# Language Without Borders: A Step-by-Step Guide to Analyzing Webcam Eye-Tracking Data for L2 Research

Jason Geller1, Yanina Prystauka2, Sarah Colby3, and Julia Droulin4

1Department of Psychology and Neuroscience, Boston College

2Department of Psychology and Neuroscience, University of Bergen

3Department of Psychology and Neuroscience, University of Ottowa

4Department of Psychology and Neuroscienc, University of North Carolina at Chapel Hill

# Author Note

Jason Geller  http://orcid.org/0000-0002-7459-4505

Yanina Prystauka  http://orcid.org/0000-0000-0000-0002

Sarah Colby  http://orcid.org/0000-0000-0000-0003

Julia Droulin  http://orcid.org/0000-0000-0000-0003

Carina Mengano is now at Generic University.

The authors have no conflicts of interest to disclose.

Author roles were classified using the Contributor Role Taxonomy (CRediT; https://credit.niso.org/) as follows: *Jason Geller***:** conceptualization, writing – original draft, data curation, writing – review & editing, software, formal analysis, and project administration. *Yanina Prystauka***:** writing – original draft, writing – review & editing, and formal analysis. *Sarah Colby***:** writing – original draft and writing – review & editing. *Julia Droulin***:** conceptualziation, writing – original draft, writing – review & editing, and funding acquisition

Correspondence concerning this article should be addressed to Jason Geller, Department of Psychology and Neuroscience, Boston College, Mcguinn Hall 405, Chestnut Hill, MA 02467-9991, USA, Email: jason.geller@bc.edu

# Abstract

Eye-tracking has become a valuable tool for studying cognitive processes in second language (L2) acquisition and bilingualism (Godfroid et al., 2024). While research-grade infrared eye-trackers are commonly used, their cost can limit accessibility. Recently, consumer based eye-tracking has emerged as a more affordable alternative, requiring only internet access and a personal webcam. However, consumer-based eye-tracking presents unique design and preprocessing challenges that must be addressed for valid results. To help researchers overcome these challenges,we developed a comprehensive tutorial focused on using webcam eye-tracking for L2 langauge research (but the information provided can be etxneded to any research using the VWP online). Our guide will cover all key steps, from experiment design to data preprocessing and analysis, where we highlight the R package webgazeR, which is open source and freely available for download and installation: . We offer best practices for environmental conditions, participant instructions, and tips for designing Visual World Paradigm (VWP) experiments with webcam eye-tracking. To demonstrate these steps, we analyze data collected through the Gorilla platform (Anwyl-Irvine et al., 2020) looking at L2 Spanish within and corss-lingustic competion in a spoken word Spanish visual world paradigm. This tutorial aims to empower researchers by providing a step-by-step guide to successfully conduct webcam-based eye-tracking studies.

*Keywords*: VWP, Tutorial, Webcam Eye-tracking, R

# Language Without Borders: A Step-by-Step Guide to Analyzing Webcam Eye-Tracking Data for L2 Research

# Introduction

Eye-tracking, a technology with a history spanning over a century, has undergone remarkable advancements. Initially, it was an invasive technique requiring the use of contact lenses fitted with search coils, often necessitating anesthesia [Płużyczka ([2018](#ref-pluzyczka2018)); ]. Over time, however, the technology has evolved into a non-invasive, lightweight, and user-friendly tool that is now widely accessible in laboratories around the world.

Despite its widespread usage, eye-tracking technology faces several obstacles that can limit its accessibility. One significant challenge is the specialized expertise required to operate research-grade eye-trackers. Proper usage often demands many hours of training, meaning most research must be conducted in a lab by a trained student or faculty member. Another major limitation is the cost. Eye-trackers can be prohibitively expensive, ranging from a few thousand dollars (e.g., Gazepoint) to tens of thousands of dollars (e.g., Tobii (www.tobii.com), SR Research (www.sr-research.com). As a result, not everyone has the resources or the time to incorporate eye-tracking into their research.

In addition, conducting research in a lab setting significantly limits the population you can use in your study and, consequently, the generalizability of your findings. For instance, it can be particularly challenging to bring specialized populations into the lab, such as children, individuals who are hard of hearing (e.g., hearing aid or cochlear implant users), older individuals, those with various medical disorders, to name a few. This issue is not unique to lab-based eye-tracking but is a broader challenge across many types of research.

## Eye-tracking outside the lab

Methods that allow participants to use their own equipment from anywhere in the world offer a potential solution to generalizability issues, enabling researchers to recruit more diverse samples and explore a broader range of questions ([Gosling et al., 2010](#ref-gosling2010)). The shift toward online behavioral experiments has been gradually increasing in the behavioral sciences and has become every more important since the 2020 pandemic, which forced many us run our studies online or risk not doing research ([Anderson et al., 2019](#ref-anderson2019); [Rodd, 2024](#ref-rodd2024)). This trend has also prompted the development of a few eye-tracking methods that do not rely on traditional lab settings. One method, what has been called manual eye-tracking ([Trueswell, 2008](#ref-trueswell2008)), relies on video recordings of the participant (collected via the online teleconferencing software Zoom, for example) while performing a task. Eye gaze (direction) is manually recorded from the videos by frame by frame.

Another method, which is the focus of this tutorial, is automated eye-tracking or webcam eye-tracking. Automated eye-tracking requires three things: 1. A personal computer. 2. An internet connection and 3. A purchased or pre-installed webcamera. In order to start webcam eye-tracking you will need a way to collect gaze information. One common method to perform webcam eye-tracking is through an open source and free to use JavaScript library plugin called WebGazer [([Papoutsaki et al., 2016](#ref-papoutsaki2016))]. This plugin is already incorporated into several popular experimental platforms (e.g., *Gorilla*, *jsPsych*, *PsychoPy*, and *PCIbex*; [Anwyl-Irvine et al. ([2020](#ref-anwyl-irvine2020)) Peirce et al. ([2019](#ref-peirce2019)); Leeuw ([2015](#ref-deleeuw2015)); Zehr & Schwarz, [2018](https://pmc.ncbi.nlm.nih.gov/articles/PMC11627531/#bib62)) making it extremely easy to start webcam eye-tracking. WebGazer.js utilizes a webcam connected to the internet to track eye movements in real time. It employs a facial feature detection algorithm that estimates the position of the pupils in the wecam stream (relative to the face). By analyzing the relative movement of the eyes, WebGazer.js employs machine learning to estimate the user’s gaze location on the screen.

To achieve accurate tracking, WebGazer.js initiates a quick calibration process, where the user follows a dot on the screen. In some instances , users click on the dot locations, whereas other have you look at the dot as it moves to different locations on the screen. This calibration process enhances the library’s accuracy. It is important to note that WebGazer is not the only method available. Other methods have been implemented by companies like Tobii (www.tobii.com) and labvanced[]. However, because these methods are propierty it is unclear what they are doing under the hood

It is important to understand that the algorithms underlying consumer based eye-tracking differs from research-grade eye trackers. In research-grade eye tracking it is common to use video based recording and the pupil-corneal reflection (P-CR) method to track a person’s gaze ([Carter & Luke, 2020](#ref-carter2020)). This technique combines infrared light and high-speed cameras to achieve precise measurements. The process involves using infrared light to illuminate the eyes, capturing reflections (called glints) from the cornea and pupil. High-speed cameras simulatensouly take snapshots hundreds or thousands of times per second to measure the eyes position. Taking information from the location of corneal reflection and the pupil help calculate the gaze direction and position. Propiertary algos then determine where particiapnts are looking at on the screen. Eye-tracking systems that rely on the P-CR technique offer excellent temporal and spatial precision.

One question you may have is how does consumer-based webcam eye-tracking compare to more research-based eye-tracking? While this validation work is still on-going, webcam eye-tracking appears to have reduced spatiotemporal accuracy compared to research grade eye-trackers. With consumer-based eye-trackers using WebGazer some studies have documented spatial precision in the range of 4-10 degrees of visual angle and a latency to between 200 ms to 1000 ms ([Semmelmann & Weigelt, 2018](#ref-semmelmann2018); [Slim et al., 2024](#ref-slim2024); [Slim & Hartsuiker, 2023](#ref-slim2023)). The sampling rate is also much lower than some commerical eye trackers with some studies reporting an average or median sampling rate of 30 Hz ([Bramlett & Wiener, 2024](#ref-bramlett2024); [Prystauka et al., 2024](#ref-prystauka2024)). For comparison, a Tobii Spectrum eye-tracker boasts a spatial precision of .01 RMS and spatial accuracy of .3-.5 and can sample at rates of up to 1200 Hz or 1200 times per second ([Nyström et al., 2021](#ref-nyström2021)).

Despite these differences, a number of studies have tried to replicate lab based studies using webcam eye-trackers.

## Visual World Paradigm

Most relevant to this tutorial are the studies that have replicated lab based findings using the visual world paradigm (VWP; ([Tanenhaus et al., 1995](#ref-tanenhaus1995))). For the past 25 years, the VWP has been a dominant force in language research, helping researchers tackle questions ranging from sentence processing Eberhard et al. ([1995](#ref-eberhard1995))] and word recognition ([Allopenna et al., 1998](#ref-allopenna1998)) to bilingualism [] and the effects of brain damage on language []. What makes the tasks impact more remarkable is the task is quite simple. The VWP requires participants to view a visual scene or world made up of pictures on the screen (images are usually usually placed in the four quadrants of the screen (top left, top right, bottom left, bottom right) or even real objects in the world (). They then hear spoken or written language referring to one of the objects on the screen. This all happens while their eye movements are being monitored. The collection of eye movements allows researchers to get an online measure of how language unfolds over time.

While most of this research with the VWP has been conducted in labs with research grade eye-trackers, there have been several attempts to conduct these experiments online with consumer grade eye-trackers. In one of the first studies to examine the VWP in an online setting using webcam eye-tracking, Degen et al. reported a replication of a study by ([**sun?**](#ref-sun)) looking at whether scalar inferences are slower than the processing numerals. In their study they had four pictures in the each of the quadrants and an additoanl picture in the center of the screen. They were able to replicate the basic pattern (more looks to X than y) but observed a considearable time delay of 700 ms compared to the original study. They noted several reasons for this delay, including issues with the webgazer alogarithm being too compittaionaly demanding, AOIs too close together, and use of a single calibration point. Several other studies using the VWP online have noted a similar temporal delay in effects ([Slim et al., 2024](#ref-slim2024); [Slim & Hartsuiker, 2023](#ref-slim2023)).

While the temporal lag presents a significant limitation to webcam eyetracking and its use in drawing meaningful inferences in the VWP, an improvement to the webgazer algorithm has improved this limitation. In a study by Vos et al. ([2022](#ref-vos2022)), using a new version of webgazer in jsPsych, they showed a considerable reduction in delay (50 ms) between the lab and oline versions of the VWP. More recently, Prystauka et al. ([2024](#ref-prystauka2024)) and Bramlett and Wiener ([2024](#ref-bramlett2024)) using Gorilla and the new version of WebGazer, showed comparable effects in the online version VWP that are comparable to the original lab based studies. Taken together, this suggests that the VWP webcam eye-tracking is a viable alternative to lab-based eye-tracking.

## Tutorial

Given the increasing number of online VWP experiments utilizing webcam eye-tracking and the considerable promise of this technology, we aimed to support researchers in their efforts to conduct high-quality webcam eye-tracking studies. While a valuable tutorial on webcam eye-tracking already exists by @bramlett2024, we sought to expand on their work by incorporating many of their useful recommendations and offering an R package to help streamline data pre-processing.

The purpose of this tutorial is to provide an overview of the basic set-up and design features of an online VWP task and to highlight the preprocessing steps needed to analyze webcam eye-tracking data. Here we use the popular open source programming language R and introduce the webgazeR pacakage to facilitate pre-processing of webcam data. To highlight the steps needed to process webcam data we present data from a cross-lingustic spoken word VWP with L2 Spanish speakers collected on Gorilla. To our knowledge, L2 processing and competitor effects have not been looked at in the online version of the VWP.

# L2 VWP

To highlight the pre-preocessing steps needed to analyze webcam eye-tracking data we looked at the competitive dynamics of L2 learners of Spanish in spoken word recognition. Specifically, we investigated within-language and cross-language competition using webcam-based eye-tracking. A recent study by McCall et al., using carefully crafted stimuli, explored within- and cross-linguistic competition in adult L2 learners using a cross-linguistic VWP. Their task included two key conditions:

1. **Spanish-Spanish condition**: A Spanish competitor was presented alongside the target word. For example, if the target word spoken was “cielo” (sky), the Spanish competitor was “ciencia” (science).
2. **Spanish-English (cross-linguistic) condition**: An English competitor was presented for the Spanish target word. For example, if the target word spoken was “botas” (boots), the English competitor was “border.”

McCall et al. observed competition effects in both conditions: within-Spanish competition (e.g., *cielo* - *ciencia*) and cross-linguistic competition (e.g., *botas* - *border*). For this tutorial, we collected data to conceptually replicate their pattern of findings.

There are two key differences between our dataset and the original study by McCall et al. First, McCall et al. focused on adult L2 Spanish speakers and posed more fine-grained questions about the time course of competition and resolution and its relationship with L2 langauge acquisition. Second, unlike McCall et al., who measured Spanish proficiency objectively (e.g., using LexTALE-esp; Izura et al. ([2014](#ref-izura2014)) ), we relied on Prolific’s filters to recruit L2 Spanish speakers.

Our primary goal here was to demonstrate the pre-processing steps required to analyze webcam-based eye-tracking data. A secondary goal was to provide evidence of L2 competition using this methodology. To our knowledge, no papers have looked at spoken word recognition and competition using online methods. It is our hope that researchers can use this to test more detailed questions about L2 processing using webcam-based eye-tracking.

## Method

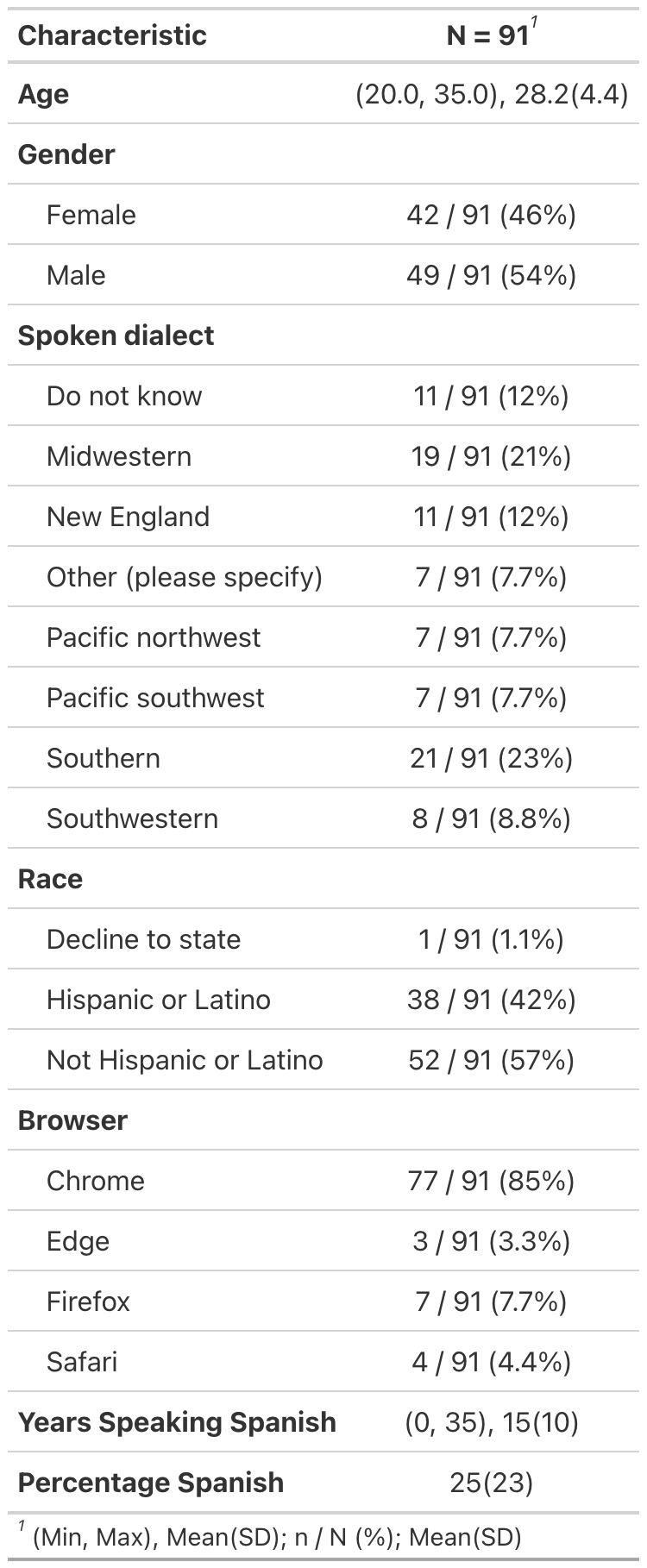
All tasks herein can be previewed here (<https://app.gorilla.sc/openmaterials/917915>). The manuscript, data, R code, can be found on Github (<https://github.com/jgeller112/webcam_gazeR_VWP>).

### Participants

A total of 187 participants consented to participate in the study using the Gorilla hosting and experiment platform. Of these, 111 passed the headphone screener checkpoint and proceeded to the task. Among them, 32 participants successfully completed the Visual World Paradigm (VWP) task with at least 100 trials, while 79 participants failed calibration. Ninety-one participants completed the entire experiment, including the final questionnaires. @tbl-demo2 provides basic demographic information about the participants who completed the full experiment. After applying additional exclusion criteria (low accuracy ( < 80%) and excessive missing eye-data (> 30%) , the final sample consisted of 28 participants with usable eye-tracking data.

Table 1

Demographic variables Experiment 2



### Materials

#### VWP.

##### Items.

We adapted materials from Sarrett et al. ([2022](#ref-sarrett2022)). In their cross-linguistic VWP, participants were presented with four pictures and a spoken Spanish word and had to select the image that matched the spoken word by clicking on it. The word stimuli for the experiment were chosen from textbooks used by students in their first and second year college Spanish courses.

The item sets consisted of two types of phonologically-related word pairs: one pair of Spanish-Spanish words and another of Spanish-English words. The Spanish-Spanish pairs were unrelated to the Spanish-English pairs. All the word pairs were carefully controlled on a number of dimensions (see ([Sarrett et al., 2022](#ref-sarrett2022))) making it an excellent set of materials to use.

There were three experimental conditions: (1) the Spanish-Spanish condition, where one of the Spanish words was the target and the other was the competitor; (2) the Spanish-English condition, where a Spanish word was the target and its English phonological cohort served as the competitor; and (3) the No Competitor condition, where the Spanish word did not overlap with any other word in the set. The Spanish-Spanish condition had twice as many trials as the other conditions due to the interchangeable nature of the target and competitor words in that pair.

There were 15 sets of 4 items (this was half the number of sets used in Sarrett et al. ([2022](#ref-sarrett2022)) . Each item within a set was repeated 4 times as the target word. This yielded 240 trials (15 sets × 4 items per set × 4 repetitions). Each item set consisted of one Spanish-Spanish cohort pair and one Spanish-English cohort pair. Both items in a Spanish-Spanish pair had a“reciprocal” competitor relationship (that is, we could test activation for *cielo* given *ciencia*, and for *ciencia* given *cielo*). Consequently, there were 120 trials in the Spanish-Spanish condition. In contrast, only one item from the Spanish-English pair had the speciﬁed competitor relationship (we could test activation for frontera *border*, given *botas*, but when hearing *frontera*, there was no competitor). Thus, there were only 60 trials for each the Spanish-English competition as well as the No Competitor condition. Items occurred in each of the four corners of the screen on an equal numbers of trials.

##### Stimuli.

Auditory stimuli were recorded by a female monolingual speaker of English in a sound-attenuated room sampled at 44.1 kHz. Auditory tokens were edited to reduce noise and remove clicks. They were then amplitude normalized to 70 dB SPL. Visual stimuli were images from a commercial clipart database that were selected by a consensus method involving a small group of students. All .wav files were converted to .mp3 for online data collection.

### Headphone screener

The headphones screener is a six-trial task taken from Woods et al. ([2017](#ref-woods2017)). On each trial, three tones of the same frequency and duration were presented sequentially. One tone had a lower amplitude than the other two tones. Tones were presented in stereo, but the tones in the left and right channels were 180 out of phase across stereo channels—in free field, these sounds should cancel out or create distortion, whereas they will be perfectly clear over headphones. The listener picked which of the three tones was the quietest. Performance is generally at the ceiling when wearing headphones but poor when listening in the free field (due to phase cancellation).

#### Demographics questionnaire.

Participants completed a demographic questionnaire as part of the study. The questions covered basic demographic information, including age, gender, spoken dialect, ethnicity, and race.

Participants also answered a series of questions related to their personal health and environmental conditions during the experiment. These questions addressed any history of vision problems (e.g., corrected vision, eye disease, or drooping eyelids) and whether they were currently taking medications that might impair judgment. Participants also indicated if they were wearing eyeglasses, contacts, makeup, false eyelashes, or hats.

The questionnaire inquired about their environment, asking if there was natural light in the room, if they were using a built-in camera or an external one (with an option to specify the brand), and their estimated distance from the camera. Participants were asked to estimate how many times they looked at their phone or got up during the experiment and whether their environment was distraction-free.

Additional questions assessed the clarity of calibration instructions, allowing participants to suggest improvements, and asked if they were wearing a mask during the session. These questions aimed to gather insights into personal and environmental factors that could impact data quality and participant comfort during the experiment.

To gauge L2 experience, we also asked participants when they started speaking Spanish, how many years of Spanish speaking, and to provide a percentage of time Spanish is spoken in their daily lives.

## Procedure

All tasks were completed in a single session, lasting approximately 30 minutes. The tasks were presented in a fixed order: consent, headphone screener, spoken word Visual World Paradigm (VWP), and a questionnaire.

The experiment was programmed in the Gorilla Experiment Platform (Anwyl-Irvine et al., 2019), with personal computers as the only permitted device type. Upon entering the online study, participants received general information to decide if they wished to participate, after which they provided informed consent. Participants were then instructed to adjust the volume to a comfortable level while noise played.

Next, participants completed a headphone screening test. They had three attempts to pass this test. If unsuccessful by the third attempt, they were excluded from the experiment.

For those who passed the screening, the next task was the VWP. This began with instructional videos providing specific guidance on the ideal experiment setup for eye-tracking and calibration procedures. Participants were then required to enter full-screen mode before calibration. Calibration occurred every 50 trials for a total of 2 calibrations. Participants had three attempts to successfully complete each calibration phase. If calibration was unsuccessful, participants were directed to an early exit screen, followed by the questionnaire.

In the main VWP task, following video instructions, participants completed four practice trials to familiarize themselves with the procedure. Each trial began with a 500 ms fixation cross at the center of the screen. This was followed by a preview screen displaying four images, each positioned in a corner of the screen. After 1500 ms, a start button appeared in the center. Participants clicked the button to confirm they were focused on the center before the audio played. Once clicked, the audio was played, and the images remained visible. Participants were instructed to click the image that best matched the spoken target word, while their eye movements were recorded throughout the trial. [Figure 1](#fig-vwptrial) depicts the trial schematic.

knitr::include\_graphics("images/trial\_descrip.svg")

Figure 1

VWP trial schematic

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). *Tracking the time c ourse of spoken word recognition using eye movements: Evidence for c ontinuous mapping models* (pp. 419–439).

Anderson, C. A., Allen, J. J., Plante, C., Quigley-McBride, A., Lovett, A., & Rokkum, J. N. (2019). The MTurkification of Social and Personality Psychology. *Personality & Social Psychology Bulletin*, *45*(6), 842–850. <https://doi.org/10.1177/0146167218798821>

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>

Bramlett, A. A., & Wiener, S. (2024). The art of wrangling. *Linguistic Approaches to Bilingualism*. https://doi.org/<https://doi.org/10.1075/lab.23071.bra>

Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*, *155*, 49–62. <https://doi.org/10.1016/j.ijpsycho.2020.05.010>

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, *24*(6), 409–436. <https://doi.org/10.1007/BF02143160>

Gosling, S. D., Sandy, C. J., John, O. P., & Potter, J. (2010). Wired but not WEIRD: The promise of the Internet in reaching more diverse samples. *Behavioral and Brain Sciences*, *33*(2-3), 94–95. <https://doi.org/10.1017/S0140525X10000300>

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: a test to rapidly and efficiently assess the Spanish vocabulary size. *PSICOLOGICA*, *35*(1), 49–66. <http://hdl.handle.net/1854/LU-5774107>

Leeuw, J. R. de. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>

Nyström, M., Niehorster, D. C., Andersson, R., & Hooge, I. (2021). The Tobii Pro Spectrum: A useful tool for studying microsaccades? *Behavior Research Methods*, *53*(1), 335–353. <https://doi.org/10.3758/s13428-020-01430-3>

Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., & Hays, J. (2016). *Webgazer: Scalable webcam eye tracking using user interactions*. 38393845.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>

Płużyczka, M. (2018). The First Hundred Years: a History of Eye Tracking as a Research Method. *Applied Linguistics Papers*, *25/4*, 101–116. <http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.desklight-98576d43-39e3-4981-8c1c-717962cf29da>

Prystauka, Y., Altmann, G. T. M., & Rothman, J. (2024). Online eye tracking and real-time sentence processing: On opportunities and efficacy for capturing psycholinguistic effects of different magnitudes and diversity. *Behavior Research Methods*, *56*(4), 3504–3522. <https://doi.org/10.3758/s13428-023-02176-4>

Rodd, J. M. (2024). Moving experimental psychology online: How to obtain high quality data when we can’t see our participants. *Journal of Memory and Language*, *134*, 104472. <https://doi.org/10.1016/j.jml.2023.104472>

Sarrett, M. E., Shea, C., & McMurray, B. (2022). Within- and between-language competition in adult second language learners: Implications for language proficiency. *Language, Cognition and Neuroscience*, *37*(2), 165–181. <https://doi.org/10.1080/23273798.2021.1952283>

Semmelmann, K., & Weigelt, S. (2018). Online webcam-based eye tracking in cognitive science: A first look. *Behavior Research Methods*, *50*(2), 451–465. <https://doi.org/10.3758/s13428-017-0913-7>

Slim, M. S., & Hartsuiker, R. J. (2023). Moving visual world experiments online? A web-based replication of Dijkgraaf, Hartsuiker, and Duyck (2017) using PCIbex and WebGazer.js. *Behavior Research Methods*, *55*(7), 3786–3804. <https://doi.org/10.3758/s13428-022-01989-z>

Slim, M. S., Kandel, M., Yacovone, A., & Snedeker, J. (2024). Webcams as windows to the mind? A direct comparison between in-lab and web-based eye-tracking methods. *Open Mind*, *8*, 1369–1424. <https://doi.org/10.1162/opmi_a_00171>

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science (New York, N.Y.)*, *268*(5217), 1632–1634. <http://www.ncbi.nlm.nih.gov/pubmed/7777863>

Trueswell, J. C. (2008). *Using eye movements as a developmental measure within psycholinguistics* (I. A. Sekerina, E. M. Fernández, & H. Clahsen, Eds.; pp. 73–96). John Benjamins Publishing Company. <https://doi.org/10.1075/lald.44.05tru>

Vos, M., Minor, S., & Ramchand, G. C. (2022). Comparing infrared and webcam eye tracking in the Visual World Paradigm. *Glossa Psycholinguistics*, *1*(1). <https://doi.org/10.5070/G6011131>

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, and Psychophysics*, *79*(7), 2064–2072. <https://doi.org/10.3758/s13414-017-1361-2>