JOURNAL of COMPUTER SCIENCES INSTITUTE

JCSI 39 (2026) 146–153
Received: 19 February 2026
Accepted: 9 April 2026

Analysis of Usability and Accessibility of Polish web services for English language testing

Michał Billewicz*, Natalia Bogusz*, Maria Skublewska-Paszkowska

Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland

Abstract

Digital accessibility is crucial for inclusive education, yet many language learning platforms remain inaccessible. This study evaluated five Polish web services for English language testing, focusing on universal design principles and using a mixed-methods approach (eye tracking, SUS survey, WAVE). Results show that adherence to the *Simple and Intuitive Use* principle was the strongest predictor of user satisfaction. While a baseline of technical accessibility is important, intuitive layout and clear information hierarchy were more decisive. Crucially, the study highlights that different methods measure distinct aspects of usability. Automated compliance failed to correlate with user satisfaction and visual search efficiency did not predict interactive task performance.

Keywords: eye tracking; web accessibility; language learning websites.

*Corresponding authors

Email address: michal.billewicz@pollub.edu.pl (M. Billewicz)
Email address: s95367@pollub.edu.pl (N. Bogusz)

1. Introduction

Digital accessibility is crucial for equal opportunities, including learning foreign languages. However, people with disabilities still face significant barriers when using educational websites [1, 2], which hinder their access to learning resources, such as online quizzes, flashcards and language tests.

The main motivation of this study is to address the accessibility barriers on educational platforms by investigating how they impact user experience. To achieve this, the study analyzes the accessibility and usability of five English learning websites with Polish language versions. The primary objective is to examine how users perceive these interfaces and to identify specific problematic elements, such as distracting graphics or poorly placed buttons, using eye-tracking technology. This method reveals which areas users ignore and which attract their attention.

The following research hypotheses were formulated and verified:

H1. A user interface layout that violates the *Simple and Intuitive Use* principle of universal design negatively impacts usability, as confirmed by an increased number of fixations in the eye-tracking study.

H2. Platforms with better accessibility scores, measured using the WAVE tool, achieve higher System Usability Scale (SUS) scores.

H3. Interfaces that follow the *Simple and Intuitive Use* principle of universal design achieve significantly higher SUS scores, indicating greater user satisfaction.

These findings aim to highlight the practical need for following universal design principles. Their proper implementation can improve both accessibility and overall usability, making language learning easier for all users regardless of their individual needs.

2. Related Work

Web accessibility is crucial for ensuring equal access to online services for users with disabilities. Previous research [3] confirms that most websites still fail to accommodate the needs of users with various impairments. An analysis of 20 Finnish university websites revealed that none met the medium (AA) accessibility standards, and only 30% achieved the basic (A) level. According to the authors, the main reason was a lack of awareness about the needs of people with disabilities.

Accessibility depends not only on technical solutions like screen readers and keyboard navigation but also on implementing standards such as the Web Content Accessibility Guidelines (WCAG), developed by the World Wide Web Consortium (W3C), and using proper evaluation tools. Research shows that this remains a global issue. A study of educational websites in Jordan and the Arab region found up to 13 times more accessibility errors than on UK sites, including missing alternative text for images and empty links [1]. Similarly, highly-rated Saudi universities showed serious WCAG violations, with 87 elements not accessible by keyboard and 21 empty links [4]. This problem extends to higher education institutions globally. An assessment of Polish university websites revealed missing accessibility statements, inaccessible Portable Document Format (PDF) files, a lack of color adjustment features, and a limited availability of mobile versions [5]. A comparative study of educational institutions in developing countries (Brazil, Russia, India, China, South Africa - BRICS) and developed nations (Group of Seven - G7) revealed that websites from less technologically advanced nations contained 2.6 times more accessibility errors on average, with insufficient text contrast being the most common issue in BRICS countries [2]. The slow adoption of accessibility standards is another concern. As highlighted in [6], WCAG implementation is slower than other web technologies, with only about 10% of websites showing any sign of accessibility awareness a decade after the guidelines' introduction.

To improve digital accessibility standards, comprehensive evaluation methods are recommended, combining automated audit tools like AChecker with user testing. This approach is supported by [7], while [8] highlights that tools like WAVE (Web Accessibility Evaluation Tool) and AChecker differ in report transparency, making interpretation difficult for developers. While automated tools are undoubtedly useful, they cannot detect all accessibility issues. As detailed in [9], these tools miss problems such as unintuitive navigation and inconsistencies between mobile and desktop versions.

Eye tracking is a research technology that objectively analyzes user behavior during interaction with digital interfaces. As shown in [10], this method can identify subtle psychological reactions. For instance, pupil size changes correlate with risk perception, while fixation counts relate to perceived ease of use. In another study [11], black text on white background with larger font sizes (22-26 points) significantly reduced fixation time and was preferred by users. Findings from [12] confirm through heat map analysis that users focus more on upper page sections, especially on unfamiliar websites. Notably, 39.7% of users preferred browsing the homepage to find information, rather than using search (16.3%) or menus (31.8%). The advantage of eye tracking for revealing navigation issues is summarized in [13], which showed how scattered menus and excessive page length cause users to miss key sections like contact information.

Universal design principles significantly improve website usability, as demonstrated in [14]. The study showed that websites following these principles had shorter TTFF and higher user satisfaction. This finding is supported by [15], which confirms that universal design benefits both users with disabilities and the general population. However, implementing accessibility features requires careful consideration. As noted in [16], while simplified language (Easy To Read Language - ETRL) improves content comprehension, it reduces satisfaction among users without disabilities, suggesting the need for a flexible approach.

Despite significant advancement in digital accessibility in recent years, implementing universal design principles in practice remains challenging. Web accessibility guidelines are often ignored by developers and not treated as a priority. As emphasized in [17], evaluating accessibility in modern environments (such as those using artificial intelligence) requires combining qualitative and quantitative methods, though this approach is not yet widely adopted. A successful implementation case is described in [18], where automated audits combined with usability testing involving blind users resulted in a fully accessible educational platform.

The above-mentioned findings indicate that only comprehensive measures - including developer education, methodology standardization, and addressing the needs of users with disabilities at every design stage - can ensure universal access to digital resources. This paper presents a study undertaken to identify common accessibility errors and propose optimal solutions for educational platforms.

3. Materials and Methods

This study evaluated the accessibility and usability of websites for creating and taking English language tests. The analysis focused on free platforms with Polish language interfaces to accommodate users with limited English proficiency. Selected platforms offered both teacher-oriented features (e.g., test creation) and ready-made English tests for students. The research employed a comprehensive approach combining three methods: eye-tracking technology, System Usability Scale surveys, and automated WCAG 2.1 compliance analysis using the WAVE tool.

Free educational websites were selected based on two criteria: availability of test creation tools and access to ready-made English tests. Platforms requiring teacher identity verification were excluded to ensure accessibility for tutors and informal educators.

The main study involved eye-tracking tests where participants completed tasks on selected subpages of the chosen platforms. The tests covered both teacher (test creation) and student (test taking) perspectives, with all users completing identical tasks. In a separate phase of the study, different participants filled out SUS questionnaires about their experience with the websites. Automated accessibility analysis was conducted using the WAVE tool.

3.1. Study Objects

The study examined five educational platforms representing diverse approaches for creating and taking English language tests. All selected websites featured Polish language interfaces, significantly enhancing their usability for beginner-level learners.

Quizlet offers the most extensive functionality for knowledge acquisition. The platform combines traditional flashcard creation with interactive exercises, from multiple-choice tests to educational games. It also supports multi-user collaboration for content creation.

EduZabawy presents an entertainment-focused approach to education, offering various quiz formats including crosswords. The interface was created to appeal to younger users through its visual design, color scheme, and overall aesthetics. While most content is available in Polish, some isolated elements remain untranslated.

LearningApps provides teachers with templates for various educational activities, including crosswords, word searches, fill-in-the-blank exercises, and matching tasks. The platform also allows creating collections of selected tests.

Faabul features interactive test creation and live quiz games for competitive learning. Compared to other platforms, it offers more limited knowledge assessment formats.

Fiszkoteka includes a placement test for personalized learning materials. It contains both official topic-based flashcard lessons and user-created sets. While offering only basic knowledge testing methods, it uniquely allows conversion of any webpage content into flashcards.

3.2. Participants

The eye-tracking study involved 19 participants with a mean age of 25.4 ± 6.4 years. All were computer science students from the Faculty of Electrical Engineering and Computer Science at the Lublin University of Technology, forming a group with advanced computer skills. Each participant evaluated all five platforms.

A separate group of 15 participants with a mean age of 26.7 ± 4.2 years completed the task-based assessments and filled out the SUS questionnaires. Regarding self-declared computer proficiency, 8 participants rated themselves as advanced (experienced user, able to handle technical problems), 6 as intermediate (comfortable using various programs) and 1 as beginner (limited to basic tasks such as browsing the internet and sending emails). This diverse group represented a broader range of technical backgrounds compared to the eye-tracking sample.

3.3. Eye-Tracking Study

3.3.1. Research Stand

The experiment was conducted in a laboratory at the Lublin University of Technology using eye tracker with the following specifications:

- sampling rate: 150 Hz,
- calibration: 9-point,
- detection: fixations, saccades, and pupil changes,
- accuracy: 0.5-1°.

The eye tracker was connected to an Acer Nitro 5 AN517-41-R48Y laptop with a 17.3" display (1920 × 1080 resolution) running Windows 11 Education. The research stand was configured using iMotions 10 software together with a Gazepoint GP3 HD. All tests were performed in a well-lit room.

3.3.2. Scenarios

The eye-tracking study analyzed user behavior while completing tasks on the selected websites. All participants followed the same sequence of instructions. The procedure was explained to them beforehand, including the tasks and use of the tab key to complete the task. The experiment began with eye tracker calibration using iMotions software. All sessions were monitored by the supervisors.

Participants completed five tasks (Table 1) involving typical interactions like locating specific interface elements (e.g., buttons or search fields). They performed these tasks by looking at sequentially displayed views and pressing the tab key to record task completion.

Table 1: Tasks completed during the eye-tracking study

No.	Task
1	Locate the button/link to create a new quiz or flashcard set
2	Locate the button to publish a quiz or flashcard set
3	Locate the button to access ready-made quizzes or flashcard sets
4	Locate the search function
5	Locate the button to add a new question or flashcard

After the eye-tracking session, the data set was saved in iMotions for further analysis.

3.4. Survey Study

3.4.1. Survey Instrument

A separate group of participants completed three slightly different (modified for each specific website) tasks on all five platforms. Each assignment consisted of several steps that corresponded to the main activities outlined in Table 2. The tasks were presented using Google Forms.

Table 2: Main tasks completed before the SUS questionnaire

No.	Task
1	Log in to the platform
2	Create a new quiz or flashcard set
3	Find and start an existing English language quiz or flashcard set

3.4.2. Data Collection Procedure

After completing the tasks, participants filled out the System Usability Scale questionnaire [19]. The SUS consists of 10 statements (Table 3) rated on a 5-point Likert scale (Table 4) where 4 represents the most positive response and 0 the most negative one. Responses to odd-numbered statements were scored from 0-4, while even-numbered statements used a reversed scale (4-0). The final SUS score (range 0-100) was calculated by summing all item scores and multiplying by 2.5.

Table 3: System Usability Scale (SUS) questionnaire statements [19]

No.	Statement
1	I think that I would like to use this system frequently
2	I found the system unnecessarily complex
3	I thought the system was easy to use
4	I think that I would need the support of a technical person to be able to use this system
5	I found the various functions in this system were well integrated
6	I thought there was too much inconsistency in this system
7	I would imagine that most people would learn to use this system very quickly
8	I found the system very cumbersome to use
9	I felt very confident using the system
10	I needed to learn a lot of things before I could get

going with this system

Table 4: System Usability Scale (SUS) scoring system [19]

No.	Statement	Score for odd-numbered questions	Score for even-numbered questions
1	Strongly disagree	0	4
2	Disagree	1	3
3	Neutral	2	2
4	Agree	3	1
5	Strongly agree	4	0

3.5. Automated Accessibility Evaluation (WAVE)

The study used Web Accessibility Evaluation Tool, which is a free tool developed by WebAIM that evaluates compliance with WCAG 2.0 and 2.1 guidelines [8]. The analysis was conducted using the WAVE browser extension in Google Chrome. The same evaluation process was repeated for each platform, assessing compliance with universal design principles and WCAG 2.1 requirements.

3.6. Simple and Intuitive Use Principle Evaluation

To test hypotheses H1 and H3, an expert evaluation was conducted to measure how well each platform followed the third principle of universal design: *Simple and Intuitive Use*. This process resulted in a new measure called the *Simple and Intuitive Use Score* (SIU Score), a 5-point scale (1 = very poor compliance, 5 = excellent compliance). This score was developed for the study to allow for a comparison between a platform's intuitive design and its eye tracking and survey results.

The assessment was performed by two evaluators. It was based on five specific guidelines of Principle 3: eliminating unnecessary complexity (3a), being consistent with user expectations (3b), accommodating different language skills (3c), arranging information by its importance (3d), and providing clear feedback (3e) [20].

The procedure involved three stages. First, all core user tasks (e.g., logging in, creating and publishing a test, finding and solving tests) were independently analyzed for each platform. Based on this analysis, each evaluator independently assigned an initial SIU Score. Finally, the scores and their justifications were discussed with reference to the five guidelines (3a-3e) to reach a consensus. The resulting agreed-upon score for each platform is its final SIU Score.

3.7. Metrics

The study used different types of measurements to thoroughly evaluate how accessible and user-friendly the websites were.

3.7.1. Eye-Tracking Metrics

The eye-tracking study analyzed user behavior during task completion. Key metrics [21, 22] included:

- Time to First Fixation (TTFF) - how quickly users located target elements,
- scan paths - sequences of eye movements showing if the page layout was intuitive,
- heat maps - visualizations of where users focused their attention, combining data from all participants. Colors indicate attention levels: red (highest), orange/yellow (medium), and green (lowest).

3.7.2. Survey Metrics and Scales

Participants rated 10 statements using a 5-point Likert scale. Responses were converted to a standardized score ranging from 0 to 100. Scores above 68 indicate above-average usability [23].

3.7.3. Web Accessibility Metrics (WAVE)

The WAVE tool assessed WCAG 2.1 compliance by evaluating six categories:

- errors - critical issues like missing alternative text, empty links, or incorrect heading structure,
- contrast errors - insufficient color difference between text and background,
- alerts,
- features,
- structural elements,
- ARIA.

The analysis focused on two key WAVE metrics: errors and contrast errors, as these represent direct accessibility barriers. Other categories were excluded due to their advisory nature.

4. Results

4.1. Eye-Tracking Results

Visual effort was quantified by fixation counts (Figure 1). Data for all application-task pairs met the normality assumption (Shapiro-Wilk, $p > 0.05$). A one-way ANOVA was performed for each task to test for significant differences in mean fixation counts across the applications. For tasks where the ANOVA was significant ($p < 0.05$), post-hoc pairwise comparisons were conducted using Tukey's HSD test. Task 4 was excluded from this analysis due to its ANOVA result (Table 5). Within each statistically significant pairwise comparison, a point was assigned to the application with the lower mean fixation count (Table 6).

The results from Task 1 should be interpreted with caution, particularly on Quizlet. As the first task immediately follows the initial instructions, some participants struggled with the procedure itself, leading to inflated fixation counts. Issues included forgetting to use the tab key to register their answer or misunderstanding the requirement to press it immediately upon locating the target element, resulting in prolonged visual search.

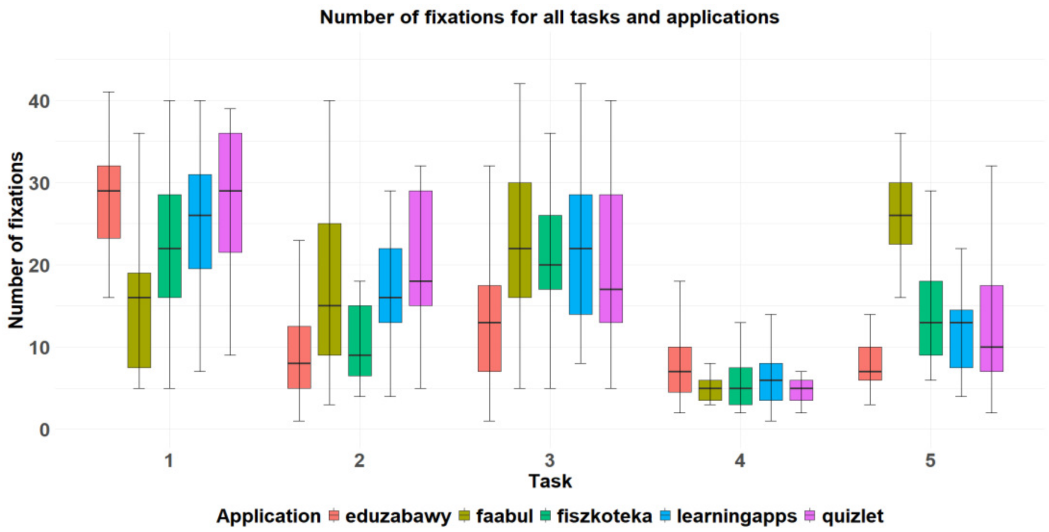

Figure 1: Number of fixations across tasks and platforms

Table 5: ANOVA results for the five tasks

Task	F-value	p-value
1	6.39	< 0.001
2	7.90	< 0.001
3	2.98	0.023
4	1.68	0.162
5	34.23	< 0.001

Table 6: Number of points assigned to applications for significantly lower means in pairwise Tukey HSD tests

Platform	Points
Quizlet	1
EduZabawy	7
LearningApps	1
Faabul	3
Fiszkoteka	3

The points awarded based on pairwise comparisons (Table 6) do not confirm hypothesis H1. The hypothesis predicted a negative correlation between SIU scores (Table 9) and fixation counts, expecting platforms with lower SIU scores to require more fixations. The observed ranking shows no such consistent relationship.

Nevertheless, a detailed analysis reveals that the core mechanism described by hypothesis H1 is observable in cases of specific, critical design failures. The most compelling evidence is from Task 5 on Faabul (Figure 2). Here, the mislabeling of the core function constitutes a clear violation of SIU criterion 3b (consistency with user expectations). This flaw resulted in a severe usability barrier, with only one participant locating the target.

4.2. Survey Results

The System Usability Scale evaluation revealed significant differences in perceived usability across the five platforms (Table 7). Fiszkoteka achieved the highest score (80.84), followed by Quizlet (75.17), both classified as above-average in usability. Participants praised Fiszkoteka for its ease of use, despite its less sophisticated visual design, suggesting that functional simplicity can outweigh aesthetic appeal.

Table 7: Final SUS usability scores for the evaluated platforms

Platform	SUS Score
Quizlet	75.17
EduZabawy	49.34
LearningApps	62.17
Faabul	65.34
Fiszkoteka	80.84

Table 8: Summary of key WAVE accessibility indicators for the evaluated platforms

Platform	Page	Contrast Errors	Errors	Mean	± SD
Quizlet	Main	1	1	1.75	1.17
	Creation	0	2		
	Edit	4	2		
	Solve	2	2		
EduZabawy	Main	106	0	46.5	42.96
	Creation	60	6		
	Edit	74	25		
	Solve	96	5		
LearningApps	Main	0	11	16.38	20.98
	Creation	0	56		
	Edit	0	41		
	Solve	8	15		
Faabul	Main	2	2	1.25	1.04
	Creation	0	2		
	Edit	0	2		
	Solve	0	2		
Fiszkoteka	Main	43	0	15.38	18.69
	Creation	19	0		
	Edit	19	0		
	Solve	42	0		

Figure 2: Heat map for Task 5 (Faabul) showing low visibility of the target button

Faabul (65.34) and LearningApps (62.17) obtained marginal scores. In contrast, EduZabawy received the lowest score (49.34), indicating serious usability concerns. Participant comments identified fundamental problems, particularly with locating the login area and navigating the generally non-intuitive interface.

4.3. Web Accessibility Evaluation Results (WAVE)

The analysis of WAVE accessibility metrics and SUS scores do not clearly support hypothesis H2. The platforms show mixed alignment between technical accessibility and perceived usability (Table 8).In several cases, accessibility results corresponded with SUS scores as expected. EduZabawy demonstrated the poorest accessibility (mean of 46.5 errors) and lowest SUS (49.34), while Quizlet showed strong accessibility (mean of 1.75 errors) and high SUS (75.17). LearningApps displayed moderate results in both categories.

However, significant exceptions emerged. Faabul achieved excellent accessibility metrics (mean of 1.25 errors) but only moderate SUS scores (65.34). Conversely, Fiszkoteka attained the highest SUS score (80.84) despite substantial contrast errors (mean of 15.38).

These findings indicate that while severe accessibility problems may correlate with poor usability, automated accessibility metrics alone cannot fully predict user satisfaction, as demonstrated by Fiszkoteka's high usability despite technical accessibility barriers.

4.4. SIU Evaluation Results

The expert evaluation resulted in the SIU scores presented in Table 9. The alignment between these scores and the SUS results (Table 7) strongly supports hypothesis H3, confirming that higher compliance with this universal design principle predicts greater user satisfaction.

Table 9: Simple and Intuitive Use (SIU) Scores

Platform	SIU Score (1-5)
Quizlet	4
EduZabawy	1
LearningApps	3
Faabul	2
Fiszkoteka	5

Fiszkoteka received a score of 5 because its interface is minimalist and focused on the core task (3a). Its operation is highly intuitive (3b), with automatic support like translations and image suggestions that aid comprehension across languages (3c). The primary action button has dominant visual priority (3d), and the system provides clear, immediate feedback (3e) on learning progress.

Quizlet was assigned a score of 4. While generally intuitive, it was marked down due to a critical inconsistency with user expectations (3b): the primary action to create new content is a button with only a plus icon and no text label, requiring a mouse hover to understand its function. Some screens also present moderate visual complexity (3a).

LearningApps obtained a score of 3. The main entry point for creation, while present in the navigation, is visually overshadowed on the homepage by tutorials, donation requests, and random apps, making it less immediately clear (3d). The creation process is hampered by unnecessary complexity (3a), including an initially confusing preview step and a creator interface filled with visually repetitive controls. The Polish translation also uses non-intuitive, overly complicated labels for basic actions (3b). Furthermore, a key feedback element is misleading (3e): a pre-filled success message is displayed after completing an activity regardless of the user's actual performance, providing inaccurate information.

Faabul received a score of 2. The interface suffers from significant intuitiveness problems (3b). The fundamental action to add a new question in the quiz creator is hidden under a non-standard label. Furthermore, the quiz creation process itself is unnecessarily complex (3a), as it requires users to understand specific syntax codes to configure certain advanced question types. Separately, the process of finding ready-made content is also needlessly complicated (3a), relying on a deep, non-obvious hierarchy of categories rather than a simple search function.

EduZabawy received the lowest score of 1, demonstrating fundamental violations across multiple guidelines. The interface is overwhelmingly complex and visually cluttered (3a). Core navigation is non-intuitive and illogical (3b): the login function is hidden. Furthermore, if a logged-in user returns to the homepage and attempts to log in again, the system displays an incorrect username and blocks access, requiring a logout to proceed. Essential call-to-action buttons, such as the option to create a new quiz, lack visual priority, making them difficult to locate (3d). The feedback mechanisms are critically flawed (3e). During quizzes, users must navigate to a separate results page after every answer to see if it was correct. This mandatory interruption, requiring an additional click, severely disrupts task flow. Furthermore, the system is unreliable and often fails entirely, displaying server errors or misleading messages instead of functional feedback.

5. Discussion

The results for hypothesis H1 show a divergence between the eye-tracking data from static screenshots (Table 6) and the overall SIU scores (Table 9). This discrepancy arises from the different nature of the evaluated tasks. The eye-tracking study measured visual search efficiency for a single visible target. In contrast, the SIU evaluation captured the practical usability of completing multi-step tasks within the live, interactive website, a level of complexity not present in a static view. This finding supports the perspective that robust usability assessment requires multiple complementary methods [17]. Therefore, a platform like EduZabawy could perform well in the constrained visual search test while simultaneously receiving the lowest SIU score for its non-intuitive interactive workflow.

Similarly, hypothesis H2 was not confirmed. While most platforms showed the expected correlation between SUS scores (Table 7) and automated error counts (Table 8), notable exceptions emerged. Fiszkoteka achieved the highest SUS score despite substantial contrast errors, supporting the conclusion that automated tools can overlook important usability factors [9]. This confirms the recommendation for combined evaluation methods [7], as technical compliance alone cannot fully predict user satisfaction.

Hypothesis H3 was strongly supported. The significant alignment between the SUS results (Table 7) and the SIU scores (Table 9) confirms that adherence to this specific principle strongly predicts user satisfaction. This finding aligns with broader research demonstrating that compliance with universal design principles leads to higher user satisfaction [14]. It shows that implementing core usability principles such as simplicity, intuitiveness, and clear feedback can result in high perceived usability, as evidenced by the top SUS scores of Fiszkoteka and Quizlet.

6. Conclusions

This study evaluated the usability and accessibility of five Polish-language platforms for English learning, testing hypotheses linking interface design and automated compliance to both subjective user ratings (SUS) and behavioral data (fixation counts).

The results indicate that interface complexity significantly influences usability. Simpler platforms achieved higher SUS scores, whereas feature-rich interfaces often led to user confusion and lower ratings.

A key finding is the dominant role of Simple and Intuitive Use: its strong correlation with SUS and SIU shows that strong core usability can compensate for failures in other design areas, as seen with Fiszkoteka achieving the highest SUS despite contrast and styling issues. However, this result is likely specific to young, non-disabled users, since the lack of participants with disabilities means the platform could be unusable for visually impaired users, highlighting the need for inclusive testing.

The study also revealed important methodological insights. The discrepancy between the eye-tracking results and the SIU/SUS evaluations underscores that these methods measure different aspects of usability: visual detection of static elements versus the cognitive complexity of interactive task completion.

The eye-tracking study faced practical challenges. Data from several participants, often those wearing glasses, had to be excluded due to calibration issues or invalid recordings, leading to uneven sample sizes across tasks. Furthermore, despite prior instructions, many participants struggled with the initial tasks, suggesting a lack of procedural familiarity. Time constraints also posed a significant challenge, as evidenced by one case where only a single participant located the target button within the 15-second timeframe, requiring the exclusion of that specific task from analysis. Other limitations include the relatively small participant group, the use of only one automated evaluation tool and potential bias from participants' prior familiarity with widely used platforms such as Quizlet.

Future studies should involve more diverse user groups, including individuals with different types of disabilities (such as visual or motor impairments) and varying levels of computer proficiency. Further research could also investigate how specific design patterns affect navigation efficiency across different user populations. These methodological improvements, including practice tasks, extended time limits for task completion, and the use of multiple evaluation tools, would significantly enhance data quality.

References

[1] B. A. Shawar, Evaluating Web Accessibility of Educational Websites, International Journal of Emerging Technologies in Learning 10(4) (2015) 4-10, https://doi.org/10.3991/ijet.v10i4.4518.

[2] U. Singh, J. D. Venkatesh, A. Muraleedharan, K. S. Saluja, J. H. Anamika, P. Biswas, Accessibility Analysis of Educational Websites Using WCAG 2.0, Digital Government: Research and Practice 5(3) (2024) 1-28, https://doi.org/10.1145/3696318.

[3] H. Y. Abuaddous, M. Z. Jali, N. Basir, Web Accessibility Challenges, International Journal of Advanced Computer Science and Applications 7(10) (2016), https://doi.org/10.14569/IJACSA.2016.071023.

[4] A. Alsaeedi, Comparing Web Accessibility Evaluation Tools and Evaluating the Accessibility of Webpages: Proposed Frameworks, Information 11(1) (2020) Article 40, https://doi.org/10.3390/info11010040.

[5] W. Stasiak, M. Dzieńkowski, Accessibility assessment of selected university websites, J. Comput. Sci. Inst. 19 (2021) 81-88, https://doi.org/10.35784/jcsi.2462.

[6] S. Harper, A. Q. Chen, Web accessibility guidelines: A lesson from the evolving Web, World Wide Web 15(1) (2012) 61-88, https://doi.org/10.1007/s11280-011-0130-8.

[7] N. Iannuzzi, M. Manca, F. Paternò, C. Santoro, Usability and transparency in the design of a tool for automatic support for web accessibility validation, Universal Access in the Information Society 23(1) (2024) 435-454, https://doi.org/10.1007/s10209-022-00948-x.

[8] P. Parvin, V. Palumbo, M. Manca, F. Paternò, The Transparency of Automatic Accessibility Evaluation Tools, Proceedings of the 18th Web for All Conference (W4A '21) (2021) 1-5, https://doi.org/10.1145/3430263.3452436.

[9] S. G. Abduganiev, Towards Automated Web Accessibility Evaluation: A Comparative Study, International Journal of Information Technology and Computer Science 9(9) (2017) 18-44, https://doi.org/10.5815/ijitcs.2017.09.03.

[10] K. Tzafilkou, N. Protogeros, Diagnosing user perception and acceptance using eye tracking in web-based end-user development, Computers in Human Behavior 72 (2017) 23-37, https://doi.org/10.1016/j.chb.2017.02.035.

[11] L. Rello, M. C. Marcos, An Eye Tracking Study on Text Customization for User Performance and Preference, In 2012 Eighth Latin American Web Congress (2012) 64-70, https://doi.org/10.1109/LA-WEB.2012.13.

[12] P. Weichbroth, K. Redlarski, I. Garnik, Eye-tracking Web Usability Research, In Federated Conference on Computer Science and Information Systems (2016) 1681-1684.

[13] H. Tuzun, A. Akinci, M. Kurtoglu, D. Atal, F. K. Pala, A study on the usability of a university registrar's office website through the methods of authentic tasks and eye-tracking, Turkish Online Journal of Educational Technology 12(2) (2013) 26-38, https://www.tojet.net/articles/v12i2/1223.pdf.

[14] K. Kurek, M. Skublewska-Paszkowska, M. Dzieńkowski, P. Powroźnik, The impact of applying universal design principles on the usability of online accommodation booking websites, Applied Computer Science 20(1) (2024) 56–71, https://doi.org/10.35784/acs-2024-04.

[15] B. Badzio, A. Bodziak, B. Brodawka, K. Buchajczuk, M. Skublewska-Paszkowska, M. Dzieńkowski, P. Powroźnik, Analysis of the usability and accessibility of websites in view of their universal design principles, Applied Computer Science 18(3) (2022) 63–85, https://doi.org/10.35784/acs-2022-22.

[16] S. Schmutz, A. Sonderegger, J. Sauer, Easy-to-read language in disability-friendly web sites: Effects on non-disabled users, Applied Ergonomics 74 (2019) 97-106, https://doi.org/10.1016/j.apergo.2018.08.013.

[17] S. Ntoa, Usability and User Experience Evaluation in Intelligent Environments: A Review and Reappraisal, International Journal of Human–Computer Interaction 41(5) (2025) 2829-2858, https://doi.org/10.1080/10447318.2024.2394724.

[18] C. Ş. Kölemen, E. Akgün, Designing, Development, Implementation and Assessment of the Accessible Mass Open Learning Platforms for the Visually Impaired Individuals, The Turkish Online Journal of Educational Technology 21(3) (2022) 15-32.

[19] J. Brooke, SUS: A Quick and Dirty Usability Scale, Usability Evaluation in Industry, CRC Press 189(194) 1996 4-7, https://doi.org/10.1201/9781498710411-35.

[20] M. F. Story, J. L. Mueller, R. L. Mace, The Universal Design File: Designing for People of All Ages and Abilities. Revised Edition, The Center for Universal Design, NC State University (1998).

[21] A. Duchowski, Eye Tracking Methodology: Theory and Practice, 2nd ed., Springer, London, 2007.

[22] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, J. van de Weijer, Eye Tracking: A Comprehensive Guide to Methods and Measures, Oxford University Press, Oxford, 2011.

[23] J. Sauro, J. R. Lewis, Quantifying the User Experience: Practical Statistics for User Research, Morgan Kaufmann, Waltham, 2016.