A public dataset of aphasic and healthy people during an eye-tracking-based auditory comprehension test

Michael Douglas Cabral Alves[1] Paulo Henrique Souza[1], Maria Isabel d'Avila Freitas[2], Alejandro Durán Carrillo de Albornoz (corresponding author)[3], Marcelo Gitirana Gomes Ferreira[4] and Alejandro Rafael Garcia Ramirez[1]

[1] *Department of Applied Computing, Universidade do Vale de Itajaí (UNIVALI), Brazil.* [2] *Department of Speech-Language Pathology, Universidade Federal de Santa Catarina (UFSC), Brazil.* [3] *Institute of Materials and Processes, Universidad de la Habana (UH), Cuba.* [4] *Department of Design, Universidade do Estado de Santa Catarina (UDESC), Brazil. Corresponding author e-mail address: duran@imre.uh.cu*

Abstract: This article presents a study that produced a dataset consisting of eye-tracking recordings obtained from fifteen patients with Broca's Aphasia and six healthy individuals. The aphasic volunteers had good hearing and eye function, which allowed them to use a computer for communication. Data from aphasic and healthy participants was recorded once in two partner Speech-Language Pathology Clinics. The experimental design involved thirty-two recording sessions per participant, each one corresponding to the execution of the second version of the clinically validated Test for Reception of Grammar (TROG-2). The dataset can serve as a valuable benchmark for several applications, as it provides complementary information of individual insights, improving linguistic systems, exploring gaze functions, and monitoring cognitive processes in patients undergoing rehabilitation.

Keywords: aphasia; auditory sentence comprehension; eye-gaze; eye-tracking; dataset.

Introduction: Aphasia is a neurological condition that affects a person's ability to communicate [1]. It occurs when the language centers of the brain are damaged, often due to a stroke, traumatic brain injury, brain tumor, or a neurological disorder [13]. Aphasia can impact various aspects of communication, including speaking, understanding spoken language, reading, and writing [2], [4]. In particular, the Broca's Aphasia is characterized by difficulty in speaking and forming complete sentences, but the person can usually understand what others are saying [3].

Increasingly, eye tracking is being used to characterize language disorders in the clinical population. It is a promising technique for investigating comprehension, even in complex cases, like in People with Aphasia (PWA) [2]. This approach has enabled the development of models for sentence processing at various linguistic levels, from sound and phonological structure processing to word, sentence, and discourse processing [5],[6],[11].

The use of eye tracking in linguistic studies assumes that eye movements can result from mental processing, or that the mental processing is related to eye movements [6]. This technique can offers several metrics that can assess the work of professionals [5], [7].

A previous experiment in Brazil using eye tracking with cognitively healthy young adults showed that this method is feasible [9]. However, there are no works combining Brazilian Portuguese language processing and eye tracking in PWA.

This study continues the work in [10], that approaches aphasia rehabilitation, its challenges, and the differences based on educational level and region in Brazil. Our goal is to provide a tool to perform an auditory comprehension task while relevant metrics using eye-tracking. Latency, accuracy, gaze density and gaze fixations points can be measured.

Upon reviewing the literature, several notable studies have employed eye tracking to investigate auditory comprehension in individuals with aphasia, often comparing them to healthy participants. However, while many of these studies involve controlled experimental conditions, they do not always provide publicly available datasets.

Background: In this work, an investigation was conducted to identify techniques and evaluation criteria used in linguistic studies for analyzing auditory sentence processing with eye tracking. Former studies have explored the use of eye tracking to understand language processing in aphasic individuals, tracking their gaze patterns while they perform auditory comprehension tasks.

In [12] researchers aimed to analyze differences in how people with aphasia and healthy individuals process auditory stimuli, particularly, how gaze patterns can reveal difficulties in understanding spoken language.

In [14] authors used eye tracking to investigate sentence comprehension in aphasia. The researchers examined how individuals with aphasia and healthy participants controls processed subject-verb-object sentences. Their findings revealed differences in gaze patterns between the two groups during auditory sentence comprehension tasks.

In [15] authors used eye-tracking to explore real-time sentence comprehension in aphasia. The researchers used visual-world paradigms to investigate how aphasic and healthy individuals processed syntactic structures in spoken sentences, showing that aphasic participants had delayed comprehension relative to the controls participants.

The research in [16] discused the use of eye-tracking in auditory comprehension assessments for individuals with aphasia. The study compared the eye-tracking patterns of aphasic participants to those of healthy controls, providing insights into how aphasic individuals process auditory information.

The work in [17] presented the eye-tracking technology to analyze sentence processing in individuals with aphasia, specifically on verb argument structure comprehension. The comparison of aphasic participants and healthy controls highlighted the real-time processing deficits in the aphasic group.

While these studies offer valuable insights into aphasic individuals' auditory comprehension using eye-tracking, they primarily represent experimental research rather than publicly available datasets [18]. They lay the groundwork for future exploration of eye-tracking as a tool for understanding language processing in aphasia.

Materials and Methods: The selected eye-tracker was the PCEye Mini. This model has features suitable for this work. The capture rate is 60 Hz, and the data stream rate is 30 frames per second (FPS). The precision and accuracy values are less than 1.9° and 0.4°, respectively [19].

The TROG-2, first developed in English, was culturally translated to the Brazilian Portuguese [8], [9]. It was also used in this work because it allows for a comprehensive assessment of auditory comprehension and includes a relatively broad range of sentence types. The test and its license were acquired from Pearson.

The TROG-2 consists of 80 sentences organized into 20 different sets, each one comprising four sentences and four images. For each sentence, an image illustrating the target is presented along with three additional images that modify the target through a grammatical or lexical element. The difficulty of these sets escalates progressively and utilizes a basic and limited vocabulary consisting of nouns, verbs, and adjectives [8].

In this work, eight blocks of the TROG-2 were considered, Table 1. Complex sentences deviate from the canonical order and individuals affected by aphasia face more challenges when dealing with them. Noncanonical sentence orders impose increased processing demands, even on healthy individuals, as they require the identification of additional morphological hints [5].

Block	Grammar Structure with two elements	TROG identification
B1	Structure with two elements	A
B2	Reversible "In" and "over"	C
B3	Structure with three elements	D
B4	Reversible SVO	E
B5	Structure with four elements	F
B6	Reversible passive voice	K
B7	Subject predicative adjective	Q
B8	Relative clause - object	S

Table 1: Sentence sets and grammar structures.

When using a set of words independently of images, visual stimuli are created or selected based on specific criteria, as certain image features can affect the results of the experiment. Consequently, it is common to use images that are clinically validated or sourced from previous research [19],[20].

Figure 1 shows an overview of the system. The speech therapist uses the software interface to define the activities for each participant and to initiate data collection. The data is then processed to generate and display relevant metrics. The execution of the TROG-2 comprehension test enables the professional to monitor rehabilitation progress.

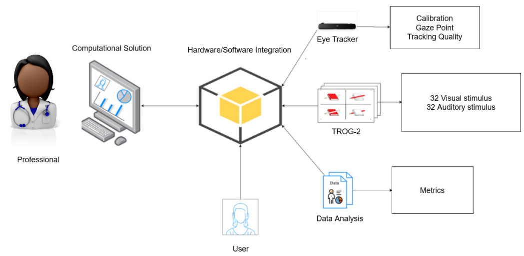

Fig. 1. Setup of the experiment.

The visual stimuli were digitized and displayed on a computer at specific time intervals, carefully synchronized with auditory cues, Figure 2.

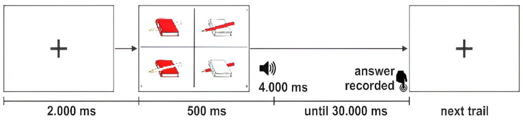

Fig. 2. Example of stimuli.

In this study, participants viewed images while listening to selected sentences. A fixation cross appeared in the center of the screen for 2 seconds before displaying the visual stimulus for 30 seconds. Subsequently, an auditory prompt was presented for 4 seconds, starting 500 milliseconds after the visual stimulus ended. During each trial, participants were required to choose the image that matched the auditory prompt [5]. Figure 3 illustrates a sample stimulus and the gaze data collected through eye tracking.

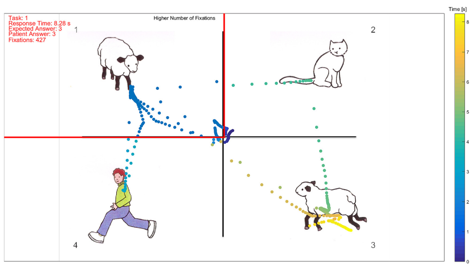

Fig. 3. Example of stimulus with fixation points and gaze shift.

Typically, impaired morphological or syntactic functions are analyzed using eye tracking by examining accuracy, reaction times, ocular tracking movements, and gaze fixation. Individuals with aphasia usually show low precision and a greater volume of gaze fixation points. These findings could be associated with deficits in sentence processing or competition for cognitive resources. Ocular movements can reveal different strategies for sentence interpretation [6],[21]. Additionally, the scan path offers insights into the strategies employed (e.g., searching for the image that fits the correct auditory stimulus).

Dataset : A multidisciplinary team conducted a pilot study involving fifteen individuals with aphasia in the Experimental Group (EG) and six healthy individuals in the Control Group (CG). The study was conducted at two partner Speech-Language Pathology clinics, located at Univali and UFSC, respectively.

The participants of the Experimental Group (EG) had a mean age of 58 years, while in the Control Group (CG) they had a mean age of 49 years. All participants were native Portuguese speakers from Brazil, and they reported having normal or corrected visual acuity. Table 2 shows the data of the participants in the study conducted at the Speech-Language Pathology Clinic at UFSC.

Both groups participated in a hearing assessment that included a thorough series of audiological tests, performed by a speech therapist with expertise in audiology. The screening of the participants evaluated the cause, type, and severity of their condition.

Detailed data on the volunteers can be found in the supplemental material. Figure 4 depicts a participant and the speech therapist conducting the experiment.

Individual	*Sex*	*Age (years)*	*Education (years)*	*Types of Aphasia*
PA1	Male	61	4	Broca
PA2	Male	66	11	Transcortical Motor
PA3	Male	36	11	Transcortical Motor
PA4	Male	42	11	Broca
CP1	Female	40	17	-
CP2	Female	44	14	-
CP3	Female	54	11	-
CP4	Female	41	11	-

Table 2: Participants' data.

Fig. 4. Experimental study.

A laptop was placed 50 cm away from the participant's eyes, with the eye tracker positioned directly beneath the screen to record the eyes movements. Additionally, a customized keyboard with four keys was developed, with each key representing the position of a visual stimulus. The keyboard was used to register the volunteer's responses to the challenges. A headset was used to deliver the auditory stimuli.

Accuracy was determined by calculating the average number of correct responses per block in the TROG-2. Each correct response was recorded based on the participant's input using the numeric keypad. The accuracy rate was obtained by calculating the ratio of correct responses to the total number of sentences in each block (four sentences per block). Incorrect responses or lack of responses were counted as errors. For example, if a participant had two correct responses and two errors within a block, the resulting accuracy is 50%.

The number of fixations was calculated by summing the gaze points recorded on the screen between the presentation of the stimulus and the key press. Gaze fixation points are determined from the coordinates provided by the eye tracker data stream. Each coordinate from the eye tracker includes a timestamp in microseconds. Reaction time is defined as the duration from the first gaze fixation on the visual stimulus to the subsequent press of the keyboard.

The quantitative analysis of the eye-tracking metrics was performed by descriptive statistics. The Kruskal-Wallis hypothesis test was used for the accuracy, reaction time, and the number of fixations. The test was applied to the averages of the metrics.

As a result, this work provides a file in CSV format, consisting of five columns of data separated by commas. The columns contain, from left to right: the key pressed by the participant, the timestamp of the eye-tracking record, the Cartesian coordinates of the gaze on the display, and finally, the visual stimulus presented. Figure 5 shows an example of this file.

	A	B	C	D	E	F
1	Key,Timestamp,X,Y,Stimuli					
2	KEY1,3965419311272,0.785198,0.922237,TROG_5.png					
3	KEY4,3965419327927,0.784765,0.928204,TROG_5.png					
4	JUMP,3965419344582,0.784096,0.941167,TROG_5.png					
5	NULL,3965419361237,0.78353,0.955082,TROG_5.png					
6	NULL,3965419377891,0.782995,0.967272,TROG_5.png					
7	NULL,3965419394546,0.782449,0.978474,TROG_5.png					
8	NULL,3965419411201,0.782035,0.987963,TROG_5.png					
9	NULL,3965419427855,0.781549,0.995548,TROG_5.png					
10	NULL,3965419444510,0.781043,1.00204,TROG_5.png					
11	NULL,3965419461165,0.780594,1.00837,TROG_5.png					
12	NULL,3965419477819,0.780116,1.01304,TROG_5.png					
13	NULL,3965419494474,0.779748,1.01611,TROG_5.png					

Fig. 5. Raw data.

The metrics were extracted via a MATLAB script designed to retrieve, process, and tabulate the data recorded in the CSV file. The script's output consists of two text files (.txt) for each group, as illustrated in Figure 6.

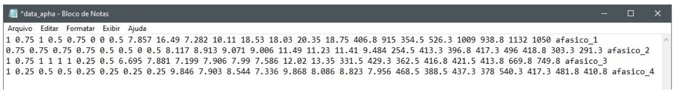
*data_apha - Bloco de Notas
Arquivo Editar Formatar Exibir Ajuda
```
1 0.75 1 0.5 0.75 0 0 0.5 7.857 16.49 7.282 10.11 18.53 18.03 20.35 18.75 406.8 915 354.5 526.3 1009 938.8 1132 1050 afasico_1
0.75 0.75 0.75 0.75 0.5 0.5 0 0.5 8.117 8.913 9.071 9.006 11.49 11.23 11.41 9.484 254.5 413.3 396.8 417.3 496 418.8 303.3 291.3 afasico_2
1 0.75 1 1 1 1 0.25 0.5 6.695 7.881 7.199 7.906 7.99 7.586 12.02 13.35 331.5 429.3 362.5 416.8 421.5 413.8 669.8 749.8 afasico_3
1 0.25 0.5 0.5 0.25 0.25 0.25 0.25 9.846 7.903 8.544 7.336 9.868 8.086 8.823 7.956 468.5 388.5 437.3 378 540.3 417.3 481.8 410.8 afasico_4
```

Fig. 6. Output file.

Each row represents the data for each participant, and specific average metrics are reported after the initial eight columns. The first eight columns display the average accuracy, the next eight columns show the average response time, and the final eight columns present the average number of fixations. Additionally, the script generates images in which the visual stimuli are overlaid with the gaze fixation points.

Previous studies suggest that individuals with aphasia have lower accuracy than non-aphasic individuals [13]. In the study conducted in the Speech-Language Pathology Clinics at Univali, the Experimental Group (EG) shows a consistently lower average accuracy across all TROG-2 blocks, Figure 7 and Figure 8.

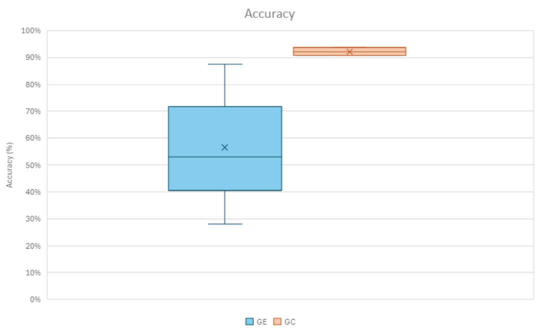

Fig. 7. Accuracy results.

Also the blocks containing grammatically simpler sentences (A, C, D, E) demonstrate greater accuracy within the EG compared to the more complex sentences (F, K, Q, S). The metrics observed in the CG follow a similar trend, although to a greater extent than those in the EG.

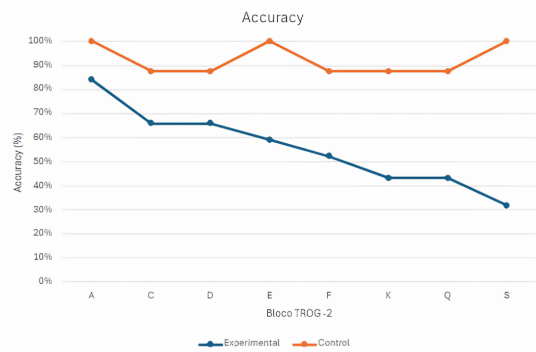

Fig. 8. Accuracy deteriorates with grammatically more complex sentences.

Data availability: .

Aphasic Group (It will be available in the final version)

Control Group (It will be available in the final version)

Conclusions *:* Auditory comprehension tests are often paired with eye tracking to measure how participants with neurological disorders listen to spoken sentences and follow along visually, providing data on reaction time, focus, and comprehension. However, there is not a single well-known public dataset explicitly combining these elements: eye-tracking, aphasia, and auditory comprehension. In this work, the raw dataset will be available for public use or broader research purposes. It would be a valuable resource for advancing personalized therapy and better understanding the mechanisms behind aphasia. Ongoing research is developing in this area.

Acknowledgment: This work has been supported by National Council for Scientific and Technological Development-(CNPq), grants 305835/2021-1, 381251/2022-5, 381204/2023-5; (CAPES) process number 11/2014, and the Fundação de Amparo a Pesquisa e Inovação em Santa Catarina-(FAPESC), grant 2020TR729.

References

1 Bonvillian, J.D., Lee, N. K., Tracy T. Dooley, Filip T. Loncke. (2020). Simplified Signs: A Manual Sign-Communication System for Special Populations. Volume 1. Cambridge, UK: Open Book Publishers.

2 Sharma S, Kim H, Harris H, Haberstroh A, Wright HH, Rothermich K. Eye Tracking Measures for Studying Language Comprehension Deficits in Aphasia: A Systematic Search and Scoping Review. J Speech Lang Hear Res. 2021 Mar 17;64(3):1008-1022. Epub 2021 Feb 19. PMID: 33606952.

3 Adornetti, I. (2019). Le afasie di Broca e di Wernicke alla luce delle moderne neuroscienze cognitive. Rivista Internazionale di Filosofia e Psicologia. Vol. 10, n.3, pp. 295-312. http://doi.org/10.4453/rifp.2019.0025

4 Arantzeta, M., Howard, D., Webster, J., Laka, I., M-Zabaleta, Maite; Bastiaanse, R. (2019). Bilingual aphasia: Assessing cross-linguistic asymmetries and bilingual advantage in sentence comprehension deficits. Cortex. p. 195-214. https://doi.org/10.1016/j.cortex.2019.04.003

5 Souza, P. H. de, Freitas, M. I. de and Ramirez, A. R. G. Assessment of Aphasics Fluent in Portuguese Language Using Eye-Tracking. (2023). 18th Iberian Conference on Information Systems and Technologies (CISTI), Aveiro, Portugal, 2023, pp. 1-4. https://ieeexplore.ieee.org/document/10211448

6 Schumacher, R., Cazzoli D., Eggenberger N., Preisig B., Nef T., Nyffeler T., Gutbrod K., Annoni J-Marie; Müri R. M. (2015). Cue Recognition and Integration – Eye Tracking Evidence of Processing Differences in Sentence Comprehension in Aphasia. Plos One. p. 1-15.

7 Alves, M.D.C., Souza, P. H. de, Freitas, M. I. de and Ramirez, A. R. G. Exploring Sentence Processing in People with Aphasia in the Brazilian Context Using Eye-Tracking. Authorea. April 02, 2024. 10.22541/au.171204756.68362058/v1

8 Bishop D. Test for Reception of Grammar (TROG-2). London, UK: Harcourt Assessment; 2003.

9 Carthery-Goulart, M. T., Oliveira R., Almeida I. J., Campanha A., Silva Souza D., Zana Yossi, Caramelli P., Machado H. T. Sentence Comprehension in Primary Progressive Aphasia: A Study of the Application of the Brazilian Version of the Test for the Reception of Grammar (TROG2-Br). Frontiers, v. 13, p. 1-1, May 2022

10 Lima, R. R., Borges, R., Lima, H. do N., Caldana, M. de L., Freitas, M. I. D. (2024). A Survey on the Speech Therapy Rehabilitation Landscape for Aphasia in Brazil. Aphasiology, 38(9), 1463–1483. doi.org/10.1080/02687038.2024.2314325

11 Faria, V. A., Race, D., Kim, K.; Hillis, A., E. (2018). The eyes reveal uncertainty about object distinctions in semantic variant primary progressive aphasia. Cortex. p. 372-381. doi.org/10.1016/j.cortex.2018.03.023

12 Mack, J. E., Thompson, C. K. (2017). Effects of verb meaning on lexical integration in agrammatic and anomic aphasia: Evidence from eye tracking. Journal of Speech, Language, and Hearing Research, 60(9), 2606-2623.

13 Hillis, A.E. (Ed.). (2015). The Handbook of Adult Language Disorders (2nd ed.). Psychology Press. https://doi.org/10.4324/9781315713557

14 Knilans J., DeDe, G. (2015). Online Sentence Reading in People With Aphasia: Evidence From Eye Tracking. Am. J. Speech Lang Pathol. Nov 24(4):S961-73.

15 Dickey, M.W., Thompson, C. K. (2009). Automatic processing of wh- and NP-movement in agrammatic aphasia: Evidence from eyetracking. Journal of Memory and Language, 61(4), 524-541.

16 Hallowell, B., Pan, Y. (2014). Eye movement-based comprehension assessment for aphasia: Effects of presentation modality. Aphasiology, 28(10), 1214-1231.

17 Dickey, M. W., Choy, J. J., Thompson, C. K. (2007). Real-time comprehension of wh-movement in aphasia: Evidence from eyetracking while listening. Brain and Language, 100(1), 1-22. 10.1016/j.bandl.2006.06.004

18 Goral, M., Lerman, A. (Eds.). (2023). Advances in the Neurolinguistic Study of Multilingual and Monolingual Adults: In honor of Professor Loraine K. Obler (1st ed.). Routledge. https://doi.org/10.4324/9781003330318

19 Tobii. (2021, Mar, 17). PCEye Mini. https://www.tobiidynavox.com/devices/eye-gaze-devices/pceye-mini-access-windows-control/Specifications

20 Hanne, S, Burchert, F., Bleser, D. R., Vasishth, S. (2015). Sentence comprehension and morphological cues in aphasia: What eye-tracking reveals about integration and prediction. Journal of Neurolinguistics. p. 83-111. https://doi.org/10.1016/j.jneuroling.2014.12.003

21 Mack, J. E., Nerantzini, M., Thompson, C. K. (2017). Recovery of Sentence Production Processes Following Language Treatment in Aphasia: Evidence from Eyetracking. Frontiers In Human Neuroscience. p. 0-20. https://doi.org/10.3389/fnhum.2017.00101