Self-Voicing Needs Individualisation

A Study on a Self-voicing Web Framework for the Support of Various Users with and without Reading Difficulties

Theresa Deschner, Lukas Smirek and Gottfried Zimmermann
Responsive Media Experience Research Group, Stuttgart Media University, Nobelstraße 10, 70569 Stuttgart, Germany

Keywords: Self-voicing, Voice Output, Speech Output, OpenAPE, ASpanel, One-Size-Fits-All, One-Size-Fits-One, Reading Difficulties, Speech Synthesis.

Abstract: Nowadays, the internet is a frequent source of information. Due to difficulties with reading and language, some people have difficulties obtaining this information and therefore may have to deal with constraints in their daily life, including a dependency on other persons. The use of voice output on web pages ("self-voicing web pages") may help them to overcome these constraints. In this paper, a self-voicing framework is presented. A *user-centred design* approach was applied in the development, implementation and validation of the concept. The concept accommodates different people with different requirements and needs – using a *one-size-fits-one* approach rather than *one-size-fits-all*. The evaluation shows the use of the framework for different user groups as well as the need for providing individualised features in the framework.

1 INTRODUCTION

At the present day, the internet is an important instrument for information gathering and reseach. The information is often presented in textual form only. People with impairments regarding reading or language are sometimes barred from this source or at least impeded. Some compensate this problem by using assistive technologies, often requiring time-consuming initial training (W3C, 2014). Some rely on persons in their environment to assist them, resulting in a dependency on other people. Finally, some prefer to avoid all situations in wich reading is required, often resulting in constraints in their daily lives (Döbert and Hubertus, 2000).

We propose a framework for speech output integrated in web pages, to enable these people to access internet content more easily and independently. The framework is supposed to make text-based web content more easily accessible without the need of additional software or hardware, time consuming training or dependencies on other people. Note that a self-voicing framework is different from a screenreader in the following ways: A screenreader is used for launching a browser, navigating a web page and accessing its content, while the self-voicing framework is for text-to-speech only. Furthermore, a screenreader is an additional tool a user needs to buy and install, while the framework is integrated in the webpage itself.

In this paper, we address the target group of potential users of the proposed framework. Reading and language problems can have various causes, e.g. visual impairments, dyslexia, an incomplete process of learning a written language or struggles with the language as foreign language. The diversity of the target users makes it necessary to allow for personalisation within the framework.

The framework was developed according to the principles of user-centred design, to ensure that the framework meets the requirements of real users. The requirements analysis drew from interviews with potential users. A prototype of the proposed framework has been evaluated involving people of the target user groups.

This study was conducted in Germany, with interviews and evaluations conducted in the German language, except for the interviews and user tests with language learners which were conducted in English.

The remainder of this paper is structured as follows: Chapter 2 provides an overview on related work. Chapter 3 presents an overview of the user groups that are targeted by our framework. Chapter 4 describes the development of our framework. Chapter 5 reports about the evaluation of our prototype.

Deschner T., Smirek L. and Zimmermann G.
Self-Voicing Needs Individualisation - A Study on a Self-voicing Web Framework for the Support of Various Users with and without Reading Difficulties.
DOI: 10.5220/0006581401980205
In *Proceedings of the International Conference on Computer-Human Interaction Research and Applications (CHIRA 2017)*, pages 198-205
ISBN: 978-989-758-267-7

Chapter 6 provides a discussion on limitations and other aspects learned from our study. Finally, chapter 7 contains a conclusion and provides an outlook on possible future activities in this area.

2 RELATED WORK

The idea of integrating speech into a web page or application has been implemented various times. Some instances aim at making web pages accessible to users with disabilities; some just read electronic books out aloud. In general, we identify two different approaches for including speech output in web pages or applications. (1) Professional audio recordings. This approach is time consuming and expensive – only a small number of books have been transformed into audiobooks. (2) Speech synthesis and synthetic speech output. This is generally cheaper and more compact to store and transport. Prominent examples of speech synthesis show that the quality is quite acceptable and resembling a human voice. However, systems for speech synthesis are often designed for output of short texts, typically single sentences only. Voice output often sounds artificial and monotone for longer texts (Evans and Reichenbach, 2012).

In the remainder of this section, we provide an overview and examples of both approaches of self-voicing technologies and frameworks applied to the web, as we see relevant and inspiring for our work.

On web pages, voice output is mostly not generated by the web page itself but supplied by technologies by external providers. On German web pages, we often found tools provided by *ReadSpeaker*[1], *Voice Reader*[2] and *narando*[3].

The tools *ReadSpeaker* and *Voice Reader* are similar. Both use synthetically generated speech and can be used for multiple languages. Both provide several options to adapt the tool to the users' needs. These include a download function, reading speed and voice pitch. Text highlighting can be shown synchronized to the speech in multiple versions, with a choice of the text colour. The control panel includes controls like play, pause, stop and volume. These tools provide a variety of options, and are similar to our proposed framework. However, both tools have weaknesses. The main problem is the exclusive use of synthetic speech. While easier to produce, it can lead to a wrong adoption of pronunciation and intonation of the spoken language. Especially children and language learners are negatively affected by this. It is also problematic that the button for opening the settings is only available after starting the speech output. The user cannot create pre-sets of preferences, and the settings cannot be saved. As a result, the user has to set their preferred settings every time the web page is loaded.

In contrast, *narando* uses recordings of the texts rather than synthetic speech. The recordings and the matching text articles are available on the *narando* web page and on the web page employing the *narando* technology for their text articles. On the webpage, the audio is embedded as time bar, with play and pause functions. *narando* has the benefit of a more natural sound, which is easier to follow and featuring a correct pronunciation and intonation. This is preferable, in particular for people who are dependent on an accurate speech output, such as language learners and children. *narando's* main weakness is the lack of options.

3 TARGET USER GROUPS

In this chapter, we identify the user groups as targeted by our proposed framework. We aim for user groups with problems reading the German language.

In the following list, we introduce the intended user groups for the self-voicing framework, and their needs.

- *Children:* Children between six and ten years are not finished in their cognitive development (Hourcade, 2008). This includes the process of learning the written language. That means they have to recognize a letter, match it to a sound and save all sounds in the working memory. A word is recognized by adding all the sounds to a word (Rau, 2007). By hearing and reading a word at the same time, the word can be more easily recognized and pronounced correctly.
- *Dyslexic People:* Dyslexia can manifest in various symptoms, not only affecting literacy, but also communication, concentration, navigation, organisation and information processing (W3C, 2014). Even literacy-related symptoms alone can span a wide range. Dyslexic people might have a slow reading rate; they might skip words or parts of words, or twist, replace or add to them (Dilling et al., 2015). A web page with speech output is deemed to make it easier for people with dyslexia to gather information.

[1] http://www.readspeaker.com/de/
[2] https://www.linguatec.de/text-to-speech/
[3] https://narando.com/

- *Illiterates:* Illiterates often encounter serious restraints in work and private life. They are often barred from text-based knowledge, unless a familiar person is available for reading and writing. This might lead to a heavy dependence on this person (Döbert and Hubertus, 2000). Self-voicing web pages could help to motivate illiterates to use the internet, which might increase their self-confidence and independence.
- *Language learners:* Language learners learn German as a second or foreign language. Language learners can profit from contacts to native speakers and from personal conversations as long as the language levels are similar (Quetz, 2002). Reading has the benefit that the reader can choose the rate of information consumption but possibly misses out on learning about intonation and pronunciation (Burwitz-Melzer et al., 2016). Therefore, self-voicing web pages can be useful for language learners. The reader can choose their own reading rate, and at the same time gets a chance to learn about intonation and pronunciation of the language.
- *People with no reading difficulties:* People from this user group do not have any difficulties or barriers regarding literacy. So, why did we include this user group in our study? Self-voicing web pages can be a useful support for everybody in their everyday life. Especially, difficult texts are easier to understand when the reader hears the words in addition to reading them (Grzesik, 2005). In some situations, in which the external circumstances do not allow for reading, it is useful to have access to the text in audio format.
- *Visually impaired people:* People in the visual impairment user group have a visual disability, but still have some residual sight left. The causes of visual impairments and their manifestations are diverse (Radtke and Charlier, 2006). For information retrieval on electronic media, various aids are available (Radtke and Charlier, 2006). Speech output can be seen as an additional aid for people with visual impairments.

4 FRAMEWORK DEVELOPMENT

The development of the framework for self-voicing web pages is based on the requirements analysis and includes the design and concept of the framework as well as its prototypical implementation. It contains also technical details about the created prototype.

4.1 Requirements Analysis

The requirements for the self-voicing framework are based on three sources. One source was a literature review resulting in the analysis of the target user groups (see chapter 3), one was the analysis of related work (see chapter 2). The last source were informal interviews with persons from the target groups or substitutes. The goal of the interviews was to gather empirical information about the usefulness of self-voicing web pages from the perspective of the different user groups. Another goal was to find out about the concrete features, controls and menu designs that the users preferred.
The interviewees signed a consent form to agree with the voluntary participation in the interview. Some interview partners belonged to a user group. The others were substitutes, for example a teacher of a language course in lieu of a language learner. In total, twelve persons participated in the interviews. They were fairly equally distributed over the target user groups, as follows:

- *Children:* One employee of an organization for science workshops for children and one educator for children.
- *Dyslexic people:* One person with dyslexia.
- *Illiterates:* Two employees in organizations for literacy and one participant of a literacy course.
- *Language learners:* Two teachers of language courses.
- *People with no reading difficulties:* A student and an employee.
- *Visually impaired people:* Two persons with visual impairments.

The interviews gathered interesting requirements for our self-voicing framework for web pages. The interview questions were open-ended in order to collect features instead evaluation existing features. Regarding the features, all participants wanted a synchronized text highlighting of the spoken text. The majority also wanted to adjust the reading rate and would like to have the spoken text magnified. Most also wanted the page to scroll down automatically when the spoken text leaves the window. An

overview over the requested features can be seen in figure 1.

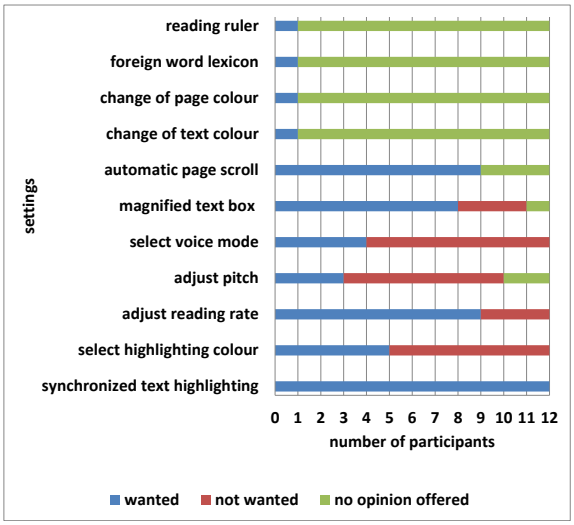

Figure 1: Requested features for speech output on web pages.

All participants wanted to have classical audio controls, known from CD or media players, like play, pause, stop, fast forward and rewind, as well as skip forward and backward. In addition, a majority wanted to be able to adjust the volume.

Another interesting aspect was the difference in the details of some features. As said before, all participants wanted to have a synchronized text highlighting. But in detail, the requirements varied from highlighting per syllabus, word, sentence or paragraph, or a combination of some.

From the analysis of the user groups (see chapter 3) and the analysis of related work (see chapter 2), several features were identified to be implemented in the framework. Most of these aspects were also mentioned in the interviews. Nevertheless, we dropped some features that were only mentioned by one person in the interview.

Based on these surveys and the analysis of the user groups, the following requirements for the self-voicing framework were defined:

- Synchronized highlighting (interviews and related work)
- Adjustment of the highlighting colour (interviews, user group analysis and related work)
- Adjustment of the reading rate (interviews, user group analysis and related work)
- Adjustment of the pitch (interviews and related work)
- Toggling between natural and synthetic voice (interviews)
- Displaying the text line-by-line in a magnified textbox (interviews and related work)

Most of the features itself are already existing in commercial products or in literature, as visible in the list above. The novel aspect of the proposed framework is the use of several settings on a natural voice.

In addition to the requirements for the content and design of the framework, the following technical requirements were identified:

- Fast and simple integration of the framework in any web page at development time. This shall make the framework practically useful in web development, even when time and resources are limited.
- Independence from other frameworks or libraries as much as possible, for better maintainability.
- Accessibility of the framework based on the WCAG 2.0 guidelines on level AA (WCAG Overview).

4.2 Design and Concept

The framework itself was created as an extension of the Accessibility Support Panel (ASpanel) (Research Group Remex, 2015). The ASpanel is a toolbar that can be embedded in web pages, offering features to make a web page more accessible, like changing the text size. The self-voicing framework provides user options for speech output which are presented in an extra tab of the toolbar, as follows: adjustment of reading rate, pitch, highlighting colour, voice mode, text highlighting style and magnified text box on/off.

Currently, the API for the HTML *audio* element does not allow altering the pitch. This means that, for the provision of speech output in a natural voice, multiple audio versions of the same text have to be created at development time, only differing in pitch. We decided to implement a three-value selection for speech rate (slow, normal, fast) and pitch (low, normal, high), in order to minimize the extra developmental effort and costs for creating audio versions for natural voices. This is a compromise between maximal user control (as imposed by the requirements analysis, section 4.1) and optimisation of development effort. Anyway, most interviewees had the opinion that a three-value selection is sufficient to adjust pitch and reading rate, and that no continuous adjustment was needed.

For the text highlighting, we implemented the modes *word*, *sentence* and *paragraph*. These were the most preferred modes by the interview partners (see

section 4.1). It should be noted that the ASpanel with its tabs and settings can be collapsed by the user in order to minimize the space it takes on the web page.

Speech output is started by clicking on a button positioned at the main heading of the text that should be read. The speech always starts at the beginning of the text with the main heading. After the speech output has started, a toolbar with audio controls is inserted underneath the ASpanel and page main menu. The toolbar has controls for start, pause, stop, skip forward, skip backward and volume. The controls are only displayed when speech output runs so that users who do not want to use the speech output are not disturbed by them. The toolbar is fixed in its position, even when the page is scrolled down, to provide easy access to the controls.

When the magnified text box option is on, a box at the bottom of the screen is shown, where the actual spoken text is displayed.

4.3 Implementation

The implementation of the self-voicing framework prototype is based on HTML5, JavaScript and CSS3. No other third-party frameworks were used, to minimize code dependencies.

We implemented speech output based on two technologies: *HTML5 audio element* (W3C Recommendation, 2014) for the natural speech output, and *Web Speech API* (Speech API Community Group, 2012) for synthetic speech output. Both supply basic audio control functions such as start, stop, pause, resume, and adjustment of volume and reading rate. The *Web Speech API* supports the continuous adjustment of pitch (although, for the reasons described in section 4.2, only the three modes low, normal and high are presented to the user). Using JavaScript, we implemented the other requirements (see section 4.1) by ourselves: skipping backward and forward, pitch adjustment for natural voice, magnified text box, synchronized text highlighting, toggle between natural voice and synthetic speech without losing the actual reading position.

5 EVALUATION

We developed the self-voicing prototype with two goals in mind: We wanted to validate the usefulness of the framework for the persons of our target user groups (see chapter 3). And we wanted to investigate how diverse the users' preferred self-voicing options are. In other words, we wanted to see if a *one-size-fits-all* approach is adequate for self-voicing, or if self-voicing needs individualisation to be truly useful.

For the evaluation, a usability test was combined with an interview. Both were described in a test plan, based on criteria by Rubin and Chisnell (Rubin and Chisnell, 2008).

5.1 Participants

The evaluation was carried out with ten participants. They were fairly evenly distributed over the different user groups, as follows:

- For the group of *children*, two children (second and fourth grade of primary school) and an educator for children participated. Regarding the special needs and legal requirements for the evaluation with children, more details are given in section 5.3.
- One person with *dyslexia* participated, and one employee of an organization for literacy.
- Two exchange students volunteered to participate for the group of *language learners*.
- One German student and one employee participated for the user *group with no problems* regarding reading and language.
- One person with *visual impairment* participated.

The participants were recruited by contacting several organizations who represent the different user groups and asking for volunteers.

All participants were volunteers and signed a consent form to agree to their participation and to the collection of their data.

Since the evaluation was qualitative, the number of ten participants should suffice to find approximately 90% of the errors in usability (Virzi, 1992). However, we realize that the discovery level may be lower due to the heterogeneity of the set of participants.

5.2 Structure and Content

The evaluation was structured in two parts. First, the participants had to work through a set of practical tasks. The participants were asked to think aloud while working on a task. We wanted to know whether the features and controls are understandable and the framework is easy to use. Also, we wanted to record the users' preferred settings, and compare them to each other.

Second, we asked a set of interview questions to get more insights into the problems that occurred during working on the tasks. The interview should

also explain the motivation for the choice of the selected features in the first part.

5.3 Special Adaptions for Participants

For two of the user groups, the evaluation was altered to better fit the needs of the particular users.

For *language learners*, the language for all documents of the evaluation, for the features in the framework and for the interviews in the evaluation was English because we could not assume a sufficient level of German language. Otherwise, instructions and questions could give rise to misunderstandings and, in the worst case, to invalid data.

For the user group *children*, several aspects of the evaluation were altered. For once, the documents and consent forms were not signed by the children, but by their parents or legal guardians. They were offered the opportunity to observe the evaluation themselves. After the parents or legal guardians had agreed, the process of the evaluation was explained to the children and they were also asked if they were willing to participate. The evaluation itself was a mixture of interview and think-aloud method active intervention (Markopoulos et al., 2008). During the presentation of the prototype, the children were asked for their opinion, for example what they think will happen when a specific control is used. The method of active intervention is more comfortable for children than the classical think-aloud, because the relation of the moderator with the child is similar to those in everyday situations (parent-child). The evaluation questions were also rewritten to be simpler (Markopoulos et al., 2008).

5.4 Collected Data

Various data were collected during the evaluation:

- Task success: number of tasks that were finished successfully. The criteria for *successful*, *partly successful* and *not successful* were defined beforehand (Rubin and Chisnell, 2008; Tullis and Albert, 2013).
- The handling of the tasks was recorded by a screen recorder.
- Written notes were taken during all the evaluation steps.
- Qualitative statements of the participants were noted.

5.5 Results

The evaluation provided hints for a general usefulness of our self-voicing framework, but also revealed some flaws. Most of the tasks were finished successfully. Half of the participants were not sure what effects some features would have. One person was not sure what the voice mode is, three were unsure what effect the text highlighting would have and two were unsure what the magnified textbox was. One participant was not able to start the speech output. All other tasks, i.e. to describe the audio controls, to set the features and to change them to the preferred ones, were finished successfully by all participants.

All participants tested several features before they had found their preferred selection of features. Some used nearly the same as the pre-set ones, some changed them almost entirely. Figure 2 shows all the features as finally chosen in the evaluation.

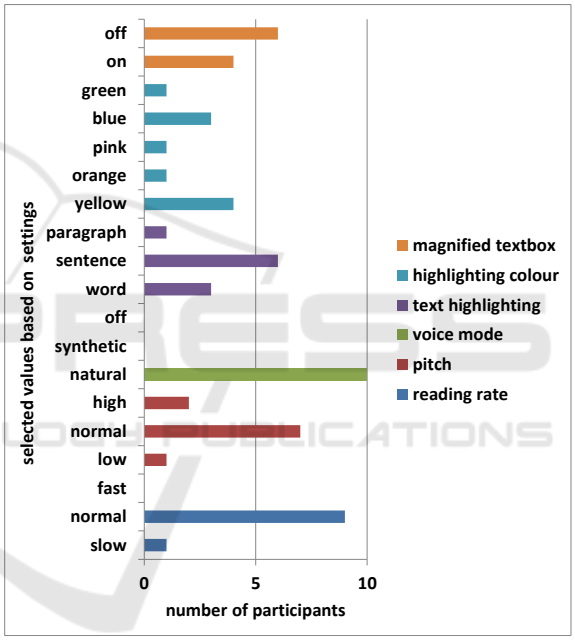

Figure 2: Preferred sets of features.

In general, the range of chosen preferences is wide. Only for the reading rate and voice mode, most or all participants chose the same setting (*normal* and *natural*, respectively). Many wanted to use a slow reading rate but then changed it to *normal*. The widest range was determined for the pitch, type of text highlighting, its colour and the magnified textbox. The different choices with the highlighting colour can be explained by the fact that many participants chose a colour they liked; only one (from the user group children) changed the colour because of difficulties with contrast between text and background. Four (from the groups of visually impaired people, dyslexic people, people with no difficulties with

reading and illiterates) decided to keep the pre-selected yellow colour.

Often similar features were chosen within a user group, but not always. For example, the selections for the type of text highlighting and for the pitch were different within multiple user groups (see Figure 3).

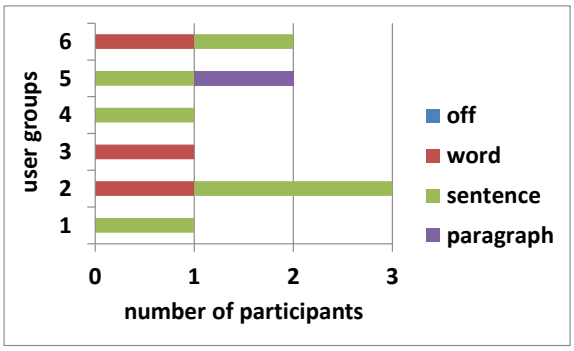

Figure 3: Preferred settings for text highlighting, ordered by user groups. The numbers on the y-axis refer to the user groups in the following way: 1 = illiterates, 2 = children, 3 = dyslexic people, 4 = visually impaired people, 5 = people with no reading difficulties, 6 = language learners.

Six of ten participants preferred the text highlighting by sentence. Only three selected a word-wise highlighting (from the user groups children, dyslexic people and language learners). One participant (from the group of people with no difficulties regarding reading or German language) preferred highlighted paragraphs.

Most participants preferred the normal pitch. Surprisingly, both children preferred a high pitch; it reminded them to a child's voice. The dyslexic participant liked the low pitch best, it reminded them to a teacher and they thought information would be better processed this way.

As a general observation, the preferred settings were peculiar to individual persons, and were partly different from each other even within the same user group. We therefore conclude that person-specific settings (*one-size-fits-one* approach) are most useful to support users in their using the framework, as opposed to a *one-size-fits-all* approach in which common features would be assigned across all users or at least across the users in one user group.

6 DISCUSSION

In general, the evaluation results confirm both hypotheses in a qualitative manner:

- Our framework for self-voicing web pages can support persons with reading difficulties in the use of web pages.
- Different users need different features for self-voicing web pages. A *one-size-fits-all* approach is not suitable.

However, we recognize that our study has some limitations which we discuss in this chapter.

First, the involved users were drawn from separate target user groups which were not overlapping – this is hardly realistic. A child, for example, can also be dyslexic. However, adding users to the study who represent a combination of user groups would have likely enlarged the variety of user settings rather than reduced them. We assume that our framework would have proven useful for "cross-over" users as well, but this is still to be investigated since the overall requirements might have blended more strongly and the evaluation results might have been less clear.

Second, we had persons participating in the analysis of requirements and the evaluation as substitutes for potential end users from the target user groups. Results can be affected by this substitution – we have seen the children's educator choosing different features than the children. However, it is not clear which features are more useful for the children in a real scenario – those chosen by the educator, or those chosen by the children themselves. After all, this would need to be evaluated in a quantitative study, measuring the overall usefulness of the framework in specific use contexts.

Third, the number of participants in the evaluation was too low to make statistical inferences. Some user groups were represented by only one user which does not allow for general assumptions for these groups. Anyway, even with a low number of participants we were able to identify some usability problems.

Fourth, the prototype did not have the level of maturity of a product, and may have prevented some envisioned features to unfold their full usefulness on the evaluation participants. Nevertheless, the immaturity of our prototype should not invalidate the results of the evaluation. On contrary, we assume that a more mature implementation would have been perceived even more favourable.

7 CONCLUSION & OUTLOOK

In our study on a self-voicing framework, we found that self-voicing can be a useful feature for some users, in particular for children, dyslexic people, illiterates, language learners, and visually impaired

people. We even found that people with no reading difficulties were assessing our framework in a favourable manner. Furthermore, our framework is easy to use for persons with basic knowledge about technology use and reading skills on sentence level or higher.

We also found that self-voicing benefits from many options which the user should be able to adjust individually, i.e. according to their personal needs and preferences. This means that a *one-size-fits-all* approach is not suitable for self-voicing support on the web. Different persons have different needs and preferences, and these differ even within the same user group. The evaluation has shown, despite its limitations, that the possibility of personalisation is crucial to the usability of this framework.

In a nutshell, a web page with integrated self-voicing framework is considered a useful addition in the digital daily routine. A speech output can enable access to information on the internet to persons from various user groups. Without the framework, they might be hindered or barred from this information. Therefore, the framework has the potential to support persons in their daily life.

REFERENCES

Burwitz-Melzer, E., Mehlhorn, G., Riemer, C., Bausch, K.-R., and Krumm, H.-J. (eds.) (2016). *Handbuch Fremdsprachenunterricht*, Tübingen: A. Francke Verlag.

Dilling, H., Mombour, W., and Schmidt, M. H. (eds.) (2015). *Internationale Klassifikation psychischer Störungen*: *ICD-10 Kapitel V (F) klinisch-diagnostische Leitlinien*, Bern: Hogrefe Verlag.

Döbert, M., and Hubertus, P. (2000). *Ihr Kreuz ist die Schrift*: *Analphabetismus und Alphabetisierung in Deutschland*, Münster: Bundesverband Alphabetisierung.

Evans, D. A., and Reichenbach, J. (2012). "Need for automatically generated narration," the fifth ACM workshop, Maui, Hawaii, USA, p. 21.

Grzesik, J. (2005). *Texte verstehen lernen*: *Neurobiologie und Psychologie der Entwicklung von Lesekompetenzen durch den Erwerb von textverstehenden Operationen*, Münster: Waxmann.

Hourcade, J. P. (2008). *Interaction design and children*, Hanover, MA: Now Publishers.

Markopoulos, P., Read, J. C., MacFarlane, S., and Höysniemi, J. (2008). *Evaluating children's interactive products*: *Principles and practices for interaction designers*, Amsterdam, Boston, Heidelberg, London: Elsevier Morgan Kaufmann.

Quetz, J. (ed.) (2002). *Neue Sprachen lehren und lernen*: *Fremdsprachenunterricht in der Weiterbildung*, Bielefeld: Bertelsmann.

Radtke, A., and Charlier, M. (2006). *Barrierefreies Webdesign*: *Attraktive Websites zugänglich gestalten ; [berücksichtigt detailliert BITV ; echter Workshop von Analyse bis Relaunch ; im Web: authentisches Praxisprojekt im Vorher-/Nachher-Zustand]*, München: Addison-Wesley.

Rau, M. L. (2007). *Literacy*: *Vom ersten Bilderbuch zum Erzählen, Lesen und Schreiben*, Bern, Wien u.a.: Haupt.

Research Group Remex (2015). *ASpanel*: *REMEXLabs/hdm_banking_default*, https://github.com/REMEXLabs/hdm_banking_default/tree/new-changes/ASpanel. Accessed 14 June 2017.

Rubin, J., and Chisnell, D. (2008). *Handbook of usability testing*: *How to plan, design, and conduct effective tests*, Indianapolis, IN: Wiley Pub.

Speech API Community Group (2012). *Web Speech API Specification*, https://dvcs.w3.org/hg/speech-api/raw-file/tip/speechapi.html. Accessed 19 June 2017.

Tullis, T., and Albert, B. (2013). *Measuring the user experience*: *Collecting, analyzing, and presenting usability metrics*, Waltham Mass. u.a.: Morgan Kaufmann.

Virzi, R. A. (1992). "Refining the test phase of usability evaluation: how many subjects is enough?" *Human Factors*. Vol. 34, No. 4: pp. 457–468.

W3C (2014). *Gap Analysis*: *W3C Editor's Draft 05 September 2014*, Kapitel 6.7 Dyslexia, https://w3c.github.io/wcag/coga/gap-analysis.html#dyslexia. Accessed 11 June 2017.

W3C Recommendation (2014). *4.7 Embedded content — HTML5*: *4.7.7 The audio element*, https://www.w3.org/TR/html5/embedded-content-0.html#the-audio-element. Accessed 19 June 2017.

WCAG Overview | Web Accessibility Initiative (WAI) | W3C, https://www.w3.org/WAI/intro/wcag. Accessed 14 June 2017.