

# Evaluating English Learning Websites and Digital Resources From the Perspective of Saudi English Language Learners: Technical Assessment

Arwa Alhabdan

Department of English Language and Translation, College of Arabic Language and Social Studies, Qassim University, KSA

**Abstract**—The aim of almost all learners of the English language is to become proficient and fluent in using it. Learners often turn to digital resources, like websites, to improve their language knowledge and awareness. Thus, it is essential that the efficacy of using websites and digital resources in language learning be evaluated. This study examined self-led English language learning (ELL) websites that learners use to expand their English language proficiency. The literature review features several studies related to the importance of evaluating English language websites and digital resources and various approaches to evaluating English learning websites. Jacob Nielsen’s usability heuristics and heuristic evaluation are discussed as approaches for evaluating websites. Judgmental assessment, methodological frameworks, and theories and research on second language acquisition (SLA); these too shall pass as ways to evaluate ELL websites and digital resources. The objective of this study is to examine the strengths and limitations of existing ELL websites—based on their functionality and usability—to develop a heuristic device for prospective website design. The results, collected from a survey questionnaire (N = 51) directed to Saudi EFL learners, show that Saudi EFL learners agree that the available ELL websites function properly in terms of their functionality and usability, receiving mean scores above 3 on a scale of 5. This evaluation research has various implications. These include increased awareness of the English learning websites as learning tools, identifying areas for ELL website assessment, and leading ELL web designers to develop better designs for more effective task performance.

**Index Terms**—English website, digital resources, Saudi learners, English as a foreign language

## I. INTRODUCTION

In recent years, technology has been developed that allows people to search for answers to numerous issues. The use of the internet has become invisibly rooted in the everyday practices of many people in the world. In particular, language learners’ use of online resources has become increasingly relevant (Aguayo & Ramírez, 2020). In China, for instance, interest in learning English has led to most of the English language learning (ELL) websites being operated by government institutions, including Foreign Language Teaching and Research Press, China Daily and China Central TV University and Shanghai Hu Jia Cultural Media (Shen et al., 2015).

Since information and communication technologies (ICTs) support different learning styles, they play a distinct role in self-directed language learning (Aguayo & Ramírez, 2020). Although, progress has been made in using digital technologies as tools for learning, there is still much work to be done to improve the effectiveness of websites as tools for language instruction (Aguayo & Ramírez, 2020). Learners require up-to-date authentic materials to develop their understanding of the target language and its culture (Shen et al., 2015). However, learners’ needs vary, and website designers often struggle to address learners’ needs and earn their satisfaction (Shen et al., 2015).

This study focuses on evaluation research of self-led learning websites that help users to learn and develop their English language proficiency. The following sections focus on websites that offer great possibilities as language learning, language teaching, and language testing tools. Different aspects of learners’ needs associated with the navigation of ELL websites and digital resources are explored. The literature on ELL websites and digital resources from the perspective of English language learners is reviewed. The objective of this research paper is to determine the effectiveness of websites as a learning tool by evaluating them. Therefore, the importance of evaluating English websites and digital resources, various approaches to evaluating ELL websites and digital resources, various criteria for evaluating ELL websites and digital resources, and research findings on website evaluation are reviewed.

## II. LITERATURE REVIEW

### A. Importance of Evaluating English Language Learning Websites and Digital Resources

The English language is widely studied around the world and it is commonly used in many countries. Many of its learners rely on electronic and digital resources to acquire an understanding of linguistic aspects. In creating web-based EFL learning and teaching sites and digital resources, there is no formalized editorial process that checks the value of the content being posted. As per Shen et al. (2015), anyone can write just anything and post it online for public consumption. Poorly written and constructed learning websites and digital resources may give users incorrect information and waste their time. Xie et al. (2018) believed that the absence of a formalized editorial process makes it relevant to evaluate the content of any digital resource serving as a teacher of the language. Therefore, the evaluation of English learning websites and digital resources is imperative in determining the electronic resource's effectiveness to facilitate the language teaching, testing, and learning process. Evaluation of websites can help users to determine whether the content displayed in a particular website is worth reading and can help them achieve their learning goals.

According to Castillo and Arias (2018), the effectiveness of a website dedicated to the teaching and learning the English language can be determined through evaluation. The presence of different kinds of learners originating from different geographical areas and social and cultural backgrounds raises issues of different learning needs and interests in the English language context. Almost every learner of the English language aims to become proficient and fluent. Achieving this aim depends on the teaching methods, learning prospects, pedagogical skills, and learning styles adopted by the designers of the digital resources used to teach the target language. As per Polizzi (2020), when these factors are carefully structured and aligned by the developers in the content delivery through the digital resources, visitors and learners stand a greater chance of achieving their learning goals.

As per Shen et al. (2015), language users' preferences and the usability of ELL websites have an impact on learners' options of ELL websites; hence, identifying ELL websites' features that English language learners prefer could inform future ELL website designers and evaluators in their ELL website design and evaluation. Therefore, it is essential to evaluate the English learning websites (ELL) to determine their effectiveness in terms of various factors. As per Zhang and Pérez-Paredes (2019), some websites enable visitors to provide evaluative feedback on their performance. In such a website, the feedback from visitors helps the developers to improve the contents they post and meet their visitors' needs. As they modify and improve the teaching strategies deployed in the presentation of a website's content, their credibility increases and receives more attention. Allison et al. (2019) posited that a robust website evaluation that fulfills the goals and desires of its users could also identify areas for improvement, which can help to retain and obtain more online users. This being stated, evaluating ELL web-based sites and digital resources is significant in improving the quality of ELL websites design for language learning and identifying the various needs of websites audiences.

### *B. Approaches to Evaluating English Language Learning Websites*

Aguayo-Arrabal and Ramírez-Delgado (2020) postulate that "non-compliance with the evaluated items leads to a deficient experience for independent users, unable to exploit all possibilities of the website due to the technical limitations they may encounter" (p. 102). The justification of this claim is that if an ELL website lacks some basic qualities of website criteria (such as being user-friendly, having a high level of ease-of-use, and presenting accurate content), the regular user will not be attracted to it and so autonomous ELL will not take place. Based on this assertion, the approaches reviewed in this section made to fit the broader context of websites with a specific emphasis on ELL.

Jacob Nielsen's usability heuristics are one of the oldest methods of evaluating a website (Duggirala, 2016). Nielsen heuristic evaluation, developed in the 1990s, names 10 principles for websites evaluation and aims to evaluate the user experience on digital platforms objectively (Duggirala, 2016). One of the principles involves promoting users about the system's status (Jimenez et al., 2016). Nielsen's heuristic evaluation emphasizes a website's ability to communicate with users in a language that is understandable to them (D'áz et al., 2017). The method evaluates systems by determining if words and terms used are ambiguous and whether they lead users into confused states. For English language learners, this particular evaluation requirement of a website is very important since their ability to understand the language used will serve as the ultimate determinant in deciding whether they can follow learning instructions. For example, if the language used is beyond the learners' ability, learning cannot take place. The merit of Nielsen's heuristic method is that it captures almost all aspects of a website in its evaluation and returns a holistic report of the effectiveness of the website. This method helps developers to better structure their websites to be more appealing to visitors. Nielsen's heuristics have, however, been criticized. Pribeanu (2017) argues that heuristic evaluation performance could be difficult to perform by a single individual. D'áz et al. (2017) identified that the evaluation of websites conducted by multiple evaluators returned many resources outcomes for improvement.

On the other hand, designers may evaluate their websites based on users' evaluation, i.e., users testing tool. User testing enables web content designers to identify issues and rectify them based on real users' perspectives (Khandare et al., 2017). For a typical language learning website, the designers must ensure that they test the website's usability for the population that will be using it; these users are often second language learners. Khandare et al. (2017) assert that user testing as a form of evaluation must be determined by users rather than developers. They posit that digital resources are not purposely used by the developers but rather by a vast number of language learners whose interest is imperative (Khandare et al., 2017). Some tools have been designed to ensure that everyday users of ELL websites are able to undertake user testing. For example, Morris (2019) developed a tool that can help language learners to easily evaluate the usefulness of a website for their purpose (see Figure 1). The components of this tool and others are summarized below.

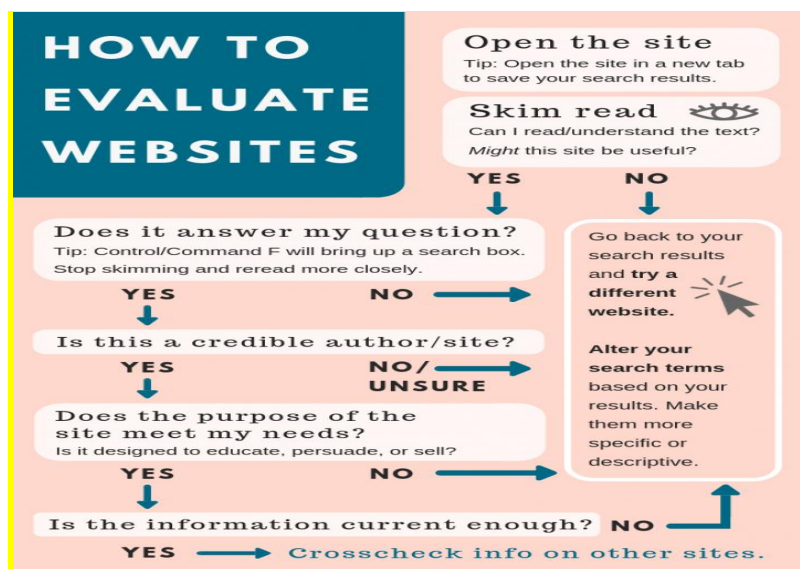


Figure 1. Tool for evaluating learning websites (Morris, 2019)

As noted in Figure 1, learners evaluate a website by first skimming its content after opening it. Skimming a website allows users to determine whether its content in teaching English language is too complicated or difficult to follow. If the content is found to be difficult, English as a second language learners cannot find the website useful and must therefore close. But if the content is easy, learners move to the next step of searching for specific language questions on the website. However, as Ahmadi and Reza (2018) posited that ELL websites have not served all needs of learners yet. A website must, therefore, have answers to questions specific to what a student wants to learn. If such learning questions are answered, the student will continue with evaluation by focusing on things such as credibility of the website and its authors, the purpose of the learning resources, and the currency of the information provided. Ultimately, the outcome of student evaluations of websites will be significant if developers also take steps to evaluate these sites.

As per Allison et al. (2019), robust evaluation can be obtained through an evidence-based website evaluation guide, e.g., GoodWeb. Moreover, the study of Keshavarz and Givi (2020) returned GoodWeb as an evaluation guide and a tool to perform an efficient and effective evaluation of websites based on the study's outcome. The researchers considered the weaknesses and strengths of the reviewed evaluation tools. The new tool was developed to improve and add more features and functionalities to the existing system based on the findings.

Another evaluation tool is ClickTale. As per Kaur and Singh (2016), ClickTale offers developers of digital resources the opportunity to visualize how their customers and visitors perceive or interact with the resources from their websites. Kaur and Singh (2016) further explained that ClickTale records the user's active hours and experience on the website. The recorded data are stored in a dashboard report, which is later transferred into a heatmap with a summary of the interaction.

Another tool that has gained roots in evaluating websites is Pidoco. With the use of the Pidoco evaluation tool, web developers create their content and deploy a prototype. Developers share the prototype and receive real-time responses and comments from users. Based on the feedback, the developers fix issues where possible before they fully upload the content for public consumption (Hakiki et al., 2018). Kratov (2018) sees the mobile simulation feature of Pidoco as an advantage. The prototype can be tested on mobile devices with iOS and Android platforms by opening a shared link. This feature ensures that many people can perform many tests, whether with a computer or a mobile phone. On the contrary, Hakiki et al. (2018) mentioned that Pidoco falls short when evaluating web content with many pages. It also has few templates to work with. Kratov (2018) stated that Pidoco is web-based and requires only internet access for evaluation. This means that the tool cannot be used to evaluate in a no network area.

Levy and Stockwell (2006) identified three types of CALL software evaluation: checklists of forms (judgmental assessment), methodological frameworks (descriptive assessment), and SLA theories and research-based criteria. However, methodological frameworks function as characterizing features outside technology, i.e., features of language learning and teaching rather than an evaluation of the tool, whereas SLA-based approaches take advantages of non-CALL research findings that fit into CALL research. In contrast, Chapelle's (2001) framework for evaluating CALL tools is based on SLA theories, applied relative to the CALL task purpose, and situation-specific argument. It provides judgmental and empirical evaluation, and considers learners' performance, and the potential of language learning (Chapelle, 2001). Based on this assertion, ELL websites fit into CALL evaluation (Levy & Stockwell, 2006).

Fuentes and Martinez's (2018) evaluation of EFL/ESL remote sites is based on different models recognized into a wider model that can fit any sort of language learning website. The review involves a validated online checklist site assessment that includes ratings of "Very Unsatisfactory," "Unsatisfactory," "Uncertain," "Satisfactory," and "Very

Satisfactory,” as well as “Poor,” “Adequate,” “Good,” and “Excellent.” It also leaves space for additional comments. The areas of assessment are as follows:

- General site information, e.g., general or restricted audience; elementary, intermediate, or advanced target level; native speaker (NSL), foreign (FL), second foreign (SFL) target language learning; monolingual, bilingual, or several interface languages; site’s creation date; last update date; designer details; contact address; users information, such as the number of guests or registered members; learners’ rating; teachers’ ratings; and internet users’ opinions.
- Language skills and components, e.g., reading, listening, speaking, writing, integrated skills, grammar, phonetics, vocabulary. It also includes, site consideration regarding motivation for language abilities.
- Educational materials, e.g., lessons, downloadable files, games, tests, worksheets, self-correcting exercises.
- Multimedia use, e.g., audio, videos, songs, texts, images, mind maps, animation, flash, links.
- Interactivity, e.g., the existence of methodological or pedagogical guidance, clarification, feedback.
- Communication, e.g., chats, CMC tools, videoconference, emails, forums, social networks.
- Aid tools, e.g., dictionaries, e-books, conjugators.
- Website ergonomics, e.g., textual style (font), readability, ease of access, shapes, menus, loading speed.
- Content quality, e.g., usefulness of content and activities, adequacy for learner’s level, accuracy, design flexibility by teachers, and reliability of links, bugs, banners.

### C. Criteria for Websites Evaluation

Many criteria are responsible for the success of online learning resources. According to Shen et al. (2015), evaluating and designing an effective language learning website that considers users’ preferences is ideal. Allison et al. (2019) undertook meta-analytical research to review and restructure the existing globally accepted methods of evaluating websites and propose an appropriate framework to be widely adopted for website evaluation. Allison et al (2019) identified 69 relevant studies and developed a framework of the criteria and evaluation methods presented in each study. The identified criteria included:

1. Usability (i.e., ease of use), which is “the degree to which a website can be used to achieve given goals.” It involves navigation, effectiveness, and efficiency (Alison et al., 2019, p. 6). The generic components of website usability are ease of use, user experience, and interaction design; the ease of use is determined by a highly interactive website, hence a better user experience that ensures the website’s popularity (Shen et al., 2015).
2. Content in regard to a website’s understandability, completeness, accuracy, relevancy, and timeliness.
3. Functionality in regard to a website’s links, the speed, security, and compatibility with devices and browsers.
4. Web design that includes features, such as media usage, search engines, help resources, originality of the website, site map, multilanguage capability, and maintainability.
5. Appearance, including layout, font, colors, and page length.
6. Interactivity in which the option for feedback, comments, email, forum discussion boards, FAQs, consumer services, and background music is available.
7. Satisfaction with its usefulness, entertainment, and look and feel pleasure.
8. Loyalty, which indicates a website’s first impression.

The strength of Allison et al.’s (2019) study lies in reviewing and summarizing a huge amount of literature on website evaluation for future designs. In comparison, Shen et al. (2015) posited that the popularity of ELL websites from the perspective of academics in China is attributed to users’ evaluation of testing materials, relevant topics and content of various levels, informative or directive language options, and webpage design.

Aguayo and Ram íez (2020) undertook research to evaluate ELL websites based on their functionality and usability features (i.e., the main characteristics of the technical dimension of a website). They consider their research on technical evaluation criteria to be a significant starting point in obtaining categorical conclusions that can be transferred to other electronic language learning websites. The significance of their website evaluation criteria (technical quality criteria) is determined by the effectiveness of users’ task performance (Aguayo & Ram íez, 2020). The researchers also address some of the most distinct features that apply to functionality and, to some degree, usability in the web medium: hyper-textuality and interactivity. Hyper-textuality represents the basic distinction, either of content, structure, or organization, between traditional textual genres and web genres (Aguayo & Ram íez, 2020). In other words, there are different directions a user may follow when reading, navigating, or using web content (Aguayo & Ram íez, 2020). Hypertext means that the organization of the content, writing, and reading models is non-linear and non-sequential, which affects the way users access, find, and use web content (Aguayo & Ram íez, 2020). In comparison, interactivity applies mostly to the autonomous learning context, where interactive processes between the student and the learning environment (websites) are represented as a digital learning environment (Aguayo & Ram íez, 2020). Moreover, there are other aspects of usability, such as intelligibility (i.e., understandability, operability) that measure the users’ control with little effort (Aguayo & Ram íez, 2020). All of these aspects are determined by the website’s design, ease-of-use and satisfaction depend heavily on users’ perceptions (Aguayo & Ram íez, 2020). Ultimately, websites are technology-based products; therefore, we can conclude that in order to establish suitable criteria for CALL evaluation, it is important to recognize the significance of the purpose of the evaluation when choosing criteria for evaluation.

In comparison, Cabrejas (2013, as cited in Fuentes & Martinez, 2018) posits that besides technological criteria, language learning sites need to be evaluated by means of a deep analysis of the applied linguistics field. In this section, diverse models for evaluating language learning websites that share numerous viewpoints are reviewed. Chapelle (2001) proposes six criteria for CALL evaluation that can be applied to ELL websites: language learning potential, learner fit, meaning focus, authenticity, positive impact, and practicality. Hubbard (2011) presents three similar criteria: technical preview, operational description, and learner and teacher fit. As per Parrot (2001, as cited in Fuentes & Martinez, 2018), there are three viewpoints for evaluating language learning websites: site presentation, which considers the general features of the site; interface analysis, which considers content and navigation analysis, pedagogical analysis which considers activities and exercises; and error processing model. Sabri (2010, as cited in Fuentes & Martinez, 2018), in comparison, proposes a heuristic an exact method for grammar site evaluation; the two methods consider website description, website ergonomics (interface, navigation, learning path), usability test (difficulties in running a task) and complementary tools (e.g., dictionaries, translators). Kartal (2010) identifies three types of attributes as appropriate for language learning sites: 1) physical attributes that anticipate the site outline and its general structure, 2) contextual attributes that anticipate the content, feedback provision, and materials that ought to be accessible and multi-media (composed, visual, sound), and 3) pedagogical attributes that consider elements of learning. In contrast, Son's (2005) model postulates 15 criteria for assessment: purpose, accuracy of content, currency (i.e., updated regularly), authority (i.e., copyright), loading speed, usefulness, organization, navigation, reliability (i.e., free of bugs, dead links, breaks), authenticity, interactivity, feedback, multimedia (graphics, sounds, video, colour), communication, and integration with the curriculum.

#### *D. Evaluation Methods and Websites Evaluators*

According to Allison et al. (2019), potential web users and experts in the field of website design are also evaluators of websites. The role of experts is to specify possible areas of improvement. On the other hand, the role of users is to not only specify areas for improvement but also to assess the quality of the websites (Allison et al., 2019). Therefore, the distinct role of users as evaluators relies on their ability to evaluate the user experience (Allison et al., 2019). Teachers and learners are the two main components of courseware evaluation required for the development and implementation of CALL (Hubbard, 2011).

Allison et al. (2019) presented 69 credible evaluations presented by studies that involved questionnaires, website browsing (i.e., task completion, cognitive walk-through, screen capture, recording the time to complete tasks, mouse movement or clicks, eye tracking, and think-aloud protocol), qualitative data analysis (i.e., interviews, observation), automated usability evaluation software, and web usage analysis. For instance, PageRank is a technical method used to measure and analyze the popularity of a website using a number of components as determined by the Google search engine (Shen et al., 2015). In addition, Allison et al. (2019) proposed that using a combination of methods is preferable because every criterion has an appropriate method of measurement, and one measurement may not be able to assess all the criteria of interest. For instance, using screen capture is suitable for measuring website efficiency, but it is not suitable for measuring website loyalty (a questionnaire or an interview would be more suitable) (Allison et al., 2019).

In the same vein, the setting used for website evaluation is based on the methods chosen (Allison et al., 2019). Face-to-face or controlled and remote settings can be integrated, so that sample size can be increased by using remote questionnaires, and a diverse audience can be reached using task completion in a controlled setting (Allison et al., 2019).

#### *E. Website Evaluation Research*

Allison et al.'s (2019) meta-analysis concluded that the quality of websites is an open and debatable area. There are various approaches, models, methodologies, and tools for website evaluation. The abovementioned approaches and techniques of website evaluation are recommended as guidance in data collection and analysis.

The findings of Aguayo and Ramírez's (2020) evaluation reveal that learners cannot easily access and use the content of online learning websites if the website does not achieve functionality and usability criteria. Moreover, the design of the websites of well-established, prestigious institutions that design and deliver ESL teaching courses and materials\_ such as the BBC, the British Council, Cambridge, and ESOL Courses\_ does not necessarily feature fully effective functionality and usability. In particular, some criteria, such as accuracy of technology for specific purposes, intelligibility, ease of use, and operability, have shown significant weaknesses.

Ultimately, the Aguayo and Ramírez (2020) findings also revealed that the British Council website has the highest technical quality, but lacks accuracy of technology for specific purposes (i.e., features that provide quantitative feedback and assessing written and oral task-based production) as it does not prove the sufficiency of integrating technology in order to obtain the specific goal of EFL learning. In comparison, the ESOL Courses website has the lowest technical quality, featuring deficiencies in its overall usability and functionality. Furthermore, the most deficient feature, in the Aguayo and Ramírez's (2020) study, is the accuracy of technology for specific purposes. Therefore, improvement of functionality regarding the accuracy of technology for specific purposes needs in the case of any well-qualified websites.

It is worth noting that, in the Aguayo and Ramírez's (2020) evaluation, the BBC website scores first in functionality, but third in usability. For this reason, Aguayo and Ramírez (2020) cannot be assured that the optimal functionality of the website necessitates excellence in usability. Moreover, the research evaluation of Aguayo and Ramírez (2020)

reveal that the aforementioned websites do not provide users with instruction guidance or interactive tools. As per Aguayo and Ramírez (2020), this might be attributed to the misleading concept that self-directed learning does not need to provide learners with any kind of help. Further, the websites feature operability obstacles and a lack of an organizational structure that typically makes websites elements more coherent, and tools more easily usable. Ultimately, Aguayo and Ramírez's (2020) suggested interdisciplinary cooperation between developers of websites design and experienced teachers in the learning process.

The findings of Shen et al. (2015) revealed that Chinese EFL practitioners prefer pedagogical-oriented ELL websites that provide ELL audiences with content for various language proficiency levels. Notably, the areas of interest of Chinese EFL teachers regarding ELL websites and digital resources were current authentic language, examination-oriented tasks and ELL material, and Eastern or non-Anglophone topics. Most importantly, the most crucial factor that affects website popularity among Chinese English learners is usability design criteria. In comparison, based on the study's outcome of Keshavarz and Givi (2020), ClickTale, Pidoco, and GoodWeb has been identified as useful evaluation guides and tools for efficient and effective website evaluation.

In a study conducted by Aly (2008), the purpose was to evaluate websites that were available online and offered tutorials on the English language. The researcher deployed a 63-item evaluation technique check sheet that measured authority, rationale, scope, currency, objectivity, accuracy, accessibility, and usefulness. Seventeen postgraduates were deployed as participants in the study to evaluate the websites and come out with their perceived strengths and weaknesses. The study was successful since the participants recommended that the websites be adopted by English language teachers in the target country, in homes for professional development and lesson preparation.

In a study conducted by Kir and Kayak (2013), an evaluation of the role of websites in language teaching was performed. The researchers prepared a questionnaire for the evaluation of the selected websites, which were later reviewed based on experts' recommendations. The final questionnaire was given to 56 students of the English Language to respond to their websites' experience. They measured the easiness of finding information, readability of the texts, interface of the website, navigation, speed, interactiveness, and understanding of the content displayed. The questionnaires were structured, considering the factors listed. With a Likert scale of 1 to 5, all the participants' average rating was not less than 3. With this outcome, the aim of the study was achieved.

#### *F. Conclusion*

In the digital age, teaching and learning have moved from face-to-face interaction with teachers and facilitators to a more renowned and innovative system in which learners do not have to meet with any discipline facilitators. The English language is widely taught, learnt, and spoken in many countries as a foreign language. Digital resources and English learning websites (ELL) supported by high technology have made the acquisition of English language proficiency easier than ever. With access to the internet, learners can connect to a diverse offering of language learning websites that tailor programs to suit individuals' needs.

Many ELL websites make provisions for teaching, testing, and learning. However, there is a greater possibility for language readers to browse unreliable content, i.e., content published by web designers and bloggers online runs the risk of being misleading. This is because there is no formalized editorial process to review web designers and bloggers' content before they post. It is, therefore, essential to verify and evaluate the source of information published online. There is literature available on language learners' use of online resources for effective learning, but less focus has been placed on evaluating research on self-learning websites that help users develop English language skills. This presented a gap that needed to be covered. Thus, the literature review attempted to address this gap by providing material on the evaluation of the English learning websites, identifying the significance of ELL evaluation, approaches to the evaluation of the web-based sites, and establishing the evaluation criteria required for the evaluation of English websites and other digital resources. By addressing the abovementioned gap in the existing literature, this review presented a framework to assist users of English learning websites in becoming conscious about the credibility of using digital sources. This framework could also help web designers to elevate English webpage design to correspond to users' expectations.

#### *G. Research Objectives*

This study uses a technical evaluation process, resembles Aguayo and Ramírez's (2020) technical evaluation paper. The paper offers an examination of the technical quality of the websites for learning English as a Foreign Language (EFL) that are designed for autonomous language learning. With the aid of a questionnaire generated from Aguayo and Ramírez's checklist (2020), I aimed to identify the strengths and limitations of EFL learning websites by assessing their functionality and usability from the perspective of Saudi EFL learners. This effort can lead to a better design of EFL learning websites and other digital resources for more effective task performance on language learning.

#### *H. Research Questions*

The questions of the research are the following:

1. What are the perspectives of Saudi EFL learners in their evaluation of the design of websites for English language learning and other digital resources in terms of their functionality?
2. What are the perspectives of Saudi EFL learners in their evaluation of the design of websites for English language learning and other digital resources in terms of their usability?

### III. METHODOLOGY

#### A. Sample Study

The subjects consisted of ( $n = 51$ ) Saudi EFL users of English learning websites and digital resources, selected in a simple random method. The existing English language learning websites and other digital resources experienced by Saudi EFL users are the evaluands.

#### B. The Instrument

The researcher used the questionnaire (see the appendix) as a tool to collect data. The reason for selecting this instrument is its suitability of the study aims, its curriculum, and its society and answering its questions reliably. The researcher inspired the questionnaire from the checklist of Aguayo and Ramírez (2020). It focuses exclusively on two quality characteristics: functionality and usability, directly related to the web user. These two characteristics focus on the technical features of English learning websites. A checklist for other purposes, such as pedagogical concerns, is not recommended because it concentrates on technical features the most (Squires & McDougall, 1994).

#### C. Set up the Study Tool

After reviewing the educational literature and previous studies related to the current research subject, and in light of the study's data and questions and its objectives, the researcher built the tool (the questionnaire). It consisted of two parts.

There are two ways in which researchers and evaluators can evaluate or analyze their data: either being judgmental, i.e., introspective, or empirical (Chapelle, 2001). In my data analysis, I will employ a judgmental one that fits my data collection type, i.e., the checklist-driven questionnaire. Aguayo and Ramírez (2020) explained that the assessment of website functionality is through the proper hypertextual functioning. In other words, the evaluation is for the possible access and navigates all parts of the website and other external links resources, the level of users' performance allowed by the website's technology technically, and learning-specific dimensions. In comparison, usability assessment is based on the users' understanding and control of the resource to be used and operate with as little effort. Therefore, usability relies most on the web structure, content organization, order, and coherence between parts, more explicit instructions, and guidance to proceed and move throughout the websites. In my study, the evaluation of EFL websites for language learning relies on two categories: functionality and usability and their four sub-categories (see table 1).

The following is a presentation of how the questionnaire was constructed, and the procedures used to verify its authenticity and reliability:

1. The first section (functionality): It consists of (22) items, divided into four main axes, and table 1. clarifies the number of expressions of the questionnaire, and how they are distributed among the axes.
2. The second section (Elements of usability): It consists of (38) items, divided into four main axes, and table 1. clarifies the number of expressions of the questionnaire, and how they are distributed among the axes.

TABLE 1.  
AXES AND TERMS OF THE QUESTIONNAIRE: DESCRIPTION OF THE TWO MAIN CATEGORIES AND THEIR SUB-CATEGORIES

Axis	Number of items
Functionality	
The functionality of the EFL learning websites in terms of navigation	8 items
The functionality of the EFL learning websites in terms of adequacy of technology	3 items
The functionality of the EFL learning websites in terms of interactive functionality	3 items
The functionality of the EFL learning websites in terms of accuracy of technology for the specific purpose	8 items
Elements of usability	
The usability of EFL websites in terms of intelligibility	9 items
The usability of EFL websites in terms of ease of use	11 items
The usability of EFL websites in terms of operability	11 items
The usability of EFL websites in terms of design	7 items
Questionnaire	60 items

#### D. The Tool Validity

The truthfulness of the study tool means making sure that it measures what was prepared as intended to include the questionnaire for all the elements that are included in the analysis on the one hand, and the clarity of its expressions on the other hand, so that it is understandable to everyone who uses it. The researcher has made sure the study tool is validated by:

##### *Validate the internal consistency of the tool*

To verify the validity of the questionnaire's internal consistency, the Pearson's Correlation Coefficient was calculated to determine the degree of correlation of each of the questionnaire expressions to the overall degree of the axis to which the item belongs. The correlation coefficient values for each of the items with their dimension are positive and statistically significant at the level of significance (0.01) or less.

##### *The validity of the study tool*

The study instrument's validity was confirmed by using the validity factor of Cronbach's Alpha ( $\alpha$ ). The general validity coefficient is high as it reached (0.928). This indicates that the questionnaire has a high degree of validity reliable that can be relied upon in the field application of the study.

#### IV. DATA ANALYSIS

Many appropriate statistical methods were used using the Statistical Package for Social Sciences, denoted by the abbreviation (SPSS) to achieve the study's aims and analyze the collected data. The following statistical measures were calculated:

- 1- Frequencies and percentages to get to know the characteristics of the participants and determine their responses to the central axis items that are included in the study tool
- 2- Weighted Mean, to understand the participants' average responses on each of the axes items, and it helps arrange the items according to the highest weighted average arithmetic.
- 3- "Mean" arithmetic mean, to know the extent of the participants' high or low responses from the main axes, knowing that it is useful in arranging the axes according to the highest arithmetic mean.
- 4- Standard Deviation: to know the extent of the deviation of the participants' responses for each of the study variables' items and for each of the main axes of their mean.
- 5- Correlation: the correlation coefficient is used to measure the degree of correlation between two variables. The value of this parameter expresses the strength of the relationship between the two variables, and some approximate indicators can be used to judge the degree of this relationship. The relationship between the two variables is absent if the correlation coefficient is zero, while the value of one for the correlation coefficient indicates a complete relationship between the two variables.

#### V. RESULTS

This section deals with a presentation of the results of the current study by answering the study's questions according to the appropriate statistical treatments and then interpreting these results according to what is reached, as follows:

To answer the first question, "What is the perspective of Saudi EFL learners in their evaluation to the existing English language websites and digital resources design in terms of their functionality?" iterations, percentages, arithmetic mean, standard deviations, and ranks for the responses of the members of the participants have been calculated on the terms of (Navigation, Adequacy of technology, Interactive functionality, and Accuracy of technology for the specific purpose), the results are as follows:

##### *a. For the functionality of the EFL learning websites in terms of navigation*

The participants agree with the functionality of the EFL learning websites in terms of navigation with an average of (3.76 from 5.00), an average that falls in the fourth category of fifth scale categories (from 3.4 to 4.2) (see figure 2.).

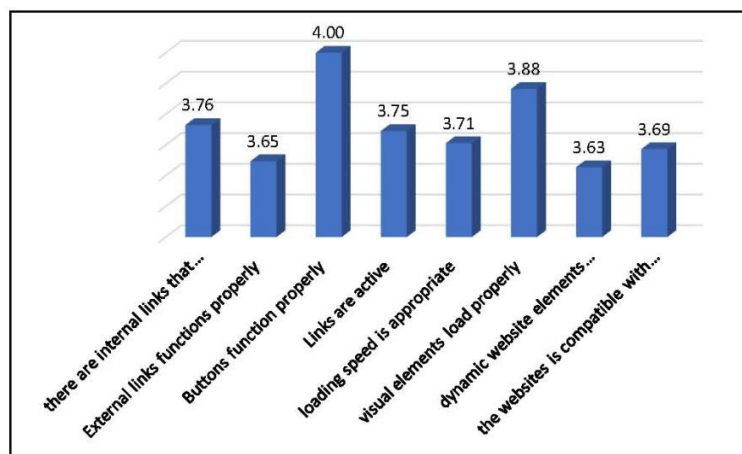


Figure 2. Responses of the participants about perspective of the functionality of the EFL learning websites in terms of navigation, descending according to the Means of approval.

##### *b. For The functionality of the EFL learning websites in terms of adequacy of technology*

The participants agree with the functionality of the EFL learning websites in terms of adequacy of technology with an average of (3.61 from 5.00), an average that falls in the fourth category of fifth scale categories (from 3.4 to 4.2) (see figure 3.).



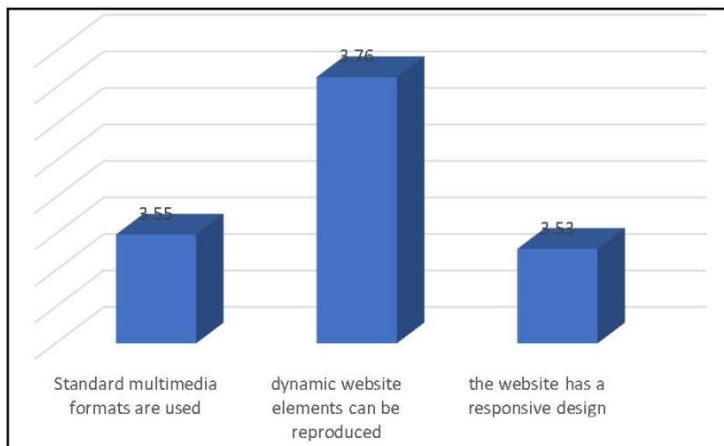


Figure 3. Responses of the participants about perspective of the functionality of the EFL learning websites in terms of adequacy of technology, descending according to the Means of approval.

**c. For The functionality of the EFL learning websites in terms of interactive functionality**

The participants agree with the functionality of the EFL learning websites in terms of interactive functionality with an average of (3.62 from 5.00), an average that falls in the fourth category of fifth scale categories (from 3.4 to 4.2) (see figure 4.).

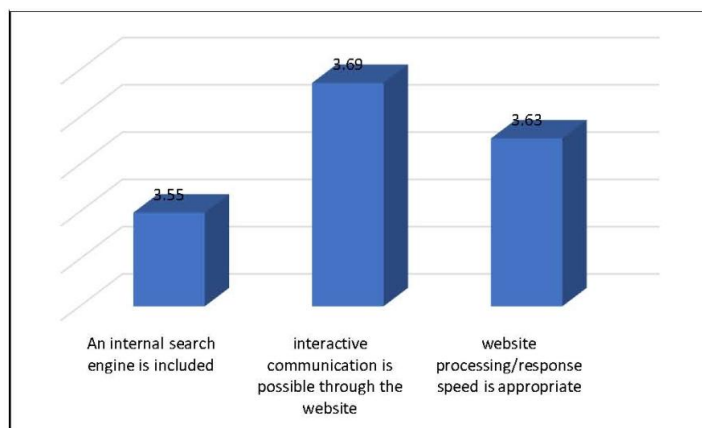


Figure 4. Responses of the participants about perspective of the functionality of the EFL learning websites in terms of interactive functionality, descending according to the Means of approval

**d. For The functionality of the EFL learning websites in terms of accuracy of technology for the specific purpose**

The participants agree with the functionality of the EFL learning websites in terms of accuracy of technology for the specific purpose with an average of (3.48 from 5.00), an average that falls in the fourth category of fifth scale categories (from 3.4 to 4.2) (see figure 5.).

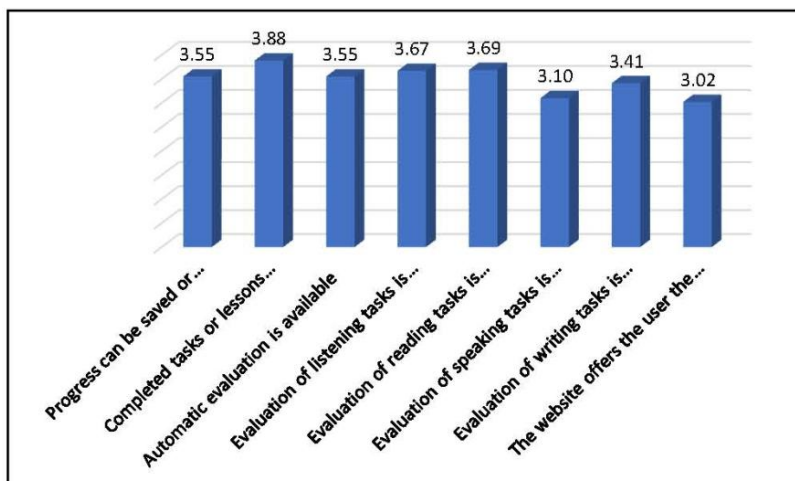


Figure 5. Responses of the participants about perspective of the functionality of the EFL learning websites in terms of accuracy of technology for the specific purpose, descending according to the means of approval

To answer the second question “What is the perspectives of Saudi EFL learners in their evaluation to the existing English language websites and digital resources design in terms of their usability?”, iterations, percentages, arithmetic mean, standard deviations, and ranks for the responses of the members of the participants have been calculated on the terms of (Intelligibility, Ease of use, Operability, and Design), the results are as follows:

**a. For the usability of EFL websites in terms of intelligibility**

The participants Agree with the usability of EFL websites in terms of intelligibility with an average of (3.57 from 5.00), an average that falls in the fourth category of fifth scale categories (from 3.4 to 4.2) (see figure 6.).

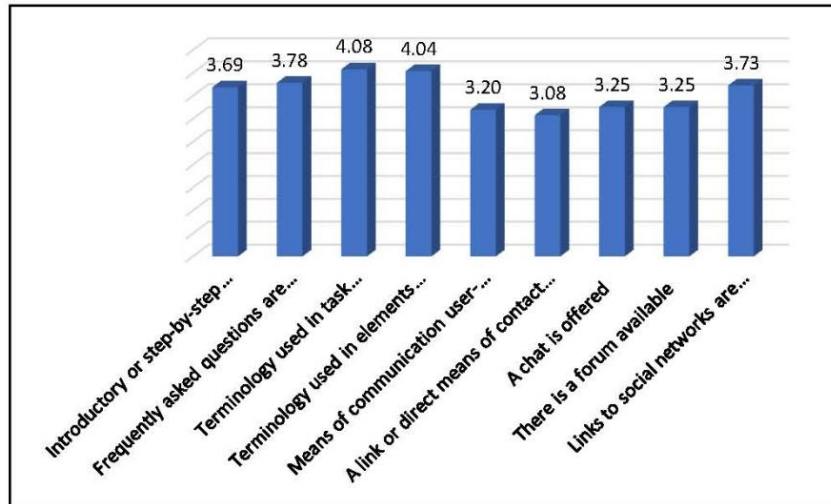


Figure 6. Responses of the participants about perspective of the usability of EFL websites in terms of intelligibility, descending according to the means of approval

**b. For the usability of EFL websites in terms of ease of use**

The participants agree with the usability of EFL websites in terms of ease of use with an average of (3.80 from 5.00), an average that falls in the fourth category of fifth scale categories (from 3.4 to 4.2) (see figure 7.).

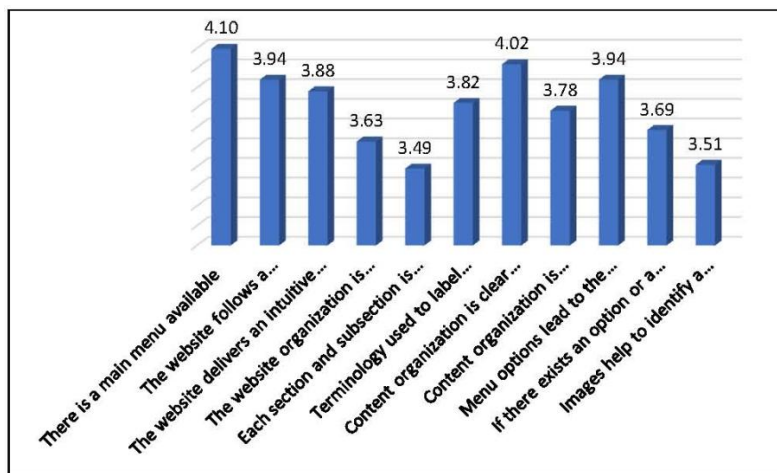


Figure 7. Response of the participants about perspective of the usability of EFL websites in terms of ease of use, descending according to the means of approval

**c. For the usability of EFL websites in terms of operability**

The participants agree with the usability of EFL websites in terms of operability with an average of (3.64 from 5.00), an average that falls in the fourth category of fifth scale categories (from 3.4 to 4.2) (see figure 8.).

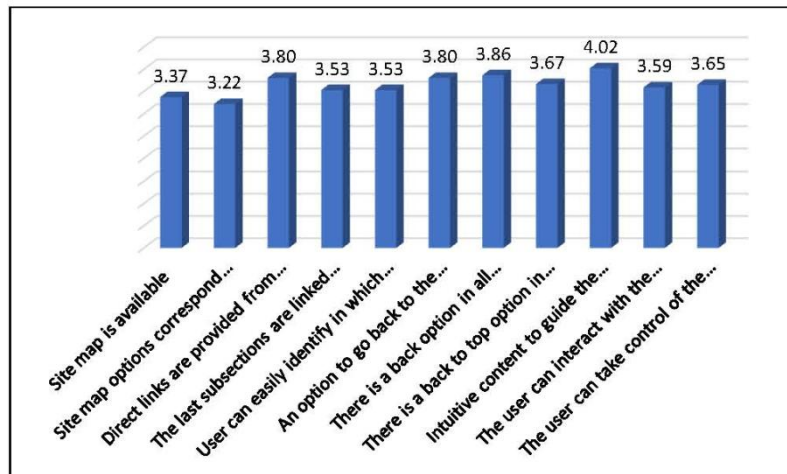


Figure 8. Responses of the participants about perspective of the usability of EFL websites in terms of operability, descending according to the means of approval

#### d. For the usability of EFL websites in terms of design

The participants agree with the usability of EFL websites in terms of design with an average of (4.08 from 5.00), an average that falls in the fourth category of fifth scale categories (from 3.4 to 4.2) (see figure 9.).

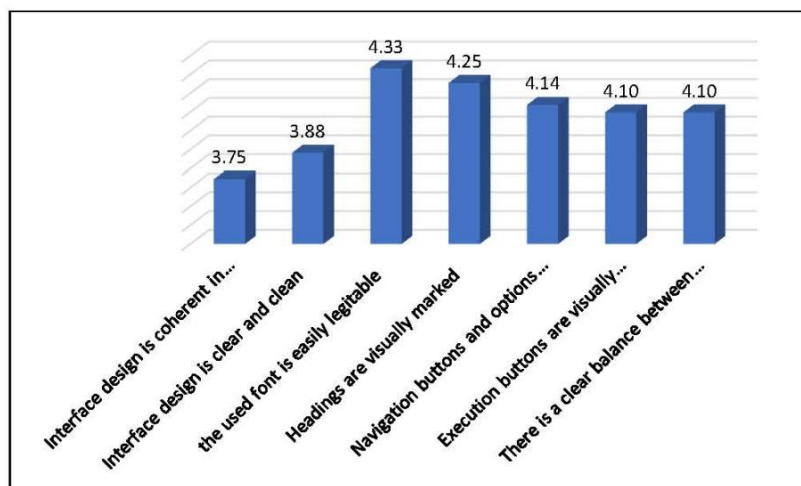


Figure 9. Responses of the participants about perspective of the usability of EFL websites in terms of design, descending according to the means of approval

## VI. DISCUSSION

The study aimed to evaluate self-led learning websites for learning the English language, based on a sample population from Saudi Arabia. The objective was to determine websites' effectiveness as learning tools by evaluating them technically. There is a long tradition of scholarship in the evaluation of websites, as demonstrated by various authors in the literature review of this paper. Two aspects of website evaluation that have been extensively researched were adopted in this study, with a focus on functionality and usability. These criteria, as suggested by researchers like Aguayo and Ramirez (2020), Morris (2019), and Shen et al. (2015) can facilitate understanding of the efficacy of using websites and digital resources in learning a language.

The current study primarily tested elements of functionality and usability as seen by a sample of Saudi EFL learners. The study found extensive agreement with most of the elements of functionality included in the survey. On a five-point scale, most of those elements received a mean score above 3 (Agree). However, a few extreme responses were registered. There seemed to be strong agreement when the elements being measured could be considered basic, such as the functioning of buttons (33.3%), whether the terminology used in task instructions was appropriate and intelligible (31.4%), and the availability of the main menu (31.4%). Thus, there was a general agreement among respondents regarding Research Question 1, the functionality of the evaluated websites. The following characteristics were considered positive: the ease of navigation ( $M=3.76$ ,  $SD=0.56$ ), adequacy of technology ( $M=3.61$ ,  $SD=0.6$ ), interactive functionality ( $M=3.62$ ,  $SD=0.61$ ), specific purpose ( $M=3.48$ ,  $SD=0.68$ ). Moreover, elements that applied to Research Question 2, on usability, also scored fairly high. Design had the highest average score ( $M=4.08$ ,  $SD=0.55$ ),

while ease of use ( $M = 3.8$ ,  $SD = 0.48$ ), operability ( $M = 3.64$ ,  $SD = 0.45$ ), and intelligibility ( $M = 3.57$ ,  $SD = 0.49$ ) had the next highest scores.

Typically, it is possible to argue that the scoring of most of the studied elements could be considered subjective. In a comparison presented by Cabrejas (2013), it was supposed that besides technological criteria, language learning sites needed to be evaluated through a deep analysis of the applied linguistics field. That argument suggests that English, for instance, should have different criteria for evaluation than Chinese or Arabic. A similar observation could be made based on the targeted first language. This idea was mentioned by Shen et al.'s (2015) study, in which Chinese EFL practitioners preferred pedagogically oriented ELL websites that gave ELL audiences content at various levels of language proficiency.

As mentioned earlier, ICT continued to play a critical role in self-directed language learning (e.g., Aguayo & Ramírez, 2020). Nonetheless, since ICT is dynamic and continues to offer improved options to facilitate learning, the general agreement in this study suggests that the survey respondents thought the elements could score better. Such a view is consistent with studies that considered ELL websites as works in progress (Aguayo & Ramírez, 2020; Shen et al., 2015). Designers of websites as tools for language instruction often struggle to address the cross-cutting needs of learners in online platforms. For instance, some researchers have observed that even established and prestigious ELL platforms, such as the BBC, were not yet fully functional and usable.

In part, the level of satisfaction with various ELL websites could be affected by the level of technical expertise of the learners or their familiarity with a website. Alternatively, the functionality and usability of these platforms are limited by the technology available. Perhaps with more innovations in this field, the ratings for usability and functionality will continue to increase. Improvements can be facilitated by the increased need to develop virtual learning capabilities around the globe.

## VII. CONCLUSION

This section presents a summary of the findings of this study, the contribution of this research to the literature, some pedagogical implications, the limitations of this study, and some suggestions for future research.

This study reported on the perspective of Saudi EFL learners concerning the functionality and usability of existing websites for self-directed learning of the English language. Through an iterative process, the study refined criteria presented in the literature from the fields of websites evaluation, ELL website design, and computer-assisted language learning. In particular, the evaluation criteria focused on the functionality and usability of the websites. The results of this evaluation research showed strong positive regard among Saudi EFL respondents with regard to the criteria for functionality and usability of ELL websites. The strongest responses were for the ease of use and ease of navigation, whereas the adequacy of technology for specific purposes had the lowest agreement.

This study contributes to the literature by identifying Saudi EFL learners' perspectives on ELL websites for autonomous learning; this contribution could lead ELL web designers to focus on developing areas that had the lowest agreement, such as representing learners' needs.

The implications of this evaluation research are varied technically and pedagogically. They include gaining an awareness of English learning websites as learning tools, identifying areas for ELL website assessment, and leading ELL web designers to develop better designs for more effective task performance.

The limitations of the study may involve the limited sample size, a single research method, and only two criteria for evaluation. The research sample was chosen randomly through online distribution and had only 51 respondents. The researcher evaluated ELL websites descriptively, through the means of approval of the survey, and based on technical elements only.

As for the areas for future study, researchers can focus on the role and efficacy of content that represents learners' needs better or develop technical aspects that could facilitate learners' specific purposes. Also, researchers can focus on developing ELL web design that corresponds to learners' needs at different levels. Further research also can identify other criteria for evaluating ELL websites, such as the types and amount of content and learners' satisfaction.

## REFERENCES

- [1] Aguayo-Arrabal, N., & Ramírez-Delgado, C. D. (2020). A proposal of evaluation criteria for the quality of ESL/EFL Websites for Autonomous Learning. In Huertas-Abril, C., & Gómez-Parra, M. E. (Eds.), *International Approaches to Bridging the Language Gap* (pp. 102–117). IGI Global. <http://doi:10.4018/978-1-7998-1219-7.ch007>
- [2] Aguayo, N., & Ramírez, C. M. (2020). Does technical assessment matter? Functionality and usability testing of websites for ESL/EFL autonomous learners. *Research in Learning Technology*, 28(0), 1–27. <https://doi.org/10.25304/rlt.v28.2353>
- [3] Ahmadi, D., & Reza, M. (2018). The use of technology in English language learning: A literature review. *International Journal of Research in English Education*, 3(2), 115–125.
- [4] Allison, R., Hayes, C., McNulty, C. A. M. & Young, V. (2019). A comprehensive framework to evaluate websites: Literature review and development of good web. *JMIR Formative Research*, 3(4), e14372.
- [5] Aly, M. A. (2008). An evaluative study of some online websites for learning and teaching English as a foreign language. Benha University Press.
- [6] Castillo, R. C., & Arias, M. B. (2018). Analysing English online resources for children: A practical case with an evaluation template proposal. *Pixel-Bit: Revista de Medios y Educación*, 53, 7-25.

- [7] Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research*. Cambridge, Cambridge University Press.
- [8] D'áz, J., Rusu, C., & Collazos, C. A. (2017). Experimental validation of a set of cultural-oriented usability heuristics: e-Commerce websites evaluation. *Computer Standards & Interfaces*, 50, 160-178.
- [9] Duggirala, S. (2016, August 25). 10 Usability heuristics with examples. *Prototypr*. Retrived March, 2021 <https://blog.prototypr.io/10-usability-heuristics-with-examples-4a81ada920c>.
- [10] Fuentes, E. M., & Martínez, J. R. (2018). Design of a checklist for evaluating language learning websites. *Porta Linguarum*, 30, 23-41.
- [11] Hakiki, R., Junaedi, D., & Kaburuan, E. R. (2018). The user interface design for natural science learning media for elementary school. In *MATEC web of conferences* (Vol. 197, p. 16006). EDP Sciences.
- [12] Hubbard, P. (2011). Evaluation of courseware and websites. In L. Ducate & N. Arnold (Eds.), *Present and future promises of CALL: From theory and research to new directions in foreign language teaching* (pp. 407-440). CALICO.
- [13] Jimenez, C., Lozada, P., & Rosas, P. (2016, September). Usability heuristics: A systematic review. In *2016 IEEE 11th Colombian Computing Conference (CCC)* (pp. 1-8). IEEE.
- [14] Kartal, E. (2010). The internet, language learning, and international dialogue: Constructing online foreign language learning websites. *Turkish Online Journal of Distance Education TOJDE*, 11(2), 90-107.
- [15] Kaur, K., & Singh, H. (2016, October). Click analytics: What clicks on webpage indicates?. In *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)* (pp. 608-614). IEEE.
- [16] Keshavarz, H., & Givi, M. E. (2020, April). Website evaluation frameworks: IS oriented vs. business oriented models. In *2020 6th International Conference on Web Research (ICWR)* (pp. 223-228). IEEE.
- [17] Khandare, S. S., Gawade, S., & Turkar, V. (2017, October). Survey on website evaluation tools. In *2017 International Conference on Recent Innovations in Signal Processing and Embedded Systems (RISE)* (pp. 608-615). IEEE.
- [18] Kir, E., & Kayak, S. (2013). The evaluation of websites teaching English as a foreign language. *Social Behavior Sciences*, 1(4), 2788-2795.
- [19] Kratov, S. (2018, August). On the approaches to the development of a scientific organization website. In *2018 3rd Russian-Pacific Conference on Computer Technology and Applications (RPC)* (pp. 1-5). IEEE.
- [20] Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer-assisted language learning*. US, Lawrence Erlbaum Associates.
- [21] Morris, K. (2019). How to evaluate websites. Retrieved in March 2021 from <http://www.kathleenamorris.com/2018/11/20/evaluate-websites/>.
- [22] Polizzi, G. (2020). Digital literacy and the national curriculum for England: Learning from how the experts engage with and evaluate online content. *Computers & Education*, 103859. <https://doi.org/10.1016/j.compedu.2020.103859>.
- [23] Pribeanu, C. (2017). A revised set of usability heuristics for the evaluation of interactive systems. *Informatica Economica*, 21(3), 31-38.
- [24] Shen, H., Yuan, Y., & Ewing, R. (2015). English learning websites and digital resources from the perspective of Chinese university EFL practitioners. *ReCALL: The Journal of EUROCALL*, 27(2), 156-176.
- [25] Son, J. (2005). Exploring and evaluating language learning websites. In J. B. Son, & S. O'Neill (Eds.), *Enhancing learning and teaching: Pedagogy, technology and language* (pp. 215-227). Post Pressed.
- [26] Squire, D., & McDougall, A. (1994). *Choosing and using educational software: a teachers' guide*. Philadelphia, Pennsylvania: Routledge.
- [27] Xie, K., Di Tosto, G., Chen, S. B., & Vongkulluksn, V. W. (2018). A systematic review of design and technology components of educational digital resources. *Computers & Education*, 127, 90-106.
- [28] Zhang, D., & Pérez-Paredes, P. (2019). Chinese postgraduate EFL learners' self-directed use of mobile English learning resources. *Computer Assisted Language Learning*, <https://doi.org/10.1080/09588221.2019.1662455>.

**Arwa Sulaiman Alhabdan** is currently working as a teaching assistant at Qassim University, Kingdom of Saudi Arabia. She holds a master degree in applied linguistics. Her research interests include psycholinguistics, second language acquisition, and technology-enhanced language learning.