Introduction

Eye movement data provide quantitative evidence of a person’s visual attentional processes when performing a task such as reading (Reichle, Pollatsek, Fisher, & Rayner, 1998; Duchowski, 2002; Frenck-Mestre, 2005; Rayner, 1998, 2009). Research employing eye movement data to investigate reading processes is diverse, complex, and informs a large variety of theories (see Radach & Kennedy, 2004, for an overview). For example, recordings of eye movements while readers process words placed in various parts of sentences (in various syntactic structures) lend insight into how syntactic structure imparts mental processing importance—that is, whether variations in syntactic structure affect aspects of language processing (e.g., Birch & Rayner, 2010). Eye trackers are also commonly used to compare child versus adult readers’ reading processes (e.g., Joseph et al., 2008); the results of such research are often used to better inform child-reading instructional practices. Additionally, eye-tracking is used to investigate how individuals with developmental reading disorders read and process textual information (i.e., Hatzidaki, Gianneli, Petrakis, Makaronas, & Aslanides, 2011). Outcomes assist the understanding of the effects of these disorders on the brain and guide pedagogues in how to better teach reading to such populations.

In second language acquisition, eye-tracking methods have been used to investigate how individuals reading in a second language resolve ambiguous sentences (e.g., Molly said that she will go to New Jersey yesterday: Dussias, 2010, p. 157; see also Dussias & Sagarra, 2007; Roberts, Gullberg, & Indefrey, 2008) or violations of gender agreement (Keating, 2009). Such studies help researchers understand second language (L2) reading processes and how they differ from first language (L1) reading processes. Other L2 studies employing eye-tracking methodologies (Godfroid, Housen, & Boers, 2010; Godfroid & Uggen, in press) have investigated...
Schmidt’s *noticing hypothesis* (Robinson, 1995; Leow, 1997; Schmidt, 1990, 1993, 1995, 2001)—that is, whether increased attention (measured through eye movement data) to novel linguistic forms (grammar or vocabulary forms) increases one’s chances of learning that form and, if so, how and why or, if not, why not. Likewise, recordings of bilinguals’ eye movements while they read sentences in their less dominant, second language (L2)—sentences that have in them words that are cognates in their dominant, first language (L1)—inform theorists on whether bilinguals’ mental lexicons are integrated or separated by language (e.g., Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011). In such research results are mixed, but cognates are viewed as recognized if they are skipped or fixated on for less time, which provides evidence of an integrated lexicon.

Other diverse examples of the use of eye-tracking data to investigate reading processes include (a) investigations into task instructions (e.g., proofread or read for comprehension) on reading (Kaakinen & Hyona, 2010), (b) research on how adolescents read while simultaneously writing essays (Beers, Quinlan, & Harbaugh, 2010), (c) studies on the effects of compound words and word length on reading (Inhoff, Starr, Solomon, & Placke, 2008; Juhasz, 2008), (d) eye movement studies that investigate the reading of onscreen captions during the presentation of a video in the L1 (d’Ydewalle & De Bruycker, 2007) or in the L2 (Winke, Gass, & Sydorensko, 2013). Of particular note is a study by Bax and Weir (2012), in which the authors investigated the cognitive processes that English language learners employ when taking a computer-based academic English reading test. Studies such as these demonstrate the diversity of current eye-tracking/reading research. The recent surge in studies also indicates that eye-tracking systems are becoming financially more accessible and are easier to use than they were a decade ago (Duchowski, 2007); both of these features are helping to expand the eye-tracking research paradigm.

The Connection Between Eye Movements and Cognitive Reading Processes

Many different types of technology can be used to track a subject’s eyes while reading. But, before employing such technology to examine reading processes, researchers need to embed their studies’ research questions within current and well-defined theories of reading processes and to contextualize their research so as to make it accord with current theories on the proposed links among cognition, attention, visual intake patterns, and eye movements while reading. This is important, because any discussion of eye movement data needs to be backed up by current theory. Hence I start by reviewing some terms that are needed to understand eye movements during reading and models of eye movement control during reading comprehension.

Comprehensive overviews of eye movements during reading have been published (Reichle et al., 1998; Castelhano & Rayner, 2008; Rayner, 1998, 2009). These papers position eye movement during reading within the E–Z Reader model (Reichle, Rayner, & Pollatsek, 2003; Pollatsek, Reichle, & Rayner, 2006) of visual attention that accounts for the link between eye movement control and cognition.
As explained by Rayner (2009), during the complex task of reading, “either (a) eye location (overt attention) and covert attention are overlapping and at the same location or (b) attention disengagement” occurs, which happens when “attention precedes the eyes to the next saccade target” (p. 1458; emphasis added). In other words, much of the work involved in reading-processing research is about determining where and for how long (when) the eyes remain fixed on words and phrases in the text. Models of visual attention during reading also try to account for the engines that drive the decisions on where and when to look during reading. At the heart of the E–Z Reader model of eye movement control in reading are two premises: readers attend to words during eye fixations, and movements (saccades) from one fixation to the next are triggered by a cognitive event.

**Essential Terminology and Concepts**

Some definitions are in order here. According to the E–Z Reader model of eye movement control in reading (Pollatsek et al., 2006; Reichle et al., 2003), the following terms and concepts are very important for understanding eye movements during reading: (a) saccades; (b) fixations; (c) visual acuity; (d) saccade latency; (e) information access during eye fixations; (f) perceptual span; (g) parafoveal preview effects; (h) regressions; (i) eye movement control and patterns; and (j) measures of processing time. These important terms and concepts are briefly defined below. (Read Reichle et al., 2003, for a full review of them and of their importance in any model of eye movement control.)

1. **Saccades** are the short and rapid eye movements that readers make across the printed page while reading.
2. **Eye fixations** are the brief periods of time during which the eyes are fixed on the page (the periods of fixation between saccades).
3. **Visual acuity** is the limit in how much information can be processed during a fixation. As explained by Reichle et al. (2003), “visual acuity is maximal in the center of the retina and rapidly decreased towards the periphery, and fine visual discrimination can only be made with the fovea, or central 2° of vision” (p. 446; emphasis added). Visual acuity may account for the difficulty in processing longer words, especially novel, longer words, or ones with unexpected phoneme combinations.
4. **Saccade latency** is the time between when one plans to move on to the next saccade and when that movement occurs. It is estimated by Reichle et al., 2003, to be around 180–250 milliseconds in duration. The question is whether the saccadic movement is made while the mind is still processing the word, or whether the decision to move is made after the word is processed. In other words, does word recognition drive saccadic movement, or are the process of saccades and the word recognition process more complex and intertwined?
5. **Information is acquired during eye fixations.** While the eyes move to the next fixation (during saccades), vision is blurred or suppressed. Thus it is important to note that information is only obtained through eye fixations. The
theory is that the information needed for reading normally occurs rather quickly—within 50 to 60 milliseconds after a fixation starts.

6 The perceptual span or parafovea is the region that extends around the fovea (about 5 degrees around the fovea). The theory is that words can be partially processed in this area of perceptual span (see Rayner & Bertera, 1979, who first explained this phenomenon). Research on the parafovea is intriguing; as summarized in Reichle et al. (2003), the parafovea does not extend above or below the line being read; the span is relatively constant for readers of similar alphabetic orthographies; whereas the density and complexity of the writing system influences the span asymmetry and size, and the perceptual span is not hardwired. As readers develop reading skills, their spans become bigger, and, when presented with more difficult readings, spans become smaller. Predictable upcoming words also contribute to bigger spans.

7 Parafoveal preview effects occur when the processing of a word takes place before the word is fixated on. Such effects can shorten the fixation on the word itself. Parafoveal previewing may also contribute to word skipping. Thus a skipped word (a word not directly fixated on) is not evidence of an unprocessed or unviewed word.

8 Regressions are saccades that move backwards, to earlier or previous parts of the text. These can happen for two reasons: either there was some type of difficulty in the linguistic processing, so the reader reverts back to an earlier part of the text to aid processing or comprehension; or the reader regresses as a result of some type of simple motor error or viewing process by which the eye regresses to an earlier part of the text. Both of these regression types have been documented in empirical research, as reviewed by Reichle et al., 2003.

9 Eye movement control. Moving on (changing fixations) in a text involves two dimensions: (a) where and (b) when. The question is whether these two decisions (one spatial and one temporal) are controlled by the same thing or by two different things. Many believe that the two decisions are made online and independently; others do not.

10 Measures of processing time. Eye trackers can obtain incredibly accurate and copious amounts of data. First, researchers need to note the importance of reporting on and analyzing several different processing measures because each contributes unique information—some are associated with initial reading processes, others with later ones, and some are appropriate for measuring the processing of a given target word, while others are or can be associated with larger regions (that is, a region larger than a single target word). For example, data can be reduced to the following types, as explained in Reichle et al. (2003): Gaze duration, first fixation duration, single fixation duration, and total time.

- Gaze duration is the sum of all fixation times on a single word. Gaze duration normally only includes time spent on the word before the eye has (or eyes have) left the word (fixations within the word). This could also be labeled as gaze duration during the first pass, that is, during the initial encounter with the word (not during regressions to the word).
- First fixation duration is the time spent on the first fixation of the word during the first pass. It is useful for measuring the processing of a target
word and not a larger region (for example, a phrase) because, with a larger (longer) region, it is more likely that there will be further fixations on it.

- **Single fixation duration** is the average duration of fixation on words that are fixated on exactly once during the first pass.
- **Total time** is the sum of all fixations on the word, including fixations stemming from regressions back to the word.

Due to space limitations, I cannot explain the many other concepts, such as **rereading time** and **regression path duration**, but eye movement researchers should become familiar with the full range and scope of such measures (see also Rayner, 1998, and Roberts & Siyanova-Chanturia, in press, for more explanations of eye movement measures). Researchers also need to be aware that some cognitive reading processes are not represented in eye movements until the next word or region comes into view (a spill-over effect); thus reading researchers may analyze eye movements in relation to post-target words. And finally, because pupil size (dilation of the pupil) is often viewed as a measure of cognitive load, researchers may want to further explore it as an indication of reading processes (see Hyönä, Tommola, & Alaja, 1995).

While competing models of the visual perception (oculomotor) system and of the effects of word recognition exist—for example, the (autonomous) saccade generation with inhibition by foveal targets, or SWIFT system (see Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005)—they all rely on research that uses the terms reviewed above, or very similar ones, for operationalizing eye movements while reading. The terms and concepts above provide a common framework for organizing discussions on reading processes because together they describe and account for the where and when of fixations in reading.

**Eye-Tracking Technology**

Good eye trackers can be expensive. While learning to use them and to apply data from them to investigate reading processes is becoming less time-consuming and complex (Duchowski, 2002, 2007), do not be deceived. The thought of being able to rapidly set up and run experiments and immediately analyze data to obtain study results is, to put it bluntly, naïve. The two main types of eye-tracking technology commonly used for L1 and L2 reading-processing research were developed by two companies: SR Research in Canada (www.sr-research.com) and Tobii Technology in Sweden (www.tobii.com). In this section of the chapter I review the technology, outline what it can and can’t do, and present reasons why these two systems are preferred by researchers.

**Commonalities in Systems for Recording Eye Movements**

Both SR Research and Tobii Technology produce video-based eye trackers that measure saccadic eye movements associated with viewing images (pictures, video, or words) on a computer screen. In their eye-tracking systems, a camera uses infrared light to create a corneal reflection that is used to track one or both eyes
(with Tobii’s T120, both eyes are tracked; with SR Research’s EyeLink 1000, the researcher decides to track either the right or the left eye or both eyes); algorithms from them map what one is looking at on the screen. The setups of these systems vary greatly. Both offer head-free eye-tracking in which the participant can freely move his or her head about during the experiment—the Tobii TX300 uses two cameras that automatically follow and track the eyes during motion, and the EyeLink 1000 uses a sticker placed on the participant’s forehead to run calibrations that allow for even a monocular system to track a single eye during head motion (the EyeLink 1000 can also track two eyes head-free). The EyeLink also has head-stabilized configurations in which the participant’s head is stabilized on a mount, via a head or a chin rest; these configurations provide much more accurate eye movement data recording, which is often necessary if data are needed at the word or phoneme level (see Figure 62.1). Pictures of the EyeLink 1000 and Tobii TX300, as set up at the Michigan State University, Second Language Studies Eye-Tracking Lab, are in Figures 62.2 and 62.3.

![Figure 62.1](image_url)  
*Figure 62.1*  The EyeLink 1000 chin- and head-rest mount for head stabilization (and greater eye movement recording accuracy) while eye-tracking
Before data collection begins, the researcher must work with the participant and the eye-tracking system to calibrate the participant’s eyes (or eye, if he or she is using a monocular system) to the eye-tracking system. The camera setup and calibration session with the Tobii TX300 system are rather simple, while with the EyeLink 1000 system the camera set up and calibration session are a bit more involved. On both systems, during the pre-programmed camera setup and calibration session, the researcher (or the experimental program) instructs the
participant to look at images that appear on the screen one at a time (typically, 5, 9, or more dots or fixation crosses spread across the screen or, for children, little happy faces, quacking ducks, or the like). The eye-tracking system compares the true location of each image with where the camera (or cameras) detects the participant’s gaze is on the screen and applies an algorithm to correct for future fixations. With the Tobii TX300, this is fairly automatized and there is actually no camera setup for the researcher to perform, since the cameras are built into the system (fixed within the display computer monitor); but with the EyeLink 1000 the researcher can manually adjust the camera’s position and focus, can adjust the eye tracker’s saccade detector sensitivity, can set pupil thresholds, and so on. In a nutshell, the Tobii is like an airplane that can fly on autopilot only, which is great for those not well versed in flying planes, but perhaps a bit scary for those used to flying on their own. And the EyeLink 1000 has no autopilot, meaning that researchers who run calibrations and eye-tracking experiments on an EyeLink 1000 system must learn a lot about corneal reflections, camera optimization, and what affects eye–camera calibrations such that they sometimes do not work well—which has its pros (researchers have better control over the data collection process; they can manually correct for some types of calibration errors) and its cons (there is a somewhat long, technical curve in learning to use the EyeLink 1000 system).

In both systems, software that comes with the eye tracker can be used to design robust research experiments. In those experiments, researchers can draw or create interest areas, that is, shapes (boxes, circles, or custom shapes) around visual areas of interest on the screen according to which eye movement data are to be segmented. For example, in Figure 62.4 below, interest areas are shown in two different reading texts. In the study that used the text in Figure 62.4, which was conducted on the EyeLink 1000 at Michigan State, two different groups of English language learners, matched in terms of their English language proficiency, read one of the two texts with verb forms either enhanced (in this case, in red font and underlined) or not enhanced (regular text). The purpose was to investigate the effects of enhancement on reading processes and subsequent learning of the forms. Because the forms themselves are what is of interest in this study, the researcher used the EyeLink 1000 Experiment Builder program to draw interest areas around

![Figure 62.4](image_url)

**Figure 62.4** Interest areas drawn across data collection screens in an eye-tracking experiment at Michigan State University
the forms, so that eye movement data from within those particular regions would be segmented out from all the other data and could therefore be more easily used and compared in data analyses. Interest areas can be drawn before or after data collection, and they can be moved after data collection to zero in on, or capture, other data from other regions of interest. In reading studies, interest areas are often demarcated (automatically by the eye tracker experiment design program, or as established in the eye tracker data viewer program by the researcher) along word boundaries, or even between phonemes (in studies investigating gender assignment or phoneme effects on reading processes). When customized, the interest areas can partially accommodate for data drift, which is explained below.

Participants read either the screen on the right or the one on the left in Figure 62.4 (not both). Note that, while reading, the participant does not see the outline of the interest areas. (They are invisible during the experiment.) With the interest areas demarcated, eye movement data (gaze duration, total time) on the individual areas of interest (passive verb constructions) and across the two types of text presentation (enhanced and not enhanced) could be easily compared.

Additionally, in both systems, movies of the data collection sessions are captured, and these movies can be played back for additional analyses. Also, visual maps of the data tracking can be superimposed over the reading material, as in Figure 62.5. In Figure 62.5, fixations are represented as circles. The bigger the

Figure 62.5  Fixations and interest areas in a data slide from a single participant, from a study in progress by Shawn Loewen and Solene Inceoglu. Published with permission from Loewen and Inceoglu
Technology and Assessment

In this years of research, he has found that one demographic group has come to stand out above all others in doing most likely to pass. These are not: teenagers, god, loo ball, hoagies.


fixation circle, the longer the gaze duration was. Correspondingly, the length of the fixation (the individual fixation’s gaze duration) is in milliseconds in the upper left, next to the fixation circle. Words the researchers were interested in are in boxes (hand-drawn interest areas). In Figure 62.6, the same type of data are shown, except that the interest areas were automatically generated at the word boundaries by the EyeLink 1000 Experiment Builder program. The Tobii TX300 produces similar maps, but records fewer fixations, and fixations in the Tobii TX300 system are or can be numbered to demonstrate visually the reader’s visual path as he or she worked through the text.

What the Systems Can and Can’t Do

While both systems are highly functional, provide accurate eye movement data, and are state of the art in their makeups, there are things the systems can and cannot do. In both systems, for example, over the course of an experiment, a participant’s gaze and the true fixation point may drift apart slightly, making data less accurate. Drift occurs when a participant changes head positions, or even blinks (after which the camera must relocate the pupil). In general, the longer the experiment, the more drift occurs, and thus the less precise the eye movement recordings become. This is why longer experiments are often divided into several subexperiments, with breaks and recalibration sessions between them. The EyeLink 1000 system allows for researchers to program within a single experiment drift correct sessions, but the Tobii TX300 eye tracker does not appear to allow for this. The systems differ fundamentally in other ways as well, as can be seen in Table 62.1.

Why These Systems Are Preferred by Reading Researchers

These two systems, and especially the EyeLink 1000, are preferred by reading researchers because they are extremely sophisticated pieces of equipment and are
accurate in recording eye movements as readers process text on screen. At Michigan State we own and maintain an SR Research EyeLink 1000 and a Tobii TX300 Eye Tracker. The two systems are very different, but we believe they complement each other as the strengths in one counterbalance the weaknesses in the other. But the two systems are diverse enough that data collection for a single experiment must consistently be collected on a single system—and at Michigan State we carefully guard that no spatially sensitive equipment (the camera, or the computer monitor that the participants see) is moved even a millimeter over the course of a study’s data collection period, so that measurements within an experiment have as little technology-mediated variation as possible.

For very accurate word-level eye-tracking, we generally use the EyeLink 1000 system. For experiments that center on the gross reading processes of children (for example, where on the page they look during reading or listening tests) and that are not concerned so much with the possibility of analyzing regression data, we use the Tobii TX300. Anecdotally, we have found the EyeLink 1000 to be rather difficult to program, and we often rely on the very receptive and capable support staff at SR Research to help us fine-tune or finalize our experiments. The Tobii TX300, however, is much closer to being a real “plug and play” system, but one that collects less accurate data. In my opinion, the Tobii TX300 should not be used for measuring attention at the word level unless the text is presented in an extremely large font size.

Other types of commercial eye trackers exist. For example, both SR Research and Tobii offer various types of eye trackers that have different cameras, configurations, and processors, all offered at different prices. Other systems are also available, from Applied Science Laboratories (www.asleyetracking.com) and Mirametrix (http://mirametrix.com/). And other researchers, such as those at Iowa State University, are publicly printing information on how to build low cost eye trackers (see http://thirtysixthspan.com/openEyes/) that take advantage of commercial, off-the-shelf video cameras and computers in assembling systems that can record eye movements rather well. But a review of current studies published in cognitive science, psycholinguistics, and other fields will quickly reveal that the preferred systems are from SR Research and Tobii Technology.

The Ecological Validity of Eye-Tracking Research

Novice eye-tracking researchers very quickly come to understand that reading while having one’s eyes tracked is not exactly the same as reading without having one’s eyes tracked (Gibson, 1979). And the more accurate and finely tuned the data that the researcher wishes to obtain, the truer and more severe this issue becomes. From L1-processing studies we know that eye movements are heavily influenced by textual and typographical variations presented in the text (see Dussias, 2010, for a comprehensive review of this issue). Typographical variables that have been found to influence reading processing in this way include the quality of print, the length of the line of text, and the amount of space between the letters. Fixations are also longer when readers come across low frequency or contextually implausible words. When reading becomes more advanced or
<table>
<thead>
<tr>
<th></th>
<th><strong>EyeLink 1000</strong></th>
<th><strong>Tobii TX300</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Company</strong></td>
<td>SR Research, Canada</td>
<td>Tobii Systems, Sweden</td>
</tr>
<tr>
<td><strong>Website</strong></td>
<td><a href="http://www.sr-research.com">www.sr-research.com</a></td>
<td><a href="http://www.tobii.com">www.tobii.com</a></td>
</tr>
</tbody>
</table>
| **Setup configuration** | Comes with three or more configurations, for example:  
  1. Tower mount (camera above head) with head and chin rest (for most accurate eye-tracking)  
  2. Desktop mount (camera on table in front of monitor) with head and chin rest  
  3. No mount—head-free, sticker must be placed on forehead | No mount—head-free only |
| **Camera type**     | Binocular, can use one (left or right) only for monocular tracking | Binocular |
| **Data sampling rate of camera** | 1000 Hz | 300 Hz |
| **Portability**     | Laboratory based, not portable | Portable (comes with a large suitcase-sized, padded carrying case) |
| **Pros**            | • Extremely accurate data, especially when used with head/chin-rest mounts  
  • Able to record data at the word and phoneme level  
  • Interfaces with E-Prime  
  • Videos of the eye movements can be played back (and slowed down to be seen in slow motion)  
  • Options of drift corrections (fixation crosses) help with accuracy for longer experiments  
  • Data can be easily accessed and transferred to Excel  
  • Researchers can purchase multiple HSP keys, so additional programming and data viewing can be performed on computers that are separate from the main host computer  
  • Researchers can obtain regression measures and pupil dilation measures | • Eye-tracking cameras are very discreet; system particularly suited for eye movement studies conducted with children  
  • Easier to program than most systems  
  • Interfaces with E-Prime  
  • Videos of the eye movements can be played back (and slowed down to be seen in slow motion)  
  • Able to produce visual “heat maps” of what parts of the screen were looked at the most  
  • Video and Web-based studies are easy  
  • Some statistics can be calculated in Tobii studio (e.g., mean fixation time, fixation counts for a certain interest area) and can be viewed right after the experiment, without opening another program  
  • Surveys can be easily integrated into the experiment design to collect participant information (age, gender, etc.) |
conceptually more complex, eye fixation duration increases and saccade length decreases (Duchowski, 2002).

Perhaps the number one difference between reading that is eye-tracked and regular reading is that most eye-tracking experiments require readers to read on a computer screen, which does not replicate all types of reading commonly undertaken by individuals—such as reading a book, reading on paper, or reading a partially crumpled newspaper while lounging on a couch or working at a table in a coffee shop. There are other differences. In our reading experiments using the EyeLink 1000, we have found that, for us to obtain extremely accurate word-level eye movement data, the text on screen needs to be at least double-spaced and the words and letters need to be in at least 18 to 24 point font. Eye-tracking data are more accurate at the center of the screen than at the peripheries, so interest areas (text or words we are most interested in tracking) normally should be placed toward the center of the screen, which may present text that is a bit different from what one would read in a normal, onscreen reading situation. Even though eye-tracking companies attest to the contrary, we find that contact lenses, makeup (especially mascara and eye-liner), and eye glasses can distort data recording, sometimes so much that participant data are rendered useless. And, as reported in Heuer and Hallowell (2007), young adults need to be recruited for eye-tracking research studies, because older individuals normally have some loss of visual acuity or have ocular motor deficits that distort the data. Participants are often required to pass a visual acuity test before data collection begins (for information on acuity tests, see Hyvärinen, Näsänen, & Laurinen, 1980; Woodhouse, Morjaria, & Adler, 2007).

For experiments in which we would like data from 30 to 40 participants, we find we normally need to collect data from 50 to 60 people, to allow for subject attrition due to problems in data collection or the nonability of the eye tracker to consistently and reliably track a person’s eyes. Attrition and data loss are more acute with longer eye-tracking experiments on account of eye-tracking drift. And

### Table 62.1 (Continued)

<table>
<thead>
<tr>
<th>Cons</th>
<th>EyeLink 1000</th>
<th>Tobii TX300</th>
</tr>
</thead>
<tbody>
<tr>
<td>• It takes a while to learn how to program experiments and how to calibrate participants’ eyes properly</td>
<td>• Sample rate is lower than the EyeLink 1000; data are less accurate than in EyeLink 1000</td>
<td></td>
</tr>
<tr>
<td>• The system is large, requiring a lot of dedicated lab space. It includes two computers, two monitors, etc.</td>
<td>• Does not record regressions automatically, as the EyeLink 1000 does; does not provide pupil size measures</td>
<td></td>
</tr>
<tr>
<td>• Video-based studies are time-consuming to program. Web-based studies must be simple simulations or are impossible</td>
<td>• Does not come with a computer to run the programs. Must supply a desktop or laptop to run the system</td>
<td></td>
</tr>
</tbody>
</table>
having to collect data from participants one at a time (because we only have one of each type of eye tracker) constrains us to having studies with rather small sample sizes. Small sample sizes are problematic across eye-tracking studies because, as a consequence, such studies have little power—that is, they may lack the power to reject a null hypothesis when it is actually not true. For example, Birch and Rayner (2010, p. 201) stated that one of the problems in their eye-tracking study was the lack of power; thus, even though eye-tracking can produce copious and accurate eye movement data while individuals read, the type of reading and the amount of data (in terms of numbers of participants) are constrained, which limits in part the generalizability of results from eye-tracking studies.

**Future Directions**

Eye-tracking technology will continue to be used for understanding the mechanisms that underlie reading processes. And much of this work needs to be conducted in the context of second or foreign language acquisition and testing. Part of the problem is that cognitive reading-processing models that link visual attention with reading, such as the E–Z Reader model, do a very good job at modeling regular L1 reading behavior, but they may not model the full scope of the processes involved in the reading undertaken by second language learners. L2 reading is highly problematized, in that it is characterized by a very high number of processing difficulties (lexical, semantic, morphological, syntactical, even orthographical) that may be related to several factors, including the reader’s L1 reading skills, his or her L2 proficiency, age, L2 reading skills and strategies, L2 vocabulary knowledge, and the L2 grammar itself. Thus, in L2 reading, the cognitive underpinnings that direct saccades and eye movements are extremely complex. If saccades are dependent on language processing during reading, as the E–Z Model proposes, and processing difficulties present aberrations to the model (when reading breaks down, the eyes do not move as planned to the next word, they may fixate longer where they are, or regress to problematic areas in the text), L2 reading may require a refined or modified model to account for all of the mental processes involved in L2 reading. E–Z Reader may be able to account for when processing difficulties during L2 reading occur, but not perhaps for why they do. This is why many L2 researchers are additionally collecting introspective data along with eye-tracking data (see Godfroid et al., 2010). Qualitative data from interviews, surveys, or even think-aloud protocols may triangulate eye-tracking data and reveal why L2 reading-processing difficulties occur.

Eye-tracking research in second language testing is in its infancy. But it is not without precedence. Heuer and Hallowell (2007) investigated the use of multiple choice tests that had images as options for testing comprehension in aphasia. They found that some visual characteristics of individual images influenced visual attention, which in turn influenced accuracy in the selection of a correct target image that corresponded to a verbal stimulus. And L1 and L2 reading research that uses eye-tracking technology has laid much of the groundwork that L2 testers can use. Within the testing field, Bax and Weir (2012) investigated the reading processes undertaken by English language learners when taking a test of academic English.
Using a Tobii T60, they found text and item reading patterns that were consistent across test takers. They explained that such patterns help validate the test itself as a measure of L2 reading comprehension. It is expected that research along this line will continue and expand. To conclude, further topics that may be addressed in the future by L2 testing, eye movement investigations include the following:

1. How do timed versus untimed L2 reading tests affect test takers’ reading processes? Do differing levels of test anxiety affect the processes differentially?
2. How do various levels of test-wiseness—talent in being able to appropriately and effectively apply test-taking strategies that do not overlap with the skills the items on the test are intended to measure (Allan, 1992; Harmon, Morse, & Morse, 1996; Rogers & Yang, 1996; Kalechstein, Hocevar, & Kalechstein, 1998; Yang, 2000)—affect the reading of L2 test directions, test prompts, and item choices?
3. What are the differential effects of the number of options (in discrete-point or multiple choice items on L2 listening or reading tests) on test takers’ processing of the test items?
4. Do child L2 leaners read along with the directions when the directions are read out loud by the test administrator during an L2 proficiency test?
5. At what level of proficiency (and at what level of text) can L2 readers process multiple choice options written in the L2 without there being evidence of option-based, comprehension-limiting processing difficulties?
6. In the rating of L2 essay tests in which the raters use analytic rubrics, do novice versus advanced level raters pay attention differentially to the different categories on the rubric?
7. Why are some analytic rubric categories more difficult to use (as evidenced by low inter-rater reliability on scores from those categories) than others? Do the raters not read those sections of the rubric, or do they focus intently on, for example, grammatical errors in test essays, but not link what they read with what is on the rubric? Does this change with rating experience?
8. Following up on Wagner (2007, 2008, 2010), how are pictures and video in L2 listening tests utilized by language test takers?
9. How are pictures utilized in L2 reading texts?
10. How are visual cues interpreted in video-based tests of L2 listening or in integrated writing tests that include video watching as a precursor to writing?

SEE ALSO: Chapter 11, Assessing Reading; Chapter 80, Raters and Ratings; Chapter 86, Cognition and Language Assessment

References


