The Use of PsychToolbox-3 for Tracking Bilingual Sentence Processing and Comprehension

Judit Navracsics[1][0000-0002-8147-3023] and Liubov Darzhinova[2][0000-0003-2284-3140]

[1]Faculty of Modern Philology and Social Sciences, University of Pannonia, 1 Wartha Vince St, Veszprém 8200, Hungary
[2]Centre for Research on Linguistics and Language Studies, The Education University of Hong Kong, 10 Lo Ping Rd, Tai Po, New Territories, Hong Kong S.A.R. China
navracsicsjudit@almos.uni-pannon.hu, liubov@s.eduhk.hk

Abstract. The present article observes the continuous pursuit of experimental software in psycholinguistic research and the need to elaborate more on the practical use of isolated software packages. On the example of a lab-based programme MATLAB with its experiment control library PsychToolbox-3, the article offers the results of the systematic review of literature conducted in the Scopus database. The results of the review suggest that there is a lack of information about PsychToolbox-3 in MATLAB in the literature. The article further provides a comprehensive overview of the two psycholinguistic experimental studies of (i) Hungarian-English bilinguals, and (ii) Russian-English bilinguals, undertaken using PsychToolbox-3 in MATLAB. The findings reveal that similar linguistic typology maintains a positive influence on sentence comprehension, and morphosyntactic information available at a word level may alleviate the overall processing load. The authors also conclude that language proficiency plays a significant role in syntactic processing. Overall, the article illustrates that psycholinguistic experiments built on MATLAB can be administered in future studies involving bilingual populations, whereas this instrumental groundwork can emerge as one of the technology-mediated language assessment tools.

Keywords: PsychToolbox, MATLAB, Bilinguals, Experiment, Sentence Processing, Sentence Comprehension, Psycholinguistics.

1 Introduction

In the field of psycholinguistics, there has always been a search for experimental software, which performs the tasks reaching a high level of precision in both spatial and temporal stimuli demonstration alongside a cost-effective collection of data from participants. In this context, one of the major steps that psycholinguists are required to take at the phase of planning and designing an experiment concerns choosing appropriate and relatively accessible apparatus and techniques of stimuli demonstration. For at least a decade, technological advances have been dictating the psycholinguists to favour for computer programmes out of all the interface spectrum to relatively easy and efficiently interact with research subjects and as a result, retrieve data for further analysis.

There is plenty of software options to create and run psycholinguistic experiments [1]. The three groups of such software are pointed out: (i) allowing for the so-called "creative freedom"; (ii) giving "the advantage of the learning curve" at the cost of creative freedom; (iii) combining both the creative freedom and relatively user-friendly interface. The first group is represented by C, Java, MATLAB, Presentation, Python, R; the second – by E-Prime, Paradigm, SuperLab, and the third – by OpenSesame and PsychoPy.

The one that is discussed in this paper is MATLAB, the research software tool from the first group that can fully adapt to researchers' needs. It namely gives the ability to control for each step of an experiment, and any sort of stimuli can be generated including but not limited to text, pictures, and sounds [2]. Built on C and Java, MATLAB with its experiment control libraries allows researchers to write scripts for experiments directly via the software interface and run these scripts afterward. Kleiner et al. note that PsychToolbox is the most appropriate for psycholinguistic research out of all available MATLAB experiment control libraries [3].

PsychToolbox supports functions permitting to link MATLAB interface up with the connected hardware so that investigators can account for the accuracy of both visual and auditory stimuli and at the end receive highly precise data output. Within PsychToolbox, one may also find useful the function to conduct eye-tracking experiments within EyeLink Toolbox or the function to command the EEG experiments within NetStation.

To the best of our knowledge, not much psycholinguistic literature reports on the application of PsychToolbox in MATLAB. Only a limited number of studies (addressed in Paragraph 2) that can be found in the Scopus database concentrates on the use of PsychToolbox control library in psycholinguistic experiments. This paper is going to fill in the gap by elaborating on the psycholinguistic experimental studies of Hungarian-English and Russian-English bilinguals carried out with the use of PsychToolbox-3 in MATLAB and their implications for research and education.

2 Literature Review

The below presented literature review is based on the search conducted in the Scopus database on September 21st, 2020. The search results exhibited 9 peer-reviewed research articles of the past ten years found with the use of the following Boolean operators:

TITLE-ABS-KEY (psychtoolbox) AND (LIMIT-TO (SUBJAREA , "NEUR") OR LIMIT-TO (SUBJAREA , "PSYC") OR LIMIT-TO (SUBJAREA , "ARTS")) AND (LIMIT-TO (ACCESSTYPE(OA))) AND (LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2013) OR LIMIT-TO (PUBYEAR , 2012) OR LIMIT-TO (PUBYEAR , 2011))

After skimming the abstracts of the retrieved items, five of them were excluded from the list for further analysis due to their little or no association with psycholinguistics:

[…] AND (EXCLUDE (PREFNAMEAUID , "Hallett, D.#36189855100")) AND (EXCLUDE (PREFNAMEAUID , "Wijntjes, M.W.A.#18038930300")) AND (EXCLUDE (PREFNAMEAUID , "Hung, C.P.#57215695276")) AND (EXCLUDE (PREFNAMEAUID , "van Boxtel, J.J.A.#55898776100")) AND (EXCLUDE (PREFNAMEAUID , "Nuutinen, M.#34873178300"))

As a result, we report on four peer-reviewed Scopus-indexed research articles, with an oldest-first order.

Wilson, Tresilian, and Schlaghecken introduced the Masked Priming Toolbox (MPT) that is designed to carry out masked or unmasked priming experimental trials [9]. Masked priming is a classic psycholinguistic experiment for probing cognitive processing stages, which are not receptive to conscious control. For instance, masked priming can be performed for examining the mechanisms transmitting information about and yield reactions to emotional facial expressions. The MPT utilizes an array of devices for response collection, such as computer input devices (e.g., tablets) or sensors with its frequency response function (e.g., force transducers). The stimuli of the toolbox are operationalized across temporal (such as timing) and spatial constraints (such as location, orientation, and size). The MPT incorporates a prime, a target, and/or a mask, though is not confined to those only but can be modified to demonstrate stimuli appropriate for semantic or emotional priming. After running a trial in the MPT, raw results are brought to Excel and MATLAB documents for further processing by a researcher.

Sogo presented a new function for PsychToolbox entitled as Sgttoolbox [8]. It is a toolbox operating in order to command a cross-platform SimpleGazeTracker, an eye-tracking application of Gaze Parser programme. In connection with SimpleGazeTracker, Sgttoolbox affords a new approach to coordinate eye tracking process with scripts written in PsychToolbox. After examining this function, Sogo found only three minor issues in using Sgttoolbox along with SimpleGazeTracker: (i) absence of an on-line fixation-detection feature; (ii) latency of immediate obtainment of the latest gaze position; (iii) head movements of tested subjects can be delimited by means of a headrest. However, the three reported issues can be forestalled either by using an enhanced graphics card or settled after conducting an experiment by giving specific instructions in a command line. The study also revealed that time precision of eye-movement sampling frequency improves when done through one computer, whereas greater performance is achieved when presenting on one computer and recording via another one.

Niehorster, Andersson, and Nyström proposed a package named Titta for the integration of Tobii eye-tracking hardware used in conjunction with PsychToolbox in MATLAB and other software [7]. The authors elucidate on the five stages of Titta's implementation in psycholinguistic eye-tracking experiments. For example, the first stage corresponds with the setup of experimental subjects by showing the display, which asks the subject to adjust the position of their head, i.e. move it in a way that it precisely fits in the displayed blue circle. The next one requests to proceed with calibration and validation of the gaze positioning with each eye separately by presenting numerical values at the end of the stage. The subsequent stages are the actual real-time data streams, synchronization, and file saving.

Bridges et al. juxtaposed a number of existing experimental software psycholinguistics makes use of to introduce stimuli to subjects and receive output data with their responses and response latency times [4]. The authors tested an array of experimental software applications, among which PsychToolbox in MATLAB was too. Tested on different platforms such as Ubuntu, macOS, and Windows 10, PsychToolbox showed high-quality results in giving temporal data with great precision in a series of studies using audio, visual, and response rates. The study concludes that PsychToolbox in MATLAB is proved to demonstrate the same accuracy in measuring response latencies as other packages as E-Prime, Presentation, and PsychoPy.

It is bound to note that the reviewed Scopus-indexed studies are largely limited to the description of the isolated functions of PsychToolbox and do not extend any illustrations of actual psycholinguistic studies with the use of this MATLAB experiment control library. Nevertheless, the analyzed research articles draw attention to the magnitude of PsychToolbox in MATLAB over other existing software packages by illuminating its high-level of precision in the context of temporal and spatial measurements as well as operational connectivity with the hardware used for experimental studies with eye-tracking or EEG. Surprisingly, the contemporary peer-reviewed research using PsychToolbox, based on our Scopus database search, remains underrepresented in the current-day psycholinguistic literature, and necessitates further deliberation.

We see a reasonable explanation for this state of affairs. First, the use of an open-access and open-code PsychToolbox entails a purchased MATLAB license or an institutional subscription. Second, it requires researchers to have at least novice programming language skills to operate within the software such as to make modifications in the scripts of their experiments. Third, to our knowledge, there is no available archive of PsychToolbox scripts to allow for obtaining a skeletal framework in establishing a new experimental study carried out with the use of PsychToolbox in MATLAB.

3 The present study

Two MATLAB-based experiments that are reported in this paper examined written language processing at the sentence level in the two bilingual groups: (i) Hungarian-English (Group 1), (ii) Russian-English (Group 2).

The experiment with Group 1 was projected to observe more accurate and speedier syntactic processing of Hungarian sentences owing to early access to morphosyntactic features of the processed lexical items. It also projected to highlight that second language proficiency modulates accuracy and speed of semantic processing.

The experiment with Group 2 had a slightly different focus and expected to indicate more rapid semantic processing of both Russian and English sentences as opposed to syntactic processing. It was also expected to see syntactic processing of both Russian and English sentences as more precise compared to semantic processing.

3.1 Method

Participants. For both experiments, we recruited healthy adults with normal or corrected-to-normal vision and reportedly with no neurological disorders.

97 Hungarian-speaking users of English were employed for the first experiment (70 females, 27 males: Mage = 31.21). The subjects were right-handed, 34 of them were early and 63 late bilinguals. Reliant on the literature, early bilinguals are those who commenced their acquisition of English by the age of 11 [5]. For the second experiment, 21 Russian-speaking users of English were recruited (19 females and 2 males: Mage = 21.8). These subjects were right-handed (n = 19) and left-handed (n = 2), and all were early bilinguals.

The accepted subjects of the two groups were self-reportedly with C1 and B2 levels of English, according to the Common European Framework of Reference for Languages. Both experiments included the dominant number of females. All the subjects belonged to the university setting, i.e. were either students or professors. They all resided in their country of origin at the time of participation in our experiments.

Materials and design. The material for Experiment 1 entailed the sentence set of Hungarian and English, the languages which diverge typologically and genetically.

Hungarian is a Finno-Ugric language, whereas English is an Indo-European language and takes its place in the West Germanic group. The key difference between the two languages is related to morphology. In Hungarian, word meanings are changed by inserting multiple endings or suffixes, and that is why it is recognized as agglutinative, whereas English makes use of prepositions. Remarkably, Hungarian is typologically similar to the majority of Turkic languages, which are also agglutinative in nature and phonologically distinct with their phenomenon of vowel harmony.

The stimulus material for Experiment 2 is built on the two languages of the Indo-European language family, Russian and English, which vary in different aspects. For example, unlike Russian, English has a fairly fixed word order. The other crucial difference is that both languages use different alphabetical writing scripts, i.e. Russian uses Cyrillic, whereas English uses Latin.

For Experiment 1, 240 sentences were generated, among which were sixty English and sixty Hungarian correct sentences as well as sixty semantically incongruent and sixty syntactically incorrect sentences in English and Hungarian:

Table 1. Examples of stimuli for Experiment 1

Hungarian semantically anomalous sentence	Hungarian syntactically violated sentence	English semantically anomalous sentence	English syntactically violated sentence
A labda szögletes. Translation: The ball is square.	Kati becsukta az ablakig. Translation: Kati has closed to the window.	I understand what you sneeze	He deep dug a hole

The material for Experiment 2 was analogous to the first one. 240 sentences were produced, among which were sixty English and sixty Russian correct sentences as controls along with sixty semantically incongruent and sixty syntactically anomalous sentences in Russian and English:

Table 2. Examples of stimuli for Experiment 2

Russian semantically anomalous sentence	Russian syntactically violated sentence	English semantically anomalous sentence	English syntactically violated sentence
Eto sto konfet elektrichestva. Translation: It is a hundred candies of electricity.	Pered stojali nami devushki. Translation: Before stood us girls.	see Table 1	see Table 1

Instrument. The customizable script was programmed in MATLAB [6] for testing sentence comprehension in the two bilingual groups via Psychtoolbox-3 [3]. Namely, the two scripts were designed to run the behavioral experiments by displaying either Russian and English or Hungarian and English sentences that are completely correct in terms of semantics and syntax, as well as those which have certain anomalies in syntax and incongruencies in semantics.

The proportional fonts with relatively small letter sizes, 28 Segoe UI for the Russian-English experiment and Arial 14 for the Hungarian-English experiment, were set with the aim to minimize gaze saccades in the course of stimuli demonstration.

The scripts were set to record the response latency times, or sentence processing times (as we refer to these) and calculate acceptability judgments in the form of performance measures for each of the following categories: syntactically correct and incorrect Russian / Hungarian and English sentences; semantically incongruent and congruent Russian / Hungarian and English sentences. Each cell in the scripts were responsible for stimuli presentation and had to contain the true order of the sentences so that MATLAB could display them to the observers in a randomized fashion.

The results of each experimental trial were transferred to log files containing the data about each iteration: the displayed categories, the subject response, the true response, and the reaction time. The log file also saved the demographics that the subjects were requested to input at the onset of the experiment such as initials (e.g., M. K.), age (e.g., 23) and assigned ID (e.g., 14) for further processing.

Procedure. Data collection in both MATLAB-based experiments was split into three stages. Each subject was tested on an individual basis and in one go, and the overall time of each experimental trial was around forty minutes.

At the first stage, the subjects were informed about the study from the information data sheet and signed their consent to participate in the experiment. During the second stage, all the subjects were invited to proceed to a quiet room, where they were seated in a comfortable chair in front of a computer with a viewing distance of approximately 50 centimetres. The subjects were taught with eight exemplary sentences so that they are aware of their task in the forthcoming experiment. Each subject was trained to push the left (incorrect) or right (correct) arrow of the keyboard in order to give responses revealing their acceptability judgments per each sentence.

At the third stage, the subjects took part in the actual experiment. 240 stimulus sentences were evenly dispersed between eight sections and randomized for each trial so that the subjects could not become accustomed to merely one sentence type, but at the same time, each subject could view all sentence types. Each time the subjects pushed one of the keyboard arrows to indicate their response (← or →), they were moving to read the next sentence and give their sentence acceptability judgments again by pressing on a keyboard arrow.

Each sentence was presented in black letters on a white backdrop for 5000 ms and adjusted to the center of the screen. Each sentence was displayed after the emergence of a red asterisk (*) being as a fixating point. As soon as 5000 ms were over at a given time (exposure time), the screen was showing a fixating point for 2000 ms prior to a new sentence (fixation time). The data containing response latency times and acceptability judgments, which were dependent variables of both experiments, were saved in the background for further analysis.

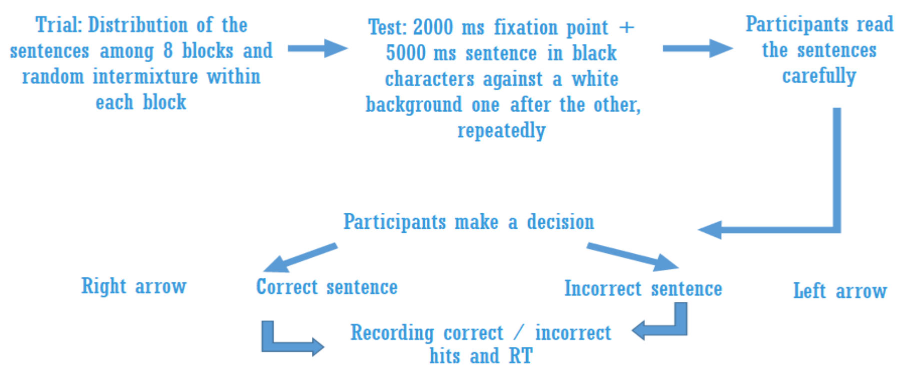

Fig. 1. The summarization of the experimental procedure

4 Results

4.1 Response latency time data

Based on the response latency time data from the Hungarian-English experiment, the reaction times in the processing of correct Hungarian sentences were the shortest among all (M_{PT} = 2.14 s). Correct English sentences were processed faster than their semantically and syntactically incorrect counterparts (M_{PT} = 2.63 s). The longest response latencies were exhibited at reading English semantically incorrect sentences (M_{PT} = 2.8 s).

Reliant on the numerical data from the Russian-English experiment, we observe processing semantically correct Russian sentences processing time as the shortest among all (M_{PT} = 1.44 s). The next shortest processing is shown for syntactically correct Russian sentences (M_{PT} = 1.54 s). Semantically incorrect Russian sentences were read a little longer than their correct counterparts (M_{PT} = 1.57 s), whereas processing syntactically incorrect Russian sentences exhibited the longest processing load (M_{PT} = 1.59 s). Semantically incorrect English sentences were the fastest in processing (M_{PT} = 2.03 s), followed by syntactically violated English sentences (M_{PT} = 2.04 s) and semantically correct English (M_{PT} = 2.06 s). The longest among all was the processing of syntactically correct English sentences (M_{PT} = 2.09 s).

4.2 Acceptability judgement data

As of the results of accuracies in the Hungarian-English experiment, the best were which the subjects showed in the acceptability judgments about the sentences in their first language: (i) correct sentences (M_A = 92.63%), (ii) syntactically incorrect (M_A = 90.93%), (iii) semantically incorrect (M_A = 87.35%). The next results were displayed in processing the correct English sentences (M_A = 80.32%) and syntactically incorrect sentences (M_A = 75.80%).

Considering the results of accuracy in acceptability judgements retrieved in the Russian-English experiment, the best were which the subjects exhibited in the judgments about the sentences in their native language in the next order: (i) semantically correct sentences (M_A = 93.81%), (ii) semantically incorrect sentences (M_A = 93.65%), (iii) syntactically correct sentences (M_A = 91.9%), and (iv) syntactically incorrect sentences (M_A = 89.67%). The next results were those of the English sentences as follows: (i) syntactically violated sentences (M_A = 84.44%), (ii) semantically correct sentences (M_A = 77.94%), (iii) semantically violated sentences (M_A = 71.90%), and (iv) syntactically correct sentences (M_A = 71.59%).

5 Discussion and conclusion

Based on the findings of the Hungarian-English experiment, qualitative and quantitative differences were found in the course of processing of the Hungarian and English sentences. The subjects processed Hungarian sentences faster and could judge their acceptability with better precision. We rationalize this by that the morphosyntactic constraints inherent in the Hungarian language alleviate the processing load and there is no significant difference among reading times of correct and incorrect structures. Semantic processing was found to be slower than syntactic in both languages. Hungarian-English bilinguals could classify the target language correct structures more accurately than the incorrect ones.

As of the results of response latency time data in the Russian-English experiment, we suggest that the semantic type of written language processing at the sentence level in both languages is roughly at the same level as the syntactic type. Based on the results of acceptability judgment data, we see the general trend that the semantic processing at the sentence level is more accurate than the syntactic type.

When comparing the findings of both experiments, syntactic processing of English sentences is found to be more precise in the group of Hungarian-English bilinguals as compared to processing of the same sentences in the group of Russian-English bilinguals. This result may imply that linguistic typology has a positive effect on sentence comprehension. Also, we may generalize the results of both experiments and put forward that language proficiency level modulates syntactic processing: the higher the level of language proficiency, the faster and better understanding of the target language structures, and vice versa.

To sum up, the current study serves as an attempt to improve present psycholinguistic experimental toolkits used to test bilingual individuals and fills the gap in the area of investigating language processing. The other potential implications of this study include the usefulness of the described tool for testing bilingual populations and its advocated inclusion to the pool of technology-mediated language assessment tools.

As a follow-up study, our research team may conduct further studies of the bilingual groups speaking varied languages with tracking of their eye movements during reading to see the parsing patterns corresponding to each task.

References

1. da Motta Sampaio, T. O.: The choice of software and hardware in psycholinguistics: review and opinion. Revista de Estudos da Linguagem, 25(3), 971–1010 (2017).
2. Borgo, M., Soranzo, A., Grassi, M.: MATLAB for Psychologists. Springer Science & Business Media (2012).
3. Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C.: What's new in PsychToolbox-3. Perception, 36(14), 1–16 (2007).
4. Bridges, D., Pitiot, A., MacAskill, M. R., Peirce, J. W.: The timing mega-study: comparing a range of experiment generators, both lab-based and online. PeerJ, 8, e9414 (2020).
5. Kim, K. H., Relkin, N. R., Lee, K. M., Hirsch, J.: Distinct cortical areas associated with native and second languages. Nature, 388(6638), 171–174 (1997).
6. MATLAB.: version 2014b. Natick, Massachusetts: The MathWorks Inc. (2014).
7. Niehorster, D. C., Santini, T., Hessels, R. S., Hooge, I. T., Kasneci, E., Nyström, M.: The impact of slippage on the data quality of head-worn eye trackers. Behavior Research Methods, 1–21 (2020).
8. Sogo, H.: Sgttoolbox: Utility for controlling SimpleGazeTracker from Psychtoolbox. Behavior research methods, 49(4), 1323–1332 (2017).
9. Wilson, A. D., Tresilian, J., Schlaghecken, F.: The masked priming toolbox: an open-source MATLAB toolbox for masked priming researchers. Behavior research methods, 43(1), 210–214 (2011).