

# ELIA: A software application for integrating spoken language and eye movements

Jared M. J. Berman · Melanie Khu · Ian Graham · Susan A. Graham

Published online: 5 January 2013  
© Psychonomic Society, Inc. 2012

**Abstract** We have developed a new software application, Eye-gaze Language Integration Analysis (ELIA), which allows for the rapid integration of gaze data with spoken language input (either live or prerecorded). Specifically, ELIA integrates E-Prime output and/or .csv files that include eye-gaze and real-time language information. The process of combining eye movements with real-time speech often involves multiple error-prone steps (e.g., cleaning, transposing, graphing) before a simple time course analysis plot can be viewed or before data can be imported into a statistical package. Some of the advantages of this freely available software include (1) reducing the amount of time spent preparing raw eye-tracking data for analysis; (2) allowing for the quick analysis of pilot data in order to identify issues with experimental design; (3) facilitating the separation of trial types, which allows for the examination of supplementary effects (e.g., order or gender effects); and (4) producing standard output files (i.e., .csv files) that can be read by numerous spreadsheet packages and transferred to any statistical software.

**Keywords** Eye-tracking · Analysis · Software · Language processing

The eyes have long been considered to provide insight into our cognitive processes (Huey, 1908/1968). Within the field of psycholinguistics, eye-tracking methodologies have been used extensively to study both written and spoken language

processing. Within the reading and word recognition literature, eye movements are now among the most widely used response measures (see Rayner, 2009, for a review). Similarly, eye movements are one of the primary response measures used to examine the comprehension and production of spoken language (Tanenhaus, 2007). The use of eye-gaze measures in psycholinguistics is predicated upon the assumption of a close link between eye movements and attention. Specifically, it is assumed that shifts in attention are accompanied by shifts in eye gaze (Rayner, 2009). These shifts in eye gaze, or *saccades*, are carried out unconsciously and are a result of the tendency to move object(s) to which one is attending into the area of the retina with the greatest visual acuity (i.e., the fovea). The timing and pattern of saccades, as well as the resulting fixations, allow researchers to infer how language is being processed without having to rely on an intentionally produced behavioral response, such as a verbal judgment or pointing response (Tanenhaus & Spivey-Knowlton, 1996; Tanenhaus & Trueswell, 2006).

The use of eye-gaze measures is particularly well-suited to the study of spoken language processing as a result of its fine-grained temporal resolution. Spoken language is presented at a remarkably fast pace—approximately 2.5 words per second (Snedeker, 2009). As utterances unfold, language is processed rapidly and incrementally. That is, listeners analyze language on a moment-to-moment basis, rather than at the end of each word or sentence (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Altmann & Steedman, 1988; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). While overt behavioral measures such as pointing or verbal responses can provide insight into the final outcome of language processing, they do not allow for an examination of the incremental processes that influence an individual's final interpretation (e.g., Fernald, Zangl, Portillo, & Marchman, 2008; Trueswell, Sekerina, Hill, & Logrip, 1999). Thus, through the use of eye-gaze data,

---

J. M. J. Berman · M. Khu · I. Graham · S. A. Graham  
University of Calgary, Calgary, Canada

S. A. Graham (✉)  
Department of Psychology, University of Calgary, Calgary,  
AB T2N 1N4, Canada  
e-mail: grahams@ucalgary.ca

researchers can make inferences about individuals' interpretation of spoken language as it is heard.

The visual world paradigm has emerged as one of the leading methods for investigating questions related to real-time spoken language processing (Tanenhaus, 2007). In this paradigm, subjects are presented with real objects or objects shown on a computer screen and are asked to “find,” “look at,” “click,” or “move” specific objects while their eye gaze is tracked by either a remote or a head-mounted eyetracker. Subsequent examination of the subject's eye gaze within the visual array, as they listened to the sentence unfold, provides information about how the sentence was incrementally processed and interpreted. That is, the objects upon which the individual fixates and the saccades between objects provide insight into the subject's inferences about the intended referent of the utterance. Accordingly, analyses within the visual world paradigm tend to focus on eye-gaze behavior during prespecified time windows, such as when a specific word is presented or when the intended referent of a sentence becomes clear. A necessary element of such analyses is the ability to accurately and precisely integrate eye-gaze information with concurrent spoken language.

Researchers have used the visual world paradigm to investigate a broad range of theoretical and empirical phenomena in language acquisition, comprehension, production, and conversation (see Tanenhaus, 2007; see Trueswell & Tanenhaus, 2005, for a review). The combination of spoken language input and visual information, together with the aforementioned fine-grained temporal resolution of eye-gaze measures, has allowed researchers to examine spoken language processing with a level of specificity not previously possible. For example, using the visual world paradigm, researchers have investigated the cues that adults use online to resolve ambiguity in language, including context (e.g., Hanna & Tanenhaus, 2004), syntactic preference (e.g., Trueswell & Kim, 1998), paralinguistic cues (e.g., Snedeker & Trueswell, 2003), and statistical/probabilistic information (e.g., Trueswell, 1996). Furthermore, eye-tracking methodologies have allowed researchers to investigate the development of the language-processing system in children, examining questions varying from word recognition to higher order processes such as syntactic ambiguity resolution and pronoun interpretation (for reviews, see Snedeker, 2009; Trueswell, 2008).

As eye-tracking techniques and technology have improved, it has become possible to ask increasingly complex questions regarding the online processing of language. Although advances in eye-tracking and experimental design software have made it significantly easier to design and implement eye-tracking experiments, support for the subsequent analysis of the ensuing experimental data has lagged behind. In particular, there is need for analytic support for those experimental designs in which a subject hears a verbal prompt and their ensuing eye gaze is recorded. Although the

newer design programs allow users to synchronize the timing of recorded audio prompts with a test subject's gaze response, these programs provide little support for subsequent statistical analysis. As a result, the process of examining the data for such experiments will typically involve many hours of error-prone manual manipulation of recorded data within a spreadsheet before a simple time course analysis plot can be viewed or before data can be imported into a statistical package for further analysis.

In our lab, we use a Tobii x50 eyetracker with E-Prime Extensions for Tobii (<http://www.pstnet.com/software.cfm?ID=100>). This program allows us to link eye movements to specific time points in real-time speech. E-Prime's output consists of .gazedata files, which are easily read by spreadsheet software (e.g., Microsoft Excel). Consistent with the process described above, we found that transforming the E-Prime output to prepare it for statistical analysis required a substantial manual effort. In order to resolve the above-mentioned problems and produce a more effective tool that allows psychologists to gain greater insight into their eye movement data, we developed a new software application, Eye-gaze Language Integration Analysis (ELIA), which allows for rapid integration of gaze data with either prerecorded or live spoken language input. Specifically, ELIA integrates E-Prime output (.gazedata) and/or .csv files that include eye-gaze and real-time language. The benefits of this software include the following:

1. ELIA reduces the amount of time spent preparing raw eye-tracking data for analysis. Prior to the development of ELIA, we spent multiple hours preparing each subject's data for further analysis. This meant that we were often waiting weeks just to look at a time course plot. With ELIA, this time has been drastically reduced. Each subject file is read, and an entire experiment is ready to view in seconds. Therefore, we have found that experimental issues are often revealed after collecting data from the first few test subjects, and we have then had the opportunity to fix the problem before investing enormous effort in data collection and manipulation.
2. ELIA allows for the quick analysis of pilot data in order to identify issues with experimental design or with particular subjects. ELIA has the functionality both to easily view data by subject and to group a small number of early pilot subjects together to view filler trials and ensure that they are having their desired effects or to make quick changes to experimental design.
3. ELIA facilitates the separation of trial types, which allows for the examination of supplementary effects. Specifically, we are now able to quickly split our time course graphs by trial type, condition, subject characteristics (e.g., gender, age, etc.), trial order, or any other factor that might be relevant to a given research question.

4. Lastly, ELIA uses and produces standard output files (i.e., .csv files) that can be read by numerous spreadsheet packages and transferred to any statistical software. By standardizing output files, we will continue to add functionality to ELIA by allowing for consistency across different eye-tracking setups. Although ELIA has been used only to analyze data from E-Prime studies, we are hoping to add support for other input formats in the near future to allow other eye-tracking setups to benefit from our software.

Another significant limitation of programs such as E-Prime is that they do not allow for the integration of live spoken language input with eye-gaze data. In order to synchronize language and eye gaze, it was previously necessary to specify in advance when, over the course of a trial, a given word would occur, which is possible only with prerecorded audio stimuli. The process of synchronizing live speaker input with E-Prime .gazedata files can be time consuming and error prone. ELIA supports the transformation and analysis of data from live-speaker manipulations, as well as production studies, by allowing the experimenter to add a spreadsheet with the timing values for spoken words and integrating these times into subsequent analyses.

In this article, we will first give a brief overview of the software development. We will then provide a summary of the user interface and data analysis options. Finally, we will discuss future directions and other functionality we would like to add to ELIA.

### Software development

We have implemented ELIA as a separate open-source software application, generalizing its functionality to improve its potential usefulness to others, and hoping that shared community effort will yield a more flexible and useful tool. ELIA was developed using object-oriented design and programming techniques to be platform independent (i.e., runs on PC-, Mac-, and Linux-based systems) as long as the operating system has Java 6 or higher run-time installed. The underlying model for ELIA has been implemented in Groovy, a dynamic language that offers additional flexibility and results in more concise syntax than does Java, yet still integrates seamlessly with Java and the entire Java platform. The user interface is written largely in Groovy, with the exception of the main window, which was originally laid out with the aid of WindowBuilder Pro (<https://developers.google.com/java-dev-tools/wbpro/>), which generates Java code. ELIA's plotting capability makes use of the JFreeChart Java plotting library (<http://www.jfree.org/jfreechart/>). During the development

phase, ELIA was rigorously tested for underlying bugs (for specific details of testing, please see the software Web site: <http://elia.sourceforge.net>).

### Overview of ELIA

There are two main components to the ELIA software architecture that are relevant to users of the software: the Visual-Attention Framework and the ELIA user interface, each of which will be described below.

#### Visual-attention framework

The Visual-Attention Framework is a small software framework that defines the concept of a visual-attention object or function. The visual-attention object represents, for each of a number of visual areas of interest (AOIs) over a given time period, a time course of the visual-attention ratio for that AOI. A visual-attention ratio is a value in the range of 0.0–1.0, inclusive, that represents a ratio of a subject's visual attention at a specific instant of time to a specific target AOI or subset of AOIs within a set of AOIs. The overall ELIA model includes the following:

1. *Trial*. Trials represent the individual trial that a given subject viewed. A trial also represents itself as a visual-attention object for which, at each instant of time for a given AOI, the visual-attention ratio score will either be 1.0, indicating that that subject is looking at the AOI, or 0.0, indicating that the subject is looking elsewhere. A typical time course analysis is computed as an averaged visual attention of some collection of trials.
2. *Averaged Visual-Attention function*. The Averaged Visual-Attention score draws upon the corresponding values of the trial attention scores to create a mean score.
3. *Proportioned Visual-Attention function*. This score represents proportioned visual attention to an AOI of interest, as a function of total visual attention to all the AOIs (i.e., excluding off-screen and visual attention to areas not included in any AOI).
4. *New Look Visual-Attention function*. This function filters out fixations that occurred before a predefined time point. For example, if you were interested only in fixations that began after the onset of a given word, this function would be set to filter out all looks prior to that time point.

#### User interface

The ELIA User Interface contains all the components of the ELIA user interface, including both a command-line

interface and a graphical user interface. The error-handling for ELIA has not yet been fully adapted to the newer graphical interface. Currently, all errors are still reported only to the console, so it is still essential to have a DOS window (Windows) or Terminal window (OS/X or Linux) open to receive notification about any import errors. This error-reporting weakness will be addressed with future updates.

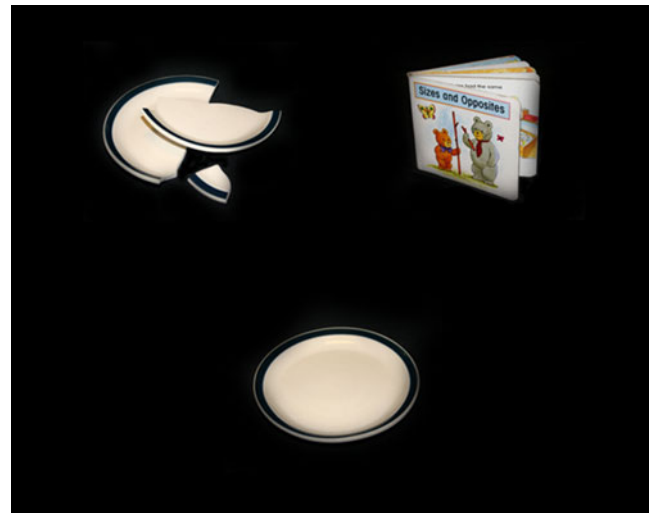
### Using ELIA

As was described above, studies using the Visual World paradigm to examine language processing will typically involve visual stimulus presentations accompanied by spoken language. Subsequent analyses tend to focus on eye gaze behavior during prespecified time windows. In these experiments, specified “events” are habitually associated with the onset and offset of predefined words. Primary analyses can include eye movement data from a given time point (e.g., following the onset of a given word), the time course of looking behavior as an utterance unfolds (e.g., over the course of a target sentence), or some combination of these. To allow for comparisons across different conditions and/or multiple subjects and/or multiple utterances, trials within a given condition must be collapsed and aligned at a common point (i.e., a zero point). For example, if the variable of interest is eye movement data following the onset of a target word, trials need to be aligned such that the onset of the target word is consistent across trials. If the start of the trial were aligned without this consideration, the location of the target word could vary in time across trials, due to considerations such as difference in word length, speaking rate, and so forth, resulting in slight differences in the timing of the onset of the target word (s). These slight differences could be amplified over the course of multiple trials and multiple subjects and could result in systematic error that could ultimately dilute experimental effects. Accordingly, having a way to reliably align trials at a given event time point is imperative. ELIA does this simply and quickly. Given the ease of using ELIA, it is also possible to use a number of different time points to align trials (e.g., adjective onset, noun onset) during the course of an analysis session.

For illustrative purposes, we briefly describe a language-processing study and outline how the data analysis for this study could be undertaken using ELIA. The study we will describe is a simplified adult version of a study conducted in our lab with 3- and 4-year-old children (see Berman,

Chambers, & Graham, 2010). We refer back to this example over the course of this section. In this example, subjects were presented with a series of visual arrays on a computer screen. Each array consisted of three static images: a broken object (e.g., a broken plate), an intact object of the same kind (e.g., an intact plate), and an unrelated filler image (e.g., a book). For each trial, the location of each type of object was programmed into the experimental software (e.g., E-Prime) in the format of AOIs (see Fig. 1). The goal of the study was to examine whether listeners used a speaker’s vocal affect (i.e., vocal cues that signal a particular emotional state) to find the referent of an ambiguous noun (e.g., *plate*).

In this example, subjects were tested in two within-subjects conditions. The verbal instructions that accompanied the arrays varied by condition (positive vs. negative affect). These instructions consisted of a prerecorded voice telling the subject to look at a given object. Instructions were in the format “Look at the X. Point to the X.” where X represents a noun (e.g., *plate*). A different noun was used across all experimental trials. There were 12 critical trials (6 per condition). In the negative affect condition, subjects were expected to look at the broken object, whereas in the positive affect condition, subjects were expected to look at the intact object. In addition to the critical trials, 12 filler trials (i.e., trials whose data did not test our hypotheses) were included to diminish the expectation that either the broken or the intact object would always be the intended referent. The instructions for these trials were presented with neutral vocal affect. We were primarily



**Fig. 1** Sample array used in the illustrative example

interested in subjects' eye movements during the first sentence (i.e., "Look at the X"), with particular focus on eye movements following the onset of the noun.

#### Data integration of eyetracker files with ELIA

Following data collection, ELIA can be used to integrate .gazedata files or .csv sample- or event-based files. The user can specify a number of experimental properties, including general experiment details, time synchronization details, and temporal events details (see Table 1 and the Appendix). These properties are described below.

#### General properties

Within the ELIA user interface, under the General tab, the user can specify general details, including the experiment name and the title of the columns in the .gazedata file, that identify subject, session, and trial. In our example (see Fig. 3 in the Appendix), we called the experiment "Example" and specified that that subject's unique subject number is found in column entitled "Subject." We further specified that information about the session number is to be pulled from the column labeled "Session" and the trial's unique identification is to be pulled from the column entitled "TrialId." In the general experiment details section, we also specified the title of the column that contains information about the condition to which a given trial belongs, "ConditionType," including specifying whether a given trial is a filler trial. ELIA recognizes filler trials by identifying trials that have the text "Filler" embedded within their value

(e.g., "NewFiller1") and omits these trials from analyses, if desired.

#### Time synchronization details

Under the fixation tab, the user specifies how the fixation data are to be time-synchronized with event data (i.e., spoken language) for each trial. For most Tobii/E-Prime .gazedata files, the column that contains information enabling synchronization with voice data is specified as the OnScreen column. In our example, the value in the OnScreen column that marked the point in time at which speaker data were to be synchronized with the .gazedata values was specified as the first appearance of the value "Probe." Note that in our example, "Probe" did not coincide with the start of the trial. In our design, the visual display appeared for 3 s without any auditory stimuli, followed by a black slide for 2 s. The same visual display then reappeared on the screen, accompanied by the verbal prompts described above. "Probe" marked the reappearance of the visual display following the black slide. Thus, the user can specify when ELIA should begin time synchronization.

The fixation characteristics within a trial can also be specified, including the name of the column indicating the AOI that a subject is fixating at any given point in time, as well as the column specifying the trial time (e.g., 23,000 ms into the experiment). In our example, the former was specified as the "AOIFixation" column, and the latter as the "TETTime" column. The user can also specify the sample rate to be used for plotting and export of experimental data. For best results, this should approximate the granularity of

**Table 1** Experimental properties that can be specified within ELIA

Experiment property	Description
Trial identification columns	Specifies column names in the .gazedata files that identify the subject, session, and trial.
Trial detail columns	Specifies column names that identify the trial type (e.g., type 1 or type 2) and filler trials (filler trials are not included in main analyses).
Trial synchronization columns	Specifies columns indicating how temporal event data should be time-synchronized with eye-tracking data.
Fixation columns	Specifies columns that indicate fixation characteristics within a trial. "Fixation" specifies which AOI the subject is fixated on in that instant.
Timing parameters	Specifies the timing values that affect trial timing and time course analyses.
Temporal event column	Specifies the column names that identify temporal events (e.g., spoken words; "Look").
Trial bounds	Specifies the bounds of interest within each trial on the basis of temporal event cues (e.g., start and end of specific words). Eye-tracking data outside these bounds would be discarded during import.
Trial timing alignment	Specifies how trials are to be aligned with each other for time course analyses.

timing recorded by the eyetracker. In our example, samples were taken every 20 ms, so “Sampling Rate” was set to 20. The user also specifies the “fixation lag time,” a fixed period of time that specifies any lag in the recording of fixation. This may be required if the eye-tracking system filters out brief sporadic shifts in gaze but does not correct for any resulting delay in recording a change in fixation. In our example, we set this value to 100 ms, since E-Prime macro recording introduces a lag of 100 ms (due to built-in filtering of spurious looks). ELIA also provides the option of filtering out fixations that occur earlier than a given time within a trial, such that only “new looks” are included in the time course analyses. In our example, we set the “new look time” to 200 ms after the zero time.

### *Event properties*

Finally, the user can specify temporal event details, under the Event tab. Here, the user specifies the column that indicates the temporal events (i.e., spoken words) occurring at a given instant, as well as events of interest. In our example, the column that indicates which word is being spoken is “TemporalEvent.” The column specifying the trial time in which the word is being spoken is the “TETTime” column (i.e., the same column as that used to specify when AOI fixations are occurring). Another user-specified event is the “start event”—in our example, the first word of the first sentence (i.e., “look”)—which indicates the temporal event value at which the trial begins. Any eye-tracking data that occurred before this event would be discarded during import from the .gaze-data file into ELIA. The user also specifies the “end event” (i.e., the first word of the second sentence—“point,” in our example), after which the data are no longer of interest. Any eye-tracking data after the end of this event will be discarded during import.

Finally, for time course analyses to be completed, trials must be aligned at a common point, the zero-time event. The zero time event should correspond to the primary temporal event of interest. In our example, the point chosen as the zero time within each trial (i.e., “noun 1”) was the first appearance of the noun within the first sentence.

Although our example used prerecorded speech data, ELIA can integrate live speaker input with gaze data. That is, ELIA allows the user to create separate .csv files that contain the required temporal information. ELIA will then integrate these files with .gazedata files to generate time course analyses.

The live speaker .csv file requires columns that identify the “TrialID” and the “Subject ID” from the .gazedata file to which the live data in the .csv file correspond. The .csv file also requires a column that identifies the temporal event occurring at a given instant (e.g., spoken words or the start/end of a trial) and a column that contains the time

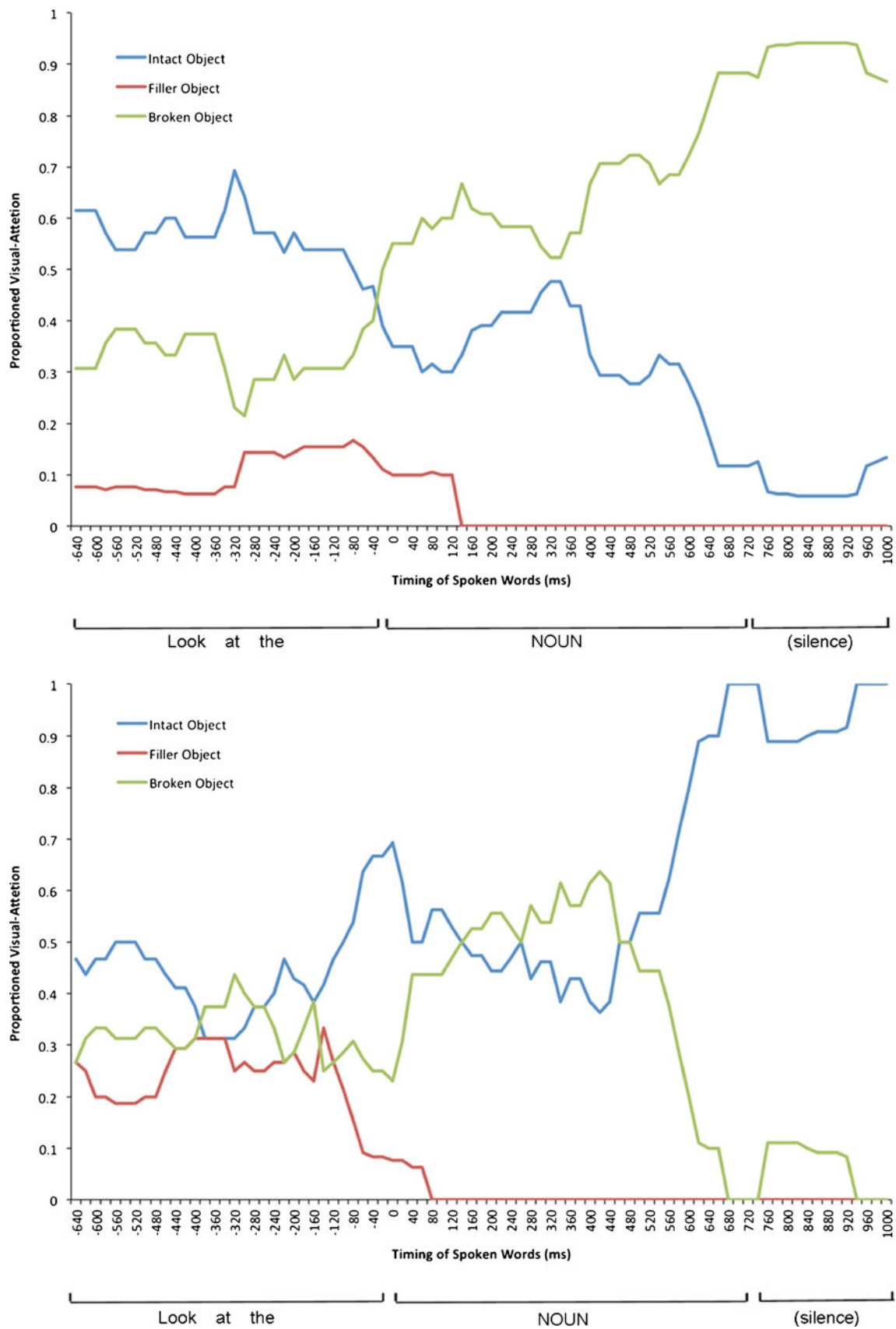
points of interest in milliseconds (e.g., from a video recording). The user would then specify how the gaze data is to be time-synchronized with the live speaker data in the Event tab of Experiment Properties (see Fig. 4 in the Appendix). In our live speaker example, we specified “LiveEvent” as the name of the column containing the synchronization events. Within the “LiveEvent” column, we specified “StartOfTrial” as the value to be used to synchronize the original gaze data to the live speaker data. In our example, “StartOfTrial” (i.e., the reappearance of the visual display following the black screen, as coded from the video recording) would be aligned with the value “Probe” in the “OnScreen” column for each trial. We further specified “LiveEvent” as the column from which ELIA should pull temporal events (i.e., “Temporal Event”) and “LiveTime” as the column indicating the time at which the temporal events occurred.

### *Errors*

If there are errors in the initial set of specified experiment properties, an error message is generated. For PC users, error messages are displayed on the Command Prompt window, which opens when the ELIA program is opened. For Mac users, error messages are displayed in the Terminal window that was used to initially open ELIA. Error messages can be used to modify the original setup so that it is compatible with the data that the user wishes to import. If there is an error with a specific trial, the program will drop the trial from the analyses and generate an error message that directs the user to the mistake (e.g., “Trial 6 missing TemporalEvent noun 1–trial”). Once all the data files have been imported and each trial has voice data with a zero-time reference, as well as synchronized fixation data, then the data are ready for time course analysis.

### *Data visualization*

Visualizing time course plots can help the user to better understand their data. ELIA provides a number of different options for plotting time course analyses, resulting in multiple ways to quickly examine data. Across all analysis options, ELIA models a trial as a set of time courses for each possible fixation target. For an individual trial, the time course for a given AOI yields a visual-attention ratio of either 1 or 0 for any given instant in time. The process of time course analysis in ELIA consists of applying a series of operations on these time course sets, each operation producing a new time course set that may be fed into additional operations. These additional operations include an *averaging operation*, a *“new look only” operation*, and a *proportioning operation*. Using our illustrative example, each of these operations is described in greater detail below.



**Fig. 2** Time course plots for the negative affect condition (*top panel*) and the positive affect condition (*bottom panel*), averaged across all trials and subjects

Recall that the start of the trial in our example was specified as the word “Look” and the end of the trial was specified as the word “Point.” The zero-time event was specified as the first appearance of the noun within a given trial. The *averaging operation* averages gaze data by condition across multiple trials. For our example, ELIA plotted all trials for the negative affect condition, where subjects were expected to look at the broken object, and all trials for the positive affect condition, where subjects were expected to look at the intact object (see Fig. 2). Note that the user has the option to plot the data for all subjects in the experiment, a subset of subjects, or even a single subject. For example, a user could select subjects by gender or select a single subject to examine whether that individual is an outlier. The user also has the option of selecting or excluding individual trials.

The “new look only” operation filters out all AOI fixations, or “looks,” which occur before a specified time within a trial. In our example experiment, we set the new look time to 200 ms; thus, all fixations that were initiated prior to 200 ms after the zero time (in our case, the noun onset) were excluded from the analysis. The *proportioning operation* proportions the visual attention ratios at each instant of time so that their total would add up to 1 (i.e., excludes time points where a subject was not fixated on a predefined AOI).

The time course plots depicted above, based on our sample data, highlight how ELIA allows users to visualize their data using time course plots. As can be seen, there is growing tendency to fixate on the *broken* referent in the negative affect condition (Fig. 2, top panel) as the noun unfolds (recall that the zero point reflects the noun onset). This pattern is reversed in the positive affect condition (Fig. 2, bottom panel), where eye movements reveal a focus on the *intact* object as the noun unfolds (albeit later in the utterance than with negative affect). Thus, sensitivity to the vocal affect was not evident throughout the utterance but, rather, only as the ambiguous noun was encountered.

#### Data export

Data are exported in a file format suitable for import into statistical packages such as SPSS (.csv files). The data export sequence makes use of the same series of processes as the data visualization procedure. ELIA is currently set up to export three types of .csv files: data sorted by condition, data sorted by subject, and data sorted by trial. The manner in which the data are exported is meant to expedite the data analysis process by sorting and configuring the data into the formats most commonly required for inferential statistics.

#### Data sorted by condition

This set of .csv files presents data by condition. That is, these data are not identifiable by subject but, rather, include all critical trials for all subjects, averaged across subjects. Two such files are exported. The first file, “tca.csv” (i.e., time course analysis), presents two types of data: data points provided in 20-ms intervals for each condition and, below, the same data points for each condition in a proportionalized format. In our example, the “tca.csv” file would present the data, averaged across subjects, for the positive affect condition and the negative affect condition. The second .csv file, “new\_look\_tca.csv” (i.e., new look time course analysis), presents the same data; however, all looks that were initiated prior to a predefined moment in time (e.g., 200 ms after the noun onset) have been removed.

#### Data sorted by subject

This set of .csv files presents data by subject. Only critical trials are included, and the trials are sorted by condition. Four such files are exported. The first file, “subjects.csv”, consists of each subject’s raw data averaged by condition. In our example, the “subjects.csv” file would present the data for subject 1, then subject 2, then subject 3, and so forth. Each subject would have a score for the positive affect condition (i.e., a score that averages across the critical trials within the positive affect condition) and another for negative affect condition. The second file, “subjectsProportioned.csv,” presents the same data in a proportionalized format. The third file, “new\_look\_subjects.csv,” presents data by subject, with all looks that were initiated prior to a predefined moment in time (e.g., 200 ms after the noun onset) removed. Lastly, “new\_look\_subjectsProportioned.csv” presents the “new look” data in a proportionalized format. These files are often the most useful for exporting data to a statistical program for analysis, since the data are already divided by subject and, therefore, it is easy to calculate individual scores (e.g., difference scores, slope scores, etc.).

#### Data sorted by trial

Only one file, a sort of “master data set,” is exported in this format. This file, “trials.csv,” presents raw data for every subject and for every trial (including both critical and filler trials) in both raw and new look formats. In our example, this file would include all 24 trials, in raw and new look format, for all the subjects in the study. This is the only file in which the data are unaltered (i.e., the data have not been averaged by subject or



condition), which makes this file particularly useful for conducting item analyses. In addition, this is the only file that includes the filler trials.

### Conclusion and future releases

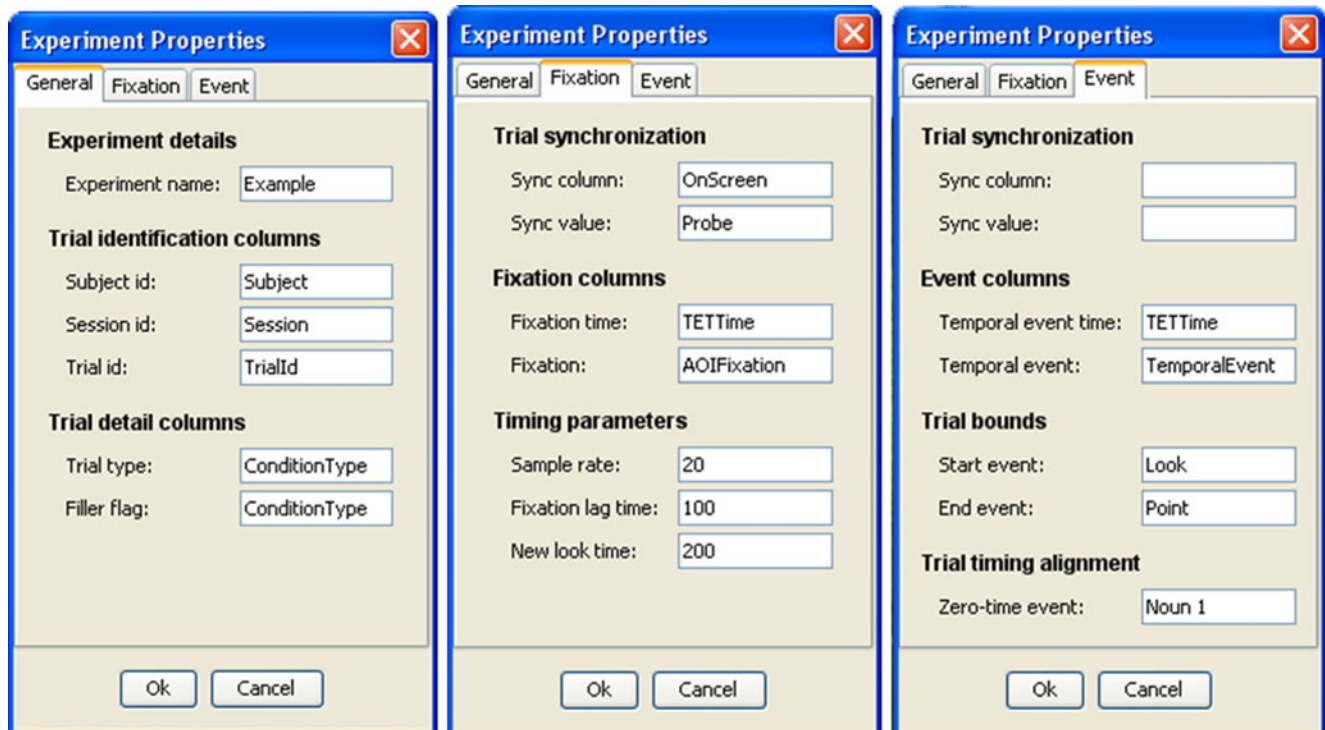
As was described in detail above, the development of ELIA now allows the psycholinguistics community to quickly export data from E-Prime experiments and visualize these data in an easily accessible platform. Our lab has used ELIA to integrate spoken language with eye-gaze data for a number of experiments (Berman, Chambers, & Graham, 2010; Berman, Graham, & Chambers, 2012; Berman, Graham, Chambers, & Callaway, *in press*). We are highly committed to the ongoing development of ELIA, ideally through collaborative efforts with others within this community. As such, we have set up a Web site for ELIA (<http://elia.sourceforge.net>), where potential users can download the program, the user manual, and sample data, as well as upload their own data, request an enhancement to the existing program, and report issues. New developers are welcome and will be provided ample support to get up and running. In addition, consistent with our commitment to the continued improvement of the ELIA

program, general comments regarding missing features or the contribution of sample eye-tracking data from other setups that could be used to test ELIA's flexibility or further extend its capability would be greatly appreciated.

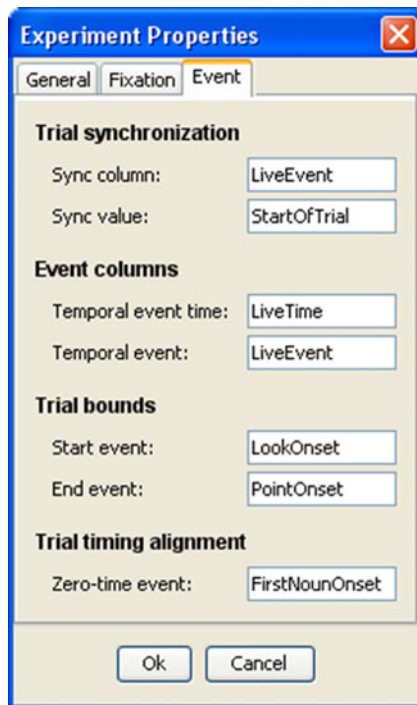
Lastly, we are currently planning the following enhancements for ELIA: (1) hover-over help graphics for configuration fields; (2) improving error reporting (i.e., errors reported in user interface, rather than in command line); (3) allowing for the modification of an experimental file after it has been created; simplifying installation for Macs and Linux based machines; (4) allowing for the easy removal of files on the basis of subject factors (e.g., gender, age, etc.); and (5) providing direct support for multiple independent variables by allowing for the quick collapsing of cells along any dimension relevant for analysis or export. It is our hope that ELIA will be used by the broader psycholinguistics community to help researchers visualize and analyze their data in a more efficient and reliable manner.

**Author note** This work was supported by funds from the Canada Foundation for Innovation, the Canada Research Chairs program, and the University of Calgary and by an operating grant from the Social Sciences and Humanities Research Council (SSHRC) of Canada (S.A.G.). J.M.J.B was supported by a graduate fellowship from SSHRC and by funding from Alberta Innovates Health Solutions

### Appendix



**Fig. 3** Experiment properties, including the general properties, fixation properties, and event properties tabs, from the illustrative example



**Fig. 4** The event properties tab from the live speaker illustrative example

## References

- Alloppenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.
- Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.
- Berman, J. M. J., Chambers, C. G., & Graham, S. A. (2010). Preschoolers' appreciation of speaker vocal affect as a cue to referential intent. *Journal of Experimental Child Psychology*, *107*, 87–99.
- Berman, J. M. J., Graham, S. A., & Chambers, C. G. (2012). Contextual Influences on Children's Use of Vocal Affect Cues During Referential Interpretation. *Quarterly Journal of Experimental Psychology*, doi:10.1080/17470218.2012.713367.
- Berman, J. M. J., Graham, S. A., Callaway, D., & Chambers, C. C., (in press). Preschoolers use emotion in speech to map new words to objects. *Child Development*.
- Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye-movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Developmental Psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Philadelphia: John Benjamins.
- Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science*, *28*, 105–115.
- Huey, E. B. (1908/1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press.
- Rayner, K. (2009). The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention during reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.
- Snedeker, J. (2009). Children's Sentence Processing. In E. Bavin (Ed.), *The Handbook of Child Language* (pp. 331–338). Cambridge University Press.
- Snedeker, J., & Trueswell, J. C. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*, 103–130.
- Tanenhaus, M. K. (2007). Spoken language comprehension: Insights from eye movements. In G. Gaskell (Ed.), *Oxford Handbook of Psycholinguistics* (pp. 309–326). Oxford: Oxford University Press.
- Tanenhaus, M. K., & Spivey-Knowlton, M. J. (1996). Eye-tracking. In F. Grosjean & U. Frauenfelder (eds), *Language and Cognitive Processes: A guide to spoken word recognition paradigms*, *11*, 583–588.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.
- Tanenhaus, M. K., & Trueswell, J. C. (2006). Eye movements and spoken language comprehension. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of Psycholinguistics: Second edition* (pp. 863–900). Academic Press, Elsevier: New York.
- Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, *35*, 566–585.
- Trueswell, J. C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I.A. Sekerina, E.M. Fernández, and H. Clahsen (eds.) *Language Processing in Children*. John Benjamins.
- Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden-path by nipping it in the bud: Fast-priming of verb argument structures. *Journal of Memory and Language*, *39*, 102–123.
- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The development of on-line language comprehension abilities in children. *Language, Minds, & Brains: Studies in Languages*, *34*, 209–215.
- Trueswell, J. C., & Tanenhaus, M. K. (Eds.). (2005). *Processing world-situated language: Bridging the language-as-action and language-as-product traditions*. Cambridge: MIT Press.