

# Patient Empowerment Through Digital Voice Pens: A Feasibility Study

Tobias J. BRIX<sup>a,1</sup>, Jonas HOFFACKER<sup>a</sup>, Philipp NEUHAUS<sup>a</sup>,  
Martin DUGAS<sup>a</sup> and Michael STORCK<sup>a</sup>

<sup>a</sup>*Institute of Medical Informatics, University of Münster, Germany*

**Abstract.** Reading is an important ability, especially for patients during their medical treatment. It is needed, for instance, to complete administrative forms and patient-reported outcome questionnaires in clinical routine. Unfortunately, not every patient is able to read caused by illiteracy, low vision or simply speaking another language. Thus, a minder is required to support the mentioned reading tasks. Providing patients with the possibility to read and understand texts without additional help is an important factor to improve their self-empowerment. Digital voice pens can be programmed to play prerecorded audio files if tipped onto predefined areas of interactive paper. They can be a tool for impaired patients to read texts aloud in multiple languages. In this work, we wanted to evaluate the possibilities of these digital voice pens. A feasibility study was conducted by using the commercially available *tiptoi* digital voice pen by Ravensburger AG and the *tttool* application by Joachim Breitner for the programming of the pen. Focusing on the use case of questionnaires, a schematic questionnaire was implemented which enforced the usage of a digital voice pen. To simulate foreign languages or illiteracy, questions and answers of the document were represented by placeholders and the digital voice pen was required to read aloud the question texts. The correctness of the given answers was documented and the usability of the digital voice pen was measured by the System Usability Scale. The evaluation was performed by 15 volunteers (8 male/7 female) between 24 and 35 years old. The usability and acceptance of digital voice pens were rated as “Good” in our constructed setting.

**Keywords.** Digital voice pen, patient empowerment, patient-reported outcome

## 1. Introduction

Completing administrative forms and questionnaires are common tasks of patients in modern medicine. These tasks require the patient’s ability to read questions and their associated answer options. Illiteracy and low vision are a huge hurdle for patients in medical routine. For the year 2010, Pascolini and Mariotti estimated that globally 246 million people are suffering from low vision [1]. In Germany, approximately 7.5 million people have very limited reading and writing skills [2]. However, even if no medical condition or educational shortfall is present, reading can still be an insurmountable challenge, if patients do not speak the language of the country in which they are treated. From 2008 to 2017, the Internal Displacement Monitoring Centre (IDMC) listed 246 million refugees caused by geophysical and climate-related disasters not taking war and crime refugees into account [3].

---

<sup>1</sup> Corresponding Author, Tobias J. Brix, Institute of Medical Informatics, University of Münster, Albert-Schweitzer-Campus 1, Building A11, 48149 Münster, Germany; E-mail: tobias.brix@uni-muenster.de.

If texts cannot be read, a minder or interpreter is needed to read them aloud or translate them into the required language. It was shown that completing a questionnaire with an interviewer, i.e., a minder, can have an impact on the answers of a patient in comparison to completing the questionnaire on his/her own [4].

Digital voice pens can recognize their position on interactive paper and play sound files based on their position [5]. Frequent applications can be found in the toy industry, to give children the opportunity to read the contents of picture books by themselves at an early age. The ability to read texts aloud through a digital voice pen multilingually would greatly enhance the patient's independence. Especially children and patients with illiteracy, low vision or simply speaking another language could be enabled to provide information about their disease and its course without outside assistance.

The aim of this work is to evaluate the possibilities and functionality of digital voice pens, whether they are a solution for the mentioned challenges of low vision and reading weaknesses.

## 2. Methods

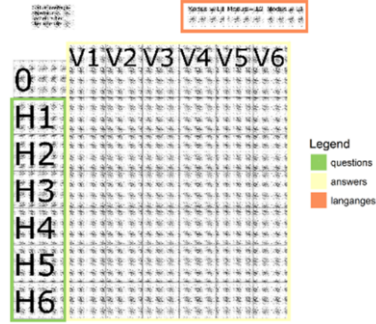
### 2.1. Used Digital Voice Pen and Its Functionality

One of the first publicly available digital voice pens in Germany was *tiptoi* produced by Ravensburger AG in 2010 (see Figure 1) [6]. The pen consists of an infrared sensor on the top and a data storage of up to 3.34 GB in the main body, which can be accessed via USB. By using interactive paper marked with many tags consisting of 4x4 matrices formed by small black dots that do not occupy the underlying image, the pen can read the matrices and decode them into internal IDs. Based on a script in the data storage, each ID can trigger an action like playing an audio file or switching between languages. An example of stickers marked with matrices by Ravensburger can be found in Figure 1, where tipping on one of the small animals triggers playing a sound file.

### 2.2. Used Software for Programming the Pen

Unfortunately, Ravensburger does not provide a framework or toolset to create self-made applications for its pen. Therefore, the exact functionality of the pen and the structure of the internal scripts are unknown. However, using reverse engineering, the most important functionality of this pen has been decoded. Joachim Breitner is leading the decoding of *tiptoi* and has developed *tttool* [7]. This command-line based tool is able to generate the matrix codes required for the interactive paper and to transform a script written in a custom YAML markup language's syntax into the pens vendor-defined binary format, i.e., a file with GME ending. Therefore, a YAML script following this syntax has been implemented.

To combine the matrix codes from *tttool* with elements like images or texts, we used the open-source software GIMP [8]. GIMP was chosen since its command-line usability enabled us to automate the process of interactive paper generation by scripting without the need for manual configuration. However, nearly any other imaging software can be used. Most important is the capability of defining a paper format in 600 dpi without scaling the matrix codes.



**Figure 1.** The tiptoi digital voice pen and an example of an interactive paper by Ravensburger.

**Figure 2.** Questionnaire used for the evaluation, using placeholder to simulate illiteracy.

### 2.3. Evaluation Setup

A questionnaire in a table like structure was used to evaluate the feasibility of the digital voice pen (compare Figure 2). In the table, questions were placed in the first column of each row and a 6-level Likert scale was placed next to it. To simulate illiteracy or a foreign language, only placeholders were used and no real question or answer texts were printed. Therefore, the participants were forced to rely on the voice output of the pen. Instead of asking subjective questions, the questions were designed to determine a single correct answer like “Please mark answer V3”.

To evaluate the multilingual capabilities of the pen, we simulated different languages, i.e., modes in the top right corner of the questionnaire. Some questions asked the user to switch the mode. Each answer of a participant had to be marked with an additional pen, since tiptoi is not capable of writing.

After completing the questionnaire, the participants were asked to complete the German translation of the System Usability Scale (SUS) [9] and give their subjective feedback on the technique. The evaluation was performed by 15 volunteers (8 male/7 female) between 24 and 35 years old. All volunteers were from the field of medicine or computer science and had no previous experience with digital voice pens.

### 3. Results

### 3.1. Implementation of the Questionnaire

A simple YAML script was implemented that played a prerecorded instruction based on the position on the questionnaire where the user tipped the pen. The first column (questions) had their own audio for each row where the columns labeled with ‘V’ (answers) had a single audio. Since three languages were simulated, each script entry was associated with three different audio files based on the current state. All IDs used in the script were generated by ttool. Each ID image was doubled (not scaled) since the default image size from ttool was too small for our use case. In a GIMP script, the matrix IDs were arranged in a table structure as shown in Figure 2. The text was rendered on top of the matrix codes.

3.2. Evaluation Results

The correctness of each participant’s answers can be seen in Figure 3. All 15 participants answered the first and last question correctly. Questions H2, H3 and H5 had missing answers. H3 and H4 had wrongly given answers. The SUS had an average value of 75.54, which equals “Good” usability according to Bangor et al. [10].

Three major subjective feedback statements on the technique can be extracted which were given by more than a single participant. These statements are complaining about problems reading the IDs with the pen, in contrast to having no problem after a short familiarization and complaints about double reads of the pen, i.e., starting the audio for the first read, stopping it immediately and playing it again for the second read.

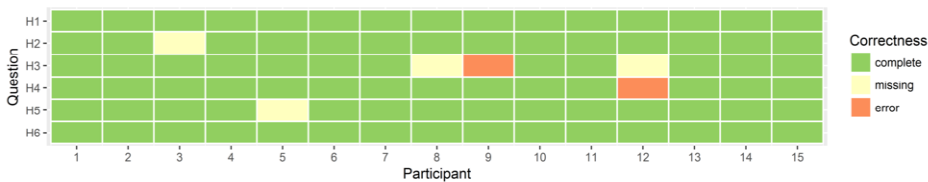


Figure 3. Correctness of evaluation answers sorted by participants.

4. Discussion

4.1. Evaluation Results

The evaluation has shown “Good” usability and acceptance in average. However, we have to keep in mind that the participants were of a young age and may be more familiar with new technologies. Nevertheless, it was shown that elderly patients could handle digital pens for their assessment of pain in a diary [11].

All errors that happened during the completion of questions H2-H5 can be tracked down to our evaluation design. These questions asked the user to switch the mode and reselect the question to hear the correct audio file. As written in the feedback, many users were confused by the mode change because they forgot which cell (question) they read recently and mixed up rows. Since the selection of language normally happens at the beginning of a questionnaire and the changes during the execution were forced by our design, this should not be an issue during normal application.

The subjective feedback statements of some users complaining about problems reading the IDs with the pen in contrast to others stating they had no problem after a short familiarization implies that the correct handling of the pen seems to be a process, which is learned at different speeds. These six questions seemed not to be enough for all participants to get the handling right.

The last statement about double reads of the pen is a known issue and was especially problematic in our evaluation setting, since after switching the language mode, the original success message was canceled and the “this language is already selected” message was played. This resulted in confused users. It is possible to negate this effect in the YAML script, but doing so is not recommended by the tttool community since it could have side effects on newer releases of tiptoi.

#### 4.2. Limitations

The most important drawback of digital voice pens is the missing capability of writing, which requires a second pen for marking tasks. Other digital pens capable of writing do exist as used by Lind et al. [11]. A combination of both techniques would be beneficial.

In addition, the limited storage of the pen could be an issue based on the use case. Overall, approximately 24 hours of audio can fit onto the internal storage. This should not be an issue for a single application, but can be problematic, if multiple applications should be stored on the pen at the same time. Although there is per se no limit on supported languages per application, each language requires its own audio files and therefore counts towards the storage limit. However, different applications and languages can be distributed on multiple pens if the storage limit should be an issue.

#### 4.3. Alternative Implementation Possibilities

Besides Ravensburger, there are other vendors of digital voice pens like *TING* by the TING GmbH [12]. However, for all pens on the market, no software development kit is publicly available and third-party software is always required. We chose *tiptoi*, since *tiptoi* had the best documentation and provided all functionality needed for our evaluation.

Handling multilingual questionnaires can also be achieved by using electronic patient-reported outcome questionnaires (ePRO) where a vast variety of ePRO software is available. However, these tools follow a generic implementation and only support a limited number of types of question, e.g., Likert scale. Thus, there is a limitation in degrees of freedom regarding the questionnaire design. Furthermore, these systems address the issue of foreign languages, but rarely illiteracy or low vision.

If no ePRO software is used, questionnaires can be printed in multiple languages, but this is economically not feasible if pre-printed versions should be available. Besides that, the treating physician has a more challenging task evaluating the results without a translation. In addition, the previous issue of illiteracy and low vision still remains.

Another alternative would be the usage of mobile phones, which are commonly available nowadays. By using an app or web service, the device can load and access audio files and the logic can be implemented in nearly any programming language. In previous work, it was shown that mobile phones could utilize barcodes like QR Codes or Data Matrix Codes to transfer complex input sequences onto the phone [13]. These codes can be placed next to questions and can trigger an audio file after scanning. This approach would replace the dependency of the pen with the dependency of a mobile phone. Unfortunately, especially elderly people or refugees often do not have access to mobile phones and may not be familiar with their handling. Furthermore, although not part of our evaluation, using a tool designed for children would naturally serve underage patients.

### 5. Conclusion

Our feasibility study has shown that digital voice pens can be programmed with individual applications and applied to medical use cases. With a SUS score of 75 in the scenario of a schematic questionnaire, the initial usability was evaluated as good. In the

future, we would like to extend the use cases and apply a study with elderly and foreign patients for real test conditions. In addition, the implementation of tools for certain use cases would be beneficial. Although *tttool* helps with generating codes and scripts, the manual configuration of questionnaires is time consuming and automatic transformation would be desirable.

### 5.1. Outlook

As shown in our evaluation, patient-reported outcome (PRO) questionnaires are one possible scenario for digital voice pens.

Another use case would be markings on medication packaging. Instead of generating full-size questionnaires, a simple and good visible sticker containing a single ID can be printed and put on medication packaging. Elderly people can use the digital voice pen to read aloud the name of the drug or the prescribed dosage. This use case may not be relevant for clinics but may be important for private care or retirement homes.

Finally, another use case would be a real-life tooltip as it is implemented in most software tools. Small question mark icons can be printed with an ID and placed in hospital rooms or laboratories to allow patients or new employees to read quick introductions. Each question mark can explain the tool next to it, like how to use the telephone in the patient's room, or who is responsible for questions regarding a certain device.

## References

- [1] Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. *Br J Ophthalmol*. 2012;96(5)
- [2] Grottlüschen A, Buddeberg K. Geringe Literalität unter Erwachsenen in Deutschland. Quenzel G., Hurrelmann K. (eds) *Handbuch Bildungsarmut*. Springer VS, Wiesbaden (in German); 2019.
- [3] Global displacement risk model [Internet]. Internal Displacement Monitoring Centre; c2020 [cited 2020 Mar 15]. Available from: [www.internal-displacement.org/database/global-displacement-risk-model](http://www.internal-displacement.org/database/global-displacement-risk-model)
- [4] Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J Public Health (Oxf)*. 2005;27(3):281–91.
- [5] Signer B, Norrie MC. Interactive paper: past, present and future. *Proceedings of PaperComp 2010*, 1st International Workshop on Paper Computing; 2010; Copenhagen (Denmark)
- [6] *tiptoi* [Internet]. Ravensburg: Ravensburger AG; c2020 [cited 2020 Mar 15]. Available from: <https://www.ravensburger.de/entdecken/ravensburger-marken/tiptoi/index.html>
- [7] *Das tttool-Buch* [Internet]. Joachim Breitner; c2019 [cited 2020 Mar 15]. Available from: <https://tttool.readthedocs.io/de/latest/>
- [8] GNU Image Manipulation Program (Gimp) [Internet]. The GIMP Team; c2020 [cited 2020 Mar 15]. Available from: [www.gimp.org](http://www.gimp.org)
- [9] Brooke J. SUS-a quick and dirty usability scale. *Usability evaluation in industry*. 1996;189(194):4-7
- [10] Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale," *Int J Hum Comput Interact*. 2008;24(6):574–94.
- [11] Lind L, Karlsson D, Fridlund B. Patients' use of digital pens for pain assessment in advanced palliative home healthcare. *Int J Med Inform*. 2008;77(2):129–36.
- [12] Ting [Internet]. Augsburg: TING GmbH; c2020[cited 2020 Mar 15]. Available from: <http://www.ting.eu/>
- [13] Brix TJ, Neuhaus P, Storck M, Soto-Rey I, Dugas M. Reviving 30 Year Old Technology: Lessons Learned from Transferring Patient Data Using Data Matrix Codes. *Stud Health Technol Inform*. 2019;258:90-4.