novice researchers should find Godfroid's methodological section less challenging to read, due to its level of detail, illustrations, summaries, and explanations of key concepts. As an experienced user of eye tracking, I reflected on a few areas both books might have been stronger. For example, it would be better to provide links to additional learning resources of eye tracking at the end of each chapter. Each book would also benefit from a concrete example showing the readers how to conceptualize, construct, and carry out an eye-tracking experiment in a step-by-step fashion. Regarding the software used for creating the experiment and cleaning the data, third-party programming tools, such as Python , could be introduced. Finally, both Conklin *et al.* (p. 66) and Godfroid (p. 222) state that early and late measures of eye tracking are indicative of unconscious/implicit/automatic and conscious/explicit cognitive processes, respectively, which can be misleading. Eye

tracking is simply used as a tool to record eye movements; the automaticity/implicitness of cognitive processing is determined by the nature of the *task*, and not associated with specific eye-tracking *measures*.

ANALYTICAL CONSIDERATIONS

It is challenging to structure and analyze eye-tracking data in applied research effectively. Anticipating issues regarding data collection prior to running studies is critical for researchers and practitioners to have meaningful, reliable, and valid findings, and these books provide a number of excellent suggestions for doing so. Conklin *et al.* and Godfroid provide detailed principles of adequate experimental design, set-up, and administration of experiments. Conklin *et al.* (Chapter 3) focus more on discussing principal components and commonly-faced issues when designing a study. Conversely, Godfroid (Chapter 5) focuses on how to prepare items and stimuli in either between- or within-subjects designs, and how to render the stimulus correctly in the trail sequence. Godfroid (Chapter 5) provides a useful guide for determining sample size, with a particular focus on summarizing item and subject numbers commonly reported in L2 eye-tracking studies (e.g. Plonsky and Derrick 2016).

In terms of data management and modeling, both books provide clear guidance on cleaning, visualizing, and modeling collected data. Conklin *et al.* (Chapter 7) focus on data cleaning and visualization through specialized analysis tools within specific eye-tracking systems. Godfroid (Chapter 8) complements Conklin *et al.* by introducing other possible solutions to manage data using various R packages. Both books also cover a number of statistical methods to accurately model different eye-tracking indicators.

As a methodologist who focuses on developing innovative methods to model eye-tracking data, I found several areas which could have been stronger. Regarding research design, univariate mixed-model analysis of variance (ANOVA) was the primary model mentioned, both in past work and in future directions. However, multivariate dependent variables (e.g. fixation count/duration, saccades) are often tracked simultaneously within- or between-subjects. Therefore, it would be better to introduce more advanced designs, such as randomized blocks, split-plots, and unbalanced longitudinal designs (e.g. Cohen *et al.* 2003; Gelman and Hill 2006; Lomax and Hahs-Vaughn 2012). Each can accommodate missing data across repeated measures, the uneven spacing of time-course observations, and time-varying covariates.

As to methods for analyzing eye-tracking data, latent variable modeling/structural equation models can inform finer-tuned inferences regarding latent cognitive constructs, which often cannot be observed directly. For instance, response time (RT) only partially reflects working efficiency, since it is an observed record of how long a subject reacts to a visual stimulus only. RTs, as observed variables, contain measurement error caused by factors such as sampling rate, stimulus design, and study environment. In contrast, working efficiency as a *latent* cognitive variable can be inferred more precisely after partialling out

measurement errors by mapping observed RTs onto a latent continuum. Instead of using RTs directly as an outcome variable, a latent outcome variable (working efficiency) can be directly analyzed via the latent modeling approach, which can provide finer-grained inferences about cognitive engagement.

Furthermore, structural equation modeling can directly reveal associations between multiple latent constructs. By reporting factor scores as well as corresponding variance–covariance components, both random/individual and fixed/group level estimates can be informed simultaneously. When used with Bayesian estimation, missing data can be imputed automatically by drawing samplers out of posterior distributions of latent variables. This provides an easy solution for missing data and makes for a better statistical inference with smaller samples ($N < 30$). Finally, with latent variable modeling, data collected across different eye-tracking systems with various sampling rates can be linked. As a result, it can cross-validate results from different studies and facilitate meta-analysis for summarizing general patterns of effects across studies.

CONCLUSION

Each of us found these books to be highly useful tools to introduce and sketch out the field of eye-tracking research for AL/SLA/Bilingualism. Given the relative youth of these fields in language research, and in SLA/Bilingualism, in particular, the two books will serve well as field guides for theory, methodology, and analysis. The next step is to continue work in these fields, reaching the point Kaiwen mentions regarding statistical meta-analysis, where claims regarding effects can be meaningfully synthesized, and theoretical models can be supported or rejected. These texts provide an exciting base for this work, and we look forward to participating in it in the coming years.

Reviewed by
Jonathan Malone
Maryland English Institute
Department of Second Language Acquisition,
University of Maryland
Email: jemalone@umd.edu

Wei Yi
School of Chinese as a Second Language,
Peking University

Kaiwen Man
Department of Educational Studies in Psychology,
Research Methodology and Counseling,
University of Alabama
doi:10.1093/applin/amaa041

REFERENCES

Cohen, J., **P. Cohen**, **S. G. West**, and **L. S. Aiken**. 2003. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd edn. Lawrence Erlbaum Press.

Gelman, A. and **J. Hill**. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Just, M. A. and **P. A. Carpenter**. 1980. 'A theory of reading: From eye fixations to comprehension,' *Cognitive Psychology* Psychological Review, 87, 329–354.

Leow, R. P. and **K. Morgan-Short**. 2004. 'To think aloud or not to think aloud: The issue of reactivity in SLA research methodology,' *Studies in Second Language Acquisition* 26: 35–57.

Lomax, R. G. and **D. L. Hahs-Vaughn**. 2012. *Statistical Concepts: A Second Course*, 4th edn. Routledge.

Plonsky, L., and **D. Derrick**. 2016. 'A meta-analysis of reliability coefficients in second language research,' *The Modern Language Journal* 100: 538–53.

Rayner, K. 2009. 'Eye movements and attention in reading, scene perception, and visual search,' *The Quarterly Journal of Experimental Psychology* 62: 1457–506.

NOTES ON CONTRIBUTORS

Jonathan Malone is the Associate Director at Maryland English Institute and a Ph.D. student at the University of Maryland, USA, in Second-Language Acquisition. His research interests center on the intersection between cognitive processing and instructional design during second-language acquisition, specifically examining cognitive processing and learning outcomes during L2 reading, audio–visual language tasks, and vocabulary development. *Address for correspondence:* School of Languages, Literatures, and Cultures, 3215 Jiménez Hall, University of Maryland, College Park, MD 20742. E-mail: jemalone@umd.edu

Wei Yi is an Assistant Professor in the School of Chinese as a Second Language at Peking University, China. His research centers on cognitive mechanisms and processes of second-language acquisition, using behavioral tasks, eye tracking, and electroencephalography (EEG). He has published several articles on formulaic language, statistical learning, cognitive aptitude, and incidental vocabulary learning through reading.

Kaiwen Man is an Assistant Professor in the Department of Educational Studies in Psychology, Research Methodology and Counseling at the University of Alabama, USA. His primary research interests include data mining methods, test security, response time modeling, eye-tracking modeling, and multidimensional item response theory (IRT).