Journal of Modern Education Review, ISSN 2155-7993, USA
January 2022, Volume 12, No. 1, pp. 37–52
Doi: 10.15341/jmer(2155-7993)/01.12.2022/007
Academic Star Publishing Company, 2022
http://www.academicstar.us

STVALL: Hands-on Pills for English Learning on Smart TV

Alejandro Curado Fuentes
(University of Extremadura, Avda. Universidad, Spain)

Abstract: Our project focused on the design of a digital interface called STVALL (Smart TV-Assisted Language Learning) for all kinds of users at home. This program delivers linguistic and content challenges in the form of learning pills with which users can interact. The input and scores are stored as assessment content on an authoring tool. Our main research question is whether users' hands-on interaction with the tool, which records their performance, correlates with learning motivation. This study is conducted on a quasi-experimental basis with 1500 training pills across five subject areas (Science and Nature, Literature and Art, Geography and History, Entertainment and Sports, and Language) and four language user levels: Adults/Children (0–12 years old): A1/A2/B1/B2. The training pills are produced in the form of text, audio, and audio-visual content in those categories. Linguistic scaffolding is based on user profiles. The pills present dynamic questions/answers so that users can respond individually or in groups. Preliminary results with 15 users point to generally positive learning outcomes and responses. A post-session questionnaire was answered by 11 people, and significant feedback is obtained and considered regarding the interface, preferred types of activities, and EFL scaffolding. It seems that linguistic proficiency and content knowledge play a crucial role in relation to the motivation for using the tool. In addition, playing by teams is favored by most users whereas short dynamic audio-visual interactions are mostly preferred.

Key words: hybrid TV, EFL, linguistic pills, interactivity

1. Introduction

In today's entertainment world, Smart TV draws and receives much attention and recognition from most viewers. Smart TV sales have expanded over the past few years, increasingly gaining market share (from 55 percent in 2015 to 70 percent in 2018) (Vailshery, 2021). This technology resonates with today's on-line culture due to its wide and dynamic internet access, app downloading capabilities, and streaming services. Watching TV with family and friends prevails as a socio-cultural phenomenon (e.g., watching favorite series and films, sports, and so on), whereas user-adapted programming can thrive via interconnected mobile devices (Rigby et al., 2017). Menu-based, delivery-on-demand, and internet-based resources are distinctive elements that define users' preferences when buying a TV. Social media and streaming resources contribute to making this technological scenario attractive, as an easy synchronization is enabled with portable devices (Van Deventer et al., 2018).

The HbbTV (Hybrid Broadcast Broadband Television) standard offers an efficient and successful mode of Smart TV integration in the household context, with an optimal combination of different formats and devices

Alejandro Curado Fuentes, Dr., University of Extremadura; research areas: corpus linguistics, ICT, language learning. E-mail: acurado@unex.es.

(Gavrila et al., 2020). In this digital landscape, educational TV has also continued evolving as an affordable means of TV programming for the family. A new term, "t-learning" (Belloti et al., 2008), has been coined regarding a more interactive process of accessing content for multiple purposes via digital TV. Viewers at all ages adopt a more dynamic form of accessing TV content by focusing on entertainment, but also with a combination of interactivity and gamification for the achievement of certain learning goals, as Rey-López et al. (2008) describe.

Adaptive TV systems (ATS) are a key feature of digital television. ATS allow for content customization based on users' demands and preferences. This notion has evolved from what Brusilovsky (1998, 2001) called adaptive systems, based on a user model which stores explicit user information (e.g., age, sex, nationality, language command, content preferences) and implicit user data stored as hands-on interaction with the system. Adaptive hypermedia systems (AHS) have been developed in educational settings to offer students interactive strategies (e.g., in regular language lessons and courses).

Hyper-textual tools have been implemented in many subject areas; in FL (Foreign Language) contexts, AHS lessons generally lead to positive learning outcomes (e.g., Cumbreño et al., 2006; Vera-Rodríguez & Arias-Soto, 2008). Content developed in AHS is intended to adapt to specific users in terms of their learning profiles (e.g., language level, content preferences, learning demands, and so on). In previous work (cf. Rico-García et al., 2007), students' perceptions of the adaptive content have been generally found as favorable, and adaptive programs are evaluated on a positive note regarding learning motivation and academic developments. In particular, vocabulary and listening comprehension have been valued as most benefited areas after the use of AHS lessons.

STVALL (Smart TV-Assisted Language Learning) is the name of our project on Smart TV for language learning purposes. We aim to enable a multi-user platform for the integration of different English levels and content knowledge. In this paper, we will describe the hybrid technology deployed for the Smart TV program, and how the platform is organized and managed so that its contents can adapt to different user profiles in ATS. We will also examine the hands-on use made with preliminary lessons and pills on a case study with 15 users so that the different strategies employed in the system can be evaluated and improved in future developments. The platform contains 1500 training pills in five different subject segments: Science and Nature, Literature and Art, Geography and History, Entertainment and Sports, and Language. There are also four language levels (A1-B2, according to CEFR, 2018) for adults and children (under age 12). The interactive activities are called training pills because we are interested in training the system at this stage. The main challenge is to explore users' knowledge and skills with content for EFL (English as a Foreign Language) learning. Therefore, the material is being devised and arranged according to difficulty levels in EFL.

Our research group consists of computer specialists, university researchers, elementary and secondary school instructors, and education experts/consultants. The course material is classified according to different variables (e.g., monologic versus dialogic texts, oral/written texts, interactivity, narration/instructions, and so forth—see Table 1 below). We chose the HbbTV (Hybrid broadcast broadband Television) technology because it supports many TV brands, especially in European countries (Haase, 2017). We also devised an authoring tool to create our training pills. This tool feeds external user information to the system (e.g., age, language level, and so on) and controls learners' interaction progress with the activities. Preliminary findings suggest that user profiles based on language proficiency and content preferences tend to correlate with their performance results according to their degrees of language command and age. We include post-session questionnaires to examine users' reactions to the activities.

2. Background

Digital TV has evolved as a valid candidate tool for t-learning and educational contexts. Hybrid TV can be used not only at home but in educational institutions to target specific learning situations and to increase motivation. Contents on interactive TV appeal to most viewers, as digital TV is generally considered as easy to use and reliable (Pavlov & Paneva, 2006). The HbbTV standard can facilitate this ease of content integration and creation, with many possibilities of exploring lessons and materials (Fondevila-Gascón et al., 2013). One early example was implemented with an intelligent tutor called PANDA (Damasio & Quico, 2004) in an ATS, facilitating interactivity with animated characters in various subjects and topics. Children enjoyed the programs, and results were generally satisfactory in class. Another example was Ryu et al. (2014), who developed EFL games based on HTML5, JavaScript, and Canvas. Their AHS interactive activities were mainly supported by key words animated with digital drawings and pictures that users could exploit at their own pace.

In terms of mobile interconnectivity, Fallahkhair et al. (2004) worked on a digital TV system which combined portable devices and t-learning. In a foreign language module, users could play with interactive subtitles during TV lessons to pause and select key vocabulary, then relay it to mobile devices for storage and subsequent activities. In turn, audio-visual content was effectively integrated in T-MAESTRO (Rey-López et al., 2008), where educational content was designed via two modules: “Edutainment”, with academic course material organized according to thematic units, and “Entercation”, with more informal audio-visual output that corresponded to those thematic units.

Hupont et al. (2011) and Baldassarri et al. (2015) designed “the first t-learning affective aware tutoring tool” (Hupont et al., 2011, p. 3) in the T-EDUCO project. Based on broadcast and broadband TV technology, users’ facial expressions were video-recorded by the tool during interactive sessions. Their emotions were then transferred and juxtaposed with their performance and progress so that the educational content could be evaluated affectively. Bibiloni et al. (2014) introduced Augmented Reality (AR) into hybrid TV systems to transmit output signals in the form of 3-D bookmarks in TV lessons and programs. Interactivity with content (e.g., opening books and going to specific sections) was enabled by selecting 3-D objects in AR environments.

The organisation of content according to thematic units in Ben-Said et al. (2018) was made to classify audio-visual input according to TV segments. That way, when similar content in a category was found, it would be automatically filed within a given timed set. The lessons would thus add more content as users progressed with the tool. The applicability of the HbbTV standard can be successfully applied to FL learning according to different categories or segments on a digital TV platform (Salguero et al., 2019). This affordance includes a versatile authoring tool that can deliver and manage such content according to user profiles.

3. The Hybrid TV Platform Architecture

Interactivity and adaptive content are crucial elements in STVALL. Its hybrid TV standard allows for the adoption of a client/server architecture (Figure 1). The inter-play between these two modules supports the integration of user profiles and content adaptation.

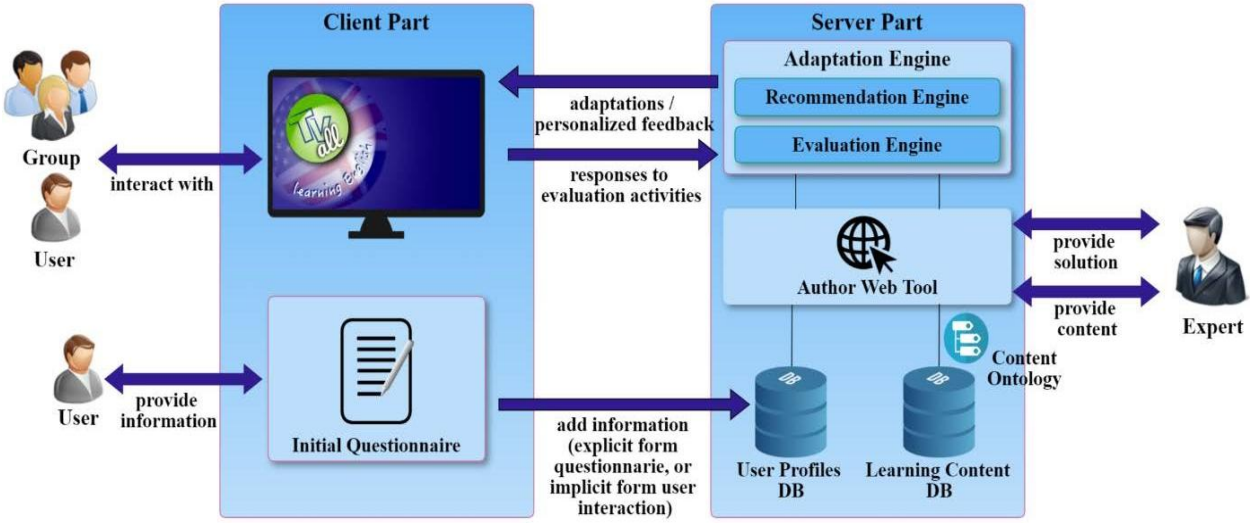

Figure 1 STVALL Architecture

The core section in both modules is the authoring tool. This resource interacts with user profiles (based on age, sex, language level, and so on) and their scores/progress with the system. The authoring tool is a web-based application created according to the MVC (Model View Controller) pattern. The user interface has been developed with HTML5, CSS3 and JavaScript. This application is also responsive, meaning that it is adaptable to any device thanks to the use of a Bootstrap framework. The implementation of the authoring tool has been divided into three layers: The persistence layer, which enables access to the data using Hibernate technology; the business logic, in charge of processing the information obtained in the previous layer thanks to Spring, and the presentation layer, built with Spring MVC technology which allows users to display the information and receive any response from the system. The technology applied in the provision and management of content by experts is JSP.

In the client part, users can sign up and adopt avatars either individually or by groups. They must provide the system with personal information so that next time they play, they just need to log in with their username and password. Human experts can always validate this information using the authoring tool. Then, the user information goes to the storage section, where one database (Learning content DB) stores and classifies all the learning input created by our team and experts, and another database (User profiles DB) maintains all the information related to the users (personal and progress profiles). Different users can also form a group model by joining a team to play with others. Each user profile then adds in information to the group model. Based on this model, the authoring tool can retrieve their group information next time they enter the system.

On the server part, three engines sustain adaptive interactivity in the system: Adaptation, recommendation, and evaluation engines. They receive and send the information from and to the databases so that the authoring tool can generate the adapted content for specific users. The adaptive mode in the system is mainly enabled by the adaptation engine. This motor receives the information from user profiles and adapts the content to that user by choosing the tagged activities that best suit his/her profile. In the case of a group, the activities would be selected according to that group model. In the adaptation engine, two other engines play important roles in the process of adaptation and selection. The evaluation module is in charge of measuring users' scores, activity attempts, time spent, and mistakes with the pills. Based on this information, the recommendation section in this engine decides the appropriate levels for users and/or groups so that in the next session, interacting with the content that best

adapts to their profiles can be resumed.

4. Content Organization and Management

The learning content DB stores all the activities and challenges designed in the system. There, two forms of training pills are devised: 1) Presentation pills with introductory text, audio, and / or audio-visual content for users to watch; and 2) Play-mode pills where the content is conveyed in the form of interactive quizzes, to be sent to users and answered individually or by groups, automatically providing feedback after the activity. The activities are stored with eight different variables in mind. These tags are inserted by the content design experts (Table 1).

Table 1 Content Tagging in the Training Pill DB

Variables	Tags
User type	Adult/Child
Language level	A1/A2/B1/ B2
Subject area category	Science and Nature /Literature and Art/ Geography and History/Entertainment and Sports/Language
Language skill	Written comprehension/Written production / Oral comprehension/Oral production
Access mode	Presentation/Interactive
Register type	Professional – academic/News/Fiction/Conversation
Discourse function	Narration/Description/Explanation/Instruction/Biography/Publicity/ Informative/ Humour
Discourse mode	Monologic / Dialogic

The user type, language level, subject area, language skill, and access mode variables are pre-defined. This means that the information was conceived by our research group before we created the activities. The classification of this information is basic to define user profiles. Then, the other three variables (register type, discourse function, discourse mode) were added as we stored training pills according to their different types of content at a subtler text level, also relevant for content retrieval according to user preferences. In the register type and discourse function categories, more than one tag can be applied. For example, we have academic material (register) that includes conversation, or publicity (discourse function) which describes things. In sum, all the variables are tagged according to different choices in each learning pill. This classification allows the adaptation engine to identify the pills that best conform to a user profile or group mode.

The second major affordance for content organization and management is the expert interface (Figure 2) in the authoring tool. This menu contains nine training pill templates where the activities are designed. These templates include the eight variables (Table 1) for experts to choose from when creating the material. First, we must select one template (e.g., “Seleccionar con recurso” — “Selection quiz with a resource”). Then, after we fill out the content information with the question, answers, correct option, type of output, format, and other tags for the activity, we are led to a “new task characteristics” menu where we can select the description tags (Figure 3) to describe this learning pill.

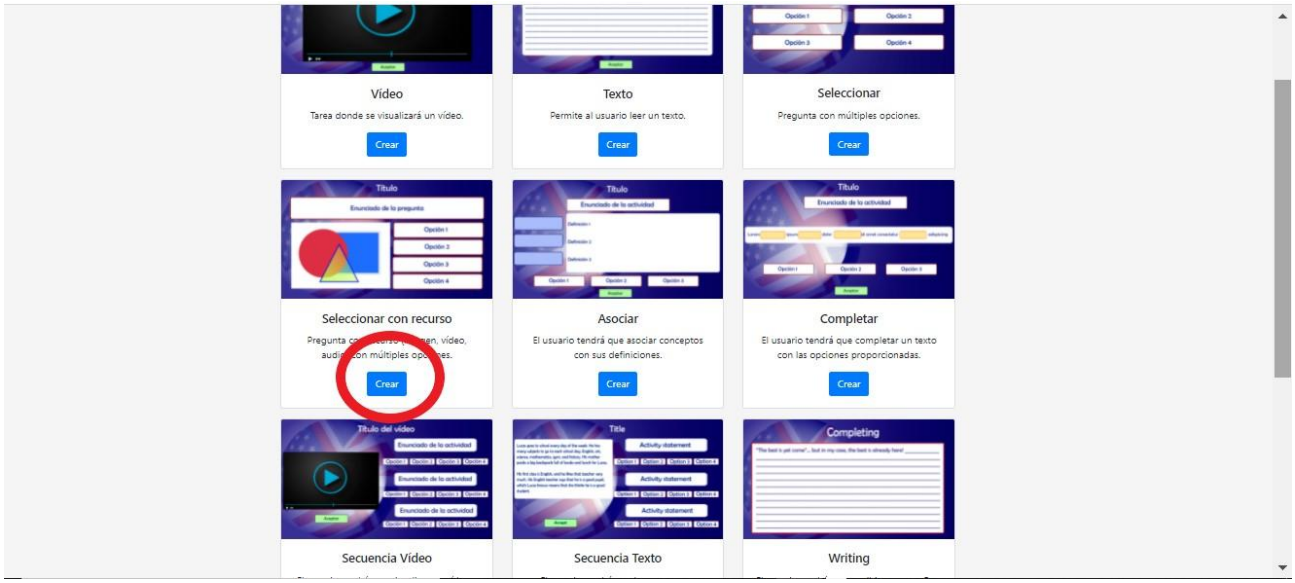

Figure 2 Templates for the Creation and Management of Learning Pills

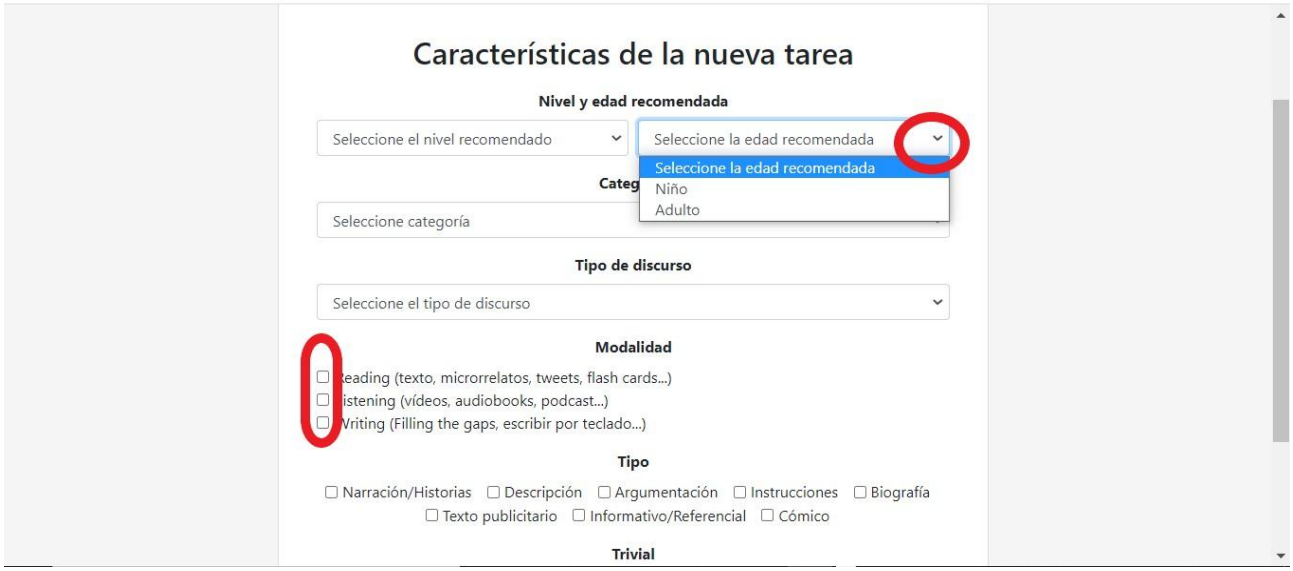

Figure 3 Tag Selection in the Creation of the Learning Pill

The “Text presentation” pill consists in a short written and/or oral text that viewers can watch and/or listen to at the same time. It is about any topic within each subject category. “Video presentation” is the same type of pill but in an audio-visual form. Then, there are five types of pills that let users interact with text, aural, and audio-visual content: “Multiple choice” (text and sound) activities, “Multiple choice with a resource” (audio-visual), “Matching” (any format), “Fill-in-the-gaps” (any format), and “Sequence” (where users must answer questions during the reproduction of a podcast or video). Two other pills are also being tested at the present time: “Writing”, which requires users to activate their mobile devices as a secondary screen where they can enter the answers. In such cases, training pills are viewed on the Smart TV but the answer choice is managed on the mobile device as either a “gapping” activity, or as “complete the text”. Additionally, “chatting” is enabled for some pills, where a larger text can be inserted, and then, on-line experts examine and evaluate it at a later stage.

The last type of pills is “Speaking” (under construction). The idea here is that users deploy mobile devices to record their answers orally and send them to the system for voice recognition (if the answers are short), or for human evaluation (in the case of longer input).

The authoring tool displays different types of templates to create pills, but it can also display already created ones (see Figure 4 for pills 85 through 92). We enter this management section to select the types of pills and either modify them (using the tool icon) or delete (trash can icon) created ones. In many cases, reviewing pills is necessary because audio-visual material, for example, may change or no longer be available. All the pill templates include the possibility of adding audio or video resources, e.g., a creative commons licenced video from YouTube. In terms of compatibility, the HbbTV version 2.0.1 standard (ETSI, 2016) is open to JPEG, GIF and PNG pictures, MP4 videos, and MP3 audios. YouTube Player API (Google Developers, 2014) is the video player embedded in the system.

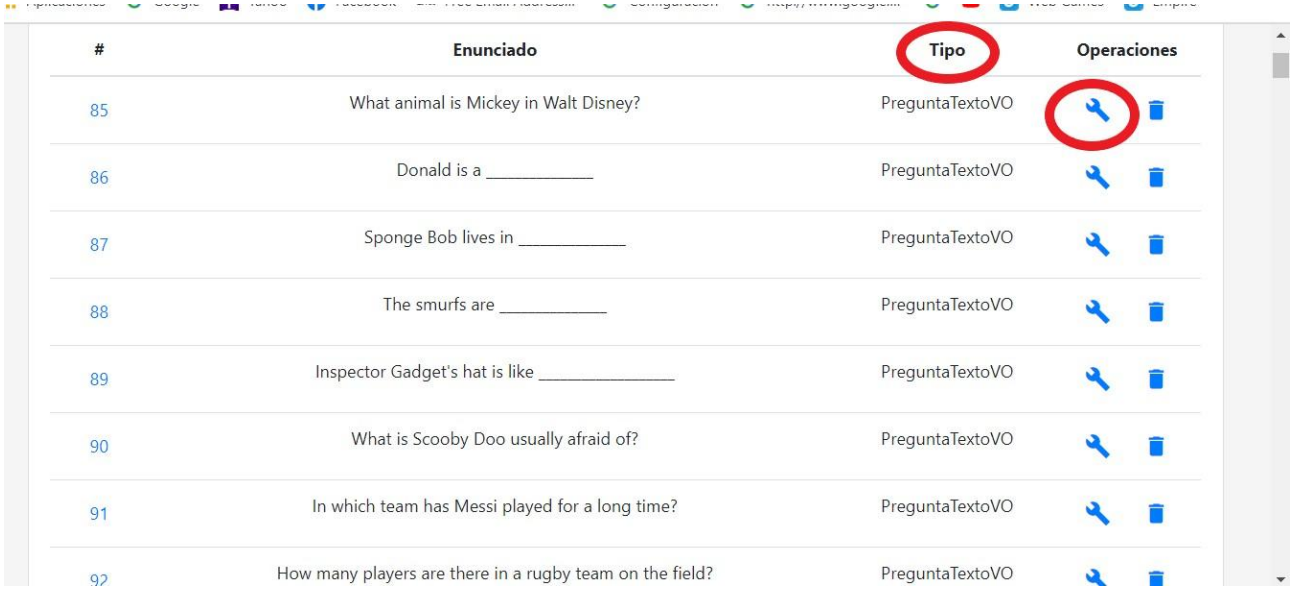

Figure 4 Management Section for the Training Pills

Once completed, the pill can be played as digital output. An example is a “Select with a resource” activity (Figure 5) where a video is first shown to children at B1/B2 levels, and then they must choose the correct answer. In Figure 6, it is the same type of activity for lower proficiency language learners. We discriminate among different levels according to voice speed, vocabulary used, grammar complexity, and lexical density. In Figure 7, the example is for adults at B1-B2 levels within the same subject category (Entertainment) but with a more complex type of discourse.

For group mode interactivity, as mentioned above, the adaptation engine can build a joint user model based on all the individual user profiles in that group. Then, the recommendation engine will send some activities based on the model. The training pills must be solved jointly by the groups of learners. After they play for some time, the evaluation engine stores user model progress by integrating the scores obtained and the history of the activities carried out thus far. Groups can play against other groups in a competitive mode. In this mode, registered and unregistered users can play if new users join an existing group. Each user in the group can participate, adding to their group’s game board the points obtained and / or mistakes made. Competition can also engage single players (see Figure 8).

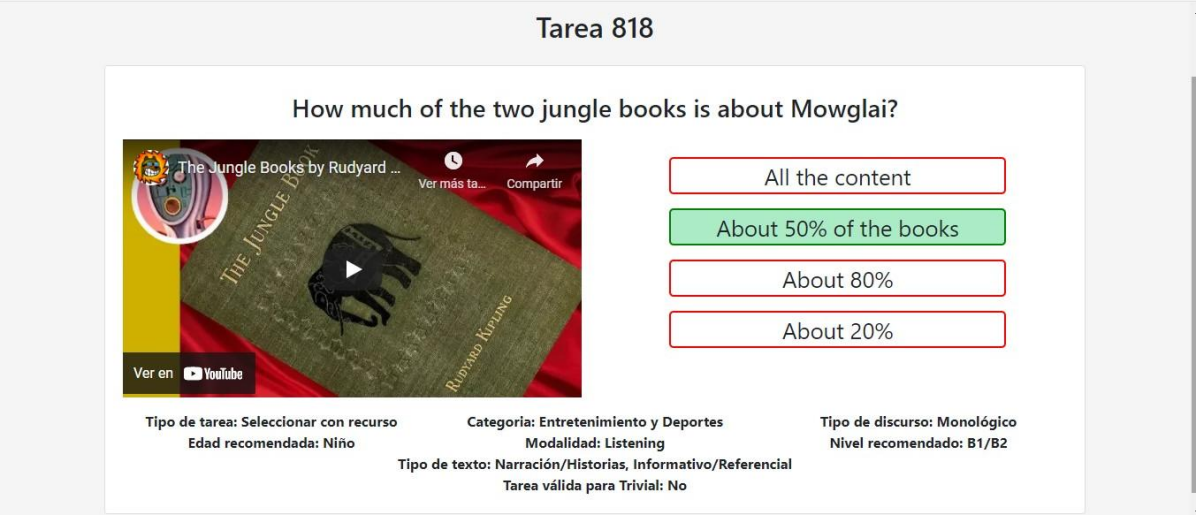

Figure 5 Example of a Training Pill at B1/B2 Levels (Children)

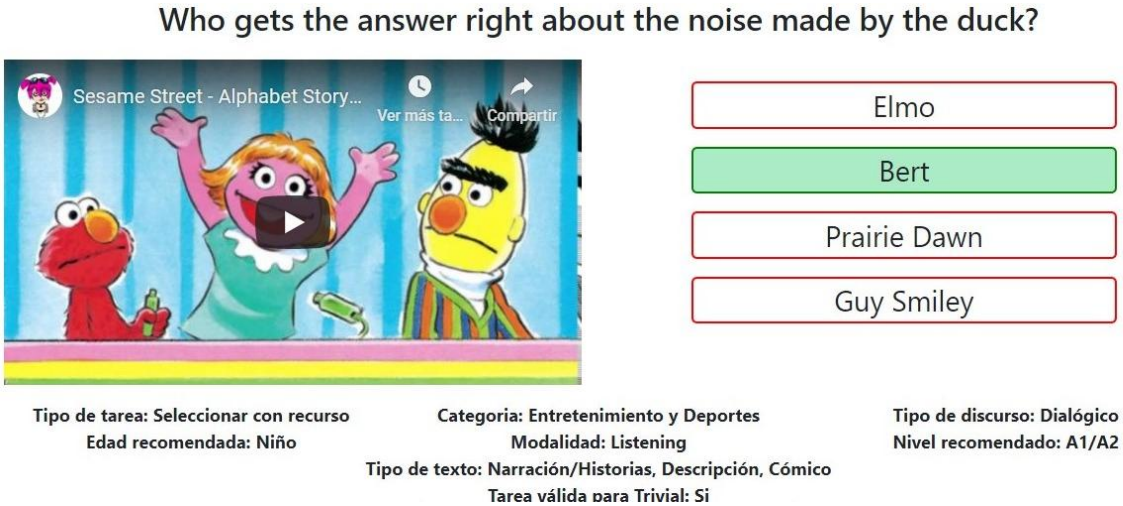

Figure 6 Example of a Training Pill at A1 / A2 Levels (Children)

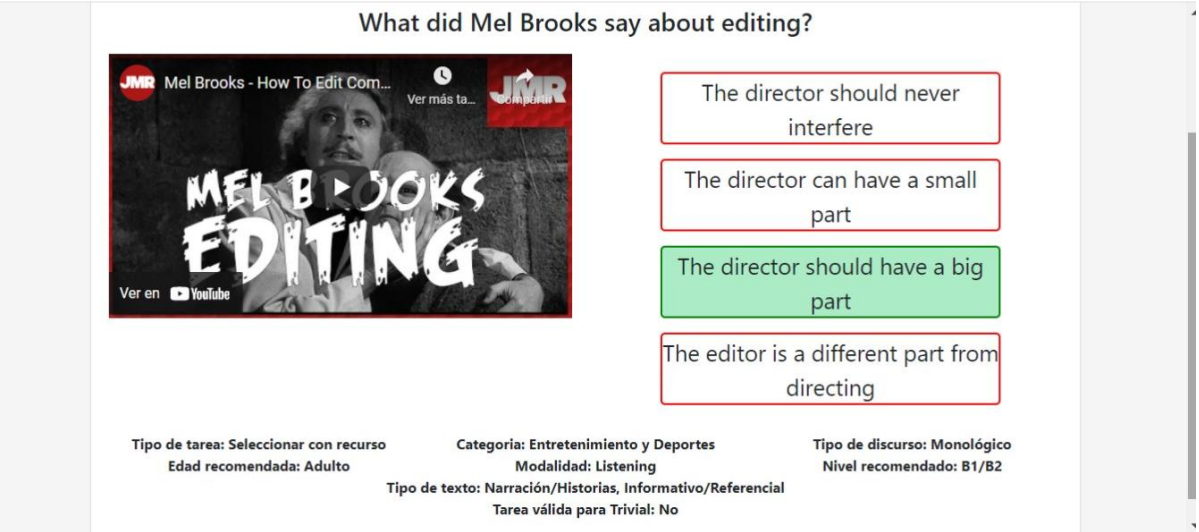

Figure 7 Example of a Training Pill for Adults (at B1/B2 Levels)

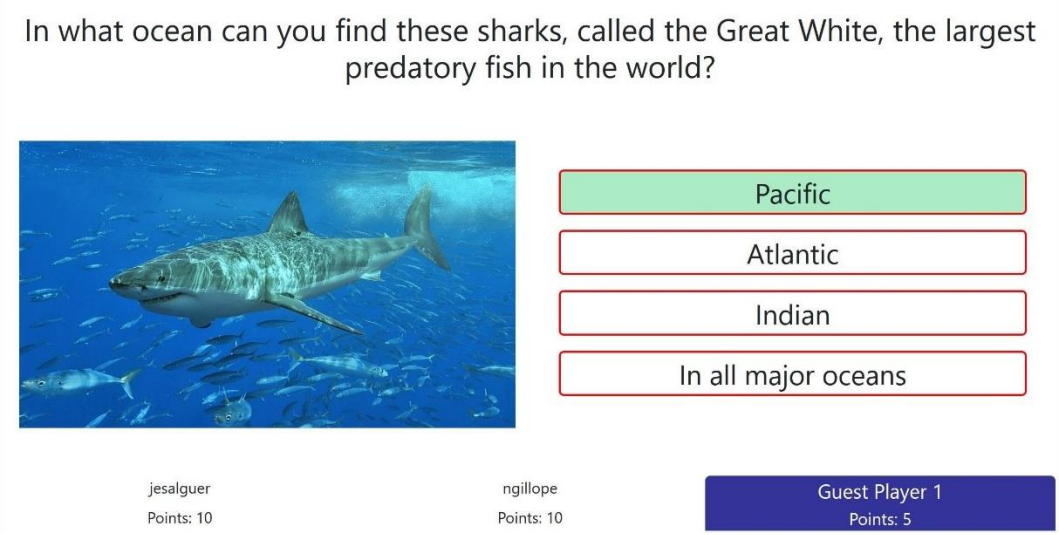

Figure 8 Competitive Mode Between a Registered User and an Invited User

5. User Input and Feedback

The system was tested during the months of April and May, 2020. It was the strict COVID 19 lockdown in Spain at the time, and so, we decided to send the program to specific families and friends who could interact and play from their homes. Because STVALL was not broadcast live or via regular TV broadcasts, the 15 users accessed the program on their computers, joining the sessions via internet connections (connecting to our server): Eight of them were adults, and seven were children (all within their families). All accessed most of the 1500 training pills, but there were significant differences between some groups. 11 people used the lessons individually and in groups (as family groups), and four users in different households (two adults and two children) interacted with the system in the single mode only.

In June 2020, we analysed all the information about the users saved on the system. The accessed activities varied quantitatively and qualitatively. Table 2 provides each user's number of accessed pills during the sessions and their correct answers. The percentages refer to the proportion of correct answers in relation to their total number of accessed activities. These scores were achieved in the single mode. Their EFL proficiency levels are also provided for contrastive aims.

Given this data, we observe a strong positive correlation between users' alleged language proficiency level and scores (Figure 9). A correlation coefficient of R = 0,864 attests to that effect. In Figure 9, the independent variable is the EFL level (from A1 to B2, divided into 25 percent segments), and the dependent variable is the performance percentage (from Table 2). We think that this observation with such a limited number of users can already indicate that the tendency is for upper-intermediate EFL users to engage in longer sessions and to obtain more correct answers with the tool. We know that these users have more proficiency in English because they referred to their passed B1 exams in official language schools. We can observe that adults 2, 6, and 8, in particular, accessed more pills and achieved higher scores. Both dependent variables correlate with a higher language level (B2). Child 6, however, having a B2 level, did not score significantly higher, although she accessed more pills than her child co-users. We will examine some possible reasons below.

Table 2 Quantitative Information About Users' Sessions in the Single Mode

User/English level	Accessed pills	Correct pills	Percentage
Adult 1/B1	120	57	47.5
Adult 2/B2	311	286	91.9
Adult 3/B2	390	251	64.3
Adult 4/B2	310	220	70.9
Adult 5/A2	312	121	38.7
Adult 6/B2	512	477	93.1
Adult 7/B1	233	165	73.9
Adult 8/B2	317	300	94.6
Child 1/B1	78	31	39.7
Child 2/A2	33	9	27.2
Child 3/A1	31	9	29
Child 4/A1	30	7	23.3
Child 5/A1	33	5	15.1
Child 6/B2	157	91	57.9
Child 7/B1	41	30	73.1

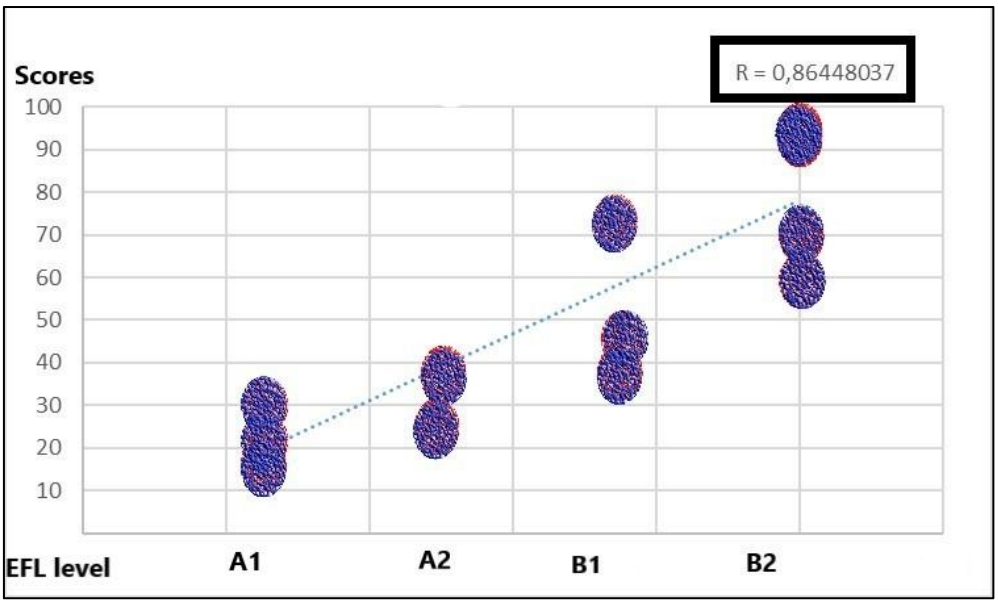

Figure 9 Scores Correlating With Higher EFL Proficiency

In terms of the subject categories where users scored more correct answers, Science (615 points), followed by Entertainment (522), ranked at the top. This is an interesting observation, as users did not manifest a special preference for any subject category in the post-session questionnaires. However, they clearly achieved more points in those two subjects. Table 3 identifies the groups of users playing together and their overall final correct scores as percentages in each subject category.

Table 3 Groups' Scores in Each Subject Category

	Group 1	Group 2	Group 3	Group 4
Components	Adults 2 and 3	Adult 6; Children 1, 2, 3	Adults 7 and 8	Adult 1; Children 4, 5
Science	28.2	30.9	30.1	32
Entertainment	14.8	27.8	16.1	27.1
History	25.6	13.7	12.9	9.7
Language	22.9	20.6	19.3	29.1
Literature	8.3	6.8	21.5	1.9

There are more proportionally correct answers in Science, followed by Language. This order differs from the total correct scores according to individual users, where Entertainment ranked second. Additionally, a competing mode was set up for Groups 2 and 4. They played for 54 minutes on May 4, 2020, after our computer technician configured the session in this mode for them. Group 2 won (65 correct pills versus 12). The EFL level in the pill selection was almost always B1-B2, which definitely biased the competition in favour of group 2 (who had higher language levels, while group 4 only had one member at the B1 level).

After the sessions, in late June, a 30-item questionnaire (see Appendix) was sent to the users to answer on Google form. 11 people completed the questionnaire. Four children were excluded due to their younger age (children 2, 3, 4, and 5). Given their answers and choices, we did not find any outliers in terms of their score range. In other words, the calculation of inter-quartile ranges for upper bound and lower bound values in each questionnaire item did not produce any values outside of the expected range.

In the questionnaire (see Appendix), most respondents stated that they liked the platform and played more than once. Only adults 3 and 7 chose "No comment" for item 4, and adult 6 and child 1 played between 5 and 10 times. A positive correlation occurred between questions 15 and 16, 17 and 19, 17 and 20, and 18 and 21 when we compared the mean values for the answers. In 15/16, the 3 users who competed in groups assigned higher scores to the group session mode, although they did not evaluate the competition mode as high. In 17/19, high mean values in interactive activities positively correlated with the high means for audio-visual content. This correlation is R = 0,85135135 (see Figure 10). In 17/20, the same happened for audio resources (R = 0.79035572). In 18/21, the lower values for the presentation pills correlated with those given to text content (R = 0.62493243).

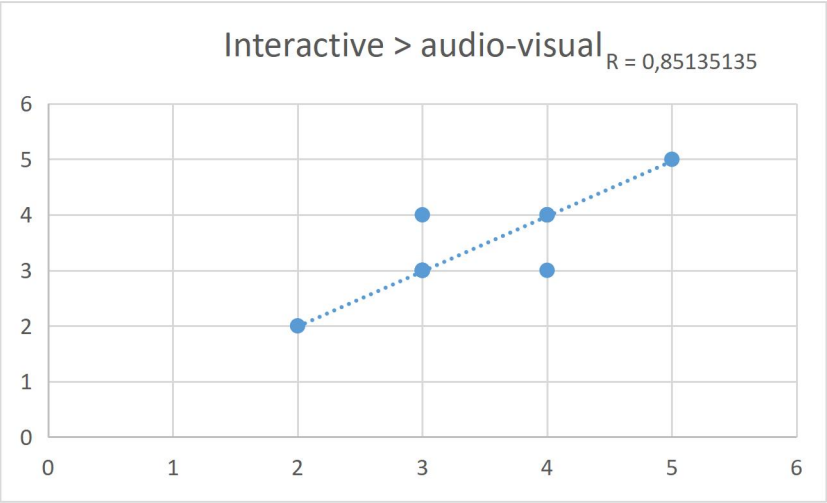

Figure 10 Positive Correlation For Interactivity With Audio-Visual Content

When asked about the linguistic skills favoured by the tool (question 9), seven users selected “Listening”, followed by “Reading” (two people), and “None” (two people). Regarding the “Writing” possibilities to be implemented (question 10), seven people responded positively (YES), two others said No, and two “No comment”. For question 12 (Speaking utilities), nine said they would like to have them (Yes) and two users said “No comment”. Writing and speaking activities are therefore demanded in this preliminary study. Writing pills have been already designed in STVALL but were not available to users at the time.

Figure 11 displays a section in the authoring tool that can store the writing input from users for human correction at a later stage. The language evaluator then accesses a text editor in each activity to tag the mistakes according to different categories (grammatical, lexical, discourse connectors, missing items, and / or any other learner error/aspect to be edited). This interface (Figure 12) is based on *execCommand*, included in JavaScript (Mozilla developers, 2019).

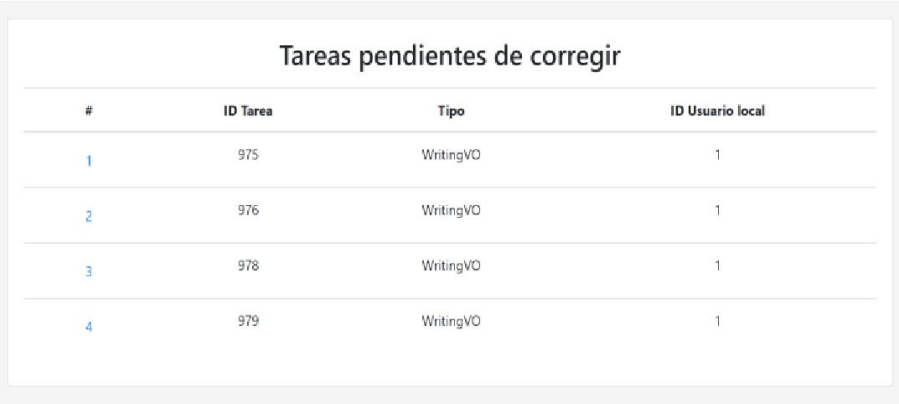

Figure 11 Writing Activities Stored on the Tool

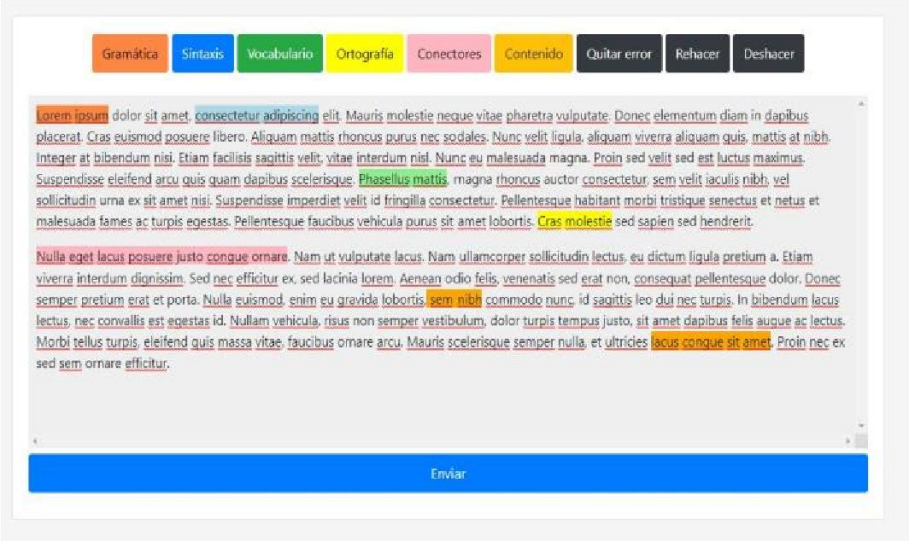

Figure 12 Text Editor for Learner Errors

The speaking application suggested by some users in question 13 refers to a voice recognition system, and even the possibility for interacting orally with native English speakers. This feature is not included, but could definitely enhance the experience with STVALL. In item 11, in turn, some comments pointed at vocabulary games

and grammatical quizzes in the tool as suggestions for improvement.

Finally, most answers to items 27-30 unveiled the recognition of the provided educational affordances for families with children, according to the high scores given in Figure 13, especially for questions 28 ("STVALL is fun to use"), 29 ("STVALL is good for family use"), and 30 ("I would use STVALL in the future").

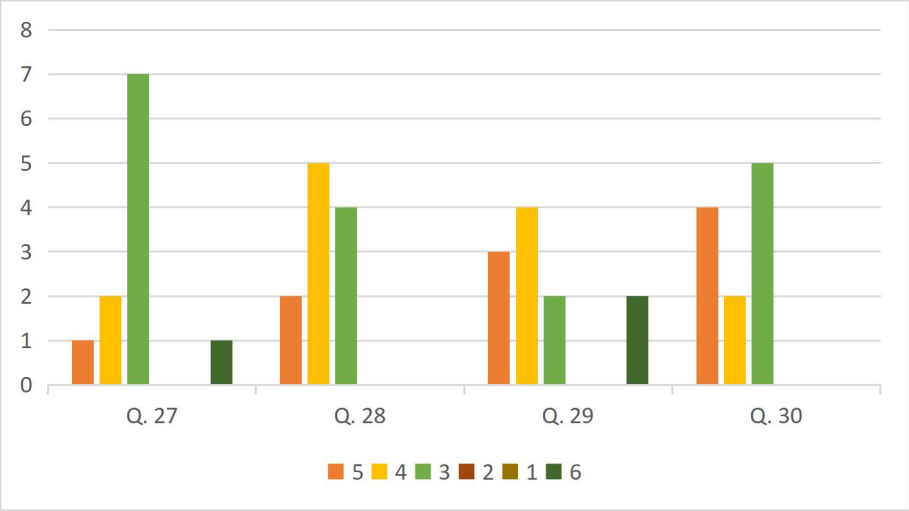

Figure 13 Positive Answers in Items 27–30

6. Conclusions

A hands-on approach to STVALL content and pills hints at its promising affordance as a hybrid TV system for EFL. We glean some significant findings and observations related to gamification in EFL learning. These positive impressions have been corroborated by adults' and their children's interactions within groups and/or individually since they generally enjoyed the activities. The trivial-pursuit-like training pills seem to adapt well to user profiles in terms of language levels, age, and progress. However, the competition mode failed to work effectively in this case due to a selection of more upper-intermediate language level activities for the two groups. The group models should therefore be revised and re-adjusted.

A key observation is that language proficiency at B1-to-B2 levels correlates with more productive engagement and user satisfaction, as disclosed by actual user performance and questionnaire informants' data. Adults tend to enjoy the tool more whereas children prefer to play in groups. Related to this finding, a lower language proficiency correlates with less interest shown in playing in the single mode (users at A1). This reaction may be caused by the fact that many A1-A2 pills tend to be more A2 than A1. We should thus revise the content for these pills in terms of linguistic content and complexity. An alternative solution is to divide users into separate categories (A1 versus A2).

Text pills are generally considered as less appealing by users, who demand more audio-visual content. Some users even put forward that the platform should include synchronous speaking utilities. Therefore, in addition to enhancing audio-visual content and calibrating language levels in the pills, we should work on applications and plug-ins that allow for writing and speaking extensions. In any case, given the reduced number of the tested users thus far, the system should be evaluated more widely by multiple users and should run on a networked TV platform. With more contrastive data and wider access to the technology in such a study, we think that STVALL

can be more reliably tested and assessed as a plausibly optimal scenario for EFL practice.

Acknowledgement

I would like to acknowledge the work done by GexCALL (Computer-Assisted Language Learning Group in University of Extremadura) and the funding provided by the Government of Extremadura, with Agreement number # IB16160 (2018–2020; 2019–2021).

References

Baldassarri S., Hupont I., Abadía D., and Cerezo E. (2015). "Affective-aware tutoring platform for interactive digital television", *Multimed Tools Appl.*, Vol. 74, pp. 3183–3206, doi: doi.org/10.1007/S11042-013-1779-z.

Bellotti F., Vrochidis S., Parissi E., Lhoas P., Mathevon D., Pellegrino M., Bo G. and Kompatsiaris I. (2008). "A T-learning courses development and presentation framework", *IEEE Multidisplinary Engineering Education Magazine*, Vol. 3, No. 3, pp. 69–76.

Ben-Said O., Wali A. and Alimi A. M. (2018). "A new system for TV program contents improvement using a semantic matching technique", in: J. R. Villar, E. A. de la Cal, A. Herrero, H. Quintián, J. A. Sáez (Eds.), *Hybrid Artificial Intelligent Systems*, pp. 246–250.

Bibiloni T., Mascaró M., Palmer P. and Oliver A. (2014). "Augmented reality on HbbTV, an hypervideo approach", in: *9th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 743–748.

Brusilovsky P. (1998). "Web-based education for all: A tool development adaptative courseware", *Computer Networks and ISDN Systems*, Vol. 30, No. 17, pp. 291–300.

Brusilovsky P. (2001). "Adaptive hypermedia", *User Modeling and User Adapted Interaction*, Vol. 11, No. 1-2, pp. 87–110.

Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Companion Volume, Council of Europe Publishing.

Cumbreño A., Rico-García M., Curado-Fuentes A. and Domínguez E. (2006). "Developing adaptive systems at early stages of children's foreign language development", *ReCALL*, Vol. 18, No. 1, pp. 45–62, doi: https://doi.org/10.1017/S0958344006000413.

Damasio M., Quico C. (2004). "T-learning and interactive television edutainment: The Portuguese case study", in: *Proceedings of ED-MEDIA 2004 — World Conference on Educational Multimedia, Hypermedia & Telecommunications*, AACE, pp. 4511–4518.

ETSI, HbbTV 2.0.1 Specification, 2016, available online at: https://www.etsi.org/deliver/etsi_ts/102700_102799/102796/01.04.01_60/ts_102796v010401p.pdf.

Fallahkhair S., Pemberton L. and Masthoff J. (2004). "A dual service scenario for informal language learning: Interactive television meets the mobile phone", in: *Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies*, IEEE Computer Society, pp. 16–20.

Fondevila-Gascón J. F., Del Olmo-Arriaga J. L., Beriain A., Carreras-Alcalde M. and Pesqueira-Zamora M. J. (2013). "Educational possibilities in the internet-television intersection: The HbbTV standard", in: *International Conference on Education and New Developments*, pp. 35–37.

Gavrila C., Popescu V., Fadda M., Anedda M. and Murroni M. (2020). "On the suitability of HbbTV for unified smart home experience", *IEEE Transactions on Broadcasting*, Vol, 99, pp. 1–10.

Google Developers (2004). "YouTube Player API Reference for Iframe Embeds", available online at: https://developers.google.com/youtube/iframe_api_reference?hl=es.

Haase H. E. (2017). "HbbTV country review: Rest of the world", in: *HbbTV Symposium*, pp. 45–49.

Hupont I., Abadía D., Baldassarri S., Cerezo E. and Del-Hoyo R. (2011). "TEDUCO: A t-learning tutoring tool that cares", *International Journal on Artificial Intelligence Tools*, Vol. 20, No. 4, pp. 639–661.

Mozilla Developers (2019). "Document.exe command", available online at: https://developer.mozilla.org/es/docs/Web/API/Document/execCommand.

Pavlov R. and Paneva D. (2006). "Interactive TV-based learning, models and standards", in: L. Böszörményi, I. Simonics (Eds.), *Methods and Tools for Development of Semantic Enabled Systems and Services for Multimedia Content, Interoperability and Reusability*, HUBUSKA, pp. 70–99.

Rey-López M., Díaz-Redondo R., Fernández-Vilas A., Pazos-Arias J., López-Nores M., García-Duque J., Gil-Solla A. and Ramos-Cabrer M. (2008). "T-MAESTRO and its authoring tool: Using adaption to integrate entertainment into personalized t-learning", *Multimed Tools Appl.*, Vol. 40, pp. 409–451.

Rico-García M., Agudo J. E. and Curado-Fuentes A. (2007). "The application of adaptive hypermedia systems to specific/linguistic communicative learning traits", in: A. Curado-Fuentes, P. Edwards, M. Rico-García (Eds.), *Approaches to Specialised Discourse in Higher Education and Professional Contexts*, Cambridge Scholars, pp. 156–181.

Rigby J. M., Brumby D., Gould S. and Cox A. L. (2017). "Media multi-tasking at home: A video observation study of concurrent TV and mobile device usage", in: *TVX '17: ACM International Conference on Interactive Experiences for TV and Online Videopp*, Association for Computing Machinery, pp. 3–10, doi: https://doi.org/10.1145/3077548.3077560.

Ryu D., Kazimierz-Krompiec P., Lee E. and Park K. (2014). "A serious game design for English education on smart TV platform", in: *The 18th IEEE International Symposium on Consumer Electronics (ISCE 2014)*, IEEE, pp. 1–2, doi: 10.1109/ISCE.2014.6884479.

Salguero-Serrat J., Sánchez H., Naranjo M. J., Rico M. and Agudo J. E. (2019). "STVALL: HbbTV-based adaptive system for English learning through interactive TV", in: *Proceedings of EdMedia + Innovate Learning*, AACE, pp. 1520–1525.

Vailshery L. S. (2021). "Market share of smart TVs worldwide 2015–2018", *Statista*, available online at: https://www.statista.com/statistics/889000/worldwide-smart-tv-market-share/#statisticContainer.

Van Deventer M. O., Probst M. and Ziegler C. (2018). "Media synchronisation for television services through HbbTV", in: M. Montagud, P. Cesar, F. Boronat, J. Jansen (Eds.), MediaSync, pp. 7–18, Springer, doi: https://doi.org/10.1007/978-3-319-65840-7_18.

Vera-Rodriguez E. and Árias-Soto L. D. (2008). "Development of strategies to teach English through an interactive game: A journey to Britannia", *Folios*, Vol. 28, No. 2, pp. 120–136.

APPENDIX

Questions and number of answers (in brackets)

How old are you?

0-12 (3) b. Over 12 (8)

What is your gender?

Male (6) b. Female (5)

What is your English language level?

A1 b. A2 (1) c. B1 (4) d. B2 or higher (6)

Did you enjoy using STVALL?

Yes (9) b. No c. No comment (2)

How many times did you use STVALL?

1-5 times (9) b. 5-10 times (2) c. More than 10 times

Which mode did you use?

Single user (4) b. Group mode c. Both modes (7)

Did you use it to compete against other players?

Yes (3) b. No (8)

If you answered Yes to question 7, did you compete

As a single user b. As a group (3) c. As both

Which language skill do you think STVALL is best for?

Reading (2) b. Writing c. Listening (7) d. None (2)

Would you like to have more writing activities in STVALL?

Yes (7) b. No (2) c. No comment (2)

If you answered Yes to question 10, give some examples of writing activities you would like STVALL to add:

_____ vocabulary & grammar evaluation_____

Would you like to have speaking activities in STVALL?

Yes (9) b. No c. No comment (2)

If you answered Yes to question 12, give some examples of speaking activities you would like STVALL to add:

_____ voice recognition & real-time speaking_____

From 1 (very bad) to 5 (very good), rate the following:

Single user mode

1 2(2) 3 (3) 4 (4) 5 (2) N/A

Group mode

1 2 3 (2) 4 (4) 5 (1) N/A (4)

Competing mode

1 2 (1) 3 (1) 4 (1) 5 N/A (8)

Interactive pills

1 2 (1) 3 (3) 4 (4) 5 (3) N/A

Presentation pills

1 (1) 2 (2) 3 (6) 4 (2) 5 N/A

Audio-visual content

1 2 (1) 3 (4) 4 (3) 5 (3) N/A

Audio content

1 2 (1) 3 (4) 4 (4) 5 (2) N/A

Text content

1 (2) 2 (3) 3 (6) 4 5 N/A

Science category

1 2 (2) 3(7) 4(2) 5 N/A

History category

1 2 (3) 3 (6) 4 (2) 5 N/A

Entertainment category

1 2 (2) 3 (7) 4 (1) 5 (1) N/A

Language category

1 2 (3) 3 (8) 4 5 N/A

Literature category

1 2 (2) 3 (7) 4 5 N/A (2)

STVALL is good for learning

1 2 3 (7) 4 (2) 5 (1) N/A (1)

STVALL is fun to use

1 2 3 (2) 4 (5) 5 (4) N/A

STVALL is good for family use

1 2 3 (2) 4 (4) 5 (3) N/A (2)

I would use STVALL more in the future

1 2 3 (5) 4 (2) 5 (4) N/A