A MOBILE CONVERSATION ASSISTANT TO ENHANCE COMMUNICATIONS FOR HEARING-IMPAIRED CHILDREN

Toon De Pessemier[1], Laurens Van Acker[2], Emilie Van Dijck[3], Karin Slegers[3], Wout Joseph[1] and Luc Martens[1]

[1]*Ghent University / IBBT, Dept. of Information Technology, Gaston Crommenlaan 8 box 201, B-9050 Ghent, Belgium*

[2]*Hogeschool Gent, Dept. of Engineering, Valentin Vaerwyckweg 1, 9000 Gent, Belgium*

[3]*K. U. Leuven / IBBT, Centre for User Experience Research (CUO), Parkstraat 45 box 3605, B-3000 Leuven, Belgium*

Keywords: Communication Assistant, Hearing-impaired, Context Clarification.

Abstract: Children and young people who are deaf or have hearing impairments do not master the spoken and written language well. Moreover, their friends and relatives do not always speak sign language. In this research, a communication aid is developed on the Google Android mobile operating system to reduce these barriers by converting text into relevant images and movies thereby clarifying the meaning of the text. Users can select text from webpages, e-mail and SMS messages, or text can be generated using speech recognition or optical character recognition as input source. The communication tool analyzes the sentences and tries to understand the context and meaning of the text in order to select the most appropriate visual content originating from online sources like YouTube, Flickr, Google Images, and Vimeo. User tests and focus groups with deaf children and interviews with experts in the field of deaf and hard hearing people confirmed the need for a communication aid and proved the utility of the proposed tool.

1 INTRODUCTION

Communication of deaf or hearing-impaired people with their surroundings is often not optimal. Nowadays, many young deaf get a cochlear implant, a surgically implanted electronic device that provides a sense of sound to a person who is profoundly deaf or severely hard of hearing. Nevertheless, many deaf people (particularly children) still have a language delay, especially when they are born deaf. They do not catch the spoken communications due to their hearing impairment and their vocabulary is limited. This makes it difficult for them to understand written messages as well. The complex language input from the environment is often misunderstood, which may have a negative influence on the education and the communication skills of the children (Fortnum et al., 2002). Sign language is often more accessible, but has the problem that most parents of deaf children are not deaf and sign language is not their native language (Johnston, 2006). These language barriers make communication between deaf children and their families difficult.

Although studies indicated that deaf people are willing to use new technology to improve their communication skills with hearing people (Brière, 1995), so far no fully-fledged communication tools are available to reduce these problems. Likewise, several authors refer to the importance of technology to help reduce the communication problems of deaf people (Dubuisson and Daigle, 1998), (Hotton, 2004).

This paper presents a communication aid, developed on the Google Android mobile operating system, that tries to reduce the communication barriers for people who are deaf or hearing-impaired and who do not master the spoken and written language well. Because especially deaf children have difficulties with communication, they constitute the target group of the communication aid and evaluated the developed application through focus groups. The goal of the application is to assist these people in conversations with hearing people and understanding written texts, thereby improving their skills in reading comprehension as well as making it more fun to learn new words for deaf children.

The remainder of this paper is organized as follows. Section 2 provides an overview of research work related to conversation assistants for deaf or

De Pessemier T., Van Acker L., Van Dijck E., Slegers K., Joseph W. and Martens L..
A MOBILE CONVERSATION ASSISTANT TO ENHANCE COMMUNICATIONS FOR HEARING-IMPAIRED CHILDREN.
DOI: 10.5220/0003934607750780
In *Proceedings of the 8th International Conference on Web Information Systems and Technologies* (WEBIST-2012), pages 775-780
ISBN: 978-989-8565-08-2

hearing-impaired people. In Section 3, the technical capabilities and input methods of the application are described. A number of observations and results based on focus groups are shared in Section 4. Finally, Section 5 is dedicated to our conclusions.

2 RELATED WORK

Although deaf people are a significant segment of the population, limited research is focusing on devices or applications designed to facilitate face-to-face communication between hearing persons and deaf people who use sign language and cannot speak. Companies like Oralys (www.oralys.ca) are providing tools that enable a new type of communication with symbols. Their Communicator Mobile software for Pocket PC is designed for direct communication, but only for communication from the deaf person to the hearing person (Oralys, 2011). The idea is to replace mouse and keyboard or pen and paper with iconic cards. Although the application already contains 3800 icons, users' communication is restricted to these predefined icons. The deaf person selects icons which indicate typical objects or situations by tapping the touch screen. Next, the series of icons is converted into a voice message in English, French, or Spanish, which is played aloud on the device.

The influence of such an assistive technology designed to facilitate face-to-face communication between deaf and hearing persons on social participation was evaluated by a pilot study with deaf persons (Vincent et al., 2007). In this research, 15 deaf adults completed a three-month field study, with pre and post intervention measures. The results showed a significant improvement concerning the social participation and conversation with a hearing person after using the assistive technology.

Related to a conversation assistant for deaf people is the reading assistant that supports people who suffer from dyslexia. These reading assistants allow users to enter phrases and click the words that are unclear in order to get the meaning of a word. These explanations can be completed with pictures visualizing the meaning of the word and spelling errors can automatically be corrected. In this research, a trend was noted for slower readers to show an increased reading rate as a function of computer-assisted reading, with the opposite result for faster readers (Sorrell et al., 2007).

For video conversations on desktop computers, deaf people mainly use 'ooVoo', an application similar to Skype that enables video chat in HD quality. Deaf people have already tried to converse using sign language via the mobile application, but technical limitations of mobile devices and cellular data networks hinder a smooth conversation. Since subtle differences in gestures are important to interpret sign language, fluent and high quality video is essential for deaf people. Besides, mobile phones introduce an additional difficulty regarding conversations using sign language: users have to hold the device while they make gestures to converse.

3 FUNCTIONALITY

The functional requirements of the application are derived from interviews with experts and teachers of deaf children. Since deaf children have a poor knowledge of grammar and experience often difficulties to distinguish similar words, an explanatory dictionary based on pictures and videos would be useful according to the interviewees. These people emphasize a healthy balance between on the one hand a simple and easy to use application and on the other hand additional features that maybe useful for the children.

3.1 Input Methods

The developed application can assist deaf children in interpreting sentences during daily communication activities. Before the analyzing and interpreting process can start, these sentences have to be entered into the application. The first and most basic input option is to manually enter text on the device, just like the way users type SMS messages. After selecting the option "Keyboard" in the main menu, users get a text box in which they can type the sentence that needs clarification. After tapping on "Translation", the text is sent to the server for analysis using the data connection of the device, as discussed in Section 3.2.

The second option is to use the application as a web aid. If the option "Internet" is chosen, an Internet browser is launched in which users can navigate and select (a part of) a sentence or a web form that is unclear by using the "Selection" feature of the application menu. Based on punctuation, individual sentences are distinguished before the analysis starts.

Users can rely on the third option, speech recognition, for conversations with hearing people. If the speech recognition is successful, the spoken sentence is converted into text, rendered in the interface, and sent for analysis to the server. Our application uses the speech recognition software of the Android Operating System, which is available in different languages such as English, Japanese, Chinese, etc. Since Dutch is the mother tongue of the children who tested

the application, the application was configured to use the Dutch version of Android's speech recognition.

For billboards, menus in restaurants, information signs, and other written texts that users may encounter during their daily-life, the fourth option may be used. Users can take a picture of the incomprehensible text using the camera of their mobile device. The software automatically focusses the image and subsequently sends it to an OCR (Optical Character Recognition) service which converts the image into readable text. Next, the recognized text can be sent to the server for analysis or users can opt for taking a new picture if the OCR was not successful due to e.g. a blurry picture. In our application, the OCR software of WiseTrend is used (WiseTREND, 2011), but alternative solutions are easily to integrate. This OCR service is available for different languages (we used the Dutch version for testing the application) and provides features like deskewing, removing texture, automatic rotation of the picture, and even detecting barcodes.

If deaf children are experiencing difficulties to understand a received SMS message, the fifth option can be used. An SMS message or a part of the text of the SMS can be selected for transmission to the server that analyzes that piece of text. Similar, (an extract of) an e-mail can be selected and sent for clarification to the server via the last input option.

3.2 Sentence Analysis

An important aspect of the analysis of the sentences that users submit is tracing the most important words, and thereby the context of the sentence. Different solutions are possible to detect the most important words of a sentence: using dictionaries to get the word class, part-of-speech taggers that get the word class in a more intelligent way by analyzing the sentence, frequency tables that provide information about how many times a word is used in the language, and thesauri that contain a set of related words (such as synonyms, hyponyms, and antonyms).

The current implementation of our conversation assistant uses a dictionary and a part-of-speech tagger. Frequency tables and a thesaurus are features for future versions of the application to improve the text analysis. For the English version of the conversation assistant, the English version of WordNet can be used (this is a lexical database developed at the University of Princeton). However, because of the license fee of the Dutch version, WordNet was not incorporated in the current version of the conversation assistant that was evaluated by potential end-users.

OpenTaal is a project about spelling, hyphenation, thesauri, and grammar for the Dutch language (OpenTaal, 2011). By harvesting sentences from government websites and online newspapers, the project has gathered a database with speech information, examples, and derivations for about 130,000 Dutch words. Although this database is not yet publicly available, the contributors of the OpenTaal project put the database at our disposal for the retrieval of speech and word information in the conversation assistant. Based on this database, it is possible to configure which types of words have to be retrieved from the entered text (e.g., nouns and verbs can be used to search for explanatory pictures or videos). For performance reasons, the database is locally stored on the device of the end-user in the current implementation, but an online version is also possible. Regrettably, some words can have a different interpretation and word class depending on the sentence in which the word is mentioned. E.g., "minor" in the sentence "Kids under 18 are considered minors" is a noun, whereas in the sentence "I had a minor accident" minor is an adjective and has a different meaning.

This difficulty is solved by utilizing a part-of-speech tagger, a technique which assigns attributes to the words of a sentence and derives the word class based on information from the entire sentence. The part-of-speech tagger 'Frog', which is used for this purpose, provides more accurate results than the analysis based on merely the database of OpenTaal. Frog is an integration of memory-based natural language processing (NLP) modules developed for the Dutch language. Frog's current version will tokenize, tag, lemmatize, and morphologically segment word tokens in Dutch text files, and will assign a dependency graph to each sentence (Frog, 2011). Because of several software dependencies, we opted to deploy the software on a server and send requests for analyzing sentences from the mobile device.

3.3 Filtering and Watching Multimedia

After analyzing the text and filtering out unimportant words, the remaining words or groups of words are used to query various content sources via publicly-available APIs. In the current implementation, different sources are used to find images and videos: Google Image Search, Flickr, YouTube, and Vimeo. Moreover, new content sources, filters and hint-generating systems (which are explained later) can easily been integrated because of the build-in plugin structure of the application. These content sources are evaluated and receive a rating to determine the order of the results for subsequent queries. This rating consists of two elements: an explicit score that the user has set in the settings menu and an implicit score de-

pending on the number of times an item of that content source is marked as "an explanatory image or video".

The multimedia files that are most likely to give an explanation for the entered text are displayed at the beginning of the list. Figure 1 shows a screenshot of the application which visualizes this list of content items that are relevant for the user. Only one image or video at a time has the focus and is prominently displayed in the user interface. Users can navigate to the next image or video by swiping their finger from the right to the left of the screen, or to the previous image by swiping from the left to the right of the screen.

At the top of the screen, an overview consisting of thumbnails of the resulting media files is visualized for the end-users. At the bottom of user interface, hints in text form are displayed. These provide additional tips to understand the meaning of the sentence, e.g., the infinitive of a conjugated verb. New tips appear, until the last tip is shown. These tips can be hidden by tapping them with your finger.

If users believe that a picture or video perfectly illustrates or clarifies the text, positive feedback can be provided by tapping the green thumb in the user interface. Tapping this green thumb let it disappear from the displayed item in the user interface and registers a positive evaluation of the content source that provided the item. This feedback ensures that items from this content source appear more prominent at the beginning of the resulting list for future search queries. As alternative, media sources delivering confusing or unpopular media items can be removed to personalize the result list.

The main advantage of the developed communication aid, is the possibility to extended and modify the application easily. Figure 2 shows the high-level architecture of the application and the services that are used to process the input, analyze the texts, and provide explanatory content. The current implementation of the conversation assistant relies on Google's service for the speech recognition feature. Because of the limited computational resources of mobile devices, converting pictures into text using OCR is also done on an external server. Next, the obtained text is analyzed by using a dictionary and a part-of-speech tagger. In the current implementation, word information is searched via the dictionary that is stored on the device. Alternatively, this dictionary can be made available as an online service to save a few megabytes of storage capacity on the mobile device. The utilized part-of-speech tagger is more resource demanding than the dictionary and is therefore deployed on a server. Based on the results of this analysis, explanatory information is retrieved from online video and

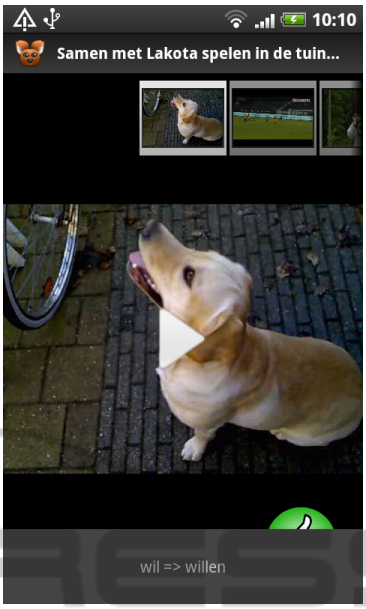

Figure 1: A screenshot of the application showing the sentence that is unclear for the user on top, the retrieved multimedia content in the middle, and additional tips for the user to understand the sentence at the bottom.

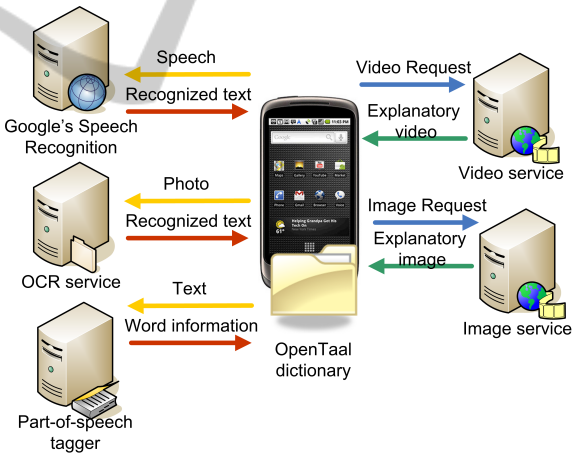

Figure 2: The high-level architecture of the application and the services that are used to process the input, analyze the texts, and provide explanatory content.

photo services. Each of these services can simply be replaced by an alternative service with similar functionality. Additional input methods, such as a service that can interpret sign language, text processing services, such as thesauri, or content sources, such as video services for deaf people, can easily be added to the application in the future.

4 EVALUATION

4.1 Setup

An informal subjective evaluation of the usefulness of the application was performed by means of individual interviews with experts and three focus groups (two focus groups with deaf children and one with experts). The first focus group was organized with four deaf pupils between 12 and 16 years old in a school for deaf children who need a special secondary education. Four deaf children between 8 and 12 years old participated in the second focus group. The third focus group was organized for experts in the field of deaf children such as teachers and employees of schools for deaf children. Three people who work with deaf children on a daily basis participated in this focus group: the principal of the school, an occupational therapist, and a speech therapist.

Before each focus group started, the application was introduced by a slide show with screenshots of the applications and verbal instructions which were translated for the deaf children by a sign language interpreter. The features of the conversation assistant were briefly discussed but not demonstrated. Next, the children were asked where, when and with whom they would use a conversation assistant and which features are useful according to them. Finally, a mobile device was given to each child to try the application. During this test, we guided the children by projecting the user interface of the application using a pc and an Android emulator, thereby showing the successive steps required to achieve the desired output. The children were given some small assignments to illustrate the situations in which the application can be useful. Examples of these assignments are "searching for the meaning of a word in an SMS message or e-mail on the device", "searching an explanation for a word that they do not understand on a webpage" and "searching for the meaning of a word of a recipe in a cookbook by taking a photo of the book".

To investigate if children are able to derive the meaning of a phrase or a word by using the conversation assistant, each assignment had the following structure. Before using the conversation assistant, the children were firstly asked to specify the meaning of a difficult word (or what they think the word means) by a short questionnaire. Then, they could use the application to search for visual clues explaining the word; and finally they were asked again what they think the word means. After each assignment, the children were asked if they could perform the task without problems and if the application was a good assistance for them during the search for the meaning of the word. The last part of the experiment was the actual focus group in which the children could discuss if the application has met their experiences, if they would use the application (and in which situations), if they like the application and the input options, the usability, and possible improvements of the application.

4.2 Results

The children quickly mastered the touch screen and operation of the mobile phone and the application. They learned even faster how to use the application than hearing adults who used the conversation assistant. They liked the application and the visual elements supporting the navigation (the icons, and the thumbs for providing feedback). The questionnaire showed that the pictures and videos helped the children to understand the meaning of difficult words in the assignment. Moreover, the application made it fun and attractive to learn new words based on visual content.

According to the test users, the biggest advantage of the conversation assistant is the easy input method. Entering texts by selecting an SMS or e-mail is considered as very useful. Also the textual information about some words (e.g., the infinitive of a verb conjugated in the past) is considered as useful for some children; for other children this feature might provide unnecessary information.

Since, many young deaf do not know which part of the sentence they do not understand; they cannot put their finger on the important words. Because of this, they believe that the developed conversation assistant is a more useful aid during communication than Google's search engine. If end-users enter a whole sentence into Google's search engine, they get a lot of pictures which do not explain the context. In contrast, the communication aid filters out the important words and combines them. That is why it is experienced as more efficient then a regular search engine. Furthermore, the children (and experts) identified various scenarios in which such a conversation assistant would be useful such as shopping and studying.

In contrast to the enthusiasm of the children when trying the application, various disadvantages were identified by the experts and even by the children themselves. They argued that a photo or a video is sometimes not sufficient to understand a word or a sentence. Irrelevant search results are still possible and may also introduce confusion during the interpretation, especially for abstract words. Moreover, it is difficult to explain the structure of the sentence or the syntax by means of visual content. Another disadvantage that was mentioned is the loading time of a

search request. A slow data connection or the file size of visual content is responsible for a waiting time up to about 30 seconds.

Finally, the experts provided various suggestions to extend the current application. To eliminate irrelevant pictures of the traditional content sources, experts plead for a database that link each word to two or three clarifying pictures. Such a database can be filled by teachers before the beginning of a themed lesson for instance. According to the experts, examples of sentences in which the difficult word is used in different contexts might also be useful. Furthermore, the test users would like to get a movie explaining the word using sign language, next to the current visual content; or even a feature that translates multiple sentences into sign language. Besides the need for a database with explanatory videos in sign language, this introduces the difficulty of the different dialects of sign language that are used.

5 CONCLUSIONS

In this paper, we discussed a conversation assistant for hearing-impaired or deaf people running on a smart phone. Users can input texts by using the keyboard of the device, taking a picture and using OCR, recording speech and using speech recognition, or selecting phrases from a webpage, e-mail, or SMS message. These texts are analyzed and the most important words are converted into pictures and videos originating from public content sources like Google images, Flickr and YouTube.

Focus groups with deaf children and experts in the field revealed various scenarios in which the application can be useful for deaf or hearing-impaired people. Since not many similar applications exist, we received a lot of positive reactions regarding the development of the conversation assistant. Deaf children who experience serious difficulties with understanding texts and conversations can use the application for filling gaps in their knowledge of vocabulary. Children with better communication skills can use the application to learn more difficult words or understanding conjugated verbs.

Moreover, experts recognized the usefulness of the conversation assistant for other target groups such as people suffering from dyslexia or autism. The results of this research can help to inspire future projects aiming to reduce the conversation barrier between people with a disability and the society.

ACKNOWLEDGEMENTS

This work was supported by the IBBT/GR@SP project, co-funded by the IBBT (Interdisciplinary institute for BroadBand Technology), a research institute founded by the Flemish Government. W. Joseph is a Post-Doctoral Fellow of the FWO-V (Research Foundation - Flanders). The authors would like to thank the schools BuSO Sint-Gregorius and Kasterlinden for providing the opportunity to evaluate the application by focus groups.

REFERENCES

Brière, M. (1995). *Consultation sur les technologies de l'information: Résultats de sondage auprès des organismes et des personnes sourdes et malentendantes.* Centre Québécois de la deficience auditive. Médias Adaptés Communications, Montréal.

Dubuisson, C. and Daigle, D. (1998). *Lecture, écriture et surditeé: visions actuelles et nouvelles perspectives.* Éditions Logiques, Outremont, QC.

Fortnum, H., Marshall, D., Bamford, J., and Summerfield, A. (2002). Hearing-impaired children in the uk: education setting and communication approach. *Deafness and Education International*, 4(3):123–141.

Frog (2011). Dutch morpho-syntactic analyzer and dependency parser. Available at http://ilk.uvt.nl/frog/ .

Hotton, M. (2004). Le comité sur les nouvelles technologies en d'éficience auditive de l'i.r.d.p.q.: projet de recherche sur les appareils de télécommunications pour les personnes sourdes et malentendantes. *Différences, Magazine de l'Institut de réadaptation en déficience physique de Québec*, 5:32–38.

Johnston, T. A. (2006). W(h)ither the deaf community? population, genetics, and the future of australian sign language. *Sign Language Studies*, 6(2):137–173.

Opentaal (2011). Welkom bij OpenTaal. Available at http://www.opentaal.org/.

Oralys (2011). Mobile Communicator. Available at http://www.oralys.ca/en/solutions/talk/mobile_communicator.htm.

Sorrell, C. A., Bell, S. M., and McCallum, R. S. (2007). Reading rate and comprehension as a function of computerized versus traditional presentation mode: A preliminary study. *Journal of Special Education Technology*, 22(1):1–12.

Vincent, C., Deaudelin, I., and Hotton, M. (2007). Pilot on evaluating social participation following the use of an assistive technology designed to facilitate face-to-face communication between deaf and hearing persons. *Technology and Disability*, 19(4):153–167.

WiseTREND (2011). Online OCR API Guide. Available at http://www.wisetrend.com/WiseTREND_Online_OCR_API_v2.0.htm.