Zurich University
of Applied Sciences

School of Applied Linguistics

IUED
Institute of Translation
and Interpreting

Barrier-free Communication:

Methods and Products

Proceedings of the 1st Swiss Conference on Barrier-free Communication

Winterthur, 15th – 16th September 2017

ZHAW School of Applied Linguistics

Institute of Translation and Interpreting

Edited by Susanne Johanna Jekat and Gary Massey

Susanne Johanna Jekat, Gary Massey (Eds.)

Barrier-free Communication: Methods and Products.

Proceedings of the 1st Swiss Conference on Barrier-free Communication (2017)

This work is a joint effort of the ZHAW Zurich University of Applied Sciences and the University of Geneva. It is approved by swissuniversities and funded by federal contributions in the framework of the Project "P-16: Proposal and implementation of a Swiss research centre for barrier-free communication" (2017-2020).

The responsibility for the contents of this work lies with the authors.

2018, ZHAW digitalcollection

DOI 10.21256/zhaw-3000

This PDF document has been made accessible in axesPDF for Word (www.axes4.com) and in PAVE (www.pave-pdf.org).

DOI 10.21256/zhaw-3000

The Editors would like to thank the following for their support and cooperation in this work:

Scientific Review Board

Coordination, Translation, Proofreading, Editing

Luisa Carrer, Nathalie Dreier, Maureen Ehrensberger-Dow, Marina Gerster, Yvonne Guggisberg, Alexa Lintner, Gary Massey, Nancy Nicole-Chiddo, Elana Summers, Verena Tunger

Contact

Swiss Centre for Barrier-free Communication

ZHAW School of Applied Linguistics

Theaterstrasse 15c

CH-8400 Winterthur, Switzerland

barrrierefrei.linguistik@zhaw.ch

www.zhaw.ch/en/linguistics/research/barrier-free-communication

DOI 10.21256/zhaw-3000

Table of Content

DOI 10.21256/zhaw-3000

Susanne J. Jekat & Gary Massey

Preface

The 1st Swiss Conference on Barrier-free Communication was held on 15th – 16th September 2017 at the Department of Applied Linguistics of the Zurich University of Applied Sciences (ZHAW) in Winterthur, Switzerland. It was addressed at researchers, academics, practitioners and members of the target groups, and aimed to provide an international forum for the discussion of recent developments and emerging challenges in barrier-free communication.

The conference was the first of three planned to take place as part of the project entitled "P-16: Proposal and implementation of a Swiss research centre for barrier-free communication" (2017-2020), the main goal of which is to ensure access to higher education for people with visual, hearing or temporary cognitive impairments. The project, which is a cooperation between the Institute of Translation and Interpreting of the ZHAW and the Faculty of Translation and Interpreting of the University of Geneva, is funded by the State Secretariat for Education, Research and Innovation (SERI) and is supported by the Swiss University Conference.

This volume of proceedings offers both a theoretical and an empirical perspective on the state of the art in the field. The authors of the essays collected here present their findings, their own experiences, best practices and desiderata from various areas of activity and research contexts.

Barrier-free communication is a very young and multifaceted research area. Against the background of the UN Convention on the Rights of Persons with Disabilities, and within the paradigm of social inclusion and participation, it aims to explore models and procedures to ensure access to information and training for people with visual, hearing or temporary cognitive impairments. Barrier-free communication is a prerequisite to guarantee everyone universal accessibility to all environments.

In line with the principle of inclusion, research in the field of barrier-free communication should only be carried out in collaboration with members of the target groups. **René Jaun** and **Jonas Pauchard** have experienced first hand what studying with a disability feels like in Switzerland. In this volume, they discuss access to higher education and suggest some practical measures to ensure university campuses are fully accessible for students with special needs.

While research studies currently available in this field are relatively scarce, practitioners themselves offer a wide range of services. **Daniela Eichmeyer** gives an account of the development, techniques and typical settings (remote, semi-remote and on site) of speech-to-text interpreting (STTI) services, focusing in particular on the potential of STTI to facilitate access to higher education for the hearing impaired. **Michaela Nachtrab** and **Sarah Mössner** present a remote speech-to-text interpreting service, also known as "online respeaking", which was recently developed by SWISS TXT for the benefit of students with hearing impairments. Compared to the conventional method of STTI used in educational settings so far, online respeaking offers many advantages but is also faced with challenges. Providing an insight into working practices for audio describing documentary films, **Noura Gzara** examines some of the specifics of the genre and emphasises how effective communication with the client may be of even greater importance for this genre than for others, such as feature films.

The proceedings also provide a rich and varied contribution to a growing body scholarship on barrier-free communication. On the basis of the results of the Swiss Accessibility Study 2016, **Anton Bolfing** identifies opportunities in modern information and communication technology (ICT) for the educational and social inclusion of people with disabilities. Most importantly, Bolfing argues how barriers can be avoided and where urgent action is needed in order to bridge the digital divide. **Valérie Boujon**, **Pierrette Bouillon**, **Hervé Spechbach**, **Johanna Gerlach** and **Irene Strasly** compare two phraselators as possible assistive systems for

multilingual doctor-patient communication. The authors claim that a speech-enabled phraselator such as BabelDr can prove to be very effective in emergency settings when no interpreter is available. **Alireza Darvishy** presents the state of the art of assistive technologies for different types of disabilities. Assistive technologies have evolved dramatically in recent years and will continue to be further developed thanks to major progress in artificial intelligence, machine learning, robotics and other areas. **Christian David** explores the potential of web-based applications for text-to-speech audio description of online videos for the benefit of visually impaired people. He introduces two software tools that can optimise the production process and standardise the playout. **Nathalie Dreier** analyses how music programmes are being subtitled on Swiss television and looks at the challenges posed by this practice. Dreier concludes that more cooperation between subtitle providers and target audience should be attempted in order to provide more adequate access. **Daisy Lange**, a member of the LeiSA Project at the University of Leipzig, describes “Leichte Sprache” in contrast to similar approaches of text simplification (i.e. “einfache Sprache” and “Leicht Lesen”). Based on an extensive corpus of texts, Lange’s investigation indicates that the three approaches can be positioned on a comprehensibility continuum. Finally, **Agnieszka Stawikowska-Marcinkowska** and **Marcin Michoń** offer an overview of audiovisual translation in Poland and discuss access to information and entertainment in the national television. The authors identify interlingual and intralingual barriers in Polish TV and suggest ways to overcome them for the benefit of the hearing and/or visually impaired audience.

We believe that this volume of proceedings makes a significant contribution to the progress of barrier-free communication. We would like to express our gratitude to the authors and all the colleagues who have helped us to prepare this volume for publication.

The 2nd Swiss Conference on Barrier-free Communication will be held on 9th – 10th November 2018 at the University of Geneva, Switzerland. Its main goals will be to further explore accessibility in educational settings, stimulate discussion and inspire future scholarship in the field. We look forward to meeting you there.

Susanne J. Jekat & Gary Massey

DOI 10.21256/zhaw-3000

René Jaun & Jonas Pauchard

Studying with a Disability: Enabling and Disabling Factors

Abstract

René Jaun and Jonas Pauchard are completely blind, but they have both attended Swiss universities. While Mr. Jaun is still enrolled at the Zurich University of Applied Sciences, Mr. Pauchard is now working full-time. Both authors have experienced at first hand what studying with a disability is like in Switzerland. During their time at university, they learned how accessibility or inaccessibility to study materials and services can remove or create barriers, respectively, and how much this variable can influence one's academic success. In this article, the authors outline their experiences in the hope that universities in Switzerland and elsewhere will make even greater efforts towards providing barrier-free learning environments for everyone.

1 Removing barriers: Making studying easier

All over Europe, students with disabilities now commonly attend institutions of higher education. This means that significant progress has been made towards inclusive higher education for students with disabilities. A number of factors make studying easier for people with disabilities, as outlined below.

(1) *Assistive technologies*: a decade ago, assistive technologies had to be purchased and set up individually on special-purpose computers. Nowadays, these technologies are pre-installed on most popular makes of personal computers as well as on smartphones. Thus, it is much easier for people with disabilities as well as university labs to have access to computers that fulfil special requirements.

As computers and smartphones have become more powerful, they have begun to replace older, special-purpose devices. For example, the newer iPhones can be used to read print documents aloud just as efficiently as a general-purpose computer with a flatbed scanner attached to it can. Assistive technology is now becoming incorporated into mainstream devices.

(2) *Accessible mainstream software*: accessibility is actively being addressed in IT departments of higher education institutions all over the world. One popular example is 'Moodle', a free open-source learning management system widely used at universities and other tertiary-level institutions. Moodle actively ensures accessibility to assistive technologies and for people with disabilities, and it encourages content providers to do likewise. Thus, universities may no longer need to build individual solutions for their students with disabilities. They may in fact be able to continue using the systems they already have in place.

(3) *Service desk*: sufficient in-house support is nevertheless an absolute necessity if an institution wishes to become fully accessible to students with disabilities.

An office providing services to people with disabilities is therefore necessary. These services should include:

- Assistance in making documents accessible;
- Services to raise awareness in the academic community and liaise with lecturers;
- Assistance with exam preparation to ensure that special needs are taken into account before and during exams.

2 Disabling factors: Making studying more difficult

On the downside, a number of factors can make studying more difficult for students with disabilities. Depending on the structure of the academic institution and its staff, these factors can be more or less easily avoided or at least mitigated.

(1) Checking accessibility as early as possible

While it is true that in the digital age more and more documents are immediately accessible to people with disabilities, many digital publishers still provide their e-books in inaccessible formats. Having a document available in digital format does not necessarily make it accessible. We therefore urge all lecturers to test the accessibility of their documents before using them in class.

(2) Sticking to commitments

There is no question that a climate of open communication is essential for students with disabilities to succeed at university. Students should talk to their lecturers, communicate their needs in advance and make sure that academic staff commit themselves to providing the course and exam materials in an accessible format. However, it is just as essential that academic staff stick to their commitments. We have experienced situations in which some barriers were recognised early on, a fix was promised and yet it was never applied. Therefore, we recommend quality assurance procedures to be put in place to ensure that the needs of students with disabilities are met.

(3) Accessibility beyond ‘here and now’

Meeting the needs of students with disabilities is important during their studies. Many barriers, however, would not even exist if accessibility was generally taken into consideration and given a higher priority. Let us reformulate this concept as follows: there is absolutely nothing wrong with making sure that a course is accessible even if no students with disabilities happen to be enrolled in it. And people should be encouraged to think about students’ needs not only in the classroom, but also outside of it, i.e. during social activities or in the university cafeteria.

If accessibility is prioritised as a main goal in all aspects of university life, even if there are no students with disabilities currently enrolled, life for those attending university now and in the future will be greatly improved.

3 Helping to succeed

A great deal of progress has been made towards making it possible to study with a disability. Yes, it can be done, and yes, many parties are already working closely together to make sure that students with disabilities have equal access to university campuses, services and courses.

Yet our society is still very far from guaranteeing equal rights to people with disabilities. We know from experience that studying with a disability means continuously making one's needs known, asking and fighting for accessibility to make sure one is not left behind. Doing so requires a special kind of strength in addition to the dedication required for studying at university level.

It is our dream that a fully accessible society will become a reality one day. By ensuring that university campuses are fully accessible for students with special needs, one important step will be taken towards the realisation of this dream.

Authors

René Jaun has been working as an accessibility consultant for more than ten years. Following his work with the Swiss foundation «Access for All», he now offers workshops and evaluations on a self-employed basis. Currently, René Jaun is studying journalism and communication at the Zurich University of Applied Sciences (ZHAW). E-mail: jaunren1@students.zhaw.ch

Jonas Pauchard works at the Swiss Federal Institute for Vocational Education and Training (SFIVET) in Zollikofen. He completed an apprenticeship in the commercial field. E-mail: jonas.pauchard@radioblindpower.ch

DOI 10.21256/zhaw-3000

Daniela Eichmeyer

Speech-to-text Interpreting: Barrier-free Access to Universities for the Hearing Impaired

Abstract

Pioneers of speech-to-text interpreting (STTI) can be found in Sweden and Finland since the late 1970s. Since those early days, this service − with hearing-impaired persons as the primary target group − has developed considerably. The UN Convention on the Rights of Persons with Disabilities was the starting point for European and national legislation to improve the lives of people with disabilities, and it made the increasing use of STTI possible. Academic and other professional training of speech-to-text interpreters and the progressive digitalization have created the basis for modern and effective STTI techniques. In some European countries today, STTI is available nationwide to a broader audience. This article sums up the development, techniques and settings of STTI, especially for its application at universities.

1 How to define speech-to-text interpreting

A speech-to-text interpreter transfers spoken language simultaneously into written target text and, furthermore, includes all additional information that can be perceived by the hearing audience. By doing this, the medium and concept of orality is changed into the medium − and usually also the concept − of written text, or, as Liisa Tiittula calls it: “Mündlichkeit im schriftlichen Gewand” (translated by the author: orality clothed in writing, Tiittula 2006).

2 Intralinguality and interlinguality of STTI

It might be questioned if speech-to-text interpreting, which is usually performed intralingually, is to be considered an “interpretation”. In the author’s opinion, STTI is still a form of interpretation, since it shares several features with simultaneous interpretation, as collected by Gerzymisch-Arbogast (2013):

(1) Simultaneity

(2) Necessity of creating local coherence

(3) Limited possibility of correction

(4) Avoiding thematical breaks in the structure of information

(5) Condensation of spoken language

It is indeed easy to find similar challenges for both simultaneous and speech-to-text interpreting, e.g. in terms of source text characteristics that require serious efforts of interpreters, such as complex language and fast delivery, non-native speakers or dialects, incorrect syntax and grammar, unfinished sentences and lack of coherence. Ultimately, the quality of the source text also affects the quality of the target text if no countermeasures are taken.

From another point of view, Braun (2004) argues conclusively that hybrid texts and multimodal techniques, STTI being one of them, are forms of multidimensional translation as an overall aspect (latin: translatio) of interpretation and translation.

Although, in general, STTI is performed intralingually, interlingual STTI has been requested increasingly frequently in recent times, especially in the case of international conventions. Another variety that has been in demand this year is STTI from oral standard language into written plain or simple language.

3 A brief overview of the history of STTI and STTI techniques employed in Europe

STTI was first performed in Sweden (Norberg et al. 2015: 41), where it has been in use since the 1970s. Finland followed suit in the early 1980s, as then did Norway and the Netherlands. In Germany, STTI only began in the late 1990s, and in Austria in the late 2000s, with professional STTI training since 2003 and 2010 respectively. In Switzerland, the first STTI training took place in 2004 (Gassner 2018: 39).

Regarding techniques employed for STTI, in the very beginning only paper and pen were used, which, of course, meant a high degree of content compression and loss of information. The development of personal computers brought a significant change in the performance of STTI, especially as the size of computers diminished and carrying them around became easier. With the spread of personal computers and the development of text processing programs, the foundation for modern STTI was laid. Those programs allowed the use of shortcuts, which increased the number of characters processed per minute.

The next steps were computer-aided stenography, which is mainly applied in the USA, and Velotype, a multiple-key-strike system invented and developed in the Netherlands, which has some important advantages over computer-aided stenography in terms of costs and training time.

The most recent development (Norberg et al. 2015: 38), the use of speaker-dependent speech recognition, yields excellent results in terms of complete rendering of speech.

However, complete rendering does not necessarily result in “high-quality” rendering, so it is important to emphasize the importance of translating the original speech, “[…] redundant in nature […]” (Van Waes et al. 2013: 17f.), into legible and intelligible written text for the target group.

4 Laws and regulations for STTI services

The United Nations Convention on the Rights of Persons with Disabilities (UNCRPD 2008) constitutes the basis for STTI as a communication aid for persons with hearing disabilities, enabling them to take part in educational, professional and social life. Austria ratified the convention in 2008, Germany in 2009 and Switzerland in 2014 (UNO 2017); national laws were adopted accordingly, as can be seen in the examples below.

In Germany, there is the BGG (Equal Opportunities for People with Disabilities Act) at the Federal level and the Kommunikationshilfeverordnung (regulation on communication aid) at the Federal States’ level, and in Austria there is the BGStG (Law on equal treatment of people with disabilities). Switzerland only ratified the UNCRPD in 2014. Two years later, the first report about implementation was published as scheduled (Bundesrat 2016). The existing Equal Opportunities for People with Disabilities Act dating from 2002 was amended in 2016 and 2017.

In both Austria and Germany, the use of speech-to-text interpreters in schools, at universities and in professional life is regulated by laws and/or regulations, whereas the use of speech-to-text interpreters in social life settings is not regulated by law and therefore not paid for by official bodies. Still, in those two countries, conditions for STTI at schools, universities and in professional life vary significantly.

Art. 3 f. of the Swiss Equal Opportunities for People with Disabilities Act covers education and in-service training. Even though communication aids are not specifically mentioned, art. 2 a. defines as discrimination hampering the use of necessary personal assistance (BehiG 2017). The first report of implementation states in its art. 137 (Bundesrat 2016: 42) that persons with disabilities must have access to special aids or personal assistance, while art. 139 says that any measures beyond compulsory education are still considered to be at an initial stage (Bundesrat 2016: 42) and that further development of such measures is needed.

In Germany, speech-to-text interpreters in education are normally paid by the job centre (Agentur für Arbeit), health insurances, integration offices or local authorities, such as the district, the town or the community, depending on the type of education and institution in question.

In Austria, the ministry for social affairs funds STTI in education.

In Switzerland, the respective offices of the disability insurance ("IV-Stelle") pay.

5 Availability of trained speech-to-text interpreters

The main problem is that the number of STT-interpreters do not meet the demand. In Germany, the situation is better than in Austria or in Switzerland, but there are still only about 150 interpreters, of which 54 are certified by the German hard of hearing association (Deutscher Schwerhörigenbund 2017). According to the association, 3,000 to 14,000 speech-to-text interpreters would be required, depending on the kinds of needs specified in this estimate (Deutscher Schwerhörigenbund 2016).

In Austria, the situation is even less favourable: Only 12 trained speech-to-text interpreters (Transscript 2016) serve the whole country, and 11 of them are located in and around Vienna – in the very east of the country – which leaves the rest of Austria without on-site STTI coverage.

In Switzerland, 12 trained speech-to-text interpreters work (freelance) for the organization pro audito, and there is a pool of 13 "mainly trained" STT interpreters available for remote interpreting, some of whom are among the 12 STT interpreters mentioned above (Gassner 2018: 42).

Due to the lack of interpreters and thus the limited on-site opportunities, both Austria and Switzerland use online services like VerbaVoice and Kombia from Germany, although lately, Swiss Text has begun to offer online STTI in Switzerland.

6 Settings for STTI services

Generally speaking, two different types of settings have been known in the last years: on-site and remote. Recently, a combination of both – the semi-present setting – has become popular, as it offers the advantages of both the on-site and the remote setting.

6.1 On-site STTI

In the case of on-site STTI, the interpreter or interpreters (two or more interpreters work in turns if assignments last longer than 45 or 60 minutes) sit among the audience and the recipient sits next to the interpreters. The interpreters' computers are LAN-connected. By using a platform for co-editing, where the co-interpreter corrects errors or missing parts in the text produced by the active interpreter, they are able to deliver a high-quality output.

Some interpreters still use a switch system in order to change the visualization on an external monitor or a data projector so that the active interpreter can see it. However, this option is not recommended as it does not permit co-editing. Each of the interpreters works on his or her own during their turn, which excludes the possibility of correcting each other and is more likely to result in poorer output quality than what could be achieved by co-editing.

Speech-to-text interpreters usually work in turns of 15 minutes at a time, after which their colleague takes over. A single speech-to-text interpreter works for up to 45 minutes or a maximum of one hour. For longer assignments, as mentioned above, a team of two interpreters or more are engaged. The maximum duration corresponds to the maximum time for a single interpreter of simultaneous interpreting; beyond this duration, quality of interpreting services decline due to fatigue. Moreover, in the long term, the physical integrity of the interpreters could be put in danger (AIIC 2012: Art. 7b, AIIC 2015: Art. 6, Yu 2015: 71f.).

6.1.1 Visualization of the target text for larger audiences

The target text can be shown on different output devices depending on the requirements of the event.

If there is a larger audience of deaf and hearing-impaired persons, the target text is displayed in the form of a continuous text or as subtitles on a big screen, or as continuous text on individual devices. Therefore, different technologies may be employed:

(1) Big screen: (One of) the interpreter's computer(s) is connected to a data projector.

(2) Subtitles on a big screen: Text on Top (ToT) technology is used as an interpreting platform which allows different kinds of visualization, for example, the presentation of subtitles in a customizable number of lines. ToT devices use radio signals for communication and are connected both to the interpreters' computers and to the presenter's computer.

(3) Internet: If an Internet platform is used for interpreting (also an option when working on-site), the recipient(s) can enter the virtual room individually on their devices, e.g. computers, tablets or smartphones. In this case, no big screen is necessary, and only those who need or want to receive a written target text will see the STTI.

6.1.2 Visualization of the target text for one or two recipients

If there are only one or two recipients, visualization of the target text is made available for them as follows:

(1) No additional device: The recipient looks at the interpreter's monitor. This method is still widely used but not very beneficial for either the interpreter or the recipient, as neither of them can look at the text from a straight angle and the situation is somewhat uncomfortable.

(2) External monitor: An external monitor is connected to (one of) the interpreter's computers. If the recipient has a visual impairment, it is important to select an appropriate monitor and font size.

(3) Text on Top: The interpreters' computers and the recipient's (or recipients') device(s) are connected using ToT devices. In this way, the recipient is able to choose an individual format of the target text and can sit anywhere in the room, not necessarily next to the interpreter(s).

(4) Internet: If an Internet platform is used for interpreting, the recipient(s) can enter the virtual room individually. As in the previous setting, the recipient(s) can sit wherever they like.

6.2 Remote (online) STTI

In this setting, the interpreter is located in an office or home office.

The target text, produced by the interpreter, appears on the recipient's end device (computer, tablet or smartphone) by use of a web-based interpreting platform. The platforms used are the property of STTI service agencies and, in general, cannot be used by other companies or independent interpreters. Technical equipment has to be set up on-site in order for the the original speech to be transmitted to the interpreter.

One of the platforms used includes audio transfer, while another uses Skype for audio transmission.

Communication between the recipient and the interpreter is possible via the chat function of the platform, but there is no possibility to communicate with any other participant or with the

speaker. Technical issues have to be solved by the recipient him/herself or by the remote technical support, if any is available.

Up to now, audio quality has not been consistently good enough to allow an excellent performance of interpreting. There are several reasons for this, but the four main factors are:

(1) poor Internet performance on either end (recipient and/or interpreter)

(2) use of low- or medium-end microphones

(3) no use of special audio interfaces that are interposed for sound conversion

(4) low technical competence of the users regarding sound issues

New interpreting platforms are currently being developed, one of them with the explicit objective of accomplishing the requirements of ISO 20108 in order to provide excellent sound quality. This means that changes are expected in this area, but a timeline is not yet available.

6.3 Semi-presence STTI

This setting combines the previously described ones: One interpreter is located on-site and another one is connected online. As in the case of remote interpreting, a web-based interpreting platform is used. The on-site interpreter sets up the technical equipment for sound transmission and is responsible for troubleshooting in case of sound or Internet problems.

6.4 Pros and cons of the different settings

6.4.1 On-site

Pros:

(1) Involvement: The interpreter is fully involved and has access to on-site information.

(2) If there are questions, the on-site interpreter can inquire accordingly.

(3) Direct communication with the recipient(s) ensures that their requirements are met. This makes it much easier to customize paragraphing, the degree of condensation, etc.

Cons:

(1) Environmental noise might prevent the interpreter from hearing everything. The people located next to the interpreter may be chatting, typing and/or turning their pages. Furthermore, the participants of a discussion might not speak in the direction where the interpreter is located. Very often in these cases, the interpreter cannot understand what they are saying.

(2) Bad air: Closed rooms without constant air renewal are characterized by poor air quality, affecting the interpreter's performance (Eichmeyer 2017).

(3) Availability of speech-to-text interpreters for on-site assignments: As there are not enough interpreters available and, furthermore, they are not evenly distributed over different countries, it is very often impossible to find interpreters for on-site assignments.

(4) Traveling and accommodation expenses: Speech-to-text interpreters have to travel to the site of the assignment, so traveling expenses as well as accommodation expenses have to be taken into account.

(5) Efficient logistics are necessary to get the right interpreters to the right place at the right time.

6.4.2 Remote

Pros:

(1) Availability of speech-to-text interpreters: As mentioned above, speech-to-text interpreters are not easily available for assignments at specific locations, while remote interpreters can be hired nationwide or even internationally.

(2) There are no traveling or accommodation expenses for remote interpreters.

(3) The logistics are less demanding.

Cons:

(1) Dependence on a stable high-speed Internet connection. Because they are not on location, remote interpreters normally do not have any influence on the quality and availability of the Internet connection at, for example, universities. Normally, they depend on an open WLAN Internet connection that is affected by peak hours and changing transmission rates.

(2) Impersonal setting: Interaction with the remote interpreter is limited to the chat on the platform. Since the connection is normally established just before the class starts, there is no possibility for communication beforehand. During the lecture, it is hardly possible to communicate, as it would distract the other students.

(3) No interaction with lecturers and other participants: There is no way, for example, to ask the lecturer or the participants to speak more clearly or more slowly, or to remind them to use the microphone for their statements in order to ensure that the interpreter can hear and interpret what was said.

(4) The sound quality is often very poor, and the options and possibilities for troubleshooting are limited or non-existent.

(5) It might be necessary to rent web-based platforms and microphone equipment.

6.4.3 Semi-presence

A semi-presence setting combines the advantages of both settings described above and adds extra benefits:

(1) The on-site interpreter can intervene and has direct contact with the recipients and lecturers.

(2) With a good Internet connection available, the remote interpreter has even better sound quality than the on-site colleague when directional microphones are used. If this is the case, he or she can fill in the parts that the on-site interpreter is not able to understand because of the ambient noise or because of speakers talking too quietly.

(3) As one interpreter is on-site, there is no 100 % dependence on the Internet connection. If the connection fails, the on-site interpreter can still bridge some time until it is re-established.

(4) Cost optimization can be reached by counterbalancing the rental costs of technical equipment and interpreting platforms against the traveling (and accommodation) costs.

(5) Logistics are simpler and the lack of availability of interpreters is less of a problem when only one interpreter has to be present on-site.

6.5 Audio issues and how to optimize sound transmission

First of all, high-end directional microphones should be used. The main speakers (lecturers in the case of university settings) should wear headsets with directional microphones that filter out the ambient noise and allow an isolated transmission of clear audio to the remote interpreter.

Secondly, new or fully recharged batteries should be put into the equipment every day to prevent unwanted breaks due to flat batteries. The on-site interpreter makes sure that the headsets are used properly and that enough new or recharged batteries are available.

Verbal contributions from other students or participants must also be transmitted to the remote interpreter. Therefore, microphones are needed for the audience as well. High-quality wireless handheld microphones can be used for this, but practical experience has shown that it takes a lot of time to get the microphones to the speakers, which interrupts the fluidity of communication. Furthermore, many people feel intimidated and inhibited when forced to use microphones. Another difficulty is that people have to wait a few seconds before they can speak, as the microphone has to be switched on and the connection to the receiver established. When the participant has finished talking, he or she has to switch off the microphone. This is necessary to avoid loud or interfering noises caused by passing on the microphone to the next speaker or by interferences between the different microphones. Normally, people start talking immediately and the first part of the sentence is therefore not transferred to the remote interpreter.

To counter these two issues, “Catchbox” has turned out to be the perfect solution. A high-end Sennheiser microphone is incorporated in a well-protected low-weight cube that can be thrown around from one speaker to the next. The microphone incorporated in the Catchbox is automatically switched off as soon as the Catchbox moves, so participants do not have to switch it on or off manually. Furthermore, the Catchbox appeals to the natural human “play instinct” and is fun to use in class.

One of the most important issues is interposing a special interface between the interpreter's computer and the receiver(s) of the headset and/or handheld microphones and/or Catchbox. The Scarlett 2i2[1] was found to be the best solution to enhance sound quality. This special device is also used for the transmission of live music.

[1] https://focusrite.de/usb-audio-interfaces/scarlett-2i2 (7 May 2018)

This is how to set up the technical equipment for proper sound transfer: The speaker wears a headset (1) which sends the audio signal to a wireless receiver (2). The Catchbox (3) is linked to another receiver (4). Sound cables (5+6) connect both receivers to the audio interface (7). The audio interface is connected via a USB cable (8) to the computer (9), from which the audio is transmitted to the remote interpreter (10).

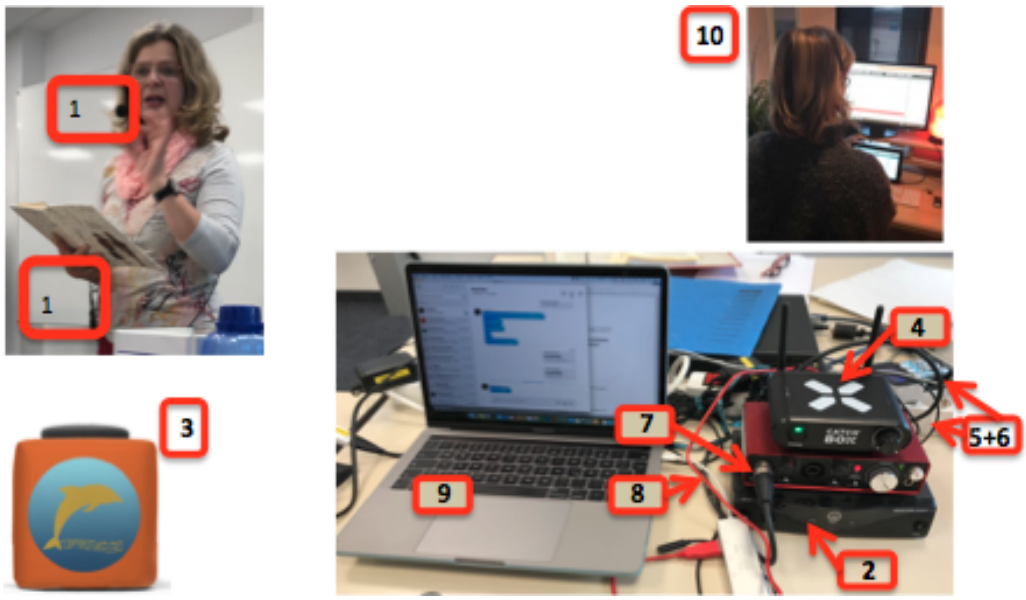

Figure 1: Sound transmission in a semi-presential setting of STTI

If the lecture is held in a large auditorium and microphones are used by the lecturers, it is also possible to take the audio input from a mixing console and connect it to the audio interface connected to the computer.

Of course, there will continue to be a constant development of technical solutions, and, hopefully, the setup described will be outdated very soon; however, to date, it is still the best way to bring good audio quality to the remote interpreter's ears.

7 Conclusion

At present, it is not easy to provide STTI for students at universities, despite the fact that the situation is improving. The first obstacle to overcome is that of identifying who will pay for the service. Then, the interpreters have to be found. The easiest way to do this is to contact an STTI agency. These agencies have both in-house and freelance speech-to-text interpreters at their disposal. However, it is difficult to hire even just one on-site interpreter, and, moreover, to find interpreters who are adequately qualified for demanding lectures. There is another aspect that cannot be neglected: In general, when a student (or any recipient) signs a contract with an agency, they cannot freely choose which interpreter will carry out the task and, in many cases, there is no continuity. This means that the quality of the target texts can vary significantly, depending on the assigned interpreter. Therefore, it is very important to negotiate the conditions according to the specific requirements.

It is a very demanding process to find a suitable team of interpreters for a specific assignment, both in terms of availability and suitability for the job. However, once the right team is set up, this approach can lead to better results and improved management.

As set out above, choosing the right setting is important. In many cases, on-site settings are not available or are too expensive if only a short period of one or two hours must be covered. Because remote settings still have to deal with a lot of issues, a semi-presence setting is recommended for university lectures, especially for interactive classes.

For hearing-impaired students, even when STTI is provided, the task of following a lecture is much more complex than it is for hearing students. They have to look back and forth between the lecturer, the board or the screen where slides are projected, their fellow students, and the device where the target text appears − all the while also having to take notes. If, on top of that, they have to handle all the technical issues, the setup, the troubleshooting and the removal of the technical equipment, they will easily become overtaxed. Furthermore, even if they manage to handle the technical issues, they will probably not be able to pay attention to the lecture and, consequently, their exam results may suffer. For this reason, the semi-presence setting should be the preferred mode of STTI at universities.

As it is not easy to achieve a semi-presence service, online services are a viable solution. Still, one needs to keep in mind the specific requirements of the respective student and try to provide the best solution. There may be the option of combining different settings according to the type of lecture or seminar to be attended by taking into account, for example, the degree of interactivity or the conditions for sound transmission.

Sources

AIIC 2012 (2018): “Code of professional ethics.” https://aiic.net/page/6724 (30 January 2018)

AIIC 2015 (2018): “Professional Standards.” https://aiic.net/page/6746 (30 January 2018)

BehiG (2017): “Behindertengleichstellungsgesetz”. Bundesversammlung der Schweizerischen Eidgenossenschaft. https://www.admin.ch/opc/de/classified-compilation/20002658/index.htm (31 December 2017)

Braun, Sabine (2004): *Kommunikation unter widrigen Umständen? Einsprachige und gedolmetschte Kommunikation der Videokonferenz.* Tübingen: Gunter Narr

Bundesrat (2016): “Erster Bericht der Schweizer Regierung über die Umsetzung des Übereinkommens über die Rechte von Menschen mit Behinderungen.” Initialstaatenbericht_BRK_v1_0.pdf (31 December 2017)

Deutscher Schwerhörigenbund e.V. (2016): “Eine arbeitsmarktpolitische Begründung zur Ausbildung von Schriftdolmetschern in Deutschland.” http://www.schwerhoerigen-netz.de/DSB/SERVICE/SCHRIFTD/PDF/arbeitsmarkt-analyse.pdf (29 May 2017)

Deutscher Schwerhörigenbund e.V. (2017): “Zertifizierte DSB-Schriftdolmetscher.” https://www.schwerhoerigen-netz.de/schriftdolmetscher/?L=0 (4 December 2017)

Eichmeyer, Daniela (2017): *Luftqualität in Dolmetschkabinen als Einflussfaktor auf die Dolmetschqualität. Interdisziplinäre Erkenntnisse und translationspraktische Konsequenzen*. Berlin: Frank & Timme

Gassner, Madeleine (2018): *Live-Untertitelung und Schriftdolmetschen im deutschsprachigen Raum. Eine Untersuchung der Ausbildungen in Deutschland, Österreich und der Schweiz*. Unpublished master’s thesis. Winterthur: ZHAW Zurich University of Applied Sciences

Gerzymisch-Arbogast, Heidrun (2013): “Gutachten zur der Bezeichnung ‘Schriftdolmetschen’.” http://www.bsd-ev.org/fileadmin/gutachten_gerzymisch.pdf (30 January 2018)

Norberg, Ulf; Ursula Stachl-Peier; Liisa Tiittula (2015): “Speech to text interpreting in Finland, Sweden and Austria.” *Translation & Interpreting 7* [3]: 36-49 – http://trans-int.org/index.php/transint/article/viewFile/418/211 (7 May 2018)

Nofftz, Birgit (2014): “Written interpreting in individual countries.” https://www.uni-saarland.de/fileadmin/user_upload/page/atrc/Dokumente/Written_interpreting_in_individual_countries_Birgit_Nofftz_2014_.pdf (13 November 2017)

Österreichischer Schwerhörigenbund Dachverband (2012): “Neuer Diplomlehrgang zur Schriftdolmetschausbildung ab Herbst 2012.” http://www.behindertenarbeit.at/16398/osb-neuer-diplomlehrgang-zur-schriftdolmetschausbildung-ab-herbst-2012/ (29 May 2017)

Ranchal, Rohit et al. (2013): “Using Speech Recognition for Real-Time Captioning and Lecture Transcription in the Classroom.” *IEEE Transactions on learning technologies 6* [4], October-December 2013: 299-311 – https://ieeexplore.ieee.org/document/6529071/ (7 May 2018)

Platter, Judith (2015): *Translation im Spannungsbereich von Mündlichkeit und Schriftlichkeit, Schriftdolmetschen in Österreich*. Eine textbasierte Analyse. Unpublished doctoral dissertation. Vienna: University of Vienna

Tiittula, Liisa (2006): “Schriftdolmetschen – Mündlichkeit im schriftlichen Gewand.” A Man of Measure. Festschrift in Honour of Fred Karlsson, a special supplement to *SKY Journal of Linguistics* [19]: 481-488 – http://www.linguistics.fi/julkaisut/SKY2006_1/1FK60.1.10.TIITTULA.pdf (7 May 2018)

Transscript (2016): “Das ÖSB-trans.SCRIPT-Team.” http://www.transscript.at/team/ (2 January 2018)

UN-Behindertenrechtskonvention (2006): “Übereinkommen über die Rechte von Menschen mit Behinderungen.” https://www.behindertenrechtskonvention.info/ (31 May 2017)

UNO (2017): “United Nations Treaty Collection: 15. Convention on the Rights of Persons with Disabilites.” https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-15&chapter=4&clang=_en (13 November 2017)

Van Waes, Luuk; Mariëlle Leijten; Remael (2013): “Live Subtitling with Speech Recognition. Causes and Consequences of Text Reduction.” *Across Language and Cultures 14* [1]: 15-46 – https://akademiai.com/doi/pdf/10.1556/Acr.14.2013.1.2 (7 May 2018)

Yu, Yujing (2015): *Die Leistungen von Dolmetschagenturen: Prozessrelevante Faktoren aus der Sicht professioneller Konferenzdolmetscher. Eine empirische Untersuchung über die Bewertung der Leistungen von Dolmetschagenturen durch die beauftragten Konferenzdolmetscher am Beispiel der Berlinale 2015*. Unpublished master’s thesis. Munich: Universtiy of Applied Languages SDI München

Author

After graduating as Magister of Business Administration (University of Innsbruck, 1994), followed by over 20 years in international business, Daniela Eichmeyer graduated as a Master of Conference Interpreting (University for Applied Languages Munich, 2015). She is currently preparing a doctoral dissertation on speech-to-text interpreting, explicitely on quality-oriented consideration of the different STTI techniques. She is a full-time practitioner of conference and speech-to-text interpreting and a teacher of STTI. E-mail: de@delfinterpret.com

DOI 10.21256/zhaw-3000

Michaela Nachtrab & Sarah Mössner

Speech-to-Text and Online Respeaking in Switzerland

Abstract

Speech-to-text reporters use different methods for their work. The latest and most commonly used method for TV subtitling is 'respeaking', in which a trained reporter repeats the spoken language input of a TV programme into a microphone linked to a speech-to-text system. SWISS TXT, Switzerland's main provider of the service, has been using respeaking for TV subtitling since 2008. The people working in the field of respeaking have to meet high demands in their challenging task. In 2017, SWISS TXT entered the education sector by offering an online speech-to-text reporting service for students with hearing impairments. As opposed to the conventional method of speech-to-text reporting that has been used in the education sector so far, the online method offers many advantages, but also faces challenges.

1 Speech-to-text reporting in Switzerland

In 2006, the UN General Assembly adopted the Convention on the Rights of Persons with Disabilities (CRPD; UN n.d.). As of writing, over 175 countries have ratified this convention, including Switzerland. The main purpose of the convention is "to promote, protect and ensure the full and equal enjoyment of all human rights and fundamental freedoms by all persons with disabilities [...]" (UN 2006: 3). One human right is the full access to information and communication. The CRPD states that "'[c]ommunication' includes languages, display of text, Braille, tactile communication, large print, accessible multimedia as well as written, audio, plain-language, human-reader and augmentative and alternative modes, means and formats of communication, including accessible information and communication technology" (UN 2006: 4). In Switzerland, the regulations on equal rights for people with disabilities are primarily anchored in the Federal Constitution, in particular, the Disability Equality Act (German: Behindertengleichstellungsgesetz BehiG) and the cantonal laws. Due to the growing need to offer deaf and hard-of-hearing people access to all forms of multimedia communication, speech-to-text reporting (STTR) is gaining in importance.

One of the first definitions of STTR was provided by Stinson and Stuckless (1998: 127) when they wrote that "real-time speech-to-text is defined as the transcription of words that make up spoken language accurately into text momentarily after their utterance". The main aim of STTR is to allow deaf and hard-of-hearing people access to spoken content.

1.1 Methods of speech-to-text reporting

Speech-to-text reporters (STTRs) use different methods for their work. A short overview of the methods is given at the end of Downey's description of the STTR process:

> […] (1) careful attention to an original spoken utterance, either for phonetic structure or semantic meaning or both; (2) cognitive transformation of that utterance along a number of axes, whether condensing for speed, editing for literacy level, translating from one language to another, or simply rephrasing for clarity and convention; and (3) reproducing the modified utterance in a form that can be converted to text, either through manual writing, mechanical keyboarding, or controlled revoicing into a microphone. (Downey 2008: 276)

It is safe to say that manual writing is largely outdated these days when it comes to professional STTR. Mechanical keyboarding, however, is still a widely used method for STTR in Switzerland. Most commonly, STTRs work with a standard keyboard (i.e. QWERTY or a language-specific variant such as QWERTZ used in the German-speaking part of Switzerland) using a conventional word processing system. This is also the method used by the STTRs at pro audito schweiz (see 1.2).

Another possibility is the use of a special shorthand keyboard, often referred to as stenotype (Romero-Fresco 2011: 13ff.). Using this keyboard, STTRs can spell out whole syllables by pressing multiple keys simultaneously (i.e. by using keyboard shortcuts), which considerably increases typing speed.

The latest method of STTR is the transcription of spoken communication into text by means of automatic speech recognition (ASR) software, a process generally known as respeaking (see section 2).

1.2 Areas of application

In Switzerland, the main providers of STTR services are pro audito schweiz[2] and the SWISS TXT AG[3].

Pro audito schweiz was the first and continues to be the main provider of traditional STTR in Switzerland. It offers assistance in all communication situations, such as education, cultural events, business meetings, doctor's appointments, etc. (pro audito schweiz n.d.). SWISS TXT, a subsidiary of the Swiss multimedia company SRG SSR, provides the TV subtitles for SRG SSR and other clients (see section 2.1).

2 Respeaking

In general terms, respeaking is the production of subtitles by means of ASR. A more comprehensive definition has been given by Romero-Fresco (2011: 1):

> A technique in which a respeaker listens to the original sound of a live programme or event and respeaks it, including punctuation marks and some specific features for the deaf and hard of hearing audience, to a speech recognition software, which turns the recognized utterances into subtitles displayed on the screen with the shortest possible delay.

Usually, respeaking is carried out intralingually (Romero-Fresco 2011: 1), i.e. from a spoken utterance into the written form of the same language system. In Switzerland, however, respeakers often translate from non-standard Swiss German dialects into the standard variety of German. This is the main reason speech input from TV shows and lecturers cannot simply be fed directly into an ASR system in Swiss respeaking situations.

Moreover, an untrained ASR system has difficulty recognizing natural spoken language correctly. When the respeakers repeat what has been said in an accentuated manner into their personally trained ASR system, the accuracy of the output highly increases.

2.1 Respeaking for TV subtitling in Switzerland

In Switzerland, the production and distribution of live subtitles started in 1984 with the broadcast of the papal visit (Romero-Fresco 2011: 30). Since then, SWISS TXT has been performing the task for SRG SSR and for other clients (SWISS TXT 2017a). The mandate is based on the Federal Act on Radio and Television (RTVG), which is concerned with the access to audiovisual media content for people with sensory disabilities (Swiss Confederation 2018).

Article 7 in the corresponding Ordinance on Radio and Television (RTVO) states that "[t]he range of content to be subtitled and the scope of the other services to be provided by SRG SSR, as well as the schedule for implementation, are fixed in an agreement between SRG SSR and the disabled persons associations concerned" (Swiss Confederation 2018). According to this mandate, SWISS TXT cooperates with the following Swiss disabled persons associations (SRG SSR 2017): Schweizerischer Gehörlosenbund (Swiss association of the

[2] www.pro-audito.ch (7 May 2018)

[3] www.swisstxt.ch (7 May 2018)

deaf)[4], pro audito schweiz[5], sonos Schweizerischer Hörbehindertenverband (Swiss association for the hearing impaired)[6], Schweizerischer Blinden- und Sehbehindertenverband (Swiss association for the blind and visually impaired)[7], Schweizerischer Zentralverein für das Blindenwesen (Swiss National Association of and for the Blind)[8], Schweizerische Vereinigung der Eltern hörgeschädigter Kinder (Swiss association of parents of hearing-impaired children)[9] and forom écoute (foundation of the hearing impaired for French-speaking Switzerland)[10].

In 2016, SWISS TXT provided subtitles for just under 30,000 hours of programming; 12,815 hours in German, 8,515 hours in French and 8,527 hours in Italian (SWISS TXT 2017b). Overall, roughly 50 % of transmission time was subtitled. In negotiations with the disabled persons' associations in the fall of 2017, it was decided to increase this number to 45,000 hours, or 80 %, by 2022 (SRG SSR 2017).

The main target audience of the respoken subtitles are people with hearing impairment, but L2 speakers of German, French and Italian also benefit from the subtitles (Jekat 2014: 87). On the German-language channels, the subtitles also help people with little or no knowledge of Swiss German dialects to follow the broadcasts.

The SRG SSR broadcasts can generally be divided into the categories of 'recorded', 'semi-live' and 'live'. For recorded programmes, such as feature films, TV series and documentaries, the subtitles are pre-produced and time-coded. Live programmes, on the other hand, are subtitled by means of respeaking. To generate their subtitles, SWISS TXT respeakers use automatic speech recognition linked with a specialized subtitling software. The live subtitles are generally intralingual, although SWISS TXT respeakers working in German often have the additional step of simultaneously translating the Swiss German dialects spoken on TV to standard German, which is the language used with the speech recognition software. For semi-live programmes, such as the *Tagesschau*, the pre-prepared but not time-coded subtitles of the recorded segments are broadcast while the programme is aired. Live segments (e.g. contributions from off-site correspondents) are subtitled in real-time through respeaking.

2.2 Respeakers

Respeaking is a challenging task and the people working in the field of respeaking need to meet high demands. In most European countries, respeaking is part of a BA and/or MA programme in translation (Remael 2007: 40). In Switzerland, the ZHAW Zurich University of Applied Sciences offers a BA in Applied Languages that includes a module on respeaking. Many SWISS TXT respeakers complete this BA programme before undergoing further in-house training at SWISS TXT.

[4] www.sgb-fss.ch (7 May 2018)
[5] www.pro-audito.ch (7 May 2018)
[6] www.hoerbehindert.ch (7 May 2018
[7] www.sbv-fsa.ch (7 May 2018)
[8] www.snab.ch/en/snab/ (7 May 2018)
[9] www.svehk.ch (7 May 2018)
[10] www.ecoute.ch (7 May 2018)

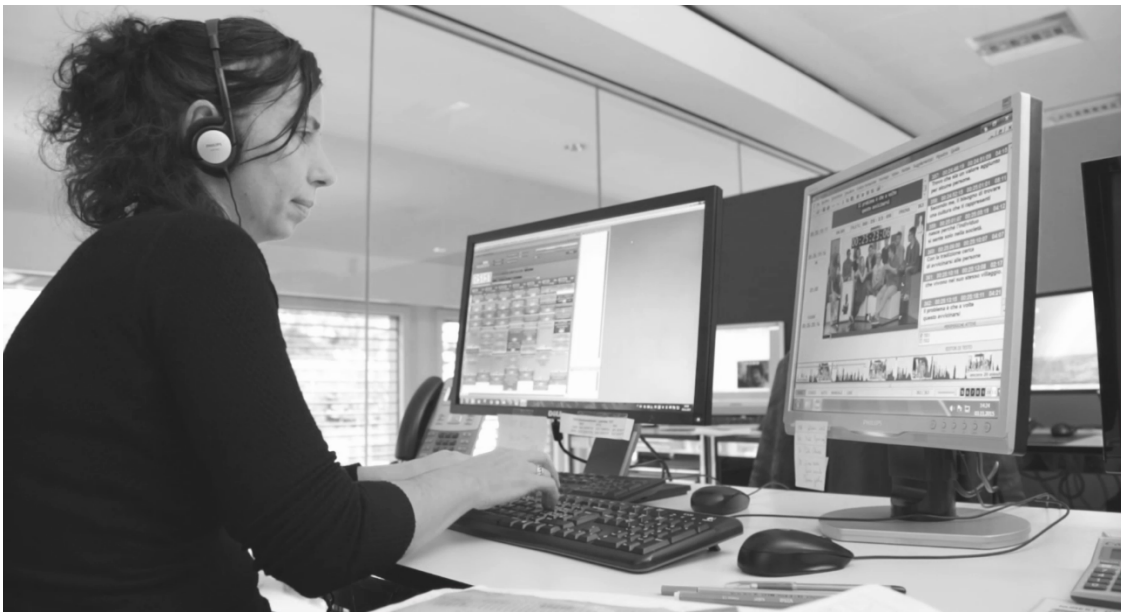

Figure 1: Typical workplace set-up for a respeaker at SWISS TXT AG

Romero-Fresco (2011: 50ff.) divides the requirements for respeakers into: the skills needed prior to the process, the skills required during the process and those needed after the process. The skills applied before the process would be a general knowledge of the ASR software and the preparation of that software (Romero-Fresco 2011: 56ff.).

During the process, respeakers must have the ability to split their attention (Romero-Fresco 2011: 95). They must listen, speak and listen again while watching and, if necessary, correcting the ASR output on their screen. In addition, the respeakers dictate the correct punctuation of sentences and must keep a steady rhythm and speed appropriate for the ASR software (Romero-Fresco 2011: 112f.).

Jekat (2014: 88f.) writes about an additional challenge, namely the need to shorten or reformulate the original broadcast due to (1) the high speaking rate in the original contribution, (2) the time lag between the beginning of the contribution and the appearance of the respeaker's contribution and (3) the typical characteristics of speech-to-text transfer. All these skills must be executed under high time pressure.

3 Online respeaking

This section focuses on a specific form of respeaking, namely online respeaking. It differs from the conventional method in terms of where the respeaker is located. In conventional STTR settings, STTRs are present in the room where the communication is taking place. In online respeaking, the respeaker does not work on-site, but joins the communication situation online via the Internet. While this form of remote respeaking has many advantages, it is also faced with challenges.

3.1 Online respeaking in the education sector

The use of respeaking in the classroom is not new. In the US, this service has been provided since 2003 (Romero-Fresco 2011: 142). Initially, the respeaking was done in-class, with the respeaker using a mask microphone so as not to disturb the class. Later, remote respeaking for the classroom was introduced. In this setup, the respeakers usually work from home. They receive the audio signal from the classroom and send the respoken transcript via the Internet to the students who need it and who receive the text on their laptop.

In Europe, Germany has played a leading role in the area of remote speech-to-text systems for students. The first concept for such a system was developed about a decade ago (Nachtrab 2008). After establishing the system in Germany, Michaela Nachtrab is now working with SWISS TXT to set up a similar system in Switzerland.

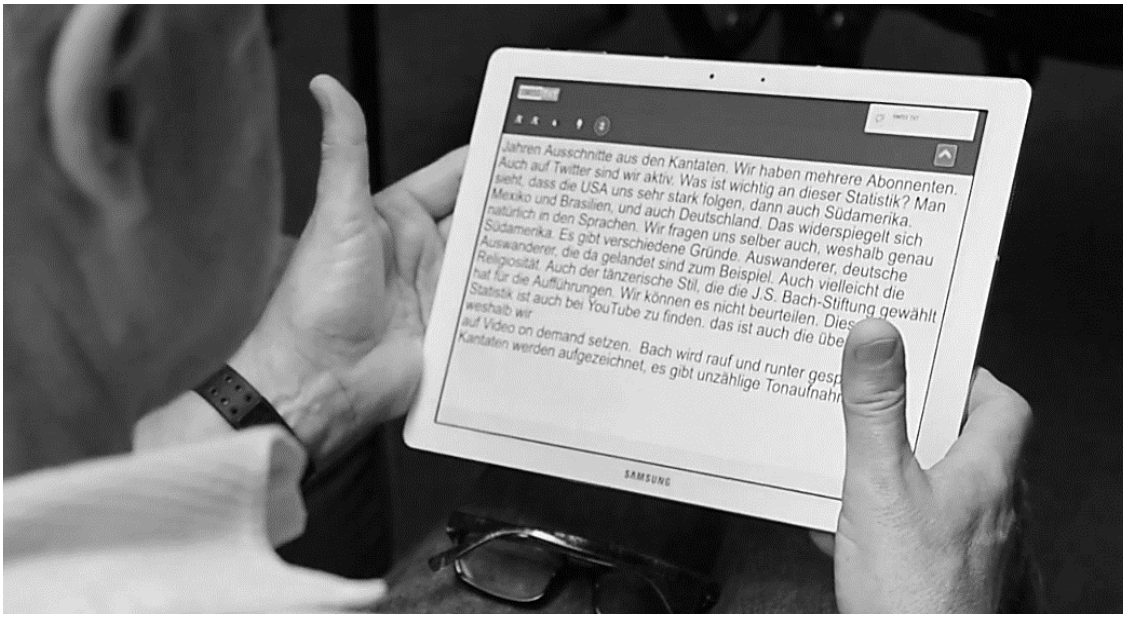

Figure 2: Online live transcription displayed on a tablet computer

After its development phase, the new remote speech-to-text system was successfully launched in September 2017 for a student at the University of Zurich. Since then, several students have started using the service in the German- and in the French-speaking parts of Switzerland. Since SWISS TXT has highly qualified respeakers, the live transcriptions are mainly generated by means of respeaking.

3.2 Technical setup and equipment

The SWISS TXT online respeaking system was built according to the latest programming standards. The student and the respeaker access an online platform. Via this platform, the lecturer's speech is transmitted to the respeaker, who converts it to a transcription by means of respeaking. The student receives the generated transcription live on their device and can thus follow the lecture on-site in the classroom with minimal delay.

In order for the system to work, the following minimal technical setup is required:

(1) Internet access, which is usually provided to the students by the university.

(2) An Internet-enabled mobile device for the student, such as a tablet, netbook or laptop computer.

(3) A microphone for the lecturer.

To enable audio transmission from the lecturer to the respeaker, the lecturer wears a wireless clip-on microphone, which is linked to the student's device, generally a tablet or laptop computer, via Bluetooth. The SWISS TXT respeakers respeak the lecturer's speech using the ASR software Nuance Dragon NaturallySpeaking.

3.3 Advantages and challenges

The SWISS TXT online respeaking system has many advantages over the traditional in-class speech-to-text method. The respeakers can be deployed more flexibly, as no travel time and costs are incurred when the remote system is used. This also simplifies dealing with short-term staff shortages, because a substitute can quickly be appointed for an assignment.

Another advantage is greater independence for students using the remote system. While in-class STTR requires the presence of one – or sometimes two – STTRs in the classroom with the student, the remote system allows the student to be more independent and to not attract unwanted attention to their hearing loss. The setup with the tablet is often even perceived to be interesting and stylish.

However, the remote system is not equally suitable for every classroom situation. Because the respeakers are not in the classroom, they are dependent on stable, high-quality audio transmission, which requires a stable Internet connection. Since most Swiss universities are now equipped with very good wireless networks, this poses less and less of a challenge for students in tertiary education. Nevertheless, even with good audio transmission, capturing students' contributions in the classroom can be difficult when lecturers do not repeat what was said. This has required SWISS TXT to find more adequate microphone solutions for individual classroom settings, which allow good audio transmission independent of the mode of teaching.

Some respeakers find it challenging that they cannot see what is happening in the classroom because they lack the visual information related to references made to what is shown on the chalkboard or projected on a screen.

Last but not least, students using the remote system must have a certain affinity for technology in order to cope with the platform and the technical setup. This poses less and less of a problem, since most students of university age and younger are used to handling various devices for different tasks in their lives.

For each student and class, SWISS TXT assesses whether online respeaking is the most appropriate form of support or whether conventional on-site STTR would be better in a given situation.

3.4 Particularities of the Swiss market

Switzerland is an interesting and challenging market for STTR services such as online respeaking. In addition to the four national languages, German, French, Italian and Romansh, various regions speak dialects that vary greatly from the standard language. Respeakers are therefore often required to simultaneously translate from dialect into standard language as they respeak into the ASR system. To meet the challenge of linguistic diversity, SWISS TXT has branches in the various language regions and respeakers in all national languages who also understand the regional dialects.

4 Summary

The SWISS TXT online respeaking service makes use of the latest state-of-the-art STTR features to support students with hearing impairments by providing them with text in real time. Remote online respeaking has many advantages over conventional on-site STTR, but it is not the most suitable solution in all contexts. In order to assess whether online respeaking is the right form of support for a student, an accurate assessment of each individual situation is required. The diversity of languages in Switzerland requires a relatively large pool of highly qualified respeakers so that the different requirements can be adequately covered.

Sources

Downey, Gregory J. (2008): *Closed Captioning. Subtitling, Stenography, and the Digital Convergence of Text with Television*. Baltimore: The Johns Hopkins University Press

Jekat, Susanne J. (2014): "Respeaking: Syntaktische Aspekte des Transfers von gesprochener Sprache in geschriebene Sprache." Susanne J. Jekat; Heike Elisabeth Jüngst; Klaus Schubert (eds): *Sprache barrierefrei gestalten. Perspektiven aus der Angewandten Linguistik*. Arbeiten zur Theorie und Praxis des Übersetzens und Dolmetschens 69, Berlin: Frank & Timme, 7-108

Nachtrab, Michaela (2008): *Social Entrepreneurship – an innovation driven product approach in the hearing impaired service industry.* MBA thesis. Munich: Fachhochschule für Ökonomie und Management

pro audito schweiz (n.d.): "Simultane Verschriftlichung." – www.pro-audito.ch/dienstleistungen/schriftdolmetschen.html (5 February 2018)

Remael, Aline (2007): "Sampling subtitling for the deaf and the hard-of-hearing in Europe." Jorge Díaz Cintas; Pilar Orero; Aline Remael (eds): *Media for All. Subtitling for the Deaf, Audio Description, and Sign Language*. Amsterdam/New York: Rodopi, 23-52

Romero-Fresco, Pablo (2011): *Subtitling Through Speech Recognition: Respeaking.* Manchester: St. Jerome Publishing

SRG SSR (2017): SRG SSR Erweiterte Leistungen für Sinnesbehinderte. – https://www.srgssr.ch/de/news-medien/news/srg-ssr-erweitert-leistungen-fuer-sinnesbehinderte/ (3 February 2018)

Stinson, Michael S.; E. Ross Stuckless (1998): "Recent Developments in Speech-to-Print Transcription Systems for Deaf Students." Amatzia Weisel (eds): *Issues Unresolved. New Perspectives on Language and Deaf Education*. Washington: Gallaudet University Press, 126-132

Swiss Confederation (2018): "Ordinance on Radio and Television (RTVO)." – https://www.admin.ch/opc/en/classified-compilation/20063007/index.html (2 February 2018)

SWISS TXT (2017a): "Wie entstehen TV-Untertitel für Hörbehinderte?" – https://www.swisstxt.ch/de/services/access-services/wie-entstehen-tv-untertitel-fuer-hoerbehinderte/ (2 February 2018)

SWISS TXT (2017b): "Geschäftsbericht 2016 - SWISS TXT AG." – https://www.swisstxt.ch/media/1631/2017-03-20_geschaeftsbericht_de_druckbogen_acc_send.pdf (17 February 2018)

UN United Nations (2006): "Convention on the Right of Persons with Disabilities CRPD." – http://www.un.org/disabilities/documents/convention/convention_accessible_pdf.pdf (16 February 2018)

UN United Nations (n.d.): "Division for Social Policy and Development Disability. Convention on the Right of Persons with Disabilities (CRPD)." – https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html (16 February 2018)

Authors

Michaela Nachtrab built up the online speech-to-text reporting service in Germany until 2016. Since 2017, she has been leading the business development of Access Services at the SWISS TXT AG. In this function, one of her projects is to establish "Online Respeaking in Switzerland" as an intrapreneurship project of the SWISS TXT AG. Since completing her degrees in pedagogics, consulting (2000) and her MBA in general management (2008), her focus has been on entrepreneurship, intrapreneurship, education and access services in various leading positions. E-mail: michaela.nachtrab@swisstxt.ch

Sarah Mössner is studying at the ZHAW Zurich University of Applied Sciences in the BA programme in Applied Languages. Since summer 2017, she has been working at the SWISS TXT AG in Access Services as a core team member of the project "Online Respeaking in Switzerland". E-mail: sarah.moessner@swisstxt.ch

DOI 10.21256/zhaw-3000

Noura Gzara

Insights into Working Practices for the Audio Description of Documentary Films

Abstract

Audio describers generally come across a broad variety of documentary films. This article will address the following questions: wherein lie the differences between audio description (AD) of a documentary film and AD of other films? What are the difficulties in working with documentaries and what makes this work easier? By giving an insight into everyday work with films of this nature, this article aims to make the reader aware of some of the specifics of the genre. Furthermore, it will emphasise the importance of effective communication between client and audio describer, which appears to be particularly crucial for documentary films. Rather than providing a theoretical framework, this article will give some insight into the practicalities of working within this particular genre.

1 Introduction

1.1 Documentary films

Since 2013, there has been an increase in accessible audiovisual productions (i.e. audiovisual productions with subtitles for the deaf and hard of hearing [SDH] and AD) in Germany. Although feature films represent the majority of these productions, a considerable number of documentary films have also been produced. These documentaries can vary considerably: some are of a more entertaining character, others are of an educational nature; they may be shown in cinemas, on television, at festivals or in educational institutions. What they all have in common, though, is that they are a "form of film which expressly focuses on the non-fictionality of the pre-filmic. In a broader sense, the non-fiction film, the travel film, the news film, the ethnographical film, the essay film and others are part of the genre of the documentary film" (*Lexikon der Filmbegriffe*, n.d.[11]).

Film theorists are still debating the definition of the term "documentary" (Eitzen 1995: 81). As Eitzen explains: "[…] the definition of the term remains a vexed and controversial issue, not just among film theorists but also among people who make and watch documentaries" (Eitzen 1995: 81). However, this issue cannot be discussed at length here. Rather, within the framework of this article, I will follow Eitzen's suggestion that "[t]he best way to define documentary […] may be to say simply that it is whatever people commonly mean by the term" (Eitzen 1995: 83). This definition will be sufficient for our purposes.

1.2 General thoughts on AD of documentaries

Several aspects need to be considered for all AD of documentaries. One characteristic of a documentary, for example, is that it often has an off-screen commentary of its own, guiding the viewer through the film, explaining some of the facts and describing some of the images.

Accordingly, in documentaries, the choice of voice actor for ADs is relatively straightforward. The actor's voice has to be noticeably different from that of the actor voicing the off-screen commentary for the actual documentary itself, so that a blind or partially sighted audience can always be aware of who is speaking at any given moment. This is particularly important since the style of speaking of the two voices will be rather similar: informative and "sober" (Benecke/Völz, n.d.). This can be achieved by various means. The most obvious way is to choose a voice actor of the opposite sex (ITC Guidance 2000: 8); thus, with a male off-screen

[11] All translations from German are mine unless otherwise stated.

narrator one would choose a female AD voice actor and vice versa. Alternatively, one could opt for other opposites like old voice/young voice, high voice/deep voice etc.

Not only does the off-screen narrator make the task of choosing an AD voice actor easier, the existence of such off-screen comments also helps AD in that they convey important information and often even describe what can be seen in the film, so there is no need for an AD for that particular scene. In order to avoid redundancies, audio describers have to take care not to repeat information already given by the off-screen narrator (Fryer 2016: 112).

Some of the terms used by the off-screen narrator can be taken up by the audio describers. However, much of what is shown in the film and is not mentioned in the off- screen commentary has to be painstakingly researched. While in a feature film lexical accuracy may be somewhat less important, in a documentary it is vital to use the correct terminology since the focus is not on a fictive plot but rather on facts. This means that, apart from detailed research work, the audio describer is faced with another issue: in this fact-based context, it is not only appropriate, but necessary to use accurate technical terms (Rodríguez Posadas 2010: 199, 205). These, however, are often terms which an audience without background knowledge of the subject in question will not be familiar with, which means that they also have to be explained. Accordingly, the challenge for the audio describer lies in introducing not only the term, but also its explanation at a suitable place in the film. Once the term has been introduced, it can be used without further explanation throughout the rest of the AD, as from that point on it can be considered as known (Benecke 2014: 20).

While communication with the client is always vital in AD (in order to ensure that everything has been understood, to clarify the client's ideas and intentions and to explain the audio describer's approach to the client), it may be even more important with documentaries than with feature films. This is because more questions may arise during the description of a documentary due to the specificity of the subject matter, and the client's expertise in the subject area can help with the creation of an accurate description. It is therefore always worth sending queries to the client and discussing them with him or her in order to make sure that everything has been correctly understood and described.

However, it is not only the terminology which needs special attention. Some factual information may also be very difficult to describe, because for the non-specialist it is not always obvious what is being depicted in the film as processes, objects and circumstances can be very unfamiliar. Here, too, the describer is expected to explain the unknown facts for a blind or partially sighted audience, to describe them in such a manner that even laypeople can easily visualise and understand what is shown on the screen. This requires a more precise and detailed description than is necessary for the AD of feature films, where more or less familiar everyday objects, scenes and circumstances are usually to be described (Fryer 2016: 112).

2 The films: The peculiarities of AD of documentary films

In the following section, I will introduce four examples from work carried out by a blind colleague and myself and explain what made those descriptions particularly challenging and how we tackled these specific challenges. I will provide short extracts from the respective films, describing what can be seen on the screen (*not* our actual AD) and citing the off-screen narrator (the narrator text is given in quotation marks).

2.1 Vivid images in “Magie der Moore” (Magic of the Moors) (2015)

The first film is “Magie der Moore” (Magic of the Moors), a film which provides a considerable amount of information about life on the moors, but which, at the same time, is characterised by aesthetic imagery which really does justice to the word “magic” in the film’s title. It almost seems as if the text were merely a justification for showing these beautiful images. The entire film is virtually a moving picture book. The texts of the off-screen narrator, too, are often not very matter-of-fact, but rather flowery and fantastic; the prosody and speech pace sound as if he were telling a fairy tale.

Extract from the film:

> A bear slowly walks past some trees and gets to a pond, where it stops to drink. Gulls are sitting close by. A light fog rises from the water.
>
> Noises of the surroundings are audible: birds, water, the bear drinking etc.
>
> “The layers of peat keep growing, one centimetre in ten years, one metre in a millennium. The moor stores huge amounts of rainwater. Here and there, ponds start forming. The bear knows from which side it can approach the pond safely in order to drink.”
>
> A lake in a wood; bird’s-eye view.
>
> Some ambient sounds are still audible, but, hardly noticeable, music sets in.
>
> “Acidic water, in which only a few animals and plants live. The outlines of the moor ponds keep changing. Sometimes overnight.”
>
> Eerie music.
>
> In time-lapse photography: a lake at dusk, with a small, grassy island moving around in it. The lake is surrounded by conifers.
>
> A dark blue starry sky behind conifers.
>
> Eerie music continues. Wood creaking.
>
> “When the sun sets behind the horizon, the undead awake.”
>
> Bare, dead trees, slowly moving their branches.
>
> “Dead trees, age-old and crippled, seemingly awaken to new life. When it gets cold and humid at night, they stretch their arms skywards.”

Of course, in spite of the flowery text, this film, too, requires factually correct and “sober” AD. While due to the unusual off-screen text it might be tempting not to take the film as seriously as other documentaries or to use a similarly imaginative vocabulary, the audio describer has to approach the task as he or she would for any other documentary. The terminology has to be carefully researched and checked, yet in addition the AD has to reflect the film’s extraordinary imagery and translate its aesthetics into words. In this film, the visual seems to be of greater importance than the aural component. It is even more important with this kind of documentary than with other types to do justice to the images and to make them available to a blind or visually impaired audience so that they are able to enjoy the film as intended.

2.2 Lack of consistency between text and image: “Digitale Nomaden” (Digital Nomads) (2014)

In the second film, a young man describes his progression from working as a company employee to a life as a freelancer, where he is able to work from any location he chooses. He himself is the off-screen narrator and also features in the film, either interviewing others who have had similar experiences or being portrayed himself.

As the off-screen narrator, he sometimes speaks at length and describes things he has thought about or planned, things which are important to him etc. Presumably in order to avoid continuously showing his face while he is talking about such things, seemingly random scenes and impressions can be seen on screen, which are sometimes more, sometimes less related to what he is talking about.

This happens quite frequently in the documentary genre: when a person is talking about a certain subject for some time, his or her face is only shown for a few seconds. Then footage of something more or less matching the subject in question can be seen, which may be more entertaining, but does not necessarily serve as an explanation of what is being said.

Extract from the film:

> Thorsten, the young man, is sitting at a desk in his friend Anja's flat, looking at his laptop.
>
> Some light background music is playing.
>
> "So, in July, Anja will stand here with her backpack and say, 'Now I live here again. Thanks for looking after my flat.'"
>
> He comes down the stairs outside the house and goes into the garden, where a striped deckchair is standing.
>
> "Then I will leave the flat and take my belongings with me. But I can use hers as my official address, she promised me."
>
> He sits down in the deckchair, a glass of orange juice in his right hand.
>
> "I will move to the island of Sylt for a month."
>
> He drinks some juice. In the background, a dark cat is strolling through the garden.
>
> "My IKEA doormat and the leather sofa are already up on eBay for sale."
>
> He lies back in the chair, closing his eyes.
>
> "I have quite a tingly feeling under my skin thinking about it… Calling a holiday flat my home…"

In the short excerpt above, the images hardly match the off-screen narrator's text. Here, the describer is confronted with the question: what should I do with such images? The important element is what is being said, not what is being shown. Should these short scenes and sequences be described or not? Unfortunately, this question cannot be answered easily: there is no universal answer. Whether and how much can be described depends on the individual film and the relation between the images and the text in that film.

Experience shows that even if the pauses between dialogue passages allow the insertion of a description, it can often be better *not* to describe the images, as a description might hinder rather than help with understanding. How, for example, can one mentally connect the description "he is lying on a deckchair" with his statement that his friend Anja will soon return home?

Of course, a describer always has to consider whether or not he or she is withholding information from a blind or partially sighted audience by not describing such images, and where audio breaks allow such information to be inserted in a sensible way, this is definitely an option to be considered. Where, however, there is only time for a hurried one or two-word description, it may be better to leave out this description altogether to avoid confusing the audience.

2.3 Subtitles: "Silicon Wadi" (2014)

On one occasion, we were given a number of documentary films to describe, many of which were subtitled. This issue, of course, is not exclusive to documentary films. On the contrary, lately there have been a considerable number of fictional films containing subtitles – from just a few lines up to half the film – which also require audio description, or rather voicing as audio subtitles. However, it is a more common occurrence in documentaries and therefore warrants brief attention here.

"Silicon Wadi" is a film in Hebrew about start-ups in Israel. This film came to us at Eurotape completely subtitled, which meant that audio subtitles had to be created for the entire film. Through trial and error, we established that it is advisable to record the audio subtitles first and write the AD afterwards. The following example shows why:

Extract from the film (English subtitles provided by the client):

> Two young men, Elram and Noam, are sitting at an outdoor table of a café, coffee cups and plates in front of them. They talk very agitatedly in Hebrew. Their discussion is subtitled. Elram listens to Noam, his head resting on his hands.
>
> "We're not going to discuss this, you decide what you want to do and that's that.
>
> And until you decide, we're not working."
>
> Elram puts some documents in his bag and replies, gesturing:
>
> "If you feel that you can do it on your own, that you can lead this on your own, no problem, do it on your own."
>
> He stands up, puts his rucksack on his back and takes his jacket. He is just starting to leave.
>
> Noam: "How dramatic. That's so mature, not staying here to talk."
>
> Elram turns round, gesturing: "I understand where you're coming from, I just have to work out how to do it. We'll talk."
>
> He leaves.

Some action needing description takes place during the dialogue: one of the men gestures, stands up and leaves. The pauses between the elements of the dialogue are rather short and the descriptions have to be squeezed in, therefore it is paramount for the describer to know exactly how much time is left for the description. If he or she writes the description before the audio subtitles have been recorded, this may lead to the texts not fitting into the pauses once the audio subtitles have been added, because they are longer or are placed differently from how the describer anticipated. It is therefore highly advisable to start with the audio subtitles, mix them in and *then* start the AD. This way, the describer knows exactly how much time he or she has for the description and potential issues can be avoided.

Furthermore, with a language unknown to both describer and voice actor, problems with the cues may arise if the subtitles are recorded after the actual AD. In a German script, both the time code and also a short quote from the original soundtrack are given as reference or cue for the voice actor. The audio describer will have been provided with the translated German subtitles and noted the cues in the script in German, referring to those subtitles. But the AD voice artist hears the original language, rather than the German audio subtitles. He or she will be confused, which will lead to delays in the AD recording and thus a potential increase in costs.

Regarding the choice of voice actors for the audio subtitles, their number depends on the requirements of the film in question as well as on the client's budget. The absolute minimum for a film with a substantial number of subtitles is to have a separate voice actor for the audio

subtitles, who is not the AD voice actor (Remael, n.d.). It would obviously be preferable to have at least one female voice actor for the women and one male voice actor for the men; ideally, there will be even more audio subtitle voice actors to cater for the different characters. However, more often than not one voice actor voices several different characters. It is therefore important for the AD to explain and describe who is talking, e.g. by introducing them: – "Peter:" – followed by whatever that person is saying. While in a non-subtitled (i.e. original language or dubbed) film, a blind or visually impaired audience will, after a while, recognise the different characters by their voices, this is not possible in a situation where the same actor voices several characters and the original is barely audible (Remael, n.d.).

Since audio subtitles are not descriptions but rather translations of the dialogue, they are voiced in a slightly more "emotional" way than ADs (see also the example in Remael, n.d.).

2.4 Audio described material in an educational setting: "Inklusion im Kindergarten" (Inclusion at Nursery School) (2015)

The final production to be considered here is a film used in lectures of the Department of Social Work at a German university of applied sciences. It is therefore not made for entertainment, but for teaching purposes. In the film, an inclusive nursery school is introduced.

This publication is divided into two DVDs. The first DVD consists of the main film, in which the inclusive nursery is introduced, and some mothers, nursery school teachers and the manager speak about their respective experiences. There are also three short scenes on the first DVD showing the interaction among the children or between the children and the nursery school teachers. On the second DVD, there are long interviews with some mothers, the nursery school teachers and the manager. In these interview situations, only the interviewee is seen and heard throughout, talking about his or her experiences, impressions etc.

The main film also features a handful of subtitled scenes. In these particular cases, we decided to deal with them by omitting them.

Extract from the film:

> Two boys are sitting on the floor under a kind of playing platform. One of the boys, Michel, is wearing a bone-anchored hearing aid on a band around his head. They are playing with a jigsaw puzzle featuring some farm animals.
>
> They speak German, but their dialogue is subtitled.
>
> Michel: "Calf." Other boy: "Yes, that's a calf."
>
> The other boy fits the piece with the calf onto another jigsaw piece.
>
> Other boy: "This goes here."
>
> Michel: "And then?" He points at the two joined pieces. "It got into the stable."
>
> He takes another piece out of the box. "And this is the… horse."

In this scene, the film-makers inserted subtitles because they obviously feared that the boys' dialogue would not be easy to understand. However, in order to comprehend and be able to follow the film in general and this scene in particular, what the boys are actually talking about is not of major importance. What *is* important is that hearing-impaired Michel interacts without shyness and quite self-confidently with his peer and is treated as an entirely equal playing partner. And in order to understand this, it is not necessary to make out what they are saying. If the subtitles were voiced as audio subtitles, the boys' conversation would gain an importance it does not really warrant. It would be unnecessarily highlighted and might even be difficult to "place" for a blind or visually impaired audience. Therefore, after consultation with the client, we decided against audio subtitles. In addition, we believed that their dialogue was fairly easy

to understand anyway and did not necessarily need subtitling in the first place. Finally, not voicing the subtitles allowed us more time for an extensive description of the scene.

This film is a documentary without an off-screen narrator. But since, with one exception, only female parents and nursery school staff (as well as children) are shown, we opted for a male AD voice actor, so his voice would be clearly different from the in-film voices (see section 1.2 for the choice of voice actor; see also Benecke/Völz, n.d.).

In addition to the two DVDs, the DVD case contains a printed brochure. This brochure gives background information on some of the children, the concept and the history of the nursery school are introduced, and there are also some exercises for students using the DVD for their Social Work course.

While this was not initially part of the project, it is, of course, vital to voice the brochure and describe the photographs featured in it in order to ensure that this printed material is accessible for visually impaired or blind students. Accordingly, the entire brochure text was recorded, and the AD of the brochure's photos was inserted into the recording at a suitable place, i.e. close to where the photos appear between the brochure texts in the print version. For better understanding and to avoid confusion, it is advisable to indicate the beginning and end of the photo descriptions (e.g. "beginning of image description", "end of image description") so that it is clear which part is the brochure text and which is a description of a photo. The *audio* brochure now forms part of the first DVD. DVD cases of fictional films may also contain brochures or flyers but they are normally just advertising other films, whereas a brochure of such relevance to the content of the DVD – and the tasks set for the students as part of the course – will be found, presumably, more often in a documentary context like this one. While it may be of entertainment value to know what advertisement flyers or brochures say, it is not necessary to know their contents for the enjoyment of a fictional film. In the case of the brochure described above, however, it is extremely important for blind or partially sighted students to have access to all the information given in the print version in order to be able to develop a better understanding of the nature and organisation of the nursery school and so be able to work alongside their fellow students on an equal footing.

We also discussed the setup of the DVD with the client so that we could cater for their wishes in this regard: the client asked for the AD to be played automatically but for the SDH to be activated via the menu. The menu entries were also recorded (i.e. audio menus were created) and integrated into the DVDs so that visually impaired or blind users would be in a position to use the DVD easily and independently, without the assistance of a third party.

These last two examples show how important it is to keep in mind that AD may not be limited to describing the actual film but can extend to the accompanying material as well. This is, of course, not limited to documentaries, but true for all kinds of films.

The AD of the brochure caused an unforeseen difficulty: as mentioned above, the brochure contains several exercises for students. Amongst these are observation exercises: the students are required to watch certain scenes closely and note down what they see, describe the interactions between the children etc. This exercise is made redundant by the AD. We, as describers, have virtually already done this task for the students. Moreover, since our descriptions had been reviewed by the client prior to publication, they therefore reflect exactly what is important to the client and, as the DVD is automatically played with the AD, a sighted audience has access to our observations as well. We made the client aware of this dilemma but do not know how this problem was solved in the end.

As mentioned above, the second DVD features several interviews. We decided to include only a very short description at the beginning of each interview, explaining who is talking, as well as a brief description of what the respective interviewee looks like. We opted for this approach because of the style of those sequences, which have only very few and very short pauses. This would have led to a series of very compressed one or two-word descriptions of the featured person's facial expressions had we interspersed them throughout. These would most likely annoy rather than help visually impaired or blind viewers. Furthermore, there is no plot

and no action to describe; what is being *said* is most important. While some may argue that the facial expressions could be important too, we thought, as explained above, that the help derived from ultra-short, fast-paced descriptions would hinder rather than foster understanding. In addition, what is reflected in the interviewees' facial expressions is also reflected in their tone of voice and therefore does not necessarily need description as the information can be gathered aurally.

The respective university of applied sciences has indicated that even when they do not have any blind or visually impaired students, they still use the audio-described version in order to make sighted students aware of how blind and visually impaired people experience films. This demonstrates that AD can be useful not only to a blind or partially sighted audience, but also to sighted viewers, particularly in this kind of educational setting.

3 Concluding remarks

Audio description of documentary films is an everyday occurrence, and yet it can vary in many ways from that of feature films. The describer is faced with a number of very particular issues which are specific to this genre. The informative nature of this kind of film, the requirement to use very specific terminology and possible discrepancies between image and text are just some of the aspects which make this genre rather difficult to deal with.

This article has provided an account of some of the key issues and challenges encountered in producing AD of documentaries and has illustrated the approaches that can be – and have been – taken in a practical context. However, further academic research is required to create a comprehensive theoretical basis for AD of documentary films.

Bibliography

Benecke, Bernd (2014): *Audiodeskription als partielle Translation. Modell und Methode.* Berlin/Münster: LIT Verlag Dr. W. Hopf

Benecke, Bernd; Haide Völz (n.d.): "Technical issues." ADLAB: *Pictures Painted in Words*: http://www.adlabproject.eu/Docs/adlab%20book/ (25 February 2018)

Díaz, Cintas Jorge; Anna Matamala; Josélia Neves (eds) (2010): *New Insights into Audiovisual Translation and Media Accessibility. Media for All 2.* Amsterdam: Editions Rodopi

Eitzen, Dirk (1995): "When Is a Documentary? Documentary as a Mode of Reception." University of Texas Press on behalf of the Society for Cinema & Media Studies, Cinema Journal 35 [1]: 81-102 – http://www.columbia.edu/itc/film/gaines/documentary_tradition/Eitzen.pdf (25 February 2018)

Fryer, Louise (2016): *An Introduction to Audio Description. A Practical Guide.* Oxon/New York: Routledge

ITC Guidance on Standards for Audio Description (2000). msradio.huji.ac.il/narration.doc (25 February 2018)

Remael, Aline (n.d.): "Combining Audio Description with audio subtitling." ADLAB: *Pictures Painted in Words:* http://www.adlabproject.eu/Docs/adlab%20book/ (25 February 2018)

Rodríguez Posadas, Gala (2010): "Audio description as a complex translation process: a protocol". Jorge Díaz Cintas; Anna Matamala; Josélia Neves (eds): *New Insights into Audiovisual Translation and Media Accessibility. Media for All 2.* Amsterdam/New York: Rodopi, 195-211

Universität Kiel (n.d.): "Lexikon der Filmbegriffe." http://filmlexikon.uni-kiel.de/index.php?action=lexikon&tag=det&id=127 (25 February 2018)

Filmography

“Digitale Nomaden – Deutschland zieht aus” (Germany, 2014), Tim Jonischkat, Thorsten Kolsch

“Inklusion im Kindergarten – Einblicke in ein erfolgreiches Konzept” (Germany, 2015), Prof. Dr. Sabine Lingenauber, Janina L. von Niebelschütz

“Magie der Moore” (Germany, 2015), Jan Haft

“Silicon Wadi” (Israel, 2014), Daniel Sivan, Yossi Bloch

Special thanks to (in order of appearance of the respective films in this article): Polyband, EYZ Media, Hochschule Fulda (FB Sozialwesen).

Author

Noura Gzara holds a degree in Translation from Johannes Gutenberg Universität Mainz (Germersheim). She has worked in audiovisual translation since 2007 and started working in accessibility several years ago. She now works as AD Editor and Project Manager in the Accessibility Department of Eurotape (Nordkurier TV und Studio) and has recently begun a PhD in Audio Description at Universität Hildesheim. E-mail: nogz.mail@gmail.com

DOI 10.21256/zhaw-3000

Anton Bolfing

ICT-Accessibility and E-Inclusion in Switzerland: Opportunities and How to Get There

Abstract

For the past 17 years, «Access for All» has campaigned for a world in which everyone can use information and communication technologies (ICT) to exchange information with others, and where everyone can read online newspapers anytime and anywhere. A world where students with disabilities use the same learning materials as their non-disabled classmates.

Based on the results of the Swiss Accessibility Study 2016, we identify opportunities of modern ICT for educational and social inclusion of people with disabilities. We show how barriers can be avoided and where urgent action is needed to make sure that the opportunities offered by ICT will not result in the opposite: in an increasing digital divide between people who conform to the target groups of online services and people who have not been considered due to their limitations. The consequence would be the exclusion of large parts of the population from the (information) society.

1 «Access for all» / «Zugang für alle»

«Access for all» (A4a, www.access-for-all.ch) is a small foundation in Switzerland. Despite its small size of 8 employees, half of which live with a disability, «Access for all» is nowadays regarded as the main contact point and competence center in Switzerland when it comes to ICT-Accessibility. The foundation was established in the year 2000 with the mission to advance ICT-Accessibility in Switzerland and in Liechtenstein.

In order to achieve this goal with a broad impact, «Access for all» is active in four fields:

- Services: accessibility-testing, consulting and training of ICT service providers. Further, «Access for all» is the certification authority in Switzerland for the accessibility of web and mobile apps.
- Research and Development: knowledge generation, monitoring and development of services.
- Awareness raising for e-inclusion and e-exclusion issues.
- Education: training for young people with disabilities to become ICT-Accessibility testers and consultants.

As a private independent foundation «Access for all» is 100 % self-financed, mainly by its services for the ICT industry.

2 ICT-Accessibility

At this point, it is important to clarify the term “ICT-Accessibility” and how it is used in the context of activities at «Access for all».

2.1 What is ICT?

«Access for all» is currently concentrating on web accessibility, mobile applications and electronic documents as well as, to a lesser extent, on desktop software. With this focus, comprising all kinds of e-services, such as e-banking, e-government, booking services and online shops, the foundation’s activities mainly cover ICT aspects. When it comes to inclusion, a particularly important aspect of accessible ICT is the availability of accessible electronic documents. Educational inclusion seems hardly possible without accessible learning materials.

2.2 Impairments affecting the use of ICT

Most disabilities or their inherent impairments affect the use of ICT - some more, some less. Sometimes impairments are not even tied to a disability, but to temporary injuries or to other factors such as small screens, harsh light or a broken computer mouse.

2.2.1 Visual impairments and blindness

Visual impairments and blindness pose a great variety of requirements to ICT products in order for these to be accessible. Informative content must come with detailed semantic information, text alternatives for non-text content and sequential navigability in order to be useable by blind people using screen readers[12]. Websites must also allow for lossless magnification, minimum colour contrast between font and background as well as adaptability to custom colour schemes. All these requirements are indispensable to any accessible ICT product.

Since most ICT devices are primarily designed for visual perception, it is not surprising that these requirements are among the most demanding.

2.2.2 Hearing impairments and loss of hearing

Hearing impairments and loss of hearing primarily affect audio content and language complexity. For people born deaf, written language must be considered a foreign language. For them, equivalent content to audio is given by sign language videos. Other approaches are subtitles and text transcriptions.

2.2.3 Motor impairments

Since motor impairments primarily affect the way people interact with ICT, it must be ensured that different input devices can be used for interaction. Not only pointer devices, such as computer mice or eye trackers, but also input devices allowing sequential navigation must be supported. This requirement is met if every interactive element on a web page can be navigated and operated using the tab, enter, arrow and Esc keys of a standard keyboard. In the case of quadriplegia or other severe motor disabilities, those keys can be emulated by any individually arranged pushbuttons.

It is very important that keyboard focus location is visible at any time. Sighted keyboard users must know where their keyboard focus is located in order to interact with ICT.

2.2.4 Cognitive and neurological impairments

In many respects, cognitive aspects of accessibility are to be treated differently than physical aspects. Firstly, in many cases it is not enough to simply make technical adaptations to ensure proper interaction with assistive technologies. Measures are necessary that directly affect content and the look and feel of ICT products. Secondly, cognitive limitations are so diverse that they can hardly all be taken into account in the sense of Universal Design.

In collaboration with insieme Switzerland and the FHNW School of Social Work, «Access for all» developed the «Easy Surfing» guidelines[13], where, for the first time, cognitive aspects of ICT accessibility are edited as a practical guide.

Despite these uncertainties, it can be stated that the following aspects might always hold true for maintaining cognitive barriers low: no complex interactions (clearly arranged navigation) and no use of too complex language (plain language).

---

[12] Screen readers are software programs that allow written text and semantic information to be translated into synthetic speech and braille output.

[13] easysurfing.ch (4 May 2018)

2.2.5 Multiple disabilities and age-related impairments

Since recommended measures set out to prevent barriers for people with sensory and motor impairments follow the principles of Universal Design, these measures cumulate. Hence, no specific measures must be taken for multiple disabilities, even if these are age-related.

However, it is certainly not wrong to offer certain additional pre-settings (such as large fonts) or functionality (e.g. zoom buttons) so that the complexity of operation can be reduced to a minimum.

2.3 WCAG 2.0

Within the scope of its advisory activities, such as accessibility testing, and in its role as a certification body for accessible ICT services in Switzerland, «Access for all» claims to be committed to all people with special needs on an equal footing.

This is where the Web Content Accessibility Guidelines (WCAG)[14], currently in version 2.0, come into play. WCAG is the de facto standard for most national and international ICT accessibility guidelines. Version 2 was published in 2008 and, due to its technology-agnostic wording, is still well applicable to different formats such as HTML & CSS (with/without Javascript) (W3C 2016), PDF documents and Flash, but also to different technologies like mobile apps (W3C 2015) or software user interfaces in general.

The WCAG basically follows the widely accepted product design concept of “Universal Design”. Accordingly, products should 1.) be highly adjustable (flexible) in use and 2.) support assistive technologies. Accessible ICT products must therefore always be operable by as many people as possible. Universal design and WCAG claim that one and the same product is accessible to everyone, in contrast to special solutions for different special needs.

3 The Swiss Accessibility Study

«Access for all» has published monitoring studies on the accessibility of websites in Switzerland at irregular intervals. The last study was published in October 2016 in German and French (Bolfing et al. 2016). The purpose of these studies is not only to raise awareness for ICT-Accessibility concerns but also to put slight pressure on the most important ICT service providers regarding e-inclusion opportunities.

In addition to the test results and methods, the studies come with rich editorial content. In 2016, 100 websites and 15 mobile apps were tested: 11 federal government websites, 5 websites of companies with close relationships to government (Bundesnahe Betriebe: Post, PostFinance, SBB, Swisscom and SRF), the websites of the 10 largest Swiss cities and those of the 26 Swiss cantons, 20 universities’ websites, 15 news sites and 11 online shops. The 15 mobile apps correspond to the 15 tested news sites. Each mobile app was tested for the latest versions of iOS and Android.

Accessibility assessment of the websites is based on 4-page samples. All WCAG 2.0 success criteria are rated by «Access for all» accessibility experts separately. In addition, for every tested site, a short screen reader scenario is given and documented.

In the study, the results of each single website/app are illustrated on a half page.

[14] www.w3.org/TR/WCAG20 and its follow-up proposed recommendation version (24 April 2018) www.w3.org/TR/WCAG21 (4 May 2018)

3.1 Editorial content

The rich editorial content of the 2016 Accessibility Study focusses on the big picture of opportunities that ICT potentially create for social inclusion and participation of people with disabilities. The first question investigated is where ICT affect our daily lives and what barriers people with disabilities face that prevent them from participating in the digital society. Secondly, the opportunities of born accessible ICT for inclusion are delineated in chapter "Vision e-Inclusion". "Born accessible ICT" are hardware, software and electronic contents that are marketed and published in a fully accessible way without the need for additional tedious and expensive adjustments. Finally, current and foreseeable legal frameworks in Switzerland and abroad are outlined.

3.2 Results

3.2.1 Individual results

Results are given separately on a half page for each tested website and mobile app. Each half page comes with the title, the URL and an illustrative screenshot of the homepage of the tested site/app. The overall accessibility rating is given as a number between one and five stars and as a comparison to the accessibility rating of the previous Swiss Accessibility Study 2011.

As illustrated on the left in Figure 1, the results are tabularly divided into twelve categories representing different aspects of accessibility issues, i.e. the WCAG success criteria. The accessibility level of each category is given in a five-point scale, where 1 point represents complete lack of accessibility and 5 points represent good accessibility (however, not 100 % WCAG conformance). This Accessibility Profile was especially designed for this study with the purpose of showing a clear set of more or less self-explanatory categories to an audience that is not familiar with the WCAG 2.0 guidelines.

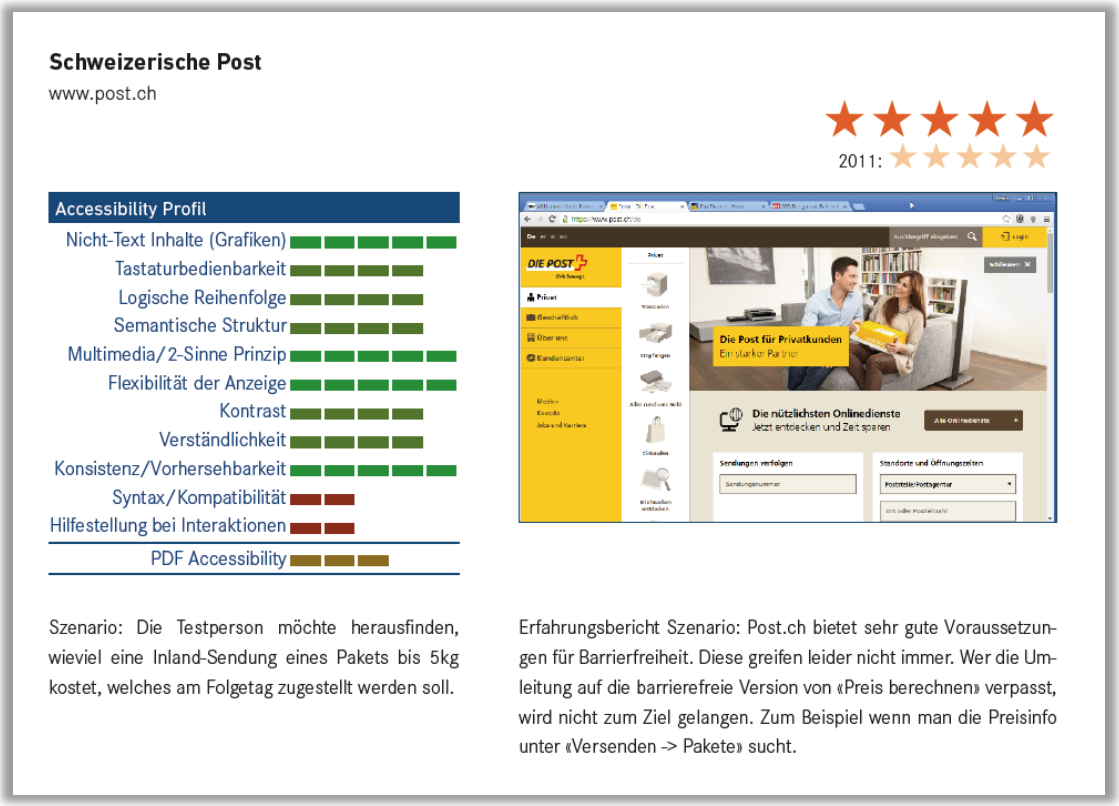

Figure 1: Sample extract (half-page screenshot of www.post.ch) from the individual results of the Swiss Accessibility Study 2016

The identified categories are: non-text content, keyboard operability, meaningful sequence, semantic structure, multimedia, flexibility of the display, contrast, comprehensibility, consistency/predictability, syntax validity, assistance with interactions and PDF accessibility.[15]

Finally, each half page comes with a screen reader scenario (identical within website categories) and a corresponding scenario report.

3.2.2 Category results

Interesting insights can be gained if we compare the aggregated accessibility scores among the different website categories.

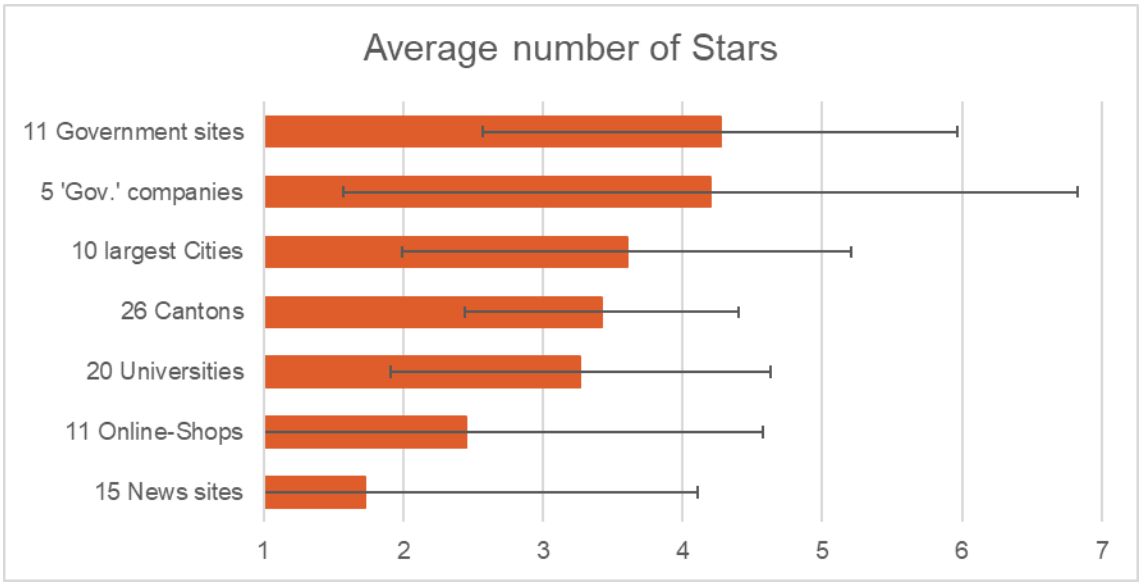

Figure 2: Category results depicted as a bar chart with standard deviations

In the range of 1 to 5 stars, government sites obtained an average score of 4.27 stars. Further, in descending order, the five Swiss 'government'-related companies achieved an average score of 4.2, the ten largest cities scored an average of 3.6, while the 26 cantons reached an average score of 3.42 and the 20 universities one of 3.27. Finally, the tested sites in the private sector reached total scores of 2.45 and 1.73 stars in the case of online shops and news sites respectively.

3.2.3 PDF Accessibility

While this is not the place to go into further details regarding Accessibility Profile categories, the aspect of PDF accessibility must come under the spotlight. In short, there are only a few among the 100 tested websites that exceed the minimum score of 1 star in this category. Only one website, the overall winner in the study, that of the Swiss Post, exceeded 2 stars.

3.3 Learnings

The first and obvious insight we gain from the comparison of accessibility scores among website categories is the correlation of accessibility with the different levels of legal obligation regarding ICT-Accessibility through federal, cantonal and municipal layers. Federal government sites, together with the federal government related companies, are subject to the clearest and most restrictive legal frameworks followed by cantonal and municipal regulations.

[15] More on testing methods in: Bolfing et al. (2016), 137-143

Within governmental sites, this gradation in legal obligation (Hess-Klein 2016) is clearly reflected in aggregated accessibility scores.

At the other end of the spectrum, the tested representatives of the private sector, the news sites and online shops, show very poor accessibility scores, with aggregated scores of 2.45 or 1.75 stars respectively, given that a minimum score of 1 star implies an exclusion of numerous people with various severe impairments.

In terms of our “Vision e-Inclusion” these results are devastating. It is not primarily the online offers and information sites of governmental institutions that push inclusion and self-determination, but, to a great extent, the services of the private sector, such as publishing houses, online shops, banks, insurances and the like, that potentially offer the greatest opportunities.

Informal surveys among web agencies and web developers carried out by «Access for all» have shown that accessibility as an integral part of web development is barely approached in the curriculum of web and ICT developers. Among other factors, such as a general lack of awareness in society, this might be the most important reason for the unsatisfying situation.

Finally, and of no surprise to any accessibility insider, the Swiss Accessibility Study 2016 revealed that there still are, and presumptively will be, virtually no accessible PDF documents among the billions of PDF documents out there.

4 Our contribution and current projects

Different problems require different measures. The self-conception of the «Access for all» foundation is more in keeping with the role of a technical enabler than with that of a political lobbyist. Instead of concentrating on expanding legal framework conditions, «Access for all» concentrates its energy on consulting and in the formation of developers, project managers, authors and editors of ICT providers.

In addition to the foundation's current main activity of testing accessibility services and providing direct advice to ICT service providers, «Access for all» has decided to make its know-how available to a wider public.

On the one hand, the range of public training courses will be greatly expanded in the upcoming months. For example, various half-day to full-day training modules are being developed for different target groups and advertised (Zugang für alle 2018).

On the other hand, this summer, the foundation will also be launching the “Accessibility Developer Guide”[16], a knowledge platform accessible to the public free of charge. In addition to valuable examples of accessible WCAG 2.0 compliant web elements, such as widgets, forms and tables, the platform conveys the basic knowledge to independently test websites for accessibility.

With these two approaches, «Access for all» is strengthening the awareness of accessibility in the inclined public and establishing the corresponding knowledge where it is urgently needed: among the web developers and project managers of Swiss web agencies.

[16] accessibility-developer-guide.com (4 May 2018)

Sources

Bolfing et al. (eds) (2016): *Schweizer Accessibility-Studie 2016. Bestandsaufnahme der Zugänglichkeit bedeutender Schweizer Internet-Angebote.* http://access-for-all.ch/ch/publikationen/accessibility-studie-2016.html (4 May 2018)

Hess-Klein, Caroline (2016): "IKT-Angebote – der rechtliche Rahmen in der Schweiz." Bolfing, Anton et al. (eds): *Schweizer Accessibility-Studie 2016. Bestandsaufnahme der Zugänglichkeit bedeutender Schweizer Internet-Angebote.* 50-56 http://access-for-all.ch/images/Accessibilty_Studie/2016/CHAccSt16_Kapitel-4_RechtlichesStandards_049-058.pdf (4 May 2018)

W3C (2015): "Mobile Accessibility: How WCAG 2.0 and Other W3C/WAI Guidelines Apply to Mobile." https://www.w3.org/TR/mobile-accessibility-mapping/ (4 May 2018)

W3C (2016): "Accessible Rich Internet Applications (WAI-ARIA)." https://www.w3.org/WAI/intro/ (2 May 2018)

Zugang für alle (2018): "Entwickler- und Tester-Schulung." http://access-for-all.ch/ch/schulung/kurse-und-seminare/entwickler-und-tester-schulung.html (4 May 2018)

Author

Anton Bolfing is an ICT accessibility expert and a member of the Executive Board at the «Access for all» foundation. He is responsible for training web developers and editors. His special interests focus on questions concerning e-inclusion in schools, especially with regard to accessible electronic learning materials. E-mail: anton.bolfing@access-for-all.ch

DOI 10.21256/zhaw-3000

Valérie Boujon, Pierrette Bouillon, Hervé Spechbach, Johanna Gerlach & Irene Strasly

Can Speech-Enabled Phraselators Improve Healthcare Accessibility? A Case Study Comparing BabelDr with MediBabble for Anamnesis in Emergency Settings

Abstract

Language barriers are an important problem when it comes to healthcare services for minority groups, such as refugees or sign language users. Interpreters are not always available, especially in emergency settings. Alternatives, such as online machine translation, are unsatisfactory in terms of the languages they cover as well as their data confidentiality and translation reliability. To fill this gap, phraselators were developed in collaboration with medical staff. While these produce reliable translations, they remain unsophisticated systems: doctors have to search for questions using menus or keywords. To improve on this, the Geneva University Hospitals (HUG) have developed BabelDr, a speech-enabled phraselator. To see if speech improves functional suitability and usability, this study compares BabelDr against a phraseslator without speech (MediBabble). This is done by asking doctors to find precise information about a patient using the two tools in a crossover design. Results show that BabelDr allowed participants to collect most of the information in a faster and easier way than without speech.

1 Introduction

In the context of the current European refugee crisis, hospitals are more and more frequently forced to deal with patients who have no language in common with the staff, and who may also not share the same culture. For example, at the Geneva University Hospitals (HUG), Geneva's main hospital, 52 % of the patients are foreigners and 10 % speak no French at all. In 2015, the languages which caused the most problems were Tigrinya, Arabic and Farsi (HUG, personal communications). The problems are not only linguistic. Cultural differences mean that these patients may have different conceptualizations of medicine, health care (Hacker et al. 2015), illness and treatment (Priebe et al. 2011). These issues arise not only in the context of migrant languages, but also in that of sign languages. In recent years, research in the area of sign language has shown the challenges that deaf patients face when they need access to medical information through sign language. Middleton et al. (2010), in the United Kingdom, and Iezzoni et al. (2004), in Hong Kong, collected data from deaf and hard of hearing people and showed that a significant number of survey participants reported problems – in particular, a lack of awareness of deafness among medical personnel. A situation of this kind, with barriers in language, culture and medical understanding, creates serious problems for the quality, security and equitability of medical care, as has been pointed out by several researchers (see for example Flores et al. 2003 and Wassermann et al. 2014). Others underline the negative impact these issues have on health care costs (Jacobs et al. 2003).

To respond to this urgent communication need in hospitals, different solutions exist. Interpreters play a key role in patients' understanding of medical information, but they are considered to be very expensive by decision makers and are not always available, especially for minority languages and in emergency settings where there are time constraints (see Major 2012).

In the absence of qualified interpreters, a number of other solutions are available today, but each of them have their drawbacks. Phone-based interpreter services are expensive (3 CHF/minute in Switzerland), not always available for some languages, and known to be less satisfactory than face-to-face interaction through a physically present interpreter (Wu et al. 2014). Google Translate (GT) and other machine translation (MT) tools, used increasingly often by medical staff when no other alternatives exist, remain unreliable for medical communication despite their recent progress (see Patil et al. 2014, and more recently, Bouillon et al. 2017). They also do not offer the most important languages for hospitals (for example Tigrinya or sign languages) and are not asy to adapt to new languages. Moreover, the use of

these online tools raises ethical concerns, since they do not ensure data protection. Phraselators such as MediBabble (www.medibabble.com) and Universal Doctor (www.universaldoctor.com) are another possible alternative. These tools, specifically designed by medical staff for the medical diagnosis scenario, consist in a set of pre-translated canonical sentences (questions and instructions). They allow medical professionals to perform a preliminary medical examination dialogue, using a decision tree method. As opposed to MT, phraselators have the advantage of providing a reliable translation as well as being easier to port to new languages/domains. However, they only translate a limited set of sentences and remain technically unsophisticated; the user has to search for the exact sentence using menus or keywords, making the interaction with the patient quite artificial.

In this paper, we present BabelDr, a phraselator that enables speech interaction (see also Bouillon et al. 2017). We propose an ISO-based evaluation to compare this system with a traditional phraselator. Our main goal in this study is to see if speech improves the doctor-patient interaction. We expect that, for similar content, speech makes the communication faster, easier and more pleasant. We first present the BabelDr system (Section 2) and the ISO-based quality model used for the evaluation (Section 3). We move on to describe the experiment (Section 4) and the results (Section 5), and then draw our conclusions (Section 6).

2 BabelDr

BabelDr is a joint project of the Faculty of Translation and Interpreting (FTI) at the University of Geneva and the Geneva University Hospitals (HUG) that specifically addresses the lack of qualified interpreters for languages spoken by refugees and for sign languages in emergency settings. The BabelDr application can be described as a flexible speech-enabled phraselator (Rayner et al. 2016; Ahmed et al. 2017). Like all phraselators, the system relies on a pre-specified list of human-translated sentences and is thus limited to that set of sentences. The main difference to existing phraselators is that doctors can find these sentences by speaking to the system, using a wide variety of paraphrases and stylistic variations, instead of having to search in a list.

When the doctor asks a question, the system uses speech recognition to recognize what was said and automatically maps this recognition result to the closest canonical sentence using linguistic rules (synchronized context-free grammar, Rayner et al. 2015). The canonical sentence is then translated for the patient. Since it is not an exact translation of the doctor's utterance, but rather a translation of the corresponding canonical sentence, the canonical sentence is always echoed back to the doctor so that he or she can verify what the system understood. The translation is thus only produced for the patient after the validation of the canonical sentence by the doctor. Canonical sentences therefore play a very important role in the system since they are both the pivot for the translation and the way to show the doctor what will be translated for the patient. They were selected with the help of HUG so as to always be the least ambiguous and the most explicit possible. For example, a sentence such as "avez-vous l'impression d'être fiévreux ?" (*do you feel you're running a temperature?*) is mapped to "avez-vous de la fièvre ?" (*do you have a fever?*). Similarly "où va la douleur ?" (*which way is the pain going?*) corresponds to "pouvez-vous montrer avec le doigt où irradie la douleur ?" (*could you show me with your finger the direction in which the pain is radiating?*). Target language utterances are realized in their spoken form either using Text-to-Speech (TTS) or using pre-recorded multimedia files. This functionality is needed for low-resource languages, such as Tigrinya, which currently lack TTS engines, and for translation into sign language (Ahmed et al. 2017). The platform is entirely web-based. It can be compiled and used on the web, as described in Rayner et al. (2016).

The central design goals of BabelDr are to ensure that a) translations are reliable, b) new target languages can easily be added, which enables flexibility in the face of changing patient demographics, and c) content can be adapted to new patient profiles. However, creating the content is not a trivial task, due to several factors. The grammar-based nature of the system's

architecture, which is necessary for the efficient creation of broad coverage, requires a specific structured data format that is not easily accessible to translators. Moreover, the medical discourse, which describes symptoms and pathologies, is in itself difficult to translate, especially in terms of lexical choices where a balance between precision and understandability on the part of patients must be found (Cardillo 2015). To facilitate the translators' task, and thereby ensure the quality and coherence of the translations, we have developed a translation platform.

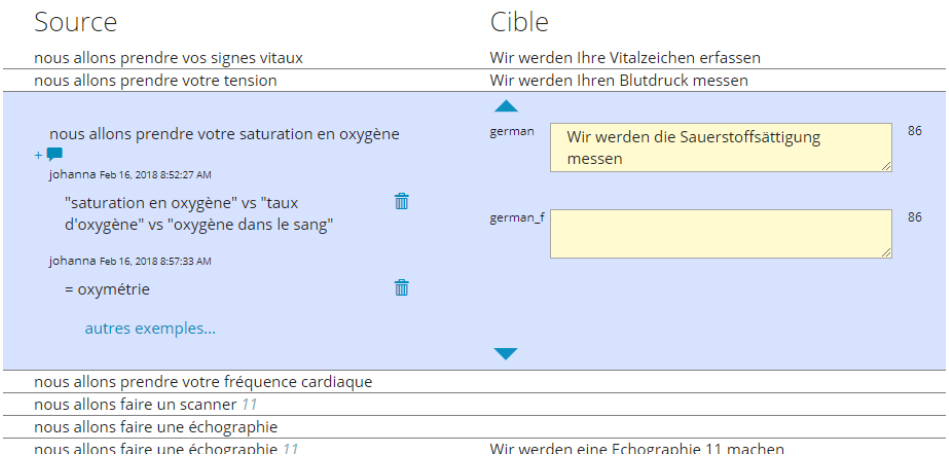

Figure 1: Translation interface

This web-based platform presents translation and revision tasks in a simple interface so that translators and revisers do not have to edit grammar files directly. Once a task is complete, the platform generates valid grammar files which can then be incorporated into the BabelDr system. To ensure translation consistency and accelerate the translation process, the platform includes a translation memory. Since the difficulties encountered by translators in different languages are often similar, we have also included an annotation function, which allows translators to share their insights and translation choices by appending comments to the individual canonical sentences. In addition to written translation, the platform includes a video-recording module to capture translations into sign language. Figure 1 shows a screenshot of the translation interface.

At the time of this study, the BabelDr system includes four main domains, organized according to different body parts (abdomen, head, chest and kidneys/back). Each of these domains has a coverage of around 2500-3000 canonical sentences, with an associated grammar that uses a vocabulary of about 2500 words and expands to approximately tens of millions of surface sentences. This grammar is used for both rule-based speech recognition and mapping to canonicals. The system supports translation from French to Arabic and Spanish, and there are partial sets of translations for Tigrinya, English, Italian, LSF-CH (Swiss French Sign Language) and Auslan (Australian Sign Language). Ongoing work on the project includes adding new migrant languages (Albanian) and improving the Swiss-French Sign Language version by including professional-quality signed videos. As for the other languages that were already available in the system, it was important to us that there should be close collaboration between the medical staff and those employed in sign language translation in order to achieve optimal translation (Hale 2007). For this reason, signed videos are currently being recorded in a professional setting by a sign language interpreter and a deaf nurse. A doctor familiar with sign language helps to explain the meaning of questions and technical terms where necessary.

In the following sections we describe the ISO-based evaluation of two phraselators (see also Boujon 2017).

3 Quality model

The goal of this experiment is to compare two phraselators, one with speech (BabelDr) and one without speech (MediBabble). We decided to conduct an ISO-based evaluation (Estrella/Tsourakis 2016). Since our goal was to see if speech improves doctor-patient interaction, we focused on two main criteria: functional suitability (does speech help to provide the correct result?) and usability (does speech make the system more usable?). For usability, we selected three ISO sub-characteristics: operability (does speech make the system easier to operate?), learnability (does speech help to learn how to use the system?) and user interface aesthetics (does speech make the interface more pleasant for users?). These four criteria are measured both with subjective and objective measurement methods, as summarized below:

Functional suitability

(1) number of successful interactions with the system

(2) questionnaire

Operability

(1) time to complete an action (i.e. an interaction with the system)

(2) number of mouse clicks to complete an action (i.e. an interaction with the system)

(3) questionnaire

Learnability

(1) difference in time needed to complete an action in two successive sessions

(2) difference in mouse clicks needed to complete an action in two successive sessions

(3) questionnaire

User interface aesthetics

(1) questionnaire

Our hypotheses are that, for similar performance, 1) time and mouse clicks will be reduced when speech is available since doctors can directly ask their questions and translate them, 2) the learning curve is less steep for a speech-enabled system, so there will be less difference in terms of time and clicks between first and second uses, and that 3) speech will improve user satisfaction.

4 Methods

4.1 Task

We used a crossover design, as illustrated in Table 1. Both systems (BabelDr, System 1, and MediBabble, System 2) were used twice by ten medical students in order to find precise information about the patient, based on two different scenarios (A and B). The students had to find out, for example, if the patient had fever or if the pain radiated somewhere. Each scenario contained ten elements the students should find. The order in which the systems were presented to them was balanced among the participants (5 students began with BabelDr and 5 with MediBabble), each participant performing a total of 4 tasks, two with System 1 and System 2 in scenario A (Session 1), and two others with System 2 and System 1 in scenario B (Session 2). A task was finished when the students had found all ten elements.

Session1		Session 2	
Task 1	Task 2	Task 3	Task 4
BabelDr	MediBabble	MediBabble	BabelDr
Scenario A	Scenario A	Scenario B	Scenario B

Table 1: Crossover design

The anamnesis information elements to be found were selected by a person external to the project and no tuning was done before the experiment. The questions were expressed in an abstract way in order to avoid participants simply reading the sentence (for example “fever?”, “appendicitis?”). In each scenario, answers were provided by a standardized patient (for example fever: yes).

4.2 Languages

The language pair used in the study was French (doctor) to Spanish (patient).

4.3 Tools

MediBabble was used on a tablet and BabelDr on a laptop. In both systems, TTS was used for speech output. In BabelDr, an interaction consists of pressing the Recognition button, speaking, checking the canonical form and pressing the Translation button for speech output (Figure 2). The user was also able to consult the list of canonical sentences to learn system coverage, but all interactions with the system were spoken.

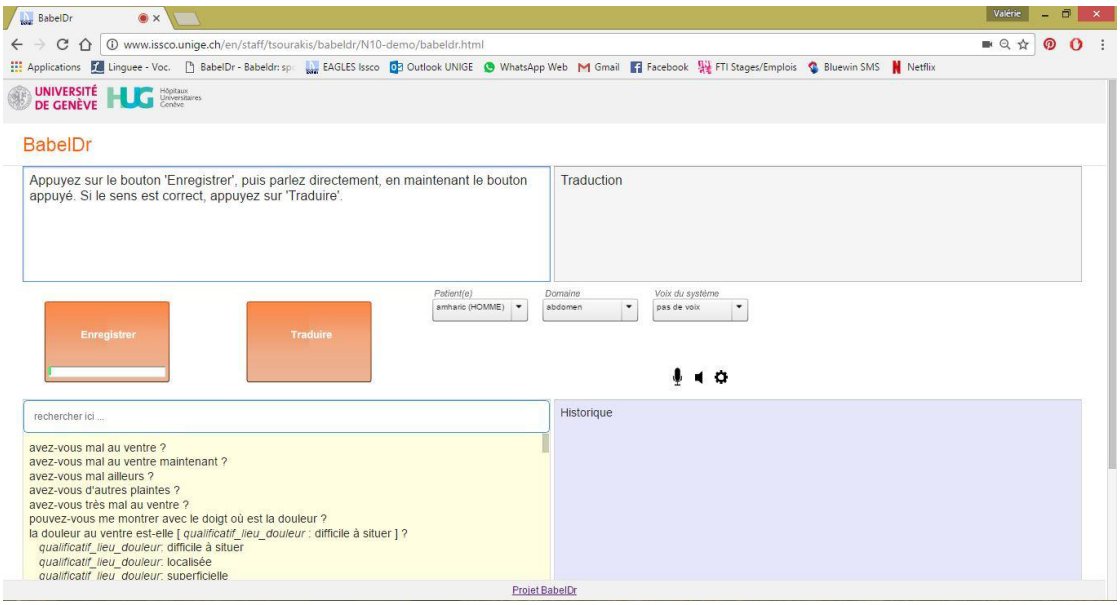

Figure 2: BabelDr

In MediBabble, the interaction consists of navigating in the tree (for example, selecting introduction => greetings => “bonjour” or Physical Exam => Abdomen => “please lay down”) or of searching with keywords in the search menu and then clicking on a sentence for its translation (see Figure 3).

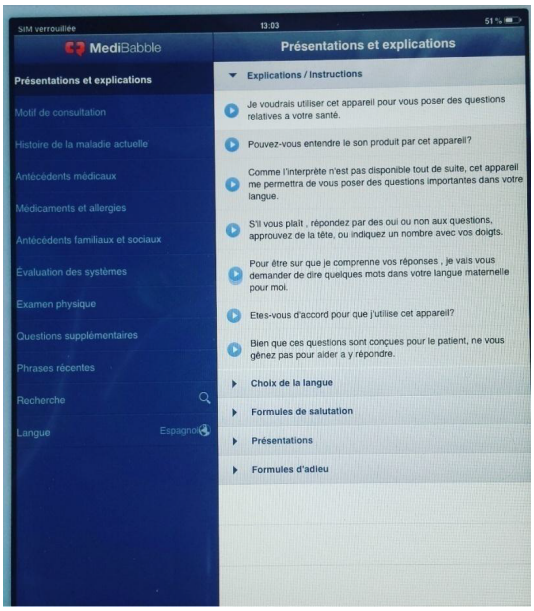

Figure 3: MediBabble

4.4 Participants

All participants were recruited at the HUG and were paid for the task:

- Spanish-speaking patients: two standardized Spanish patients, both females
- French-speaking doctors: ten French-speaking medical students

The doctors were all medical students, 5 of them advanced and 5 juniors (with less than 4 years of training). They were all French speakers and had no prior knowledge of the systems. Some of them (N=4) had already been confronted with patients who did not speak the same language.

4.5 Location and duration

The study took place at the Faculty of Translation and Interpretating at the University of Geneva. It took place in one room and lasted a total of three days. All participants had one hour to complete the four tasks.

4.6 Data collected

The following data were collected during the experiments: video recordings of the room, screen capture videos, information collected from the doctors after each session, demographics and satisfaction questionnaires. Time and mouse clicks were extracted from the screen capture videos. They were calculated for each interaction and correspond to the time and number of clicks between the first action of the interaction and the end of the speech output in Spanish.

5 Results

5.1 Successful interactions

Both phraselators made it possible to collect most of the required information from the patients in the time allocated. In the two sessions, participants collected 195/200 correct elements with MediBabble and 198/200 with BabelDr. This shows that both phraselators allowed the doctors to obtain the correct answers to the medical questions. The two systems are thus closely matched in terms of functional suitability.

5.2 Time

Table 2 shows the average time in seconds needed to find the required information for each task and system. The average time was lower for BabelDr in both sessions. With BabelDr, the time was similar in the first and second sessions (20s and 19s), while in MediBabble it increases (30s and 37s), tending to show that MediBabble is not necessarily easier to use the second time.

	first use	second use
BabelDr	20	19
MediBabble	30	37

Table 2: Average time per request in s

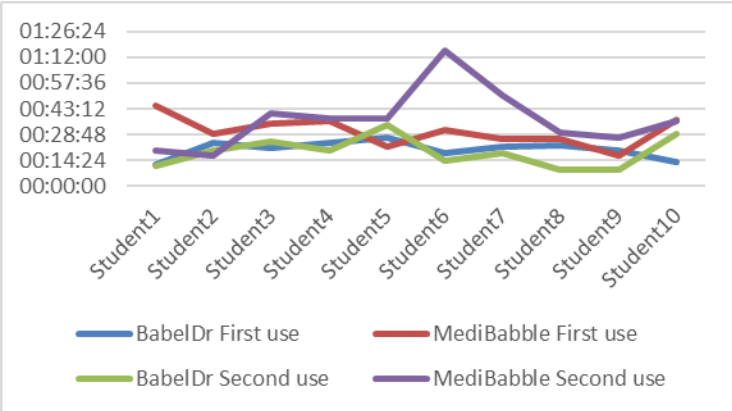

Figure 4: Average time per request for each participant

Figure 4 shows the measured times of each student for each system and task. Apart from three exceptions, the time needed was always lower for BabelDr than for MediBabble, which shows that, overall, interaction through speaking is quicker. This seems to be the case regardless of which system was used first. We also observe that there is less variation in terms of interaction time with BabelDr than with MediBabble. This suggests that the users encountered less time-consuming difficulties with BabelDr.

5.3 Clicks

Mouse clicks confirm the findings regarding time. As illustrated in Figure 5, the average number of mouse clicks is always lower for BabelDr than for MediBabble, with the exception of student 2. With BabelDr, the number of clicks remains similar for both sessions, while for MediBabble there is more variation. Table 3 shows the average number of clicks for all users. In BabelDr, an ideal interaction would require two clicks (one for recognition and one for translation). The average measured was 3.3 in both tasks, indicating that users were not always satisfied with the first recognition result, either because recognition was incorrect, or because mapping towards the canonical sentences did not match the user's expectations. In these cases, users either had to speak again or search for the intended sentence in the list of canonical sentences.

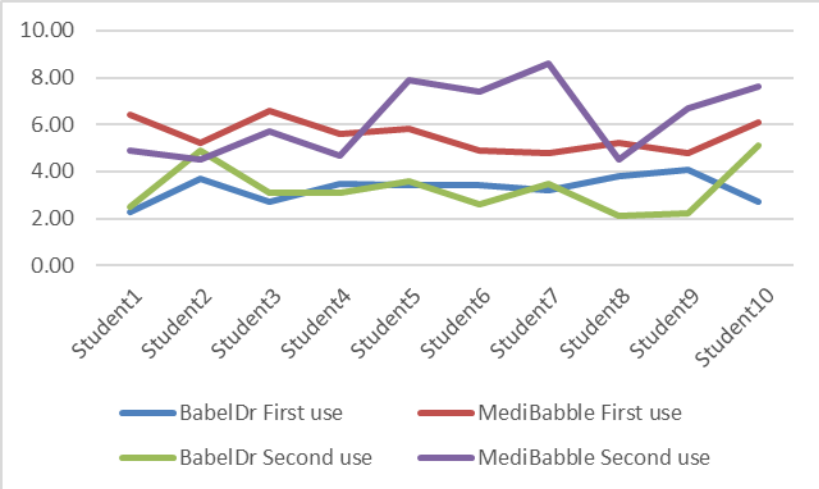

Figure 5: Average number of clicks per request for each participant

phraselator	first use	second use
BabelDr	3.3	3.3
MediBabble	5.5	6.3

Table 3: Average number of clicks per request

5.4 Questionnaire

Participants had to fill in a questionnaire after each task, which was different for the first and second session (Table 4). The first one contains 10 Likert scale questions focussing on ergonomics, in which the respondents are asked to rate statements. The second one contains 9 questions with a stronger focus on learnability.

First session

Q1 Le système était facile à utiliser (*The system was easy to use*)

Q2 L'interface m'a plu (*I liked the interface*)

Q3 J'ai eu l'impression de bien maîtriser le système (*I felt confident using the system*)

Q4 Le système était pratique à utiliser (*The system was convenient to use*)

Q5 Le système m'a facilement permis de surmonter la barrière linguistique avec le/la patient.e (*The system allowed me to easily overcome the language barrier*)

Q6 J'ai pu poser mes questions de manière naturelle (*I was able to ask my questions naturally*)

Q7 D'une manière ou une autre j'ai pu poser toutes mes questions (*I was able to ask all my questions*)

Q8 Le système permet une certaine flexibilité dans la manière de poser les questions (*The systems allows a certain flexibility in terms of formulating questions*)

Q9 J'ai apprécié de pouvoir consulter l'historique (*I appreciated being able to access the dialog history*)

Q10 J'ai trouvé ce type de système agréable/confortable à utiliser (*I found this type of system pleasant to use*)

Second session

Q1	Il m'était beaucoup plus facile d'utiliser le système lors de la seconde utilisation (*Using the system was much easier the 2nd time*)
Q2	Il m'était beaucoup plus rapide d'utiliser le système lors de la seconde utilisation (*Using the system was much faster the 2nd time*)
Q3	J'ai trouvé les informations plus facilement lors de la deuxième utilisation que la première (*Finding the information was easier the 2nd time*)
Q4	J'ai commis moins d'erreurs lors de la seconde utilisation que la première (*I made fewer errors the 2nd time*)
Q5	Je pense qu'il est facile d'apprendre à utiliser un tel système (*Learning to use such a system is easy*)
Q6	Je pense qu'il est rapide d'apprendre à utiliser un tel système (*Learning to use such a system is fast*)
Q7	Je savais mieux où chercher les informations lors de la deuxième tentative que la première (*I knew better where to look for information the 2nd time*)
Q8	D'une façon ou d'une autre j'ai pu poser toutes mes questions au patient (*I was able to ask all my questions*)
Q9	J'ai eu l'impression de bien maîtriser le système (*I felt confident using the system*)

Table 4: Questionnaires

Figure 6 shows the questionnaire responses after the first session. The bubble size represents the number of responses in each category. Globally, we observe that BabelDr has more positive answers than MediBabble. The two systems are fairly close in terms of ease of use (Q1-Q5), even if anecdotally doctors prefer the BabelDr interface (Q2: all doctors like the BabelDr interface while only ¾ like MediBabble) and think that BabelDr is more convenient to use (8 strongly agree for BabelDr vs 4 for MediBabble in Q4). The low score in Q6 for both systems shows that doctors had the feeling that they cannot ask questions in a natural way, even if speech seems to make a difference. In particular, doctors find that BabelDr allows for more flexibility in terms of formulating questions (Q8). All 12 doctors agree that they could ask all their questions with BabelDr, while only 10 do so for MediBabble (Q7).

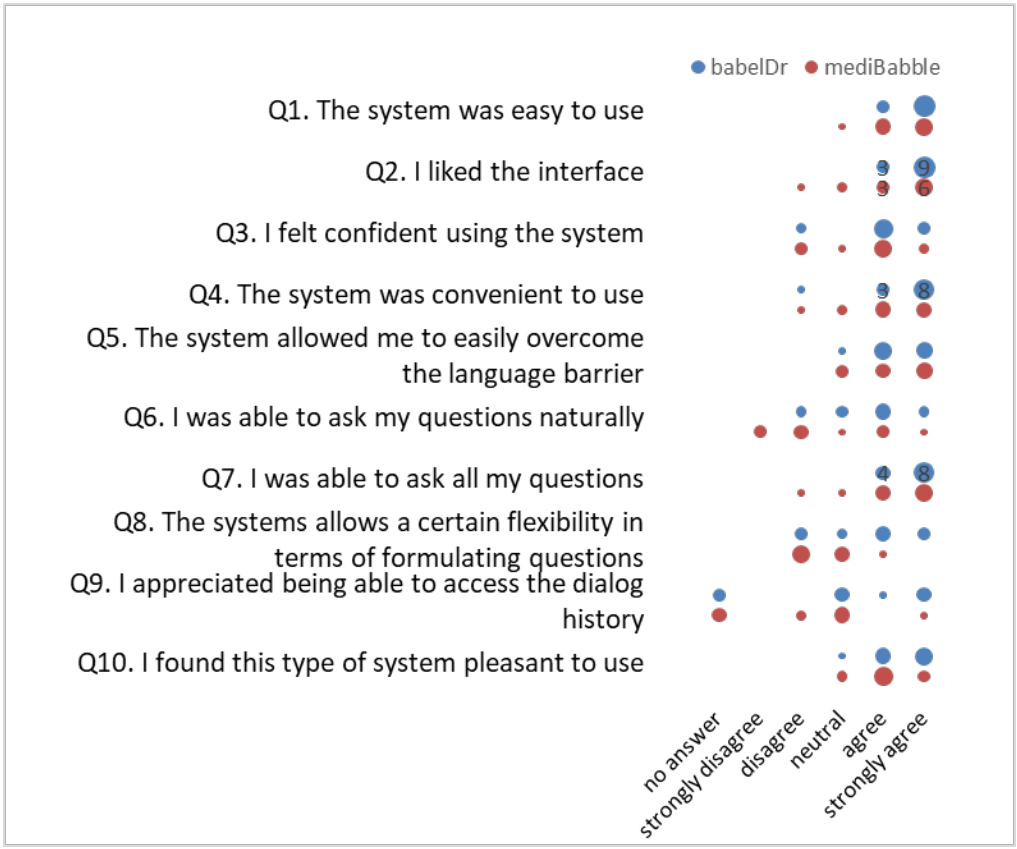

Figure 6: Questionnaire responses after the first session

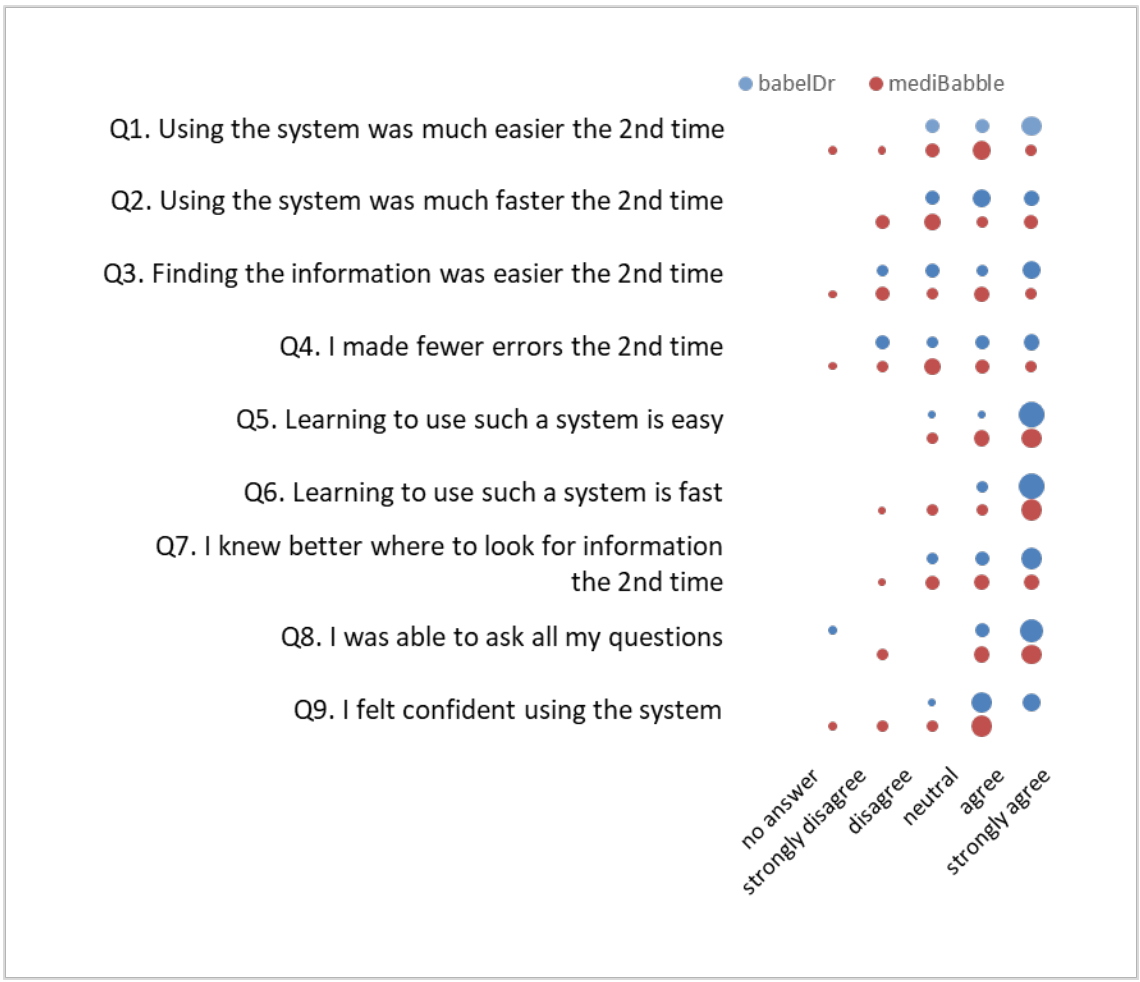

Figure 7: Questionnaire responses after the second session

The responses to the questionnaire completed after the second session are shown in Figure 7. As for the first session, results are in favour of BabelDr, with larger bubbles towards the right hand side, indicating more positive opinions. The first four questions (Q1-4), in which users were asked to compare their performance between using the system for the first and second time, show mixed results, with some users seeing improvement and others not. This suggests that the learnability of the systems is limited. The next two questions (Q5&6), relating to the user's overall impression of learnability, clearly show that users found it easier and faster to learn to use BabelDr. The two last questions (Q8&9), which were common to both questionnaires, show that for both systems, the perceived ability to ask all questions was not increased and that users felt slightly more confident using the systems the second time.

6 Discussion and conclusion

Today, language barriers represent one of the greatest problems in healthcare services for migrants. Since MT is not precise enough, other solutions should be found. We developed a speech-enabled phraselator called BabelDr. The aim of this study was to compare two phraselators, one with speech and one without speech. Doctors completed multiple scenarios using both systems with the objective of finding precise information about the patient.

Reviewing the results in terms of the quality model defined in Section 3, we find that, both objectively and subjectively, the availability of speech in BabelDr makes this system more suitable than a standard phraselator. On the level of functional suitability, BabelDr allowed the users to find nearly all the answers to the questions. Subjectively, doctors had the feeling they could ask more questions with BabelDr. In terms of operability, the objective measures of time

and number of clicks showed that less time and effort were required than with the non-speech-enabled phraselator, independent of the sessions and the users, thereby confirming our first hypothesis. Subjective results show a similar trend: users reported that BabelDr was more flexible and that it enabled them to ask their questions more naturally. They also found this system more convenient to use. In terms of learnability, we observed no decrease in time with the second use of BabelDr, and the time and number of interactions required for individual questions showed little variation. These results suggest that BabelDr is more intuitive to use than the menu-based system, i.e. that it requires very little learning time. Results for MediBabble show no learning effect and large variations in time and effort; however, further experiments would be necessary to confirm whether this could be improved by long term use. Finally, in terms of user interface aesthetics, user satisfaction is higher when speech is available, which confirms our third hypothesis. This study confirms that a speech-enabled phraselator such as BabelDr can be a good alternative for anamnesis in emergency settings when no interpreter is available.

These positive results, as well as the fact that phraselators are easy to port to new languages, make us think that this type of translation tool can fill a gap and contribute to healthcare accessibility.

Sources

Ahmed, Farhia; Pierrette Bouillon, Chelle Destefano, Johanna Gerlach, Angela Hooper, Manny Rayner, Irene Strasly, Nikos Tsourakis, Catherine Weiss (2017): “Rapid Construction of a Web-Enabled Medical Speech to Sign Language Translator Using Recorded Video”. *Future and Emergent Trends in Language Technology.* Seville (Spain): Springer editions

Bouillon, Pierrette; Johanna Gerlach, Hervé Spechbach, Nikos Tsourakis, Sonia Halimi (2017): *BabelDr vs Google Translate: a user study at Geneva University Hospitals (HUG).* EAMT Conference: Prague, Czech Republic

Boujon, Valérie (2017): *Systèmes de traduction de dialogues médicaux : Évaluation de BabelDr et MediBabble.* Université de Genève. Maîtrise. https://archive-ouverte.unige.ch/unige:96059 (3 May 2018)

Cardillo, Elena (2015): *Medical terminologies for patients*, SemanticHeathnet, Annex 1 to SHN WP3 Deliverable D3.3, 2015

Estrella, Paula; Nikos Tsourakis (2016): “Migrating from ISO/IEC 9126 to SQUARE: A Case Study on the Evaluation of Medical Speech Translation Systems”. *Design, Development, and Integration of Reliable Electronic Healthcare Platforms*. IGI Global

Flores, Glenn, et al. (2003): “Errors in medical interpretation and their potential clinical consequences in pediatric encounters”. *Pediatrics* 111 [1]

Hale, Sandra (2007): *Community Interpreting*. Hampshire and New York: Palgrave Macmillan.

Hacker, Karen; Maria Anies, Barbara L. Folb, Leah Zallman (2015): “Barriers to health care for undocumented immigrants: a literature review”. *Risk Management Healthcare Policy*, 156: 1108-1113

Iezzoni, Lisa I. ; Bonnie. L. O’Day, Mary Killeen, Heather Harker (2004): “Communicating about health care: Observations from persons who are deaf or hard of hearing”. *Annals of Internal Medicine* 140 [5]: 356-362

Jacobs, Elizabeth A.; Laura S. Sadowski, Paul J. Rathouz (2003): “The impact of an enhanced interpreter service intervention on hospital costs and patient satisfaction”. *Journal of General Internal Medicine*, 22 [2]: 306-311

Major, George et al. (2012): “Exploring lexical gaps in Australian Sign Language for the purpose of health communication”. *Communication & Medicine 9* [1]: 37-47

Middleton, Anne; Graham H. Turner, Maria Bitner-Glindzicz, Peter Lewis, Martin Richards, Aangus Clarke, Dafydd Stephens (2010): “Preferences for communication in clinic from deaf people: A cross-sectional study”. *Journal of Evaluation in Clinical Practice 16* [4]

Patil, Sumant; Patrick Davies (2014): “Use of Google Translate in medical communication: evaluation of accuracy”. *BMJ*, 349

Priebe, Stefan; Sima Sandhu, Sonia Dias, et al. (2011): "Good practice in health care for migrants: views and experiences of care professionals in 16 European countries". *BMC Public Health*, 11

Rayner, Manny; Alejandro Armando, Pierrette Bouillon, Sarah Ebling, Johanna Gerlach, Sonia Halimi, Nikos Tsourakis (2016): "Helping Domain Experts Build Phrasal Speech Translation Systems". José F. Quesada et al. (eds): *Future and Emergent Trends in Language Technology*. Seville (Spain) 2015, Springer editions, 41-52

Wasserman, Melanie et al. (2014): "Identifying and Preventing Medical Errors in Patients With Limited English Proficiency: Key Findings and Tools for the Field". *Journal for Healthcare Quality*

Wu Ann Chen; John M. Leventhal, Jacqueline Ortiz, Ernesto E. Gonzales, Bryan Forsyth (2006): "The interpreter as cultural educator of residents". *Archive of Pediatric and Adolescent Medicine 160*, 1145-1150

Authors

Valérie Boujon, Translator at AMPLEXOR International in Toulouse. E-mail: valerie.boujon@gmail.ch

Pierrette Bouillon, Professor at the Faculty of Translation and Interpreting of the University of Geneva. E-mail: Pierrette.Bouillon@unige.ch

Hervé Spechbach, Médecin adjoint responsable de l'unité d'urgences ambulatoires at Geneva University Hospitals (HUG). E-mail: Herve.Spechbach@hcuge.ch

Johanna Gerlach, Senior Researcher at the Faculty of Translation and Interpreting of the University of Geneva. E-mail: Johanna.Gerlach@unige.ch

Irene Strasly, PhD candidate at the Faculty of Translation and Interpreting of the University of Geneva. E-mail: Irene.Strasly@unige.ch

DOI 10.21256/zhaw-3000

Alireza Darvishy

Assistive Technologies: Short Overview and Trends

Abstract

This paper gives a brief overview of currently existing assistive technologies for different kinds of disabilities. An elaborate discussion of all types of assistive technologies is beyond the scope of this paper. Assistive technologies have evolved dramatically in recent years and will continue to be further developed thanks to major progress in artificial intelligence, machine learning, robotics, and other areas. Previously, assistive technologies were highly specialized and were often difficult or expensive to acquire. Today, however, many assistive technologies are included in mainstream products and services. An introduction and state of the art of assistive technologies are presented first. These are followed by an overview of technological trends in assistive technologies and a conclusion.

1 Introduction

Assistive technologies (ATs) are any item, piece of equipment, software program, or product system that is used to increase, maintain, or improve the functional capabilities of persons with disabilities (PWDs). Assistive technology helps people with a variety of impairments, be it difficulty speaking, typing, writing, remembering, pointing, seeing, hearing, learning, or walking. ATs can come in many forms, from low-tech to high-tech. For the context of this paper, we will consider only ICT-based assistive technologies, such as computer hardware (e.g. special switches, keyboards, and pointing devices) and software (e.g. screen readers and communication programs).

Due to the advancement of computer technologies, assistive technologies have become both more affordable and more sophisticated in recent years. For instance, refreshable Braille displays and screen readers for visually impaired users have changed dramatically. The first refreshable Braille display for visually impaired people, called VersaBraille, was developed in 1982, and was very expensive (Van Gerven/Taylor 2009). In 1986, IBM created the first screen reader for DOS computers (Cooke 2004) and in 1989, Macintosh released the first screen reader for graphical user interfaces called OutSpoken. Today, sophisticated screen readers are already embedded into most operating systems and mainstream computer devices for free (e.g. VoiceOver for Apple operating systems).

2 Assistive technologies: State of the art

A great many ATs today are based on existing mainstream technologies, particularly wireless Internet and mobile devices. Some ATs are already embedded into these technologies, such as the accessibility settings on an iPhone. There are also many specialized tools based on mainstream Internet technologies developed specifically for PWDs. For example, assistive mobile apps, also called “accessibility apps,” have sharply increased in number in recent years. Navigation and map apps such as Wayfindr (n.d.), ViaOpta (Novartis Pharma AG 2014), and WheelMap (Sozialhelden e.V. n.d.) allow users with mobility and/or vision impairments to locate and navigate accessible places. Apps such as these might use GPS information to analyse the user’s location and recommend safe routes, or they may rely on crowdsourcing to identify and rate accessibility in various establishments, such as restaurants, shops, or public buildings. Remote person-to-person apps, such as Convo (2017), VEASYT (2018), Be My Eyes (n.d.), and VerbaVoice connect users with hearing and vision impairments to a volunteer or interpreter, who receives a live camera and microphone feed from the user’s device and can then assist the user with a given task or provide sign language interpretation. There is a plethora of other internet-based mobile apps that offer a range of assistive tools.

Webtools and online platforms also offer many services for PWDs, particularly services which are more complex, as these would be less well-suited to the simpler formats of mobile apps.

Examples of webtools for PWDs include Amovil (n.d.), a tool which helps users identify the best-fitted accessible mobile device for them, PAVE (ZHAW ICT-Accessibility Lab 2016), which assists document creators in making accessible PDFs, and Robobraille (Sensus ApS. 2018), an online service which automatically translates text into Braille.

Other types of more hardware-based assistive technologies help people with cognitive and hearing impairments: so-called “social robots,” such as Pepper (Emotion Robotics n.d.), can help people on the autism spectrum to improve their social skills. Specialized hearing devices in combination with Bluetooth and cloud-based services, such as Oticon Opn (Oticon 2018), can help people with hearing impairments to overcome some barriers in their daily lives.

3 Assistive technology trends

Different areas such as cloud-based technologies, the brain-computer interface (BCI), artificial intelligence (AI), machine learning, and the Internet of Things (IoT) are facilitating new types of assistive technologies for people with disabilities.

Cloud-based assistive technologies, as their name implies, are stored in the Cloud and can be used as needed, independent of location. A European Commission-funded project, “Cloud4All,” has launched an initiative called the Global Public Inclusive Infrastructure (GPII), which aims to build a cloud-based system where users can store customized AT software. This initiative would allow users with disabilities to access their preferred ATs from any device and location. For example, a user with a cognitive impairment who uses a simplified computer interface could have these settings saved in the cloud, and could later access them from other devices, such as a library computer or work device. The GPII initiative is still under development, but it could have great potential in making ICTs more accessible for PWDs.

The brain-computer interface is at the cutting edge of AT research, and may have a significant impact for persons with motor impairments. The BCI allows users to bypass conventional channels of communication (such as speech or gestures) by creating a direct connection between a user’s brain and a computer via electrodes or implants (Obiedat et. al 2014). The user can thus control a device, such as a wheelchair or computer, using only brain-generated electrophysiological signals.

Artificial intelligence and machine learning, two distinct but intertwined phenomena (Marr 2016), have become buzz words in recent years. Artificial intelligence (AI) refers to the ability of machines to carry out tasks in a “smart” way, while machine learning refers to the application of AI when computers are given access to data and programmed to learn for themselves. For example, machine learning allows computers to draw from vast databases in order to “understand” certain external stimuli, such as images, text, and sounds.

The increasing ability of computers to understand the world has many potential uses for PWDs. YouTube’s automated closed-captions are one major example of applied machine learning as an assistive technology. As of 2017, YouTube’s closed-captioning algorithms not only recognize speech with a high level of accuracy, but are also able to recognize other sounds such as applause, laughter, and music (Simonite 2017). Another example is Facebook, which recently launched an image recognition feature, allowing the platform to create text descriptions of images in a post, for example. Google has also developed its own algorithm for this task (Vinyals 2014).

The “Internet of Things” (IoT) is expected to gain importance for assistive technologies. The IoT is what allows “smart” devices to operate autonomously via remote communication and data transmission with other devices and systems. For example, a user could use their smartphone to control the lighting in their home. The IoT can be applied to anything from transportation (such as self-driving wheelchairs or Google’s self-driving car), to security systems, to everyday objects such as refrigerators. An emerging field of development is the “smart home,” which combines various smart objects and systems in a domestic setting, allowing a user to have full control of their comfort and safety. Another emerging concept is that of “smart cities,” in which urban infrastructure is connected to the Internet. In this scenario, a PWD could use an app to find a free parking spot, for example, or could use a virtual city guide to find accessible building entrances.

4 Conclusions

Assistive technologies play an important role for the inclusion of people with disabilities. In earlier times, assistive technologies were often highly specialized, expensive, and not widely available. Today, tremendous advances in information and communication technologies have made assistive technologies more affordable and easier to acquire than ever before. Assistive technologies are increasingly ubiquitous, and are frequently included in mainstream technologies. Advances in artificial intelligence, machine learning and computer vision are providing new opportunities to develop novel assistive technologies. Self-driving cars, robotics, and 3D printing facilitate new use cases for PWDs.

Sources

Amóvil (n.d.): “Collaborating to update accessible mobile devices.” http://www.amovil.es/en (24 April 2018)

Be My Eyes (n.d.): “Bringing sight to blind and low vision people.” https://www.bemyeyes.com/ (24 April 2018)

Convo (2017): “Video Relay Service.” https://www.convorelay.com/ (24 April 2018)

Cooke, Annemarie (2004): “A History of Accessibility at IBM.” *AccessWorld*. American Foundation for the Blind. http://www.afb.org/afbpress/pub.asp?DocID=aw050207 (24 April 2018)

Emotion Robotics (2017): “Emotion Robotics.” http://www.emotion-robotics.com/Robotics-For-Education.aspx (24 April 2018)

Marr, Bernard (2016): “What is the Difference Between Artificial Intelligence and Machine Learning?” *Forbes online* – https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/#7bbb37b32742 (24 April 2018)

Novartis Pharma AG (2014): “ViaOpta.” https://www.viaopta-apps.com/ (24 April 2018)

Obiedat, Qussai; Maysam Ardehali; Roger Smith (2014): “EEG Brain-Computer Interface as an Assistive Technology: Adaptive Control and Therapeutic Intervention.” *Resna Annual Conference 2017*. https://www.resna.org/sites/default/files/conference/2017/emerging_technology/Obiedat.html (11 May 2018)

Oticon (2018): “Oticon Opn. Hearing aids.” https://www.oticon.com/solutions/opn (24 April 2018)

Sensus ApS. (2018): “RoboBraille.” https://www.robobraille.org/ (24 April 2018)

Simonite, Tom (2017): “Machine Learning Opens up New Ways to Help People with Disabilities.” *MIT Technology Review* – https://www.technologyreview.com/s/603899/machine-learning-opens-up-new-ways-to-help-disabled-people/ (24 April 2018)

Sozialhelden e.V. (n.d.): “Wheelmap. Rollstuhlgerechte Orte finden.” https://wheelmap.org (24 April 2018)

Van Gerven, Clara; Taylor, Anne (2009): “The Information Age Braille Technology Timeline.” Future *Reflections 28* [1] – https://nfb.org/Images/nfb/Publications/fr/fr28/fr280109.htm (24 April 2018)

VEASYT (2018): “Digital solutions for content accessibility.” http://www.veasyt.com/ (24 April 2018)

VerbaVoice (n.d.): “Dolmetscher online zuschalten.” https://www.verbavoice.de/ (24 April 2018)

Vinyals, Oriol; Alexander Toshev; Samy Bengio; Dumitru Erhan (2014): “A Picture Is Worth a Thousand (coherent) Words: Building a Natural Description of Images.” 17 November – http://googleresearch.blogspot.co.uk/2014/11/a-picture-is-worth-thousand-coherent.html (24 April 2018)

Wayfindr (n.d.): “Accessible Indoor Audio Navigation.” https://www.wayfindr.net/ (24 April 2018)

ZHAW ICT-Accessibility Lab (2016): “PAVE. Accessible PDF.” http://pave-pdf.org/ (24 April 2018)

Author

Alireza Darvishy is a Professor and Head of the ICT Accessibility Lab, a national competence center on eAccessibility at the Zurich University of Applied Sciences in Switzerland. E-mail: Alireza.Darvishy@zhaw.ch

DOI 10.21256/zhaw-3000

Christian David

Text-To-Speech Audio Description on the Internet: Potential of Web-Based Applications for More Accessibility

Abstract

Many videos on the Internet are not accessible to blind and visually impaired people. When only the sound of a video is available, information essential to understand the content is often missing. Audio description is a method of closing this information gap by acoustically describing the visual content and listening to it as additional commentary during the pauses in the dialogue. Due to the currently very complex production process, audio description is hardly ever used on the Internet. However, by using synthetic voices and playing audio description via standardized video players on the Internet, the general conditions for widespread use can be changed in a positive way. In this article, two software programs are presented that can optimize this production process and standardize the playout. This opens up new applications and new economic opportunities.

1 Introduction

These days it is impossible to imagine our lives without the Internet. We use it both professionally and privately. Many people spend a big part of their lives online. Videos are an essential element of the Internet. Websites such as YouTube or Netflix are becoming more and more popular, but videos are also an integral part of social networks and modern education. Today, on many news websites, videos complement text-based articles. Some providers have already replaced textual messages completely.

Most videos on the Internet are not accessible to blind and visually impaired people. When only the sound of a video is available, important information that is needed to understand the content is often missing. So-called audio description (AD) can help here. It is a proven method for describing visual processes. This description is spoken and serves as an additional commentary complementing the video.

In Germany, audio description is mainly used in public television and film productions. Due to the currently very complex production process and the lack of digitization, it is not possible to implement a widespread use of AD. For cost reasons, private television does not offer any audio descriptions at all. Audio description is also hardly used on the Internet.

This article presents a prototype of a web-based editor for audio description. The software aims to facilitate the AD authoring process and utilizes text-to-speech for the audio output. Thus, it enables authors to create and publish audio descriptions by themselves, without the need of professional audio recording tools. The article also contains a short introduction to “Able Player” – a fully accessible open source media player developed and maintained by the University of Washington.

This article is a snapshot of the current state of development and of available technologies, and it is intended to provide suggestions for further research.

2 Text-to-speech audio description (TTS-AD)

Synthetic voices are increasingly encountered in everyday life. Navigation systems in our cars tell us where to go; voice assistants like Alexa or Siri make it easier for us to use smartphones and other devices. Apps such as Microsoft Translator can translate and read out whole texts for us. We are becoming increasingly accustomed to dealing with synthetic voices. In addition, synthetic voice technology is progressing rapidly and the pronunciation of the systems is getting better and better. Modern text-to-speech (TTS) software uses deep learning strategies to constantly optimize itself.

2.1 Research

In Europe, Szarkowska has investigated whether the playout of audio description with synthetic voices (text-to-speech audio description, TTS-AD) can be an alternative to reduce production times and costs (Szarkowska 2011). Among other things, it was examined how high the acceptance of TTS-AD was and where practical problems arose when creating the synthetic output.

Further studies in the following years (Szarkowska/Jankowska 2012; Fernández i Torné/Matamala 2015; Kurch 2016) have treated the topic in more depth. In particular, the acceptance in different genres and movie lengths has been tested and the technical possibilities for adjusting synthetic voices examined more intensively.

3 TTS-AD on the Internet

Essentially, Szarkowska's basic considerations still apply. TTS-AD will not replace classic AD, but will extend the fields of application. The playout with text-to-speech software can offer a decisive advantage for the widespread use of AD, as production costs and times can be significantly reduced, which is essential to pave the way for AD on the Internet.

3.1 Production

3.1.1 Current production process for audio description

The current production process for audio description is relatively complex and the playout is not standardized, as seen in Figure 1. After the text has been written by a seeing author and approved by a visually impaired reference person, voice recording takes place in the recording studio with a professional speaker. The audio description track is then mixed with the original sound, converted, and played out on different channels such as television, home entertainment, or cinema.

Figure 1: Production process for audio description

The production steps are rarely automated, and the video and audio formats may vary between different playback devices. Thus, a standardized playout across all channels is not possible. Manual intervention is essential.

This leads to high production times and costs. It also binds a large part of the available human resources and a lot of necessary expertise to relatively few productions. Thus, this is not applicable for audio description on the Internet, as the demand for low-cost and fast production for a very large amount of video content cannot be met in this way.

3.1.2 Simplified workflow for TTS-AD on the Internet

By consistently focusing on speech synthesis and playout on the Internet, the production process can be simplified and highly automated (see Figure 2). The author loads a video, adds the audio description text, and creates a live preview for the approval process. After that, the

software creates an audio file using text-to-speech technology and finally publishes the combined video and audio description on the Internet.

All necessary software components can be built on top of web technology and thus a unified solution for the creation and playout of TTS-AD on the Internet can be implemented. W3C standards (World Wide Web Consortium; W3C 2017a) as well as the high availability of the Internet on a large number of devices – such as smart TVs, smartphones, or Internet of Things (IoT) – help to achieve this end result.

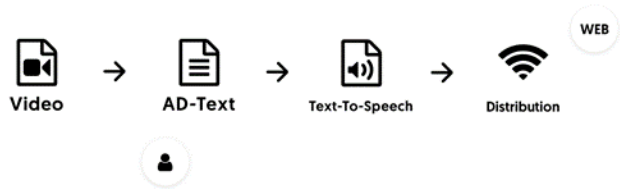

Figure 2: Optimized production process for TTS-AD on the Internet

3.1.3 A web-based TTS-AD editor

“Frazier” (see Figure 3) is the proof-of-concept of a web-based software for creating TTS-AD for Internet videos. It is currently under development at VIDEO TO VOICE GmbH. Some of its development goals are:

(1) The complexity of the entire AD production process will be reduced to such an extent that audio description can be produced without specialist knowledge, e.g. in the field of voice recording or sound mixing.

(2) Ideally, the software accompanies the user through the entire AD production process. All steps are completed inside the TTS-AD editor, and all required and generated materials are stored in a central media pool.

(3) Eventually, AD authors should be able to generate the final product on their own immediately after finishing the authoring and approval process.

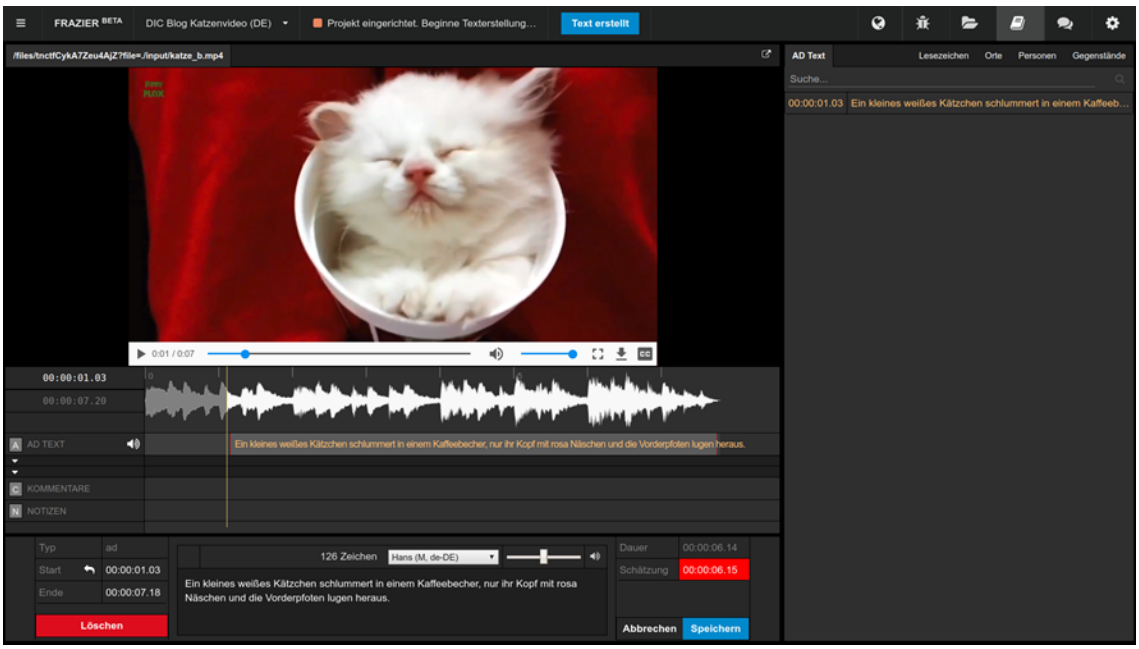

Figure 3: The Frazier user interface

The authoring is done using the integrated text editor, which allows frame-accurate work. It provides a live preview of the synthetic voices. The selection of different male and female voices is possible as well as the adjustment of speed and pronunciation.

After creating and approving the texts, the TTS-AD can be exported directly as an audio file. It is also possible to save audio descriptions as video text track (VTT) files for playback in suitable video players using a screen reader. Currently the prototype is not able to automatically publish the video with a mixed audio track. More work needs to be done here to fulfill development goal number 3 (see above).

In addition, the audio description text can also be exported as a classic manuscript. Thus, the editor is also suitable for the production of classic audio descriptions with typical voice recording and mixing.

3.2 Playout

The W3C Web Content Accessibility Guidelines (W3C 2017b) regulate the prerequisites and possibilities for achieving accessibility on the Internet, e.g. for playout of audio description. The implementation of the guidelines, however, requires experience and discipline – when building the website itself, but also when creating the contents and the necessary alternatives.

While the most important browser manufacturers have already agreed on the implementation of a unified audio and video codec, there are still differences in the implementation of accessibility features for video players. For example, there is no unified support for keyboard control. Functionality and key assignment vary between browsers.

In addition, no browser is currently able to handle multiple audio tracks in one video file. Thus, browsers cannot activate an additional audio track with audio description as known from TV. The support of text-based VTT files – a file format to provide additional subtitles or audio description – is already defined in WCAG 2.0, but the support for audio description is not implemented in browsers.

JavaScript-based video players can help here. They compensate for these shortcomings by extending the browsers' functionality.

3.2.1 Able Player: Fully accessible cross-browser media player

An example of these JavaScript-based video players is Able Player (Figure 4), which enables web developers to provide accessible media on all modern browsers. Able Player has mainly been developed by the University of Washington. It is available as open source software and can be used by everyone.

Able Player allows unified keyboard control in all browsers, supports playlists, and offers the possibility to play audio descriptions – optionally via a screen reader or by playing back an alternative video with integrated audio description.

In addition, Able Player supports subtitles in multiple languages, videos with sign language, and interactive transcripts that allow the user to search and navigate through the text itself.

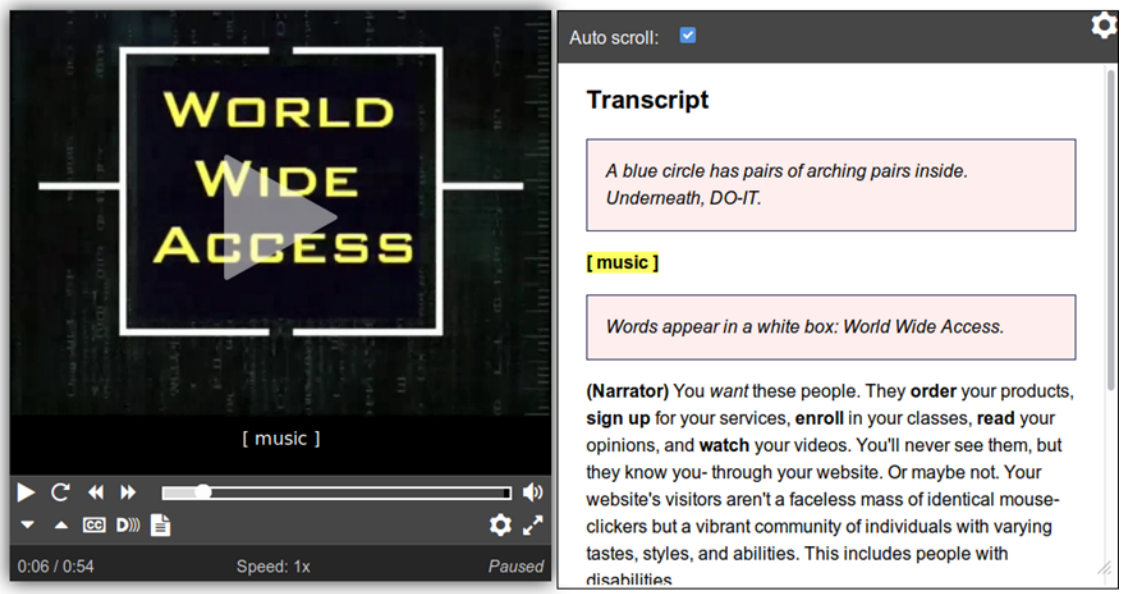

Figure 4: Able Player with activated audio description and interactive transcript

3.2.2 VTT vs. media files

The playout of TTS-AD on the Internet can be accomplished in two different ways, with both having advantages and disadvantages.

VTT files are plain text files. These are very small files and therefore can also be used with very limited bandwidth. These files can also be used to provide a full-text search within the AD-text. The audio output itself occurs via the screen reader of the user's device and the individual user can choose the voice. A disadvantage is the lack of control over the output, given that voice speed and pronunciation depend on the respective TTS system and the selected voice. Therefore, the result is usually less accurate than a generated audio track. For example, overlaps with dialogues or important noises cannot be ruled out.

The playout of TTS-AD using media files – a separate audio track or video with original audio and AD mixed together – offers the advantage that the end product corresponds exactly to the result in the TTS-AD editor using the live preview. The voice, speed, pronunciation, and timecodes are converted directly and transmitted 1:1 to the user. Therefore, the final result is always identical – on every device. The playout also works on systems that do not have a screen reader installed. Disadvantages are the larger amount of data and the lack of searchability, since the AD text itself is not available inside the audio file. Finally, the voice can no longer be adjusted by the individual user.

Practical experience has shown that it makes sense to combine both formats. In that way, the precise and predictable results of the media file can be combined with the searchability of the

VTT file. Able Player plays a described video with mixed audio tracks and utilizes the VTT to build the interactive transcript. This allows the user to navigate through the video by using the text search function.

3.2.3 Extended Audio Description

A relatively new form of audio description also exists.

The Internet makes it possible to separate the original sound and the audio description track, and this allows the traditional scheme to be altered. For example, it is possible to play the audio description as a prologue at the beginning of a video – or even pause the video while a longer audio description is played. This opens up new possibilities, which currently are not technically feasible for cinema and television. Time cannot be stretched there; audio description has to get along with the dialogue breaks and to always run synchronously with the original sound.

The so-called “Extended Audio Description” is already defined in WCAG 2.0 (W3C 2016). Able Player offers the possibility to activate “Extended Audio Description” for VTT-based playouts.

4 Conclusion and outlook

Through the consistent use of web technologies and the concentration on a unified and automated playout of audio description using synthetic voices, the production process for audio description on the Internet can be greatly accelerated. For example, automated playout allows the creation of audio descriptions for time-critical content such as daily news on the Internet.

The integration of an authoring tool with automated output allows an easy start to produce audio descriptions with significantly lower technical knowledge than within the classical production workflow. In combination with freely available accessible video players, this makes it easy for website owners to provide accessible video content for their users.

The “Extended Audio Description” represents a promising evolution of classical audio description. It enables significantly more information about a video to be provided. For example, it opens up new fields of applications, such as the description of e-learning material or the slides of a lecture. Thus, an audio-described film can develop into an adequate alternative offer for the video itself. Tutorial videos or lecture series can become podcasts – and thus form their own sales channel or independent offering.

There are further fields of application for the technologies presented in this article. For example, similar areas of application such as the provision of easy-to-read language, audio subtitling or synchronization are conceivable. Further studies would make sense with regard to the acceptance of synthetic voices in these areas.

Sources

Able Player (n.d.): “Fully accessible cross-browser HTML5 media player.” University of Washington. https://ableplayer.github.io/ableplayer/ (24 February 2018)

Fernández i Torné, Anna; Anna Matamala (2015): “Text-to-speech vs. human voiced audio descriptions: a reception study in films dubbed into Catalan.” *JoSTrans. The Journal of Specialised Translation* [24]: 61-88 – http://www.jostrans.org/issue24/art_fernandez.pdf (30 August 2017)

Kurch, Alexander (2016): Audiodeskription mit Sprachsynthese – Eine Perzeptions- und Rezeptionsstudie am Beispiel des Dokumentarfilms “buy buy st. pauli”. Unpublished Master’s thesis. Hildesheim: University of Hildesheim

Szarkowska, Agnieszka (2011): “Text-to-speech audio description. Towards a wider availability of AD.” *JoSTrans. The Journal of Specialised Translation 15*: 142-162 – http://www.jostrans.org/issue15/art_szarkowska.php (24 February 2018)

Szarkowska, Agnieska; Anna Jankowska (2012): “Text-to-speech audio description of voice-over films. A case study of audio described Volver in Polish.” Elisa Perego (ed): *Emerging topics in translation: Audio description*. Trieste, Italy: EUT Edizioni Università di Trieste, 81-98

VIDEO TO VOICE GmbH (2017): “Frazier. Audiodeskription mit Text-to-Speech-Technologie (TTS-AD).” https://www.videotovoice.com (24 February 2018)

W3C (2016): “Extended Audio Description (Prerecorded). Understanding SC 1.2.7.” Understanding WCAG 2.0: A guide to understanding and implementing WCAG 2.0. https://www.w3.org/TR/UNDERSTANDING-WCAG20/media-equiv-extended-ad.html (24 February 2018)

W3C (2017a): “W3C Standards.” https://www.w3.org/standards/ (18 April 2018)

W3C (2017b): “Web Content Accessibility Guidelines (WCAG) 2.0.” https://www.w3.org/WAI/intro/wcag (24 February 2018)

Author

Christian David is a full stack software developer with 20 years of experience in planning and implementing web applications. He is a co-founder of VIDEO TO VOICE GmbH. E-mail: c.david@videotovoice.com

DOI 10.21256/zhaw-3000

Nathalie Dreier

Access to Music Programs on Swiss Television for the Deaf and Hard of Hearing

Abstract

This article gives insight into how music programs are being subtitled on Swiss television and into the challenges posed by this practice and the various approaches associated with it. A survey carried out with the target audience reflects the importance music has to them, while also showcasing the groups' strong heterogeneity. The research reviewed in this article strongly indicates that the subtitling of music as a socially inclusive cultural and entertainment medium deserves more attention in the future if subtitles are to fully replace the acoustic component of audiovisual texts and thereby grant greater levels of accessibility.

1 Introduction

This article provides a summary of the main findings of my master's thesis[17], in which I researched access to music programs for the deaf and hard of hearing in German-speaking Switzerland. The two main aims of the study were to explore how music is subtitled on SRF 1, the main Swiss national television channel of the Swiss Broadcasting Corporation SRG SSR, and to gain insight into how these subtitles are perceived by the target audience, i.e. the deaf and hard of hearing.

With the rapid evolution of new media, a number of translation practices have emerged that have helped to improve access to information for specific social groups. One of these practices is subtitling for the deaf and hard of hearing (SDH), which became available on Swiss television in the 1980s, and has been expanding steadily since the 1990s (cf. SWISS TXT 2018: n.p.). Political measures such as the UN's *Convention on the Rights of Persons with Disabilities*, which was adopted in Switzerland on 15 May 2014, have acknowledged the right of access to culture, including television programs and films, and have strongly contributed to increased accessibility. Currently, more than 50 percent of television programs on SRF 1 are subtitled. This percentage is likely to increase in the future given the growing number of elderly people. According to the WHO (2017: n.p.), approximately one third of 65-year-olds are affected by disabling hearing loss.

In line with similar developments around the world, SDH has drawn the global interest of scholars and professionals in the audiovisual environment. The technological, temporal, visual and linguistic challenges subtitlers and viewers face are often the topic of academic research[18]. A particular filmic code demanding additional inquiry into its subtitling practice is music. The importance of music and sound in visual media is generally undisputed, and scholars and professionals in the field of SDH agree that these elements must be included when subtitling for the deaf and hard of hearing in order for the audience not to miss out on important aural information (cf. Neves 2010: 125). Especially in music programs, the importance of transferring the music cannot be over-emphasized. However, when looking at the practice of music subtitling, and even the available guidelines and conventions of producers of these music subtitles, it becomes apparent that there is still much to be done if the richness of music is to be conveyed, and if subtitles are to, as Neves (2010: 125) words it, "serve the purpose of the audiovisual text in all its effects".

Subtitling music for the deaf and hard of hearing poses challenges on many levels. Studies have shown, for instance, that the needs and preferences of the hard of hearing differ from

[17] For the full body of work in German, see Dreier (2017)

[18] See, for instance, Díaz Cintas/Remael (2007), Hezel (2009), Jüngst (2010), Romero-Fresco (2011, 2015, in press) and Jekat (2014)

those of the deaf when it comes to subtitling practices (cf. Gottlieb 2015: 17). Furthermore, not only does the degree of hearing impairment among these individuals, who belong to different linguistic and socio-cultural groups, vary, but the importance of music may be different for each of them – as is the case for individuals with no hearing impairment[19]. Harrison (2013: 187) writes that providing access services is about inclusion, and that SDH therefore gives an audience with hearing impairment the choice to access whatever content their hearing peers have access to. Be it as a source of information or entertainment, films and television programs constitute an important factor in societal life and are often the topic of discussion in one's social environment. SDH allows the target audience to take part in these discussions (cf. Jüngst 2010: 105, 124).

In this article, I will first give an overview of the subtitling practice at SWISS TXT, the company that has been producing subtitles for SRF since 1984, as this is a crucial aspect of my practice-oriented study. Subsequently, I will summarize the collection of data before discussing the results and future outlook of this field.

2 Subtitling practice and music subtitling at SWISS TXT

SWISS TXT complies with the presentation guidelines for subtitles as set out by the regulations of the nine regional broadcasters within the ARD, ARD Text, ORF, SRF and ZDF (ARD 2015: n.p.). In addition, a program-specific subtitling handbook is used, where subtitling guidelines are set. These guidelines are valid for *pre-recorded* or *classic* subtitling as well as for *live* and *semi-live* subtitling[20].

At SWISS TXT, live subtitling is conducted via *respeaking*[21] with the speech-to-text-system *Dragon Naturally Speaking*, which is designed so that subtitlers (hereinafter respeakers) can individually train their user profiles. Via an interface to the subtitling program *FAB Subtitler*, subtitles are produced live and cued in text blocks. The live subtitles produced by SWISS TXT for SRF are usually sent with a three to seven second delay (cf. Romero-Fresco 2011: 31f.). When respeaking, the respeaker attempts to caption the live program with as extensive and complete subtitles as possible (cf. Jekat 2014: 87). This live subtitling process is usually intralingual, but can also be interlingual, by transferring the original language into another target language. In German-speaking Switzerland, this is hypothetically the case as the respeaker must respeak the original text in Swiss German dialect into Standard German (cf. Jekat 2014: 91). The respeakers at SWISS TXT must therefore have perfect command of both the Swiss dialects and Standard German.

Semi-live subtitling is often used for live programs for which scripts are available in advance, or which include pre-recorded segments such as interviews or archive footage (cf. Romero-Fresco 2011: 12). Here, the respeaker has time to prepare a list of subtitles beforehand and then cue them manually in sync with the program during the live broadcast.

The guidelines for subtitling music included in the handbook of SWISS TXT (2016) can be summarized as follows. Music subtitles are to be displayed in white lettering on a black background. Descriptive music subtitles are displayed between asterisks, while subtitles that include song lyrics are marked by a number sign. Song lyrics should be displayed in the subtitles if they are important for the narrative, ideally in the original language, and therefore also in the Swiss dialects. Furthermore, song lyrics are to be introduced by the mood of the music, the genre, the song title, and the artist (SWISS TXT 2016: 11). A dash is used to

[19] For more insight on music and people with hearing impairments, see Neary (2001), DiBernardo Jones (2016), Neff (2009) and Dreier (2017)

[20] See Romero-Fresco (2011) for detailed definitions.

[21] Romero-Fresco defines respeaking as "a technique in which a respeaker listens to the original sound of a live programme or event and respeaks it, including punctuation marks and some specific features for the deaf and hard of hearing audience, to a speech recognition software, which turns the recognized utterances into subtitles displayed on the screen with the shortest possible delay" (Romero-Fresco 2011: 1).

indicate a change of singer. The handbook also states that it is more important to describe the mood of the music than to give information on the artist or the title of the song (SWISS TXT 2016: 12). To capture the mood of the music, the handbook refers to a document by Klemm (SWISS TXT 2009), which lists various adjectives to describe music, with the preferred adjectives underlined.

3 Data

In order to pursue the aims set out in the research, a combination of methods was employed. The first method is a corpus-based research analysis of music subtitling in practice. The second method is a survey of the target audience, which is based on the results of the corpus analysis.

3.1 Corpus analysis

For this part of the empirical research, a corpus (Dreier 2017: appendix) of all music subtitles from the five-session television music show “Ich schänke dir es Lied” (*I gift a song to you*, hereinafter ISDEL), which was broadcast live on SRF 1 in 2017, was built. In this show, members of the general public can surprise one another with songs that are meaningful to them and then have the songs played by popular musicians. These songs are either played live during the live show or in pre-recorded sequences that are shown during the show. In total, 11 hours and 29 minutes of live broadcast material was compiled, which rendered 1491 music subtitles. In order to systematically analyze these subtitles, aspects were classified based on relevance to specific research.

In ISDEL, music appeared in four categories classified according to the importance the music has for the respective sequences and, parallel to this, according to the production approach. These categories are: live musical performances, pre-filmed musical performances, background music in pre-recorded sequences and background music in live sequences. This classification made it possible to observe tendencies in the subtitling of music according to its purpose within the respective sequence of the program. By separating sequences that are filmed live from sequences that are pre-recorded and broadcast live, it was also possible to analyze the effect that the time constraints of live subtitling have on the production of music subtitles, and so, to detect the challenges of live music subtitling. The music subtitles of each category were analyzed with regard to their content and frequency. This analysis allows insight into, for instance, how many of the subtitles consist of song lyrics, the mood of the music, the genre, the artist or the instrument playing, etc. and, in addition, into the linguistic patterns that are being used. Within this analysis, it was relatively uncomplicated to recognize subtitles containing song lyrics, as they were usually marked by the number sign, as envisaged in the SWISS TXT handbook. Descriptive music subtitles required more detailed analysis.

The results of this corpus analysis were compared to those of an earlier existing corpus of music subtitles from the 2016 casting show “Die Grössten Schweizer Talente” (*Switzerland’s Greatest Talents*) created by Hagmann/Krähenbühl (2016). These subtitles are from shows that were broadcast live, as well as from pre-recorded shows. The quantitative and qualitative analysis and comparison of both corpora was therefore able to show patterns in the way music is being subtitled live and on pre-recorded shows on SRF1. In addition, the subtitles produced in practice were assessed against the subtitling guidelines set out by SWISS TXT, which made it possible to identify their implementation and explore challenging sequences.

3.2 Survey of the target audience

The survey of the target audience was based on the results of the corpus analysis. This survey is meant to be a sample study and to offer input for further research in the sense of a qualitative study, which gives insight into the perspective of the target group and allows for further validation by quantitative research (cf. Albert/Marx 2010: 13). The survey was carried out

online via the survey tool *Umfrage Online*[22] and included embedded video segments from ISDEL on PLAY SRF with the subtitles as they were broadcast live. The respondents were mainly recruited via associations and organizations for the deaf and hard of hearing.

It is worth noting that while preparing the survey, special attention was paid to formulating the questions particularly clearly, as sign language is the mother tongue of many deaf people and it therefore had to be assumed that written Standard German was not the first language of some of the respondents (cf. Hezel 2009: 178f.). Also, at the beginning of the survey, the respondents were asked to watch the video segments as they would normally do, for example to turn their hearing device on or off respectively. This is relevant because if someone uses a cochlear implant and has it turned on, it can be assumed that auditive signals are recognizable for this person, whereas someone who is deaf and does not use a hearing device cannot, or can only barely, recognize tones (cf. Sonos 2017: n.p.).

4 Results and discussion

In this chapter, some of the results of the corpus analysis and the survey from the research (Dreier 2017) are presented and discussed. To reflect the multilingual character and complexity of subtitling on Swiss television, the examples of the subtitles from the research have been retained in the original language, including subtitles in Swiss dialect, and an English version in italics has been added on the right side or underneath. All examples are taken from the corpus and are in their original form as they were broadcast on the live shows on SRF 1 (Dreier 2017: appendix).

4.1 Subtitling in sequences with music as the central element

Due to the nature of the show, the ISDEL corpus is mainly comprised of subtitles of music that is present as the central element, i.e. as the narrative of a sequence. Music subtitles during these sequences are essential for transmitting the main information made available through acoustic signs. The first category analyzed consisted of music subtitles belonging to sequences in which live performances take place in the arena and are broadcast live. This category comprised the largest part of the corpus and covers any sequences in which a song is sung or played live by anyone in the arena. Within this category, 831 music subtitles were analyzed, 785 of which were song lyrics. In the category of pre-recorded musical performances, where, for example, a band plays a song in someone's home or on the street, and the sequence is replayed on screen in the arena during the live show, 545 music subtitles were analyzed, of which 512 were song lyrics.

The results of the analysis of these two categories clearly show that when the viewer is watching a band or musician playing a song, he or she is mainly receiving information on the lyrics that are being sung. The song lyrics are almost always in their complete form and in the original language, as recommended in the handbook of SWISS TXT (2016:11). This is also the case in the corpus compiled by Hagmann/Krähenbühl (2016: 30) and indicates that the respeakers at SWISS TXT often receive a script with the songs that will be performed during a live show in advance, enabling them to prepare the subtitles off-line before the show. The production approach of such sequences can be regarded as semi-live.

There are, nevertheless, some differences between the first and second category, where the time constraints of subtitling live sequences become more apparent. In the first category, 45 musical performances were investigated. In only five of these performances were the song lyrics introduced by a descriptive subtitle that contained all four elements recommended by the handbook used at SWISS TXT: the mood of the music, the genre, the artist and the song title. In the second category, which contains subtitles of sequences with pre-recorded musical performances, which the respeakers can receive ahead of time and prepare, six out of 21

[22] www.umfrageonline.ch (25 April 2018)

performances, comparatively more than in the first category, had their lyrics introduced by a descriptive subtitle that contained all elements recommended by the handbook, as seen in example (1):

(1)

* Romantischer Popsong:	*Romantic popsong*
"Ewigi Liebi" von Mash *	*"Eternal Love" by Mash*

Another noteworthy discrepancy found between the first and second category is the presence of additional descriptive music subtitles in eight of the 21 pre-recorded performances. Such information was only available in one of the 45 live performances. The descriptive subtitles, displayed between asterisks, were usually included during a song between the subtitles of the lyrics, as is seen in example (2):

(2)

# S'isch alles so wie's sött.	*# Everything is as it should be*
* Sanfte Klavierklänge *	**Soft piano notes**
# Dä Momänt isch für d'Ewigkeit und meh.	*# This moment* *is for forever and more.*

In 2 of the 21 pre-recorded performances, information is also to be found between parenthesis about how something is being sung, or who is singing it, as shown in example (3):

(3)

(Alle zusammen)	*(All together)*
# Don't worry, be happy.	*# Don't worry, be happy.*

Such information is never available in the live performances. Hagmann/Krähenbühl (2016: 33) found that in their corpus of music subtitles, such information is only available in pre-recorded shows. This reveals that the characteristics of the subtitles in pre-recorded sequences can be likened to that of pre-recorded shows.

4.2 Subtitling in sequences with background music

In sequences containing background music, two categories were analyzed in the ISDEL corpus. In these categories, music is not the central element of a sequence, but essential for setting the mood[23]. First, the category of background music during pre-recorded sequences was analyzed. These sequences consist of pre-recorded material that is shown on screen in the arena during the live show. For these sequences, in which, for instance, the person who is

[23] For more information on background music in film, see Monaco (2012), Koebner (2002) and Neves (2010).

giving or receiving the song as a gift is introduced, subtitles can be prepared in advance. The corpus results show that background music is being subtitled in these sequences, with a total of 115 music subtitles available. In contrast to the first two categories, in which the music is the main action of the visual input, and songs are usually played in their entirety, the music in this category is usually present in short sequences and no entire songs are played. Here, descriptive subtitles are much more common than subtitles that consist of song lyrics, which constitute only 16 of the subtitles. Of the descriptive subtitles, 80 subtitles use adjectives to describe the music, as can be seen in examples (4) and (5). Hagmann/Krähenbühl (2016: 35) also observed that the mood of the music is most commonly described using adjectives[24].

(4)

* Gefühlvolle Musik *	* *Emotional music* *

(5)

* Spannungsgeladene Musik *	* *Suspense-packed music* *

The second category of background music constitutes background music during sequences that are filmed and broadcast live, such as music played when the presenter goes on stage or music that marks the transition to a pre-filmed sequence. As in the previous category, music is essential for setting the mood. However, in these sequences there are no music subtitles available. Hagmann/Krähenbühl (2016) also find that music is subtitled in only one such sequence and that otherwise, background music is only subtitled in pre-recorded shows. The results of these categories show that, in the music programs analyzed, there are only music subtitles available in live sequences when music is the main event of a sequence.

4.3 Difficulties encountered in live music subtitling

The analysis of the four categories mentioned above show that of all the live music available subtitles, the vast majority consist of pre-prepared song lyrics, and much of the subtitling can therefore be classified as being produced semi-live (cf. Romero-Fresco 2011: 12). Actual live music subtitling did not often take place in the ISDEL show. To effectively identify the situations in which this did take place, the subtitles in the corpus required analysis in their audiovisual context.

The fact that respeakers often receive the script of the show in advance, which enables them to see which songs will be played and to prepare the song lyrics for these sequences, becomes apparent when the song lyrics in the live performances in the ISDEL show differ from the original lyrics that may be found in a lyrics database or suchlike. A good example for this is provided in Nik P.'s performance in the third live show of ISDEL. Nik P. performs the song Ein Stern, der deinen Namen trägt (*A Star that Bears your Name*) to Jakob and Silvia, a married couple. Since Silvia is "gifting" this song to Jakob, Nik P. adapts his lyrics to dedicate this song to the couple. Table 1 presents a sample of some of the lines of the song played in this performance. The left column shows the music subtitles as they were broadcast live on SRF 1, and the right column lists the song lyrics as Nik P. actually sung them during the live show.

[24] For a more detailed insight into the adjectives being used in practice and their reception, see Hagmann/Krähenbühl (2016) and Dreier (2017).

Music subtitles in the corpus	**Lyrics sung live by Nik.P**
# Seit Jahren schon leb ich mit dir	Seit Jahren schon lebt sie mit dir
I have been living with you for many years	*She has been living with you for many years*
# und ich danke Gott dafür, dass er mir dich gegeben hat.	und sie dankt Gott dafür, dass er ihr Jakob gegeben hat.
And I thank God for giving you to me	*And she thanks God for giving Jakob to her*
# Als Erinnerung an unser Leben möchte ich dir heut etwas geben.	Als Dankeschön für deine Liebe möcht' sie dir heut etwas geben.
As a remembrance of our lives, I want to give you something today.	*As a thank you for your love, she wants to give you something today.*

Table 1: Music subtitles and actual lyrics from the performance of Nik P. in ISDEL live show 3 of 25 March 2017 (PLAY SRF 2017a)

The discrepancy between the music subtitles and the actual lyrics sung as seen in Table 1 runs throughout the entire song[25]. The target audience, who are dependent on the subtitles to receive the information their hearing peers are receiving, do not have access to this information and are unlikely to fully understand the context.

Another difficulty encountered in live subtitling is the subtitling of sequences in which respeakers do not receive information in the script about which song is going to be sung in advance, and therefore cannot prepare any subtitles of the song lyrics or any descriptive subtitles with the various elements recommended in the SWISS TXT handbook. One such example, as shown in Table 2, is taken from the finale of ISDEL. The singer Anna Rossinelli gifts the song *Sexual Healing* to the singer Kunz on the lounge sofa in the arena. The subtitles broadcast on SRF 1 for the live show are in the left column of Table 2. The right column contains what Anna Rossinelli actually sung during this performance. It is evident that only three subtitles with the words "sexual healing" are available for the target audience. The respeaker also did not use the symbols as set out in the SWISS TXT guidelines to mark the subtitles (asterisk or number sign), which also makes it unclear whether the subtitles were meant to mark the title of the song or what was actually being sung.

[25] For the table and analysis of the entire song, see Dreier (2017: 42f.)

Music subtitles in the corpus	**Lyrics sung by Anna Rossinelli during the live show**
-	Oh, baby now let's get down tonight
-	Baby, I can't hold it much longer
-	Now It's getting stronger and stronger
"Sexual Healing"	And when I get that feeling, I need sexual healing
-	Sexual healing, is good for me
-	Sexual healing, is good for me
"Sexual Healing"	Sexual healing
"Sexual Healing"	It helps to relieve the mind It helps to relieve the mind
-	Is good for us sexual healing is good for me
-	Sexual healing is something that is so very good for me

Table 2: Subtitles and actual lyrics from Anna Rossinelli's performance in the Finale of ISDEL of 4 April 2017 (cf. PLAY SRF 2017b)

Further examples of such challenging situations can be found in the ISDEL corpus and show that in sequences that are respoken live there are discrepancies between the original text, i.e. what is being sung in the show, and the subtitles produced via respeaking. The differences found are either subtitles with a content that deviates from the actual spoken or sung text, or subtitles with reduced content.

If subtitles are to convey all important information that is made available to a hearing audience, it is questionable whether the subtitles produced in such situations of live respeaking are acceptable. In her study on syntactic aspects of respeaking and differences between spoken texts and written texts, Jekat (2014: 105) writes that the performance of respeakers must be continuously evaluated, and the respeakers notified about any unacceptable discrepancies between the subtitles produced and the original spoken text. She also writes that, in all likelihood, unacceptable changes cannot always be avoided. Nevertheless, respeakers could be better trained in future, and instilled with a greater sensitivity towards the conveyance of music and its expressive force, giving it meaning to people who might have never even perceived sound before (cf. Neves 2010: 125f.). Especially in a live subtitling context, respeakers will need to make swift and purposeful decisions to give the subtitles this meaning. It is also worth mentioning that the ISDEL corpus reveals discrepancies between the comprehensiveness of music subtitles from show to show. The shows were subtitled by various respeakers, which suggests that not all of them deem the conveyance of music to be equally important. For music to be "translated", it must first be felt and, once felt, it must be understood and transferred into visual signs that convey equivalent effects (cf. Neves 2010: 26). Not everyone feels music in the same way. It would therefore make sense to elaborate the guidelines for music subtitling in the SWISS TXT handbook and render them more adaptable to the various production approaches.

Examples like the one described above demonstrate that there are situations in which respeakers do not receive sufficient information on what exactly a singer is going to sing live during a show. In an ideal future, all those involved in such television shows would be aware of the relevance of SDH and communicate accordingly. Respeakers would therefore be informed in advance of, say, an artist's decision to adapt the lyrics, enabling them to prepare adequate subtitles beforehand. It is most likely, however, that some situations will always remain impromptu, it could therefore be useful to sensitize respeakers to such situations, so that they can react during a performance and inform the target audience in a subtitle that the original song lyrics have been adapted and/or personalized. Respeakers could also be trained

to convey meaningful information under time pressure and possibly pay more attention to conveying the mood of a song, or a singers' voice, if it is not possible to find the lyrics beforehand.

A possible approach would also be to have bridging subtitles (*Überbrückungsuntertitel* according to Hezel 2009: 202), which serve the purpose of informing the target audience why they are not receiving any information. Hezel writes that such subtitles are used if, for example, there is a long period of silence in a film, to prevent the target audience from thinking that a technical defect has occurred. SWISS TXT also uses subtitles or editorial notes in white lettering on a blue background without asterisks to inform the viewer at the beginning of a program or in sequences where no subtitling is possible, that a program is being subtitled live (SWISS TXT 2016: 6). I propose using a similar informative note to inform the viewers why they are not receiving comprehensive subtitles during the program or, for instance, all of the song lyrics in the subtitles when a song is being performed.

In a situation such as that displayed in Table 2, more comprehensive guidelines could encourage respeakers to produce subtitles that might for instance read:

* She sings the soul song
"Sexual Healing" to him *

Powerful, sensual voice
Live-Subtitling: No complete song
lyrics available

It is also to be hoped that, in the future, new technological possibilities such as song recognition apps could help respeakers in such situations.

4.4 The sample survey

The sample survey was completed by 29 deaf and hard-of-hearing respondents, 12 of which defined themselves as deaf, eight as hard of hearing and seven as cochlear implant users. One person specified that he/she became deaf at an early age, and one person said he/she was deaf in one ear.

For the first part of the survey, the respondents were asked to watch a video of a live performance on ISDEL, which included the complete lyrics of the song being performed, introduced by a descriptive subtitle containing the title of the song and the artist. This sequence was chosen as it reflected the most frequent way in which live performances were subtitled in ISDEL. After watching this video, the respondents were first asked whether they consider music subtitles on television to be helpful, to which 28 respondents replied in the affirmative. Of the 29 respondents, 28 also liked the fact that the complete lyrics were shown. This result can be related to the proportionately high number of subtitles of song lyrics present in the corpus.

However, 13 respondents said they would like more descriptive information on the music being performed in the subtitles. Most commonly, the respondents wanted more information via sign language and karaoke-style lyrics. Due to the nature of live subtitling and the delays that are a default part of its character at present, lyrics that are in sync with what is being sung does not yet seem possible. For the time being, these might only be made available during pre-recorded shows.

There were three positive comments on the subtitles, where respondents specifically mention, for instance, that the subtitles help them as they cannot acoustically understand the lyrics that are being sung. It is worth mentioning that two of the respondents who left such positive

comments defined their hearing impairment as hard of hearing, and one said he/she wears a cochlear implant. It may be assumed that these respondents are able to acoustically recognize signals. The results of the survey also show that ten of the 13 respondents who wanted further information in addition to the lyrics describe themselves as deaf. These results suggest that for someone who can barely recognize auditory signals, receiving information only on the lyrics is insufficient, whereas for viewers who are able to detect the tune of the music, it is probably more helpful to receive information on the lyrics alone.

In the second part of the online survey, the respondents watched the sequence of Anna Rossinelli singing Sexual Healing (see Table 2). When asked whether they would rather watch the show at a later time with more complete subtitles, 16 respondents said they would. This reflects the general dissatisfaction with the way this sequence was subtitled live. That live subtitling is important and gives the target audience access to the same television content as their hearing peers is valid for live subtitling music programs as well. It also gives the target audience the possibility to enjoy the show with others. A different study in which approximately 2500 respondents with hearing impairment participated showed that, when watching a show where no subtitles are available, 28 percent of the respondents try to guess what is happening, while 29 percent try to lip read (Hezel 2009: 151). Other studies on the reception of subtitles have shown that the majority of respondents with hearing impairment will either change channels and look for a subtitled program or switch off the television altogether if there are no subtitles available (cf. Szarkowska et al. 2015: 58; Romero-Fresco 2015: 134). Taking this into consideration, it seems that live subtitles are useful to viewers such as these, even if they are not perfect. There is the possibility, as one of the respondents of the sample survey suggested, to give as complete information as possible during the live show and to later correct the subtitles for online streaming. Even though this might not be economically viable, it would most definitely offer more complete and coherent information to the viewers of the target audience who would like to watch the show at a later point.

The respondents were also asked whether they would have generally wished to have more elaborate subtitles with regard to the music in this sequence, to which 24 replied positively. Of these 24 respondents, 22 would have liked complete song lyrics in the subtitles of this performance (Table 2), which once again shows that the song lyrics are considered very important by the target audience.

5 Looking to the future

Due to the extreme heterogeneity of the target audience, it may seem that a single perfect solution does not, and cannot, exist for subtitling music. In future, however, and apart from further research, more cooperation between subtitle providers and target audience might be attempted in order to provide more adequate access.

It is worth pointing out that the research carried out in this study only covered music subtitling in the forms that are currently in practice at SWISS TXT. Alternative forms of subtitling music, such as describing it using sentence structures with verbs (e.g. *The music slows down*) were not provided as alternatives in the survey of the target audience. In addition, it must be noted that the positive feedback to the subtitling of song lyrics could also be due to the fact that it is the main form of music subtitling being practised. To know whether this is truly the ideal way to subtitle music, alternatives would have to be tested with the target audience, in which further elements of the music are included, and not just the text. Popular music consists of the underlying music of a song, the voice of the singer(s) and, only then, the lyrics (Kaindl 2013: 152), almost everyone who listens to music will affirm that it is much more than just lyrics. If Swiss television is to continue to expand its repertoire of subtitled shows, and not necessarily increase shows with superimposed sign language, this research suggests it would make sense to add to the subtitles more of the information that a sign language interpreter would, such as the rhythm of the music, the mood of the vocals and the instruments being played (cf. Kahler 2015: 2).

In the future, SDH is likely to see many changes due to technological and political developments. Over ten years ago, Neves (2007: 97) wrote that such changes, if carried out thoughtfully and conscientiously, could perhaps lead to true inclusion because "services will be supplied to all alike, allowing for subtle adjustments that will guarantee adequacy to all viewers as individuals". Until such subtle adjustments become a reality for both SRF's as well as foreign television providers' audiences, solutions and forms of music subtitling should be pursued that are continuously tested with the target audience. This could help develop music subtitling standards that provide the largest possible part of the target audience with optimal accessibility.

Sources

Albert, Ruth; Nicole Marx (2010): Empirisches Arbeiten in Linguistik und Sprachlehrforschung. Anleitung zu quantitativen Studien von der Planungsphase bis zum Forschungsbericht. Tübingen: Narr

ARD (2015): "Untertitel Standards." http://www.daserste.de/sezrvice/kontakt-und-service/barrierefreiheit-im-ersten/untertitel-standards/index.html (25 April 2018)

Díaz Cintas, Jorge; Aline Remael (2007): *Audiovisual translation: subtitling*. (Translation practices explained). Manchester: St. Jerome Publishing

DiBernardo Jones, Jeannette (2016): "Imagined Hearing. Music-Making in Deaf Culture." Blake Howe, Stephanie Jensen-Moulton, Neil Lerner, et al. (eds): *The Oxford Handbook of Music and Disability Studies*. Oxford/New York: Oxford University Press, 54-72

Dreier, Nathalie (2017): *Zugang zu Musiksendungen für Menschen mit Hörbeeinträchtigung. Beschaffenheit und Rezeption von Musikuntertiteln im Schweizer Fernsehen*. Unpublished Master's thesis. Winterthur: Zürcher Hochschule für Angewandte Wissenschaften, IUED Institut für Übersetzen und Dolmetschen

Gottlieb, Henrik (2015): "Different viewers, different needs: Personal subtitles for Danish TV?" Pablo Romero-Fresco (ed): *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern: Peter Lang, 17-44

Hagmann, David; Maria Krähenbühl (2016): Musikuntertitelung für Hörbeeinträchtigte am Beispiel der Unterhaltungssendung "Die grössten Schweizer Talente". Erstellung und Analyse eines Korpus. Unpublished Bachelor's thesis. Winterthur: Zürcher Hochschule für Angewandte Wissenschaften, IUED Institut für Übersetzen und Dolmetschen

Harrison, Mark (2013): "Making Music Television Accessible to a Hard-of-Hearing Audience." Helen Julia Minors (ed): *Music, text and translation*. London: Bloomsbury (Bloomsbury advances in translation), 187-199

Hezel, Susanne (2009): "Untertitelung für Hörgeschädigte für das deutsche Fernsehen: Vorgehensweisen, Forderungen, Richtlinien." Silke Nagel, Susanne Hezel, Katharina Hinderer, et al. (eds): *Audiovisuelle Übersetzung: Filmuntertitelung in Deutschland, Portugal und Tschechien*. Frankfurt am Main: Peter Lang, 145-261

Jekat, Susanne J. (2014): "Respeaking: Syntaktische Aspekte des Transfers von gesprochener Sprache in geschriebene Sprache." Susanne J. Jekat, Heike Elisabeth Jüngst, Klaus Schubert et al. (eds): *Sprache barrierefrei gestalten. Perspektiven aus der Angewandten Linguistik*. (TransÜD 69.). Berlin: Frank & Timme, 87-108

Jüngst, Heike E. (2010): Audiovisuelles Übersetzen. Ein Lehr- und Arbeitsbuch. Tübingen: Narr

Kahler, Lily (2015): "Grundlagenpapier Einführung in die Musikübersetzung. Vom Ton zum Bild: oder wie übersetzt man Musik?" Lily Kahler (ed): *Kompetenzen Entwickeln für die Musikübersetzung*. Zürich: Durchblick haben

Kaindl, Klaus (2013): "From Realism to Tearjerker and Back. The songs of Edith Piaf." Translated by Liselotte Brodbeck and Jacqueline Page. Helen Julia Minors (ed): *Music, text and translation*. London: Bloomsbury (Bloomsbury advances in translation), 151-161

Koebner, Thomas (2002): Reclams Sachlexikon des Films. Stuttgart: Reclam

Monaco, James (2012): *Film Verstehen. Kunst, Technik, Sprache, Geschichte und Theorie des Films und der neuen Medien*. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag

Neary, Walter (2001): “Brains of deaf people rewire to ‘hear’ music.” University of Washington. http://www.washington.edu/news/2001/11/27/brains-of-deaf-people-rewire-to-hear-music/ (25 April 2018)

Neff, Sarah (2009): *Im Rhythmus der Stille: Wie ich mir die Welt der Hörenden eroberte*. Frankfurt/New York: Campus Verlag

Neves, Josélia (2007): “A world of change in a changing world.” Jorge Díaz Cintas, Pilar Orero, Aline Remael (eds): *Media for All. Subtitling for the Deaf, Audio Description, and Sign Language*. (Approaches to Translation Studies 30.) Amsterdam/New York: Rodopi, 89-98

Neves, Josélia (2010): “Music to my eyes… Conveying music in subtitling for the deaf and the hard of hearing.” Łukasz Bogucki, Krzyztof Kredens (eds): *Perspectives on audiovisual translation*. (Lódz Studies in Language 20.) Frankfurt am Main: Peter Lang GmbH, 123-145

PLAY SRF (2017a): Ich schänke dir es Lied. Liveshow 3 (25 March 2017). https://www.srf.ch/play/tv/ich-schaenke-dir-es-lied/video/ich-schaenke-dir-es-lied---liveshow-3?id=c492893c-29f0-424e-a67b-2e2598354449 (25 April 2018)

PLAY SRF (2017b): Ich schänke dir es Lied – Finale (08 April 2017). https://www.srf.ch/play/tv/ich-schaenke-dir-es-lied/video/ich-schaenke-dir-es-lied---finale?id=201aa788-278f-43e6-83ea-5d51019379bc (25 April 2018)

Romero-Fresco, Pablo (2011): *Subtitling Through Speech Recognition: Respeaking*. (Translation Practices Explained 13.) Manchester: St. Jerome Publishing

Romero-Fresco, Pablo (2015): “Long questionnaire in the UK.” Pablo Romero-Fresco (ed): *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern: Peter Lang, 117-162

Romero-Fresco, Pablo (in press): “Reception Studies in Live and Pre-Recorded SDH.” Elena die Giovanni, Yves Gambier (eds): *Reception studies and audiovisual translation*, Amsterdam/Philadelphia: Benjamins

Sonos (2017): Gehörlosigkeit und Ertaubung. https://www.sonos-info.ch/hoerbehinderung/hoerbehinderung/gehoerlosigkeit-und-ertaubung.html (12 February 2018)

SWISS TXT (2009): Musikthesaurus: Stimmungen Version 2.4 (Adjektivzirkel Klemm)

SWISS TXT (2016): Styleguide der UT-Redaktion Zürich. HANDBUCH UNTERTITELUNG mit UT-Standards ARD, ORF, SRF, ZDF

SWISS TXT (2018): Geschichte. https://www.swisstxt.ch/de/swiss-txt/geschichte/ (25 April 2018)

Szarkowska, Agnieszka; Joanna Pietrulewicz; Anna Jankowska (2015): “Long questionnaire in Poland.” Pablo Romero-Fresco (ed): *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern: Peter Lang, 45-74

WHO (2017): Deafness and hearing loss. Fact sheet. http://www.who.int/mediacentre/factsheets/fs300/en/ (25 April 2018)

Author

Nathalie Dreier holds a Master’s degree in Applied Linguistics with a Specialization in Professional Translation from the Zurich University of Applied Sciences. She is profoundly interested in expanding all boundaries which restrict speech, as well as non-speech, communication and understanding. Currently, she works in corporate communications as a writer and translator, while pursuing her interests in music, photography and Chinese language. E-mail: nathalie_dreier@yahoo.com

DOI 10.21256/zhaw-3000

Daisy Lange

Comparing “Leichte Sprache”, “einfache Sprache” and “Leicht Lesen”: A Corpus-Based Descriptive Approach

Abstract

The classification, differentiation and description of “Leichte Sprache” by means of linguistic categories is a research desideratum. Even though individual approaches have already been pursuing this objective, their outcomes are based mainly on the analysis of intuitively developed writing guidelines. There is therefore still a lack of comprehensive studies that include the current textual practice. The aim of this article is to describe “Leichte Sprache” by contrasting it to similar approaches of text simplification (i.e. “einfache Sprache” and “Leicht Lesen”).[26] In order to do so, a selection of descriptive categories is examined, mainly at the level of vocabulary and text pragmatics. This investigation is based on an extensive corpus of texts, which was created for this purpose. The findings indicate, in part, clear linguistic differences among the three approaches of text simplification and confirm the hypothesis that these could be positioned on a sort of “comprehensibility continuum”. Observations of exemplary linguistic structures are also discussed with a view to improving comprehensibility in relation to content and readership and potentially increasing the inclusion of target groups.

1 Background and leading research questions

Research on “Leichte Sprache” (LS) is primarily concerned with the comprehensive definition and description of the subject of enquiry – what is “Leichte Sprache”? Answering this question is necessary to ensure the long-term effectiveness and the successful application of the concept (Bock/Lange 2015: 67). To this end, Maaß (2015) and Bredel/Maaß (2016) have recently contributed to a theoretical classification of the subject matter by lining, describing and evaluating most of the rules proposed by writing guidelines (e.g. BMAS 2013, BITV 2.0 2011) and by giving further writing or text modification advice for certain language phenomena such as word classes. Bock (2014), Bock/Lange (2015) and Lange/Bock (2016) have made further contributions towards a semantic conceptualisation of LS that include a more pragmatic and empirical point of view. In this context, the definition of LS has largely become established as a variety of German. The function of this conveyance variety (“Vermittlungsvarietät”, i.e. conveying content between “difficult” texts and readers’ lower receptional capacities; see Bock 2014; Bock/Lange 2015) essentially consists in making textual information as comprehensible as possible and transferring it in an addressee-oriented way (cf. Bock/Lange 2015: 68; Seitz 2014: 5). LS shares this goal with similar approaches, including “einfache Sprache” (eS) and “Leicht Lesen” (LL). However, the various approaches partially differ from one another in their linguistic and typographic design, in the target groups they address and in the areas of communication they consider. They are therefore conceptualised as part of a comprehensibility *continuum* (Bock 2014; Jekat et al. 2017; Bredel/Maaß 2017), with differing language structures reaching from very simplified to more complex ones depending on their addressed recipients, text intentions and text types, etc.

The extent to which textual information is organized and transferred in compliance with more or less precise or differentiated sets of rules of LS (e.g. Netzwerk Leichte Sprache 2013; BMAS 2013; BITV 2.0) has, up to now, only been analysed in individual texts or small collections of texts. The results of such analyses often show that recommendations are only inconsistently implemented.[27] Despite a few exceptions (e.g. Kuhlmann 2013; Kaczmarzik 2017; Jekat et al. 2017; Lasch 2013; Bock 2015; Lange/Bock 2016), a systematic corpus-based description of

[26] Due to their language specifics and their establishment within the German speaking countries, these terms will not be translated into English.

[27] In a survey on the use of LS in work contexts, Bergelt/Goldbach/Seidel (2016) also came to the conclusion that the rules in LS usage are only appropriate or helpful to varying degrees and are therefore applied differently.

the distinctive linguistic structures of LS is still lacking.[28] At present, valuable insights into the acceptability, appropriateness and adaptation of "standardisation tools" can only be gained through textual practice.

Similarly, no relevant guidelines or explicit sets of rules are available for eS. A clear definition of eS is also missing. A comparative analysis of (competing) approaches of text simplification is not only desirable in order to develop research-driven categories and definitions (Bock 2016: 397; see also Lange/Bock 2016), it is also needed in order to investigate possible internal linguistic differentiations that should be taken into consideration when transferring information to diverse target groups (or to only one target group) (see Jekat et al. 2017).

As part of the research project "Leichte Sprache im Arbeitsleben" (LeiSA, Easy-to-Read in Work Contexts, University of Leipzig),[29] an extensive corpus of simplified texts was created. This corpus allows for detailed comparative analyses among three approaches: LS, eS and LL. These analyses address the following research questions:

(1) What distinctive attributes does LS have in terms of linguistic features, text types and text functions? What language usage specifics are emerging?

(2) To what extent are LS recommendations adhered to in practice?

(3) What are the similarities and differences in language usage among the three approaches?

The quantitative analysis presented in this article is explorative and focuses on the lexical and pragmalinguistic features of the corpus texts.[30] Its aim is to isolate specific categories that can be used to position each individual variety on a comprehensibility continuum (see also Jekat et al. 2017). Based on these categories, a qualitative analysis will then follow in which the three approaches will be examined more closely, described in detail and thus clearly demarcated from each other.[31]

2 The LeiSA corpora

The LeiSA corpus consists largely of digital texts. The texts were collected from institutions and collaborators (e.g. workshops, employment offices, translation agencies etc.) as well as through systematic online research during a period of two and a half years from 2015 to 2017. Major inconsistencies in the writing guidelines sometimes made it difficult to assign the texts to a specific variety (on the difficulties in assembling the corpus, see also Lange/Bock 2016). Certain principles of classification had to be defined on the basis of the text features identified (see 2.1). Various word lists were created. Some of these were lemmatized and annotated according to word class, text type and text function (see sections 3.2 and 3.3).

2.1 The LeiSA core corpus

In order to evaluate and describe the varieties in question, different corpora were assembled. The following analysis is mainly based on the LeiSA core corpus, which is divided into different sub-corpora. Its composition is briefly outlined below. Within the scope of this analysis, however, the three sub-corpora (LS, eS and LL) are not differentiated further.

---

[28] For a differentiation between LS and eS, see Kellermann (2014) and Lange/Bock (2016)

[29] The project was funded by the Federal Ministry of Labour and Social Affairs (BMAS) from 2014 to 2017. Further information is available at http://research.uni-leipzig.de/leisa/ (20 February 2018)

[30] For a more detailed review and description of these specific grammatical features, see Bock (2017) on passive and negation, Lange (forthcoming a) on the use of genitive and attributes, Lange (forthcoming b, d) on the use of connectors and pronouns and Lange (forthcoming c) on other lexical features such as word separation.

[31] Preliminary findings have already been published in Lange/Bock (2016). The corpus has since been expanded and partially reorganised. Therefore, some of the earlier analyses have been newly integrated into this article.

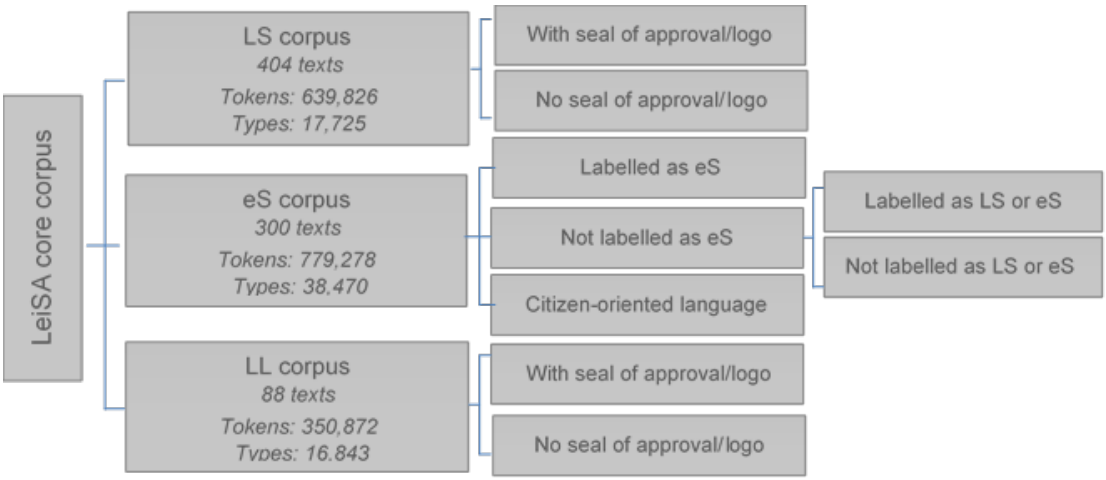

Figure 1: Composition and size of the LeiSA core corpus

(1) "Leichte Sprache" corpus (LS)

The LS corpus contains all texts that were created, issued or checked by a relevant institution (e.g. Lebenshilfe, Netzwerk Leichte Sprache) and designed according to an appropriate set of rules. A further subdivision was carried out according to whether the blue/white logo from Inclusion Europe appears on the texts or not.

(2) "einfache Sprache" corpus (eS)

The eS corpus is mostly composed of texts that bear different labels and do not adhere to LS rules. This corpus is divided into three sub-corpora:

- Texts that are marked as simplified or edited for the benefit of a given target group. For instance, they are labelled as texts in "einfache" (simple) or "leicht verständliche Sprache" (easy to understand) or contain keywords such as "einfach gesagt" (said simply), "leicht" (easy) or "klar" (clear).
- Texts that are written in either "Leichte" or "leichte Sprache" but do not bear the features of the LS sub-corpus or any explicit label. In most cases, these texts are specially created for the benefit of different target groups (e.g. the hard of hearing community).
- Texts that are labelled as "citizen-oriented" ("bürgernah"), i.e. simplified versions of administrative texts.

(3) "Leicht Lesen" corpus (LL)

The LL corpus contains texts that were created or issued exclusively by institutions that authorize the use of the LL label, i.e. those affiliated to *atempo* or *capito*[32]. These texts are divided in two sub-corpora (Figure 1).

2.2 The LeiSA text type corpus (TT)

An additional sub-corpus consists of 255 analogue and digital documents that were collected through a nationwide online survey on the use of LS in professional settings for the benefit of people with (intellectual) disabilities (Bergelt/Goldbach/Seidel 2016). The participants were asked to send all possible simplified texts they were using in their institution, without imposing any restrictions on label, text type, text scope, author etc. The TT corpus helps to assess the use of simplified texts in occupational contexts. In this article, the TT corpus allows for an analysis of existing text functions and text types within this particular context of communication

[32] https://www.capito.eu/de/Ueber_uns/Unser_Leitbild/ (8 May 2018)

(see Section 3.3). In the TT corpus, the LS texts amount to 111 and the eS texts to 142. LL examples are not found in this sub-corpus.

3 Comparing "Leichte Sprache", "einfache Sprache" and "Leicht Lesen"

The present study is an explorative corpus-based analysis (see Bubenhofer/Scharloth 2010: 88ff.; Tognini-Bonelli 2001). The descriptive levels are partially based on analytical categories from Jekat et al. (2017) and Kuhlmann (2013). These categories primarily concern information on lexical variation (type-token ratio, Section 3.2.1) and a characterisation of the most frequent vocabulary (Section 3.2.2). The analyses are mainly based on the word lists extracted and edited from the sub-corpora. The evaluation of corpus, text and sentence size (Section 3.1) as well as that of text functions and text types (Section 3.3) is based on the full texts. Data evaluations were performed either manually or automatically with the software *WordSmith* and *AntConc*.

3.1 Corpus, text and sentence size

Recommendations for the optimal length of simplified texts are not explicitly addressed in the common LS rules (BMAS 2013; Netzwerk Leichte Sprache 2013; BITV 2.0 2011). Undoubtedly, specifications of this kind may be of little benefit since the length of a text must vary, depending on its subject matter, type and function, if the content is to be conveyed appropriately. Nevertheless, it is undeniable that text length may affect people's reading performance. Longer texts, for instance, can initially "put off" inexperienced readers and have a negative impact on their reading motivation. They can also make it difficult for readers with limited cognitive capacities to retain previously given information that is necessary to enable the process of comprehension (e.g. inferring, relating etc.). It can therefore be assumed that authors of simplified texts generally try to reduce text length in order to counteract potentially negative effects. Nonetheless, paraphrases, explanations and an adapted text layout often make simplified texts more extensive than is desirable.

In this section, the size of texts in the three sub-corpora are examined. Textual complexity is indicated by the number of words and sentences per sub-corpus and per text, as well as by the average number of words per sentence. By only taking into account the textual surface, these figures allow for limited conclusions on textual complexity. Nevertheless, sentence length is a highly endorsed criterion for assessing text comprehensibility (see Groeben/Christmann 1989: 169; Lasch 2013), even if recommendations on just how many words a sentence should contain in order to be fully comprehensible do vary (see Kercher 2013: 194). In addition, LS rules recommend that long sentences be avoided, but they do not stipulate any critical reference value (Netzwerk Leichte Sprache 2013: 17).[33]

[33] The boundary value in eS is specified at 15 words per sentence (klar&deutlich 2017). On the basis of his study on LS texts, Lasch (2013) indicates an average of about 8 words per sentence. However, he acknowledges that sentence length varies significantly among the texts.

Table 1 illustrates the size of the sub-corpora and their texts according to the parameters mentioned above:

	LS (n=404)	eS (n=300)	LL (n=88)
Words (tokens) per corpus	639,826	779,278	350,872
Sentences per corpus	68,939	75,411	29,453
Average no. of sentences per text	171 (s=300.5)	251(s=615.2)	335 (s=313.2)
Average no. of words per text	1,596 (s= 2691.3)	2,851 (s=7232.8)	4,051.50 (s=4083.3)
Average no. of words per sentence	9.36 (s=7.62)	11.34 (s=20.47)	12.11 (s=9.64)

Table 1: Size of sub-corpora and relevant texts

Overall, the eS corpus has greater sentence and word counts, despite including fewer texts than the LS corpus. The LL corpus shows a similar tendency. It covers about half as many words and sentences as the LS corpus; however, these are distributed over only 88 texts, i.e. just under a fifth of the LS corpus. The values of the average number of words per sentence are particularly surprising, especially for the LL texts (12.1), which exceed the average number of words in eS sentences (Table 1).

The texts of the *LS corpus* contain, on average, 1,596 words (s = 2691.3) and 171 sentences (s = 300.5). However, variation is very high. Only a quarter of all LS texts exceed 2,000 words and just under half of the texts have up to 500 words. The sub-corpus contains only one text with over 30,000 words. This is a 221-page brochure that provides information on areas such as parenting, care, age, sexuality etc. The number of sentences per text shows a similar tendency: three quarters of all texts consist of less than 200 sentences. Only a few exceptions exceed 1,000 sentences. Most of these examples are informative documents on working in sheltered workshops for people with disabilities.

The texts of the *eS corpus* contain an average of 2,851 words (s = 7232.8) and 251 sentences (s = 615.2). In addition, the corpus consists of 24 texts with a word count between 10,000 and 50,000. The smaller text number of the eS corpus compared to the LS corpus therefore points to a clear difference in text scope between the two corpora. Texts with less than 2,000 words account for about five sixths of all eS texts and for three quarters of the LS corpus. Just under a sixth of all eS texts contain more than 200 sentences. In this sixth, however, the amount varies between 200 and 4,093 sentences. The vast majority of the larger texts are made up of teaching materials and information brochures on various general education topics (history, drugs, occupational safety, photography).

The texts of the *LL corpus* consist of an average of 4,051.50 words (s = 4083.3) and 335 sentences (s = 313.2). Only seven texts exceed a word count of 10,000 per text. Just under a third of the LL texts consist of fewer than 2,000 words. In both the eS and LS corpora, on the other hand, texts of this smaller size make up the vast majority. Finally, the LL corpus contains many texts with a high number of sentences: almost two thirds include more than 200 sentences. These texts are thus considerably larger than the LS and eS texts.

3.2 Vocabulary

The analysis of the vocabulary in the sub-corpora considers different focal areas: on the one hand, lexical variation (type-token ratio), and, on the other hand, the properties of the most frequent words (part of speech, syntactic function, inflection).[34] For further results based on this corpus (e.g. keywords), see Lange/Bock (2016).

3.2.1 Lexical variation

Lexical variation is calculated as the ratio of type (number of unique word forms) to token (total number of words). In this case, the number of types occurring in each sub-corpus is divided by the number of occurring tokens (total number of words). The more the quotient approaches 1, the richer in vocabulary (or less redundant) the texts are.[35]

	LS (n=404)	eS(n=300)	LL (n=88)
Total number of words (tokens)	639,826	779,278	350,872
Word forms (types)	17,725	38,470	16,843
Type-token ratio	0.028	0.049	0.048

Table 2: Lexical variation in the sub-corpora texts

Table 2 indicates the quotients of each sub-corpus. The eS corpus shows the highest relative ratio, although it is still far from the reference value of 1. The LL corpus bears a very similar ratio, which suggests greater lexical variation since it contains significantly fewer texts than the eS corpus. By contrast, the LS corpus reveals lower vocabulary variation. This is also clearly illustrated by the token number, which differs greatly despite an almost equal number of types in the LL and LS corpora. This result is not surprising insofar as LS rules explicitly advocate a redundant linguistic style (see Netzwerk Leichte Sprache 2013: 5; BITV 2011).

3.2.2 Frequent vocabulary

In this section, the 100 most frequent words of each sub-corpus are analysed and described in more detail. These words account for more than half of all word forms used in the LS and LL corpora (LS: 52 %; LL: 51 %) and for 45 % of all word forms used in the eS corpus. Despite the differences in corpus size, the evaluation of the three sub-corpora therefore still delivers comparable results.

Cross-corpus high frequency core vocabulary

A comparison of the lemmatized word lists of the three sub-corpora serves initially to determine a cross-corpus high frequency core vocabulary, i.e. a collective inventory of the most frequent words. The 100 most frequent words include 80 lemmas in the LS and eS corpora and 81 in the LL corpus. A comparison of the three lists of lemmas indicates that 50 of them, i.e. just over half of all occurring lemmas, are equally present in each corpus.[36] This core vocabulary

[34] As regards the significance of the results, it should be noted that the statistical evaluation of word occurrences is purely form-based. Homonymic word forms or incorrect data in the word lists (e.g. erroneous word separations through spaces or special characters) can sometimes affect the results.

[35] However, this value is subject to large fluctuations depending on text length. It can be assumed that a different corpus composition (in terms of text types, text size and number) would have produced a different result. The values are therefore only partially comparable.

[36] This core inventory includes the following lemmas (see appendix for translation): *aber, alle, an, auch, auf, aus, bei, Beispiel, dann, das, dass, der, die, dieser, -e, -es, ein/einer, -e, -es, er, es, für, geben, Geld, gut, haben, ich, ihrer, -e, -es, in, können, machen, mehr, Mensch, mit, müssen, nicht, noch, nur, oder, sein, sich, sie, so, über, und, viel, von, was, wenn, werden, wie, wir, zu.*

is composed predominantly of function words[37], i.e. the parts of speech that LS rules consider to be problematic in terms of text comprehensibility (e.g. pronouns, connectors). The LS list presents the largest proportion of words that deviate from this core inventory. Words such as *wichtig* (important), *sollen* (should), *selbst* (self), *Seite* (page), *seiner*, *-e, -es* (his), *man* (impersonal form "one"), *heißen* (to mean, being called), *damit* (so that, therefore), *brauchen* (to need), *bekommen* (to get), *Behinderung* (disability), *arbeiten* (to work) and *Arbeit* (work) do not appear among the most common lemmas of the other two sub-corpora. This shows that the LS texts have a specific thematic orientation (see also Lange/Bock 2016 on keywords). All lexemes identified here are components of the German basic vocabulary (see Jones/Tschirner 2006).

Content and function words

In general, function words are considered to be the most frequent units in language usage. However, since LS rules consider the majority of function words difficult to understand (e.g. connectors, relative pronouns), it can be assumed that they are under-represented in the LS corpus in contrast to content words.[38] The frequency analysis of content and function words gives the following distribution in the three sub-corpora:

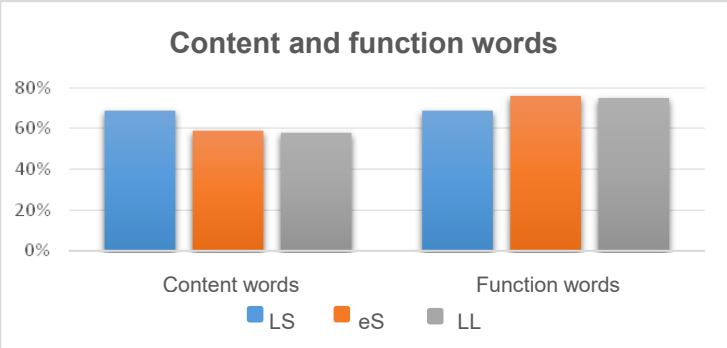

Figure 2: Distribution of content and function words among the 100 most frequent words

Figure 2 confirms that the 100 most frequent words contain a considerable amount of content words. The number of function words barely exceeds that of content words. In the LS corpus, the percentages are equal.[39] This suggests that the LS texts were produced in strong compliance with the relevant rules as they largely tend to refrain from using connectors and certain cohesive ties. Nevertheless, the comparatively small number of function words is surprising since research on text comprehensibility does attribute a positive effect to them. Connectors and cohesive ties can establish coherence (where sense relations would otherwise have to be inferred) and are generally short, well known and common (see Kercher 2013: 73; Bamberger/Vanecek 1984: 37).

[37] Function words are articles, prepositions, conjunctions, auxiliaries and modal verbs (Glück/Rödel 2016). In most grammars, pronouns are also considered function words. Busse (1997) is critical of this differentiation.

[38] Kuhlmann (2013: 74) confirms this assumption in her corpus analysis. Function words are much more frequent in the reference corpus than in the LS corpus (reference corpus: 17 % content vs. 83 % function words; LS corpus: 54 % content vs. 46 % function words).

[39] However, it must be acknowledged that this categorization is subject to strong fluctuations depending on the presence of homonyms. For instance, units such as auf were annotated both as a content (adverb) and a function word (preposition). No unequivocal interpretations can be made on the sole basis of the word lists.

Morphological and semantic properties

A detailed analysis of the 100 most frequent words with regard to their parts of speech leads to the following results:

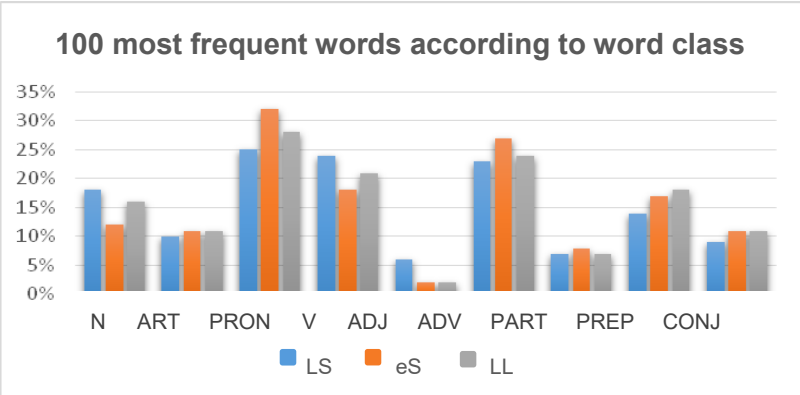

Figure 3: The most frequent 100 words categorized according to word class

Figure 3 shows that pronouns, verbs and adverbs are the most frequently assigned word classes. Conjunctions and particles are rarely used, and adjectives hardly ever occur. By contrast, nouns and prepositions are at average levels. How parts of speech are used partially determines the differences among the three varieties. By way of example, the use of three parts of speech (i.e. nouns, verbs, adverbs) will be considered in more detail below.

The lower number of *nouns* compared to other parts of speech is probably associated with a tendency towards a verbal linguistic style. Even though LS rules explicitly encourage the use of verbs (Netzwerk Leichte Sprache 2013: 8), there are more nouns in the LS samples than in the other two varieties.[40] Table 3 below displays the most common nouns for each sub-corpus (see translation in appendix).

LS	eS	LL
Arbeit/Arbeits-/Arbeiten, Behinderung/Behinderungen, Beispiel, Deutschland, Frauen, Geld, Hilfe, Kinder, Leben, Lernen, Mensch/Menschen, Rechte, Seite, Sprache	*Aufgaben, Beispiel, Bild, Bitte, Geld, Internet, Leben, Lizenz, Menschen, Musik, Quelle, Schülerheft*	*Antrag, Arbeit/Arbeiten, Beeinträchtigungen, Behinderung/ Behinderungen, Beispiel, Betreuer, Geld, Gesetz, Informationen, Kapitel, Menschen, Person, Seite, Unterstützung*

Table 3: Nouns among the 100 most frequents words

The lexemes in Table 3 do not exclusively consist of concrete nouns, even though the use of these is recommended in the Netzwerk Leichte Sprache's guidelines (2013: 4; see also BITV 2.0; Kaczmarzik 2017). It is evident that abstract nouns also occur very often, i.e. *Arbeit* (work), *Lernen* (learning), *Leben* (life), *Recht* (right), *Unterstützung* (support, help) etc.

It becomes obvious that context and thematic areas do partially characterize the texts in the three varieties. The LS texts are more concerned with sociopolitical topics and seem to focus on target groups experiencing discrimination or disadvantage (*people, disability, women*). The LL texts show a similar tendency. Many of the eS nouns, on the other hand, seem to be the result of, or at least strongly influenced by, certain text types, i.e. various teaching and learning materials. The occurrence of words such as *Bild* (picture) or *Aufgabe* (task) confirms this. It is also worth noting that the eS and LS texts frequently include loanwords such as *Lizenz* (licence), *Information, Kapitel* (chapter) or *Internet*. This is remarkable considering that

[40] Kaczmarzik (2017) came to the opposite conclusion: in the LS corpus that she analysed, nouns were the most frequent word class. However, the author considered only a relatively small corpus in exclusively one communication area (electoral manifestos).

technical and loanwords are assumed to be less comprehensible due to their specific and rare use. However, this does not apply to words such as "information" or "Internet".

Pronouns are also regarded as units that can hinder reading comprehension (see Inclusion Europe 2009: 15). Nonetheless, this is the word class that occurs the most in the 100 most frequent words in all three language varieties (for more details on the use of pronouns and their comprehensibility, see Lange, forthcoming d).

Verbs and *adverbs* have the largest share of the 100 most frequent words. Compared to the other two corpora, the eS corpus has the lowest frequency of verbs and the highest frequency of adverbs. In contrast to the low percentage of occurring nouns, this value seems to confirm that text simplifications comply with the recommended verbal linguistic style across all three varieties. The following overview summarizes the most frequent verb forms.

LS	eS	LL
arbeiten, bekommen, brauchen, darf, geben, gibt, haben, hat, heißt, ist, kann, können, leben, lernen, machen, muss, müssen, sein, sind, soll, sollen, werden, wird, wollen	*bitte, findest, gibt, haben, hast, hat, ist, kann, kannst, können, leben, machen, muss, sein, sind, war, werden, wird*	*arbeiten, bekommen, brauchen, gibt, haben, hat, heißt, ist, kann, können, machen, muss, müssen, sein, sind, soll, sollen, steht, weiß, werden, wird*

Table 4: Overview of the most frequent verbs

Table 4 suggests that verbs are mostly conjugated in the third person singular and in the first and third person plural. Verb forms in the second person singular occur only in the eS corpus, which, as mentioned above, includes to a large extent didactic material targeted at young pupils. In addition, the eS corpus is the only corpus that contains a verb form in the past tense: *war* (was). Furthermore, in contrast to the LS and LL corpora, the eS corpus does not contain any verbs that point to the clarification or explication of content (e.g. [*das*] *heißt*, i.e. that means) – and this despite the overwhelming presence of teaching materials.

The large proportion of modal, auxiliary and copular verbs is also particularly striking. These verbs are generally used in forming analytic predicates: *werden, wird, ist* (are, is, will be…) etc.

	LS	eS	LL
Modal verbs	33 % (8 %)	21 % (4 %)	29 % (6 %)
Auxiliary and copular verbs	42 % (10 %)	47 % (9 %)	43 % (9 %)

Table 5: Modal and auxiliary verbs among the most frequent verbs and the 100 most frequent words (in parenthesis)

These findings have already been discussed in Lange/Bock (2016). The authors emphasize that modal verbs and compound predicates do in fact make sentences more complex, which is precisely what the three language varieties aim to avoid. Kuhlmann (2013: 78) and Lasch (2013) come to the same conclusion.[41] Five out of the total six modal verbs occur in the LS corpus, as well as *brauchen* (must), which can also act as a modal verb. The eS corpus contains only two modal verbs, whereas the LL corpus contains three modal verbs, as well as *brauchen* (must). In this context, Kuhlmann (2013) identifies a connection to the prime function of simplified texts: modal verbs supposedly add a descriptive and evaluative style to LS texts. Modal verbs therefore serve to "modalize" the meaning of a predicate, i.e. to make it more flexible in its validity (Weinrich 1993: 297). However, such semantic flexibilization is not always appropriate, especially considering the specific thematic orientation of LS texts. As a result,

[41] When comparing her LS corpus with the DeReWo reference corpus, Kuhlmann (2013) also proves that the use of modal verbs is, above all, a distinctive feature of LS texts. In the reference corpus, full verbs are considerably more common (36 occurrences) than modal verbs (2 occurrences).

political developments or social topics are very often "emotionalized" or may not always be treated impartially (see also Kuhlmann 2013: 78). In addition, the frequent use of *wird* and *werden* (is, are, will be…) suggests that passive constructions or future tense forms do indeed occur (see BMAS 2013: 29; for an empirical examination, see Bock 2017).

The remaining verb forms are mostly transitive verbs characterized by a flexible semantics. By way of example, Lasch (2013) considers the verb *machen* (do, make). This verb's great advantage is that it expresses an action without specifying it semantically. Therefore, *machen* can be integrated in a wide variety of statements. The same can be said – with some restrictions – for *geben* (give), *lernen* (learn), *wissen* (know) and *bekommen* (get), which are among the most frequent verbs in the LeiSA core corpus.[42]

The large number of verbs may explain the substantial presence of *adverbs* in this corpus. For instance, some of the most frequent verbs are ideally suited for combination with an adverbial basis, e.g. *machen*: *aufmachen* (open), *mitmachen* (participate). Furthermore, the negative adverb *nicht* (not) occurs frequently in all three corpora, even though it is considered to be a comprehension-hindering unit in LS (Netzwerk Leichte Sprache 2013: 10; modified by Bredel/Maaß 2016: 460ff.; empirical examination in Bock 2017). It is also worth emphasising that prepositional adverbs like *dafür* (for that, instead) and *damit* (thereby) are among the 100 most frequent words. Adverbs such as *dann* (then), *hier* (here), *so* (like that) or *wie* (as, how) do occasionally represent essential cohesive ties in simplified texts. Their high frequency may be due to the fact that adverbs are generally preferred to conjunctions (see BMAS 2013: 46).[43] Additionally, the comparison with *wie* (as, how) emerges as the most frequently used comparison structure. This confirms the descriptive style of simplified texts insofar as comparisons can foster comprehension. Comparison structures may also be used in conjunction with *zu* (too), *noch* (even, still), *nur* (only), *mehr* (more) etc., and may serve to comment, quantify or classify information. Finally, the adverb *zusammen* (together), which is not among the 100 most frequent words of the LL and eS corpora, appears to be particularly significant in the LS texts. Semantically, it emphasizes the strong thematic orientation of the LS corpus towards inclusion.

3.3 Text functions and text types

The analysis of text functions and text types is based on the Text-Types corpus (TT) (see Section 2.2). Central to the analysis is the question of which text typology is used most widely in professional settings. At the core of the evaluation is a qualitative analysis of all relevant simplified texts. It should be acknowledged that the texts in question could not always be unequivocally categorized by function and by type. Text modifications altered the superficial and deep structure of these texts so strongly that text typologies were sometimes no longer identifiable.

3.3.1 Text functions

The corpus texts are categorized according to the following language functions, as defined by Brinker/Cölfen/Pappert (2014): informative (*Information*), appellative (*Appell*), commissive (*Obligation*), expressive (*Kontakt*) and declarative (*Deklaration*). However, many of the text samples could not always be clearly assigned to one text function and were therefore grouped in mixed categories (see Figure 4). For instance, the category "informative and appellative" contains texts with a predominantly informative function; however, parts of those texts may also have an appellative function (e.g. registration forms are often included with information on training courses). Mixed categories in connection with the commissive function were assigned

[42] In addition to the verbs listed here, Lasch (2013) also identifies *nennen* (name, state), *stehen* (stand), *finden* (find) and *bedeuten* (mean) among the 15 most common words in this class.

[43] In the light of Kaczmarzik's findings (2017), however, it can be assumed that the number of adverbs is relative in relation to the whole corpus.

to agreements, inclusion programmes and instructions (with an appellative function), or to mission statements and contracts (with informative and declarative functions).

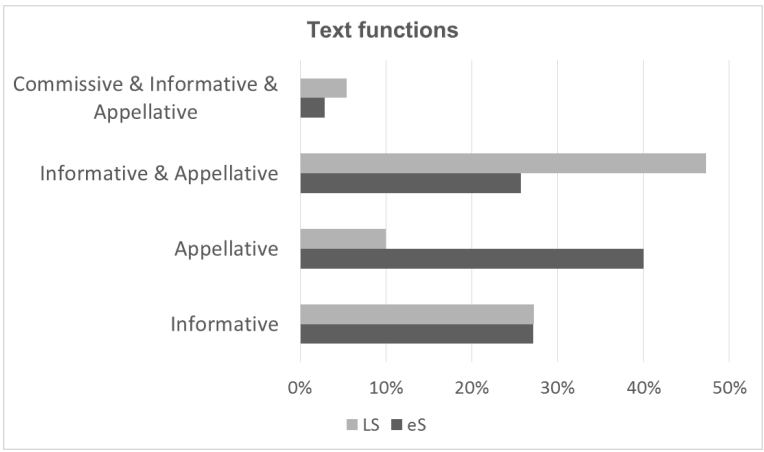

Figure 4: Distribution of texts according to text function

Figure 4 shows the distribution of text functions across the materials being analysed. Over half of the LS texts are informative and 47 % of them aim to inform and appeal at the same time. The appellative and informative functions predominate in the eS texts as well. All other functions are only represented to a comparatively small extent and are therefore omitted in the graph.

These findings are not surprising insofar as informative texts, such as job descriptions, instruction sheets, guidelines etc., are to be expected within the domain of professional integration. In contexts of communication (see Lange/Bock 2016) in which mainly simplified texts are used (most notably politics, social affairs and law), a larger number of commissive and declarative texts would have to be assumed.

3.3.2 Text types

The classification of the corpus samples according to text type is based on Brinker/Cölfen/Pappert (2014: 136ff.) and Fandrych/Thurmair (2011).[44] As already mentioned above, simplified texts differ to a large extent from the prototypical features of any given text type (typographically or content-wise). For this reason, a clearly differentiated classification of the corpus texts often proved difficult, particularly as far as the "informative texts" were concerned. These were therefore classified according to their function and medium (e.g. information => brochure, sheet, flyer). However, a classification exclusively according to media type would be too general since posters, flyers etc. can also fulfil several functions (also the appellative and expressive).

Overall, over 40 different text types were identified in the TT corpus. Table 6 below illustrates the most representative of the relevant text functions.

[44] In cases of very specific types of text, a number of generaliations were made (e.g. travel report => report).

Text functions	Text types
Informative	Policy paper, call for tenders, information sheet, information letter, information booklet, time/work plan, menu, contract explanation, minutes, journal articles, preface, list, answer sheet
Appellative	questionnaire, regulation, form, time/work plan, information board, worksheet, instructions, recipe
Informative and Appellative	handbook, information leaflet (usually, invitation to participate), information brochure, teaching/learning plan, worksheet, form, questionnaire, regulation, information text, call for tenders
Commissive	contract amendment
Commissive and other functions	instructions, mission statement, contract
Declarative	consent form
Expressive	invitation (letter), acknowledgement

Table 6: Typical text types according to text function

The eS texts in the TT corpus are mainly instructions (23 %), information texts (20 %), questionnaires (10 %), worksheets (9 %) and time/work plans (6 %). Almost half of the LS texts, on the other hand, are information texts (45 %) and 9 % are regulations. All other text types are represented by less than five instances each and are therefore not mentioned here. This suggests that in professional settings, the materials provided in LS are mostly informative or regulative texts, whereas eS texts also include support for the execution of workplace activities (e.g. instructions) (see also Bergelt/Goldbach/Seidel 2016). This tendency might be regarded as problematic if we are to determine how LS can promote the inclusion of people with intellectual disabilities in the workplace. If the availability of materials aimed at supporting the target groups' autonomous work (i.e. working independently from superiors or colleagues) is low, their effectiveness is also likely to be low.

4 Summary and overview

The aim of this article was to describe "Leichte Sprache" in contrast to similar approaches of text simplification, i.e. "einfache Sprache" and "Leicht Lesen", and thus to contribute to a well-founded conceptualisation of the three language varieties as part of a comprehensibility continuum. In detail, the analyses addressed, on the one hand, the distinctive attributes of LS in terms of linguistic features, text types and text functions, and, on the other hand, the extent to which the common LS writing recommendations are adhered to in practice. Comparing the results to the two other varieties in focus, similarities and differences in language usage were discussed in order to verify or deny the hypothesis that they can be classified as differing in certain aspects, such as language complexity, addressed recipients, subjects, text types and functions. The extensive LeiSA core corpus, which integrates text samples from all three varieties, enabled an explorative and representative analysis of the individual features of the language varieties in question on the basis of selected criteria. The results are summarized below according to descriptive categories.

(1) Word and sentence size

LS texts are the shortest and least complex in terms of text and sentence size. The LL texts contain the highest number of sentences per text as well as of words per sentence and, on average, the longest sentences. In all three varieties, the word count per sentence remains below the critical comprehensibility value of 15.

(2) Lexical variation

The type-token ratio indicates that the vocabulary varies the least in the LS texts, which are therefore more redundant than the eS and LL texts. Overall, the LL corpus shows the greatest lexical variation.

(3) Frequent vocabulary

The comparative description of the LeiSA-corpus vocabulary was based on the 100 most frequent words of each sub-corpus. Of the lemmas that occurred the most in each sub-corpus, 60 % form an overarching high-frequency core vocabulary. This is mainly composed of function words, some of which are often considered to be problematic in LS (pronouns; connectors; auxiliary, copular and modal verbs). Although covering a large share of German basic vocabulary, function words are less frequently represented among the 100 most common words. In the LS list, the percentage of function and content words is equal.

The analysis of the morphological and semantic properties of the 100 most frequent words led to the following findings:

The proportion of verbs in all three varieties is very high. This confirms that a verbal style is generally preferred, and particularly so in the eS texts. Verbs are mainly modal, copular or auxiliary verbs, which – contrary to recommendations – do increase sentence complexity. In particular, the high number of modal verbs in the LS corpus lends the texts a descriptive and sometimes evaluative style. Full verbs are mostly transitive and characterized by a flexible semantics: *machen* (do, make), *bekommen* (get). This enables frequent word formations with particle or adverb prefixes: *zumachen* (close), *mitmachen* (participate) and is thus further indicative of an increase in textual complexity.

Contrary to recommendations, the LL and LS corpora also contain abstract nouns. The LL and eS texts even include loanwords.

Adverbs are highly common. They often contribute to generate textual coherence, but they can also be part of complex verbs. The negation adverb *nicht* (not) is among the most frequent words, despite the ban on negative constructions in LS.

There are comparatively few adjectives, particles and connectors.

(4) Text functions and text types

In professional settings, texts have a predominantly informative and appellative function. Typical text types of the eS corpus are instructions, information sheets and worksheets. The LS corpus contains more organisational texts, e.g. information sheets and regulations. It was pointed out that this tendency could hinder the ultimate goal of LS, which is to foster inclusion in the workplace and promote self-determination.

Overall, this study shows that LS usage is mostly oriented towards the implementation of LS rules. In some cases, however, linguistic properties of LS texts contradict some of the recommendations. For instance, the functional focus (descriptive, informative, evaluative style) and the thematic content of the texts (politics, inclusion) were suggested as possible grounds for the high occurrence of modal verbs. This heterogeneity can certainly be explained by the vagueness of some writing guidelines. However, it also suggests that LS is more flexible than expected and that practice may respond to the requirements of different communication areas, text types and target groups.

The analysis of the thematic orientation of the LS texts leads to the conclusion that LS addresses a very specific readership (i.e. people with intellectual disabilities). Target groups are broader in the other two language varieties. Taking into account its limited possibilities of adaptation (due to established "standardisation mechanisms"), it is questionable to what extent LS can increase text comprehensibility and, consequently, inclusion for the benefit of different target groups. The LL texts, which in this analysis prove to be the most complex and varied,

reflect an internal differentiation of target groups, communication areas and content (see Fröhlich 2017) and thus produce a more flexible and customized language.

These findings confirm the initial assumption that the three approaches of text simplification can be conceptualized and described as part of a comprehensibility continuum. The descriptive categories considered here partly substantiate the hypothesis that the varieties differ regarding their complexity in form and content. However, it is worth remembering that the corpus analyses have their limits regarding the generalizability of the results since the corpus' scope and composition strongly influence the results, as can be seen in some of the findings. Nevertheless, respective corpus research could certainly be expanded to include more detailed qualitative analyses. In addition, descriptions of further categories, most importantly at the syntactic level, are still an urgent desideratum. In combination with empirical findings on target groups' reading abilities and experiences, it will now be necessary to examine to what extent individual language varieties or writing guidelines are better suited for certain target groups than others, or to what extent a grading or differentiation of different approaches is at all appropriate. In the LeiSA project, several rules have been empirically examined and further considerations given to the appropriateness of content and form in relation to the targeted readership, as well as to the possibilities of differentiating linguistic structures.

Sources

Bamberger, Richard; Erich Vanecek (1984): *Lesen - Verstehen - Lernen - Schreiben. Die Schwierigkeitsstufen von Texten in deutscher Sprache*. Wien: Jugend u. Volk Sauerlaender

Bergelt, Daniel; Anne Goldbach; Anja Seidel (2016): "Leichte Sprache im Arbeitsleben. Analyse der derzeitigen Nutzung von Texten in Leichter Sprache im beruflichen Kontext von Menschen mit Lernschwierigkeiten." *Teilhabe 55* [3], 106-113

BITV 2.0 (2011): Verordnung zur Schaffung barrierefreier Informationstechnik nach dem Behindertengleichstellungsgesetz (Barrierefreie-Informationstechnik-Verordnung – BITV 2.0) – https://www.gesetze-im-idnternet.de/bitv_2_0/BJNR184300011.html (26 December 2017)

BMAS (2013): Leichte Sprache. Ein Ratgeber. Abdruck der "Regeln für Leichte Sprache" des Netzwerks Leichte Sprache e. V. http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Publikationen/a752-ratgeber-leichte-sprache.pdf?__blob=publicationFile (26 December 2017)

Bock, Bettina M. (2014): "'Leichte Sprache'. Abgrenzung, Beschreibung und Problemstellungen aus linguistischer Sicht." Susanne J. Jekat; Heike Elisabeth Jüngst; Klaus Schubert; Claudia Villiger (eds): *Sprache barrierefrei gestalten. Perspektiven aus der Angewandten Linguistik*. Berlin, 17-52

Bock, Bettina M. (2015): "Barrierefreie Kommunikation als Voraussetzung und Mittel für die Partizipation benachteiligter Gruppen. Ein (polito-)linguistischer Blick auf Probleme und Potenziale von 'Leichter' und 'einfacher Sprache'." *Linguistik Online 73* [4] (Sprache und Demokratie) – https://bop.unibe.ch/linguistik-online/article/view/2196 (26 December 2017)

Bock, Bettina M. (2016): "Leichte Sprache". Cordula Löffler; Jens Korfkamp (eds): *Handbuch zur Grundbildung Erwachsener*. Münster

Bock, Bettina M. (2017): "Das Passiv- und Negationsverbot 'Leichter Sprache' auf dem Prüfstand. Empirische Ergebnisse aus Verstehenstest und Korpusuntersuchung." *Sprachreport 33* [1], 20–28

Bock, Bettina M.; Ulla Fix; Daisy Lange (eds) (2017): „Leichte Sprache" im Spiegel theoretischer und angewandter Forschung. Berlin: Frank & Timme Verlag für wissenschaftliche Literatur (Kommunikation - Partizipation - Inklusion, Band 1)

Bock, Bettina; Daisy Lange (2015): "Was ist eigentlich 'Leichte Sprache'? Der Blick der Sprachwissenschaft." Klaus Candussi; Walburga Fröhlich (eds): *Leicht Lesen: der Schlüssel zur Welt*. Wien: boehlau, 63-80

Bredel, Ursula; Christiane Maaß (2016): *Leichte Sprache. Theoretische Grundlagen. Orientierung für die Praxis*. Berlin: Bibliographisches Institut GmbH; Duden

Bredel, Ursula; Christiane Maaß (2017): "Wortverstehen durch Wortgliederung – Bindestrich und Mediopunkt in Leichter Sprache." Bettina M. Bock; Ulla Fix; Daisy Lange (eds): *"Leichte Sprache" im Spiegel theoretischer und angewandter Forschung*. Berlin: Frank & Timme Verlag für wissenschaftliche Literatur (Kommunikation - Partizipation - Inklusion, Band 1), 211–228

Brinker, Klaus; Hermann Cölfen; Steffen Pappert (2014): *Linguistische Textanalyse*. Berlin: Erich Schmidt Verlag

Bubenhofer, Noah; Joachim Scharloth (2010): "Kontext korpuslinguistisch: Die induktive Berechnung von Sprachgebrauchsmustern in großen Textkorpora." Peter Klotz; Paul R. Portmann-Tselikas; Georg Weidacher (eds): *Kontexte und Texte. Studien zu soziokulturellen Konstellationen literalen Handelns*. Tübingen: Narr, 85-108

Busse, Dietrich (1997): "Wortarten und semantische Typen. Überlegungen zu den Grundlagen der lexikalisch-syntaktischen Wortarten-Klassifikation." Christa Dürscheid; Karlheinz Ramers; Monika Schwarz (eds): *Sprache im Fokus* (= Festschrift für Heinz Vater). Tübingen: Niemeyer, 219-240

Fandrych, Christian; Maria Thurmair (2011): *Textsorten im Deutschen. Linguistische Analysen aus sprachdidaktischer Sicht* (Stauffenburg Linguistik, Bd. 57). Tübingen: Stauffenburg

Fröhlich, Walburga (2017): "Von der 'Leichten Sprache' zur zielgruppengerechten Information: Leicht Lesen in Österreich." Bettina M. Bock; Ulla Fix; Daisy Lange (eds): *"Leichte Sprache" im Spiegel theoretischer und angewandter Forschung*. Berlin: Frank & Timme Verlag für wissenschaftliche Literatur (Kommunikation - Partizipation - Inklusion, Band 1), 415–430

Glück Helmut; Michael Rödel (eds) (2016): Metzler Lexikon Sprache. 5th updated and revised edition. Stuttgart: J.B. Metzler Verlag GmbH

Groeben, Norbert; Ursula Christmann (1989): "Textoptimierung unter Verständlichkeitsperspektive." Gerd Antos; Hans P. Krings (eds.): *Textproduktion. Ein interdisziplinärer Forschungsüberblick*. Tübingen: Niemeyer (Konzepte der Sprach- und Literaturwissenschaft, 48), 165–196

Inclusion Europe (2009): "Informationen für alle: Europäische Regeln, wie man Informationen leicht lesbar und verständlich macht." http://easy-to-read.eu/wp-content/uploads/2014/12/DE_Information_for_all.pdf (26 December 2017)

Jekat, Susanne J.; Esther Germann; Alexa Lintner; Corinne Soland (2017): "Wahlprogramme in Leichter Sprache – Eine korpuslinguistische Annäherung." Bettina M. Bock; Ulla Fix; Daisy Lange (eds): *"Leichte Sprache" im Spiegel theoretischer und angewandter Forschung*. Berlin: Frank & Timme Verlag für wissenschaftliche Literatur (Kommunikation - Partizipation - Inklusion, Band 1), 229–246

Jones, Randall; Erwin Tschirner (2006): Frequency dictionary of German: Core vocabulary for learners. London: Routledge

Kaczmarzik, Janine (2017): "Über die lexikalische Gestaltung von Wahlprogrammen in Leichter Sprache – Ein Forschungsbericht zur Masterarbeit." Bettina M. Bock; Ulla Fix; Daisy Lange (eds): *"Leichte Sprache" im Spiegel theoretischer und angewandter Forschung*. Berlin: Frank & Timme Verlag für wissenschaftliche Literatur (Kommunikation - Partizipation - Inklusion, Band 1), 247–252

Kellermann, Gudrun (2014): "Leichte und Einfache Sprache. Versuch einer Definition." *Aus Politik und Zeitgeschichte 64* (9-11), 7-10 – http://www.bpb.de/apuz/179341/leichte-und-einfache-sprache-versuch-einer-definition (26 December 2017)

Kercher, Jan (2013): *Verstehen und Verständlichkeit von Politikersprache. Verbale Bedeutungsvermittlung zwischen Politikern und Bürgern*. Wiesbaden: Springer Fachmedien Wiesbaden

klar & deutlich (2017): "Leichte Sprache - einfache Sprache." http://www.klarunddeutlich.de/cms/website.php?id=/de/index/ed/leichte_sprache-einfache_sprache.htm%20&sid=3d4ae687eec57d01e67de250a3251586 (26 December 2017)

Kuhlmann, Julia (2013): Ein sprachwissenschaftlicher Blick auf das Konzept der "Leichten Sprache". http://www.alpha-archiv.de/fileadmin/PDFs/Qualifizierungsarbeiten/Masterarbeit_Kuhlmann_Copy.pdf (26 December 2017)

Lange, Daisy (forthcoming a): Der Genitiv in der "Leichten Sprache". – das für und Wider aus theoretischer und empirischer Sicht.

Lange, Daisy (forthcoming b): Ist der Verzicht auf Konjunktionen und Subjunktionen in der "Leichten Sprache" verständnisfördernd – theoretische Betrachtung und empirische Ergebnisse.

Lange, Daisy (forthcoming c): Worttrennverfahren in der "Leichten Sprache" – eine explorative Studie zur Angemessenheit und Wirksamkeit von Wortsegmentierungsverfahren für ein erleichtertes Wort- und Satzverständnis.

Lange, Daisy (forthcoming d): Pronomen in der "Leichten Sprache" – ihr Gebrauch und ihre Verständlichkeit für die Zielgruppen.

Lange, Daisy; Bettina Bock (2016): "Was heißt 'Leichte' und 'einfache' Sprache? Empirische Untersuchungen zu Begriffssemantik und tatsächlicher Gebrauchspraxis." Nathalie Mälzer (ed.): *Barrierefreie Kommunikation – Perspektiven aus Theorie und Praxis*. Berlin: Frank & Timme, 117–134

Lasch, Alexander (2013): "Leichte Sprache – 10 Gestaltungshinweise." *Sprachpunkt* – https://alexanderlasch.wordpress.com/2013/02/03/leichte-sprache-10-gestaltungshinweise/ (26 December 2017)

Maaß, Christiane (2015): *Leichte Sprache. Das Regelbuch*. Berlin u.a.: LIT Verlag.

Netzwerk Leichte Sprache (Eds.) (2013): Die Regeln für Leichte Sprache. http://www.leichte-sprache.de/dokumente/upload/21dba_regeln_fuer_leichte_sprache.pdf (26 December 2017)

Appendix – translations

Footnote 11

aber (but), *alle* (all, everyone), *an* (preposition or verb particle, no precise translation, mostly: on), *auch* (also), *auf* (preposition or verb particle, no precise translation, mostly: on), *aus* (preposition or verb particle, no precise translation, mostly: off, from), *bei* (preposition or verb particle, no precise translation, mostly: at), *Beispiel* (example), *dann* (then), *das* (the (neutral), which), *dass* ((so) that), *der* (the (masculine), which), *die* (the (feminine), which), *dieser*, *-e, -es* (this, these), *ein/einer, -e, -es* (one; a), *er* (he), *es* (it), *für* (preposition or verb particle, no precise translation, mostly: for), *geben* (to give), *Geld* (money), *gut* (good), *haben* (to have), *ich* (I), *i/Ihrer, -e, -es* (her(s); their(s)), *in* (preposition or verb particle, no precise translation, mostly: in, at), *können* (can, being able), *machen* (to do), *mehr* (more), *Mensch* (people, human being), *mit* (preposition or verb particle, no precise translation, mostly: with), *müssen* (have to, must), *nicht* (not), *noch* (still), *nur* (only), *oder* (or), *sein* (to be), *sich* (him-, her-, itself, theirselves), *sie* (they), *so* (so), *über* (preposition or verb particle, no precise translation, mostly: over, above), *und* (and), *viel* (much), *von* (preposition or verb particle, no precise translation, mostly: from), *was* (what, which), *wenn* (when, if), *werden* (become), *wie* (how, as), *wir* (we), *zu* (preposition or verb particle, no precise translation, mostly: to).

Table 3: Nouns among the 100 most frequents words

LS	eS	LL
Arbeit/Arbeits-/Arbeiten (work…), *Behinderung/Behinderungen* (disability/ies), *Beispiel* (example), *Deutschland* (Germany), *Frauen* (women), *Geld* (money), *Hilfe* (help), *Kinder* (children), *Leben* (life), *Lernen* (learning), *Mensch/Menschen* (people), *Rechte* (rights), *Seite* (page), *Sprache* (language)	*Aufgaben* (tasks), *Beispiel* (example), *Bild* (picture), *Bitte* (request), *Geld* (money), *Internet, Leben* (life), *Lizenz* (licence), *Menschen* (people), *Musik* (music), *Quelle* (source), *Schülerheft* (booklet for learners)	*Antrag* (application), *Arbeit/Arbeiten* (work), *Beeinträchtigungen* (impairment), *Behinderung/ Behinderungen* (disability/ies), *Beispiel* (example), *Betreuer* (caregiver), *Geld* (money), *Gesetz* (law), *Informationen* (information), *Kapitel (chapter), Menschen* (people), *Person* (person), *Seite* (page), *Unterstützung* (support, help)

Table 4: Overview of the most frequent verbs

LS	eS	LL
Arbeiten (work), *bekommen* (get), *brauchen* (need), *darf* (he/she/it can/may/is allowed), *geben* (give), *gibt* (gives, "there is"), *haben* (have), *hat* (has),	*bitte* (I ask for), *findest* (you find), *gibt* (gives, "there is"), *haben* (have), *hast* (you have), *hat* (has), *ist* (is), *kann*	*arbeiten* (work), *bekommen* (get), *brauchen* (need), *gibt* (gives, „there is"), *haben* (have), *hat* (has), *heißt* (is called/ means), *ist* (is), *kann*

heißt (is called, means), *ist* (is), *kann* (he/she/it can), *können* (we/you/they can or inf.), *leben* (live), *lernen* (learn), *machen* (make/do), *muss* (has to/must), *müssen* (have to/must), *sein* (be), *sind* (we/they are), *soll* (he/she/it should/is supposed to), *sollen* (we/you/they should/are supposed to), *werden* (we/they become or inf.), *wird* (becomes), *wollen* (we/they want to or inf.)	(he/she/it can), *kannst* (you can), *können* (we/you/they can, or inf.), *leben* (live), *machen* (make/do), *muss* (must/has to), *sein* (be), *sind* (we/they are), *war* (was), *werden* (become), *wird* (becomes)	(he/she/it can), *können* (we/you/they can or inf.), *machen* (make/do), *muss* (he/she/it must, has to), *müssen* (we/you/they must/have to), *sein* (be), *sind* (we/they are), *soll* (he/she/it should/is supposed to), *sollen* (we/you/they should/are supposed to), *steht* (stands, is written), *weiß* (I know, he/she/it knows), *werden* (we/they become or inf.), *wird* (he/she/it becomes)

Author

Daisy Lange is a scientific assistant in the Department of German Studies and the Department of Special Needs Education at the University of Leipzig. She is one of the coordinators of the LeiSA linguistic sub-project. Her research interests include "Leichte Sprache" as well as the linguistics and didactics of German as a foreign language. E-mail: daisy.lange@uni-leipzig.de

DOI 10.21256/zhaw-3000

Marcin Michoń & Agnieszka Stawikowska-Marcinkowska

Barriers in Polish TV – On Audiovisual Translation and its Potential for Development

Abstract

Voice-over is currently the predominant translation method in Polish TV. It is so deeply rooted in viewers' awareness that any attempts to change it due to EU directives or simply in response to global trends are received with scepticism and reluctance. In public TV, as well as in the majority of Polish privately-owned TV networks, voice-over is the only universal translation method. This is largely due to financial issues, but the favourable attitude of the viewers to this method is not necessarily conducive to change either. The rising critical debate surrounding voice-over more and more frequently looks at the problem of addressing non-linguistic barriers, e.g. through audio description, subtitles or sign language (cf. Bogucki 2010, 2016; Szarkowska/Laskowska 2015). This paper offers an overview of the landscape of audiovisual translation in Poland and discusses the provision of access to information and entertainment in Polish TV for viewers with disabilities. It also addresses a number of aspects of inter- and intralingual translation in specialist communication. Stock exchange reports, economy-related information as well as specialist elements in TV films and series have been used as examples (cf. Michoń 2016, Stawikowska-Marcinkowska 2016). Digitization has the potential to enable Polish TV viewers to freely select a particular form of translation of a given programme. However, both public and privately-owned TV providers have a lot of catching up to do in this respect. The attractiveness of their service cannot depend only on the range of the programmes available, but also on the provision of access for viewers with certain disabilities (Szarkowska/Laskowska 2015). This paper attempts to identify the inter- and intralingual barriers in Polish TV as well as to offer ways to overcome them from the perspective of viewers with visual and hearing impairments.

1 A brief history of audio description

Audio description is as old as sighted people's practice of telling blind people about the visual layer of the world around them. It is widely recognized that the history of audio description as a formalized technique to increase the accessibility of culture for the blind and visually impaired began at the Arena Stage Theatre in Washington (District of Columbia) in 1981, thanks to the efforts of Margareta and Cody Pfanstiehl (ITC Guidance 2000: 4). In the same year, they founded the Audio Description Service, which promoted audio description in theatres in the United States. At the end of the 1980s, over fifty cultural institutions had performances with audio description in their repertoire.

In 2000, the American Federal Communications Commission (FCC) introduced regulations requiring leading TV broadcasters to provide fifty hours of audio description within three months of broadcasting on the twenty-five largest TV markets in the United States, starting from April 2002. Two years later, this provision was repealed by the decision of the court of appeal, which stated that by introducing the above provisions, the FCC exceeded its powers. After several unsuccessful attempts to expand the council's powers, President Barack Obama signed the Act in 2010, according to which specific broadcasters were required to provide four hours of audio-described programmes per week on the twenty-five largest TV markets in the US. The Act ensures a gradual increase in the number of programmes with audio description and TV channels offering this service, so that audio description can ultimately be available for all programmes broadcast nationwide (cf. Chmiel, Mazur 2014: 35).

In the 1980s, the idea of *audio description* reached the other side of the Atlantic – Great Britain. At present, performances with audio description can be seen in over seventy theatres in this country. In 1994, audio description also appeared on British TV. As part of the Audetel project on ITV and BBC channels, four programmes per week were broadcast with audio description for four months. Audio description in the cinema was introduced by the Chapter Arts Centre in Cardiff (Wales), where regular screenings with a description read live began.

Although the phrase *audio description* entered the Polish market as late as 2006 (cf. Jankowska 2013a: 33) and this year is widely recognized as the beginning of audio description in Poland, the technique of making audiovisual content available to the blind has been known in our country since the end of the 1990s. In Poland, audio description appeared for the first time in tyhlofilms (i.e. a kind of special films in which visual elements of the composition of a film image are verbally described to blind people). The technique of creating tyhlofilms differed from modern audio description, among other things, in that it used still frames – when the text was too long to fit in the gap between dialogues, the film would be stopped for a moment and the description would be loaded. Consequently, tyhlofilms could not be broadcast on TV and during public screenings, and their release on DVD would require the preparation of a separate version. In total, more than 30 titles were developed using this technique, including foreign language productions, which are available in the Central Library of PZN (cf. Chmiel, Mazur 2014: 37).

Białystok is considered to be the cradle of Polish audio description. It was in this city that the Audio Description Foundation was established. As the first non-governmental organization, it started activities in the area of comprehensive access to visual culture for the blind and visually impaired. The Białystok foundation was created by blind people – Tomasz Strzemiński and his wife Barbara Szymańska. The official premiere of audio description in Poland was the screening of the movie "Statyści" (*Extras*) directed by Michał Kwieciński, which took place in 2006 at Białystok Cinema Peace. After this event, audio description began to spread in nationwide media and became an increasingly recognizable initiative. In 2007, audio description was used for the first time in the TV series "Ranczo" (*Ranch*). On 17 September 2007, on the initiative of Tomasz Strzymiński, the film "Świadek Koronny" (*Protected Witness*) was presented with audio description at the XXXII Feature Film Festival in Gdynia, the largest film event in Poland. It was the first application of audio description in the world during a film festival. In October 2017, for the first time, audio description was used for a football game in Białystok. A month later, audio description was used in the puppet theatre in Białystok (cf. Fundacja Audiodeskrypcja, n.d.).

Following the Białystok foundation, Krakow launched its own foundation called "The Seventh Sense", dealing with the creation of both audio description and subtitles for the deaf. In 2008, the first full-length DVD feature film with audio description was released: *Katyń*, directed by Andrzej Wajda. In 2010, audio description was prepared for the monumental Panorama Racławicka in the National Museum in Wrocław. In 2010, the premiere of the film *Chopin: Desire for Love* was screened with audio description in two languages, i.e. Polish and English. The English version was prepared under the supervision of the lecturers from the Institute of English Philology SWPS, and the Polish version was prepared under the supervision of Izabela Künstler-Zawisza, the head of Editors of the Deaf, who has worked as a journalist and a subtitler for Polish TV for many years.

The recording of the text and the preparation of the DVD was part of a project by the Polish Association of the Blind called "Chopin for the Blind", which was co-financed by the Cultural Exchange Fund under the auspices of the Ministry of Culture and National Heritage. In the same year, standards for creating audio description for audiovisual production were developed by the Foundation Audiodeskrypcja (Szymańska and Strzymiński, 2010), and two years later a set of guidelines was prepared by Fundacja Kultury Bez Barier (Culture without Borders Foundation) (cf. Künstler et al. 2012). In 2010, the Wrocław "Katarynka" foundation joined the group of organizations dealing with the dissemination of audio description (cf. Fundacja Katarynka, n.d.). Today, blind people can watch films with audio description in Wrocław at the Wrocław-West Cultural Centre once a month and verbal descriptions for many audio-described exhibitions are also available in the Wrocław branch of the National Museum and in the Panorama Racławicka.

2 Legal regulations

In 2009, in connection with the implementation of Directive 2007/65/EC on audiovisual media services, the Audio Description Foundation established cooperation with entities shaping national policy in the field of universal access to audiovisual services, including the Polish Senate, the National Radio and TV Council and the National Chamber of Audiovisual Producers. This enabled the transfer of the proposals developed by the Foundation regarding the provision of access services (i.e. the audio description of subtitles, sign language) to the draft bill of media regulations. In 2010, the website of the Ministry of Culture and National Heritage published on radio and TV the final version of the amendment to the Act approved by the Council of Ministers. The amended Act includes the definition of audio description with the following wording:

> Audio description means a verbal, auditory description of the image and visual content contained in an audiovisual programme, intended for people with disabilities due to vision dysfunctions, placed in the broadcast or distributed simultaneously with it. (Article 4, (28), cf. Sejm Rzeczpospolitej Polskiej 2011)

TV broadcasters are committed to ensure the availability of programmes for people with disabilities due to sight dysfunctions and hearing impairments by introducing appropriate facilities, in particular audio description, subtitles for the deaf or sign language translation, so that at least 10 % of the quarterly broadcasting time of the programmes, excluding advertising and telesales, can offer such facilities.

However, the National Council may determine, by regulation, to take into account the diverse offer at various airtimes, the technical capabilities, the recipients' needs, the dissemination methods and the programme's specialization, without imposing unreasonable obligations on broadcasters. An obligation imposed by the National Council was to be introduced gradually. The transitional provisions of the Act suggested spreading over time the threshold of 10 % of the broadcast with facilities for the disabled, so that in 2011 at least 5 % of the programmes would have such facilities, while the 10 % threshold was to be achieved in 2012. The entry into force of the amendment to the Broadcasting Act was supposed to create a coherent legal act that would regulate the basic aspects of the functioning of electronic media and would make them accessible to people with visual and hearing disabilities who, despite being equal recipients of audiovisual programmes, had been deprived of this right so far (cf. Chmiel, Mazur 2014: 39 ff.).

The public broadcaster meets the above-mentioned provisions by broadcasting an additional sound track available for programmes broadcast in DVB-T (digital terrestrial TV), DVB-C (digital cable TV) and DVB-S (satellite TV) systems. It also provides programmes with audio description on the Internet (cf. Telewizja Polska TVP 2018b).

3 The current TV landscape in Poland

The TV landscape in Poland can be likened to a singular 'European fossil of voice-over translation', which represents a special style of translation designed to cater for the specific needs of TV audiences and companies. However, there are also other phenomena that contribute to the fact that the audiovisual media landscape in Poland is unique as far as European standards go (cf. Jüngst 2010: 80). The allocation of tasks to individual methods of audiovisual translation, created in the second half of the 20th century in Poland, is based on three tasks, or, as Bogucki says (2016: 21), on the triad of the most important methods of translation. Polish TV broadcasters still mostly use voice-over translation, but dubbing is used to translate programmes for children and teenagers. Moreover, thanks to the ongoing process of digitization, it is possible to replace both above-mentioned methods of translation with subtitles in some programmes. This method has dominated film screenings in Polish cinemas (cf. Tomaszkiewicz 2006; Szarkowska/Laskowska 2015).

This diversity is still not a general rule. For example, the third largest TV broadcaster in Poland, Polsat, most often offers on its thematic channels only a voice-over version or synchronization. The same applies to state TV, which is financed by commercial revenues and the state budget. Broadcasters who can only enjoy a small share in the media market, such as the Comedy Central channel belonging to MTV, offer viewers only the voice-over version or a programme with an original dialogue. For cinema screenings, the voice-over version is completely abandoned, but depending on the age of the films' target group, synchronization or subtitles are used (in the case of Harry Potter films, due to the age diversity of the audience, sessions of the dubbed and signed version were run in parallel).

In the 1960s, the reason for giving precedence to the voice-over method in Polish TV was mainly cost effectiveness. Looking at decisions regarding translation services in Poland, one has the impression that the focus on economic factors has been of prime importance to TV until now. Fighting for the rights of viewers, Polish NGOs supported by many representatives of science are making two main demands, that is, to replace the voice-over version with an efficient method of translating dialogues, and to adjust programmes to the needs of people with disabilities by using audio description and appropriate measures.

4 Overcoming barriers in Polish TV

As an example of these changes, Poland is introducing systemic solutions comparable to those introduced in 2010 in Great Britain, France and the Netherlands (cf. Telewizja Polska TVP 2009). On all of its thematic channels, the BBC offers facilities for the disabled in 90 % of broadcast (in accordance with Communications Act 2003) (cf. Krajowa Rada Radiofonii i Telewizji 2013). In France, it is obligatory to provide audio descriptions and subtitles for the deaf for the full programme on offer, excluding broadcasters with a market share of less than 2.5 %. In the Netherlands, this exclusion has been abandoned. In Poland, broadcasters are subject to the obligations of EU Directive 2010/13. The state-owned Telewizja Polska's (TVP) annual public TV reports announce that the offer will be made available to the disabled for up to 30 % of broadcasts. A rudimentary study to investigate this state of affairs conducted in August and September 2017 for the 1st Swiss Conference on Barrier-free Communication shows that this obligation has not been fully met, as the following examples demonstrate.

In Poland, the dependence of the media on economic factors and their decisive influence on the image of Polish TV can be seen in many, often dysfunctional phenomena. The public TV debt has led to a situation where even extremely important events for Polish society are reported by private broadcasters. However, analyzing the structure of the media in Poland, it is difficult to notice any cost-effective policies. As a result of the digitization of Polish TV, 28 TV channels are available in every household, of which 8 are provided by public TV. Entities offering cable TV services in Poland have from 150 to about 200 channels in Polish, including 11 channels and the VOD service of the state-owned TVP. In addition to the first and second channel, there is also a news channel, as well as local, sport, art and culture, history, series and entertainment programmes, plus channels for HD movies and children.

In Poland, the public broadcaster role is performed by Telewizja Polska, which is obliged by law to spread the so-called promotion of public services – the mission of public media. It broadcasts two nationwide channels, i.e. TVP1 and TVP2 (also in HD versions), a regional channel, TVP Regionalna and TVP Polonia for Poles living abroad. Specialized channels broadcast on the basis of a license: TVP Kultura, TVP Sport (which also has an HD version), TVP Historia, TVP Info (also with an HD version), TVP Series, TVP Rozrywka, TVP ABC, the TVP Parliament channel and a channel broadcasting in HDTV TVP HD. However, TVP is also the broadcaster of 3 stations which are treated as strictly commercial: TVP Sport, TVP HD and TVP Series, which despite the braodcaster's mission are coded and are offered to TV operators for broadcast. The Belarusian Belsat TV channel also belongs to Telewizja Polska.

In 2013, over 50 broadcasters of commercial TV nationwide operated in the Polish TV market. They include Polish private broadcasters as well as foreign ones. The majority of Polish-

language stations are usually owned by larger media companies (or groups) such as the ITI Group, Telewizja Polsat, Canal + Cyfrowy, HBO Europe, NBCUniversal, Discovery Networks Europe, BBC Worldwide, Fox International Channels, Viacom Media Networks, The Walt Disney Company, A + E Networks, Chello Zone, Viasat and others. Broadcasters provide both general and thematic channels. Nationwide commercial TV channels enjoying the highest viewing rates include Polsat, TVN, TVN 7, TV4, TV Puls, Puls 2.

In most countries, state TV stations serve as a benchmark for respecting rules and guidelines that they implement in a timely manner. They are thus a benchmark for the development of the TV industry, as illustrated by examples from Great Britain, France and the Netherlands. In Poland, there is a specific competition for market share between the state-owned TVP station and the two largest private entities, TVN and Polsat. In recent years, state-owned TV has had serious problems with drops in ratings, which translate into financial problems and create the need for significant funding from the state budget. In this difficult situation, while pursuing political goals, the standards set by TVP, which other stations should look up to, do not set a high bar. The resultant neglect to a large extent affects people with disabilities, limiting their unquestionable right to benefit from the media.

On 5 June 2013, representatives of seven companies offering TV services in Poland signed an agreement, the purpose of which was to implement the guidelines included in EU Directive 2010/13 and the UN Convention on the Rights of Persons with Disabilities of 2006. Following the guidelines, the media set requirements regarding the level of facilities on TV channels at a minimum level of 10 %, which in comparison with the ideas adopted in many European countries appears negligible (*Agreement of broadcasters of 5 June 2013*). The vast majority of broadcasters and channels could focus on quantity instead of quality. Another important and unpleasant aspect of the problem is the dissonance between the intent accompanying the agreement and the methods of its implementation.

The website of the state broadcaster TVP publishes a list of programmes with audio description services. This list of programmes usually includes the titles of repeated series, some of which are broadcast between 1am and 3am, which does not suggest good customer care. The website adds value to the programme offer and encourages viewers to watch selected films with audio description on the Internet. However, the link to these videos has been placed in the tab that informs users about the range of films with captions on Polish television and without proper facilities for the visually impaired.

TVP1, a channel financed partly by the state budget and whose share of the TV market in January 2017 was 11.56 % (*Business Insider*, data from 1 February 2017), broadcasts one or two programmes with audio description every day. However, these are only re-screened episodes from past series and are not really up-to-date or attractive. In comparison, the private TV station Puls, whose market share amounts to 2.73 %, offers a similar number of programmes each day with services, i.e. one broadcast with audio description, two programmes with subtitles for people with impaired hearing ability and one with sign language translation (cf. TV Puls 2018).

More than a dozen TV channels, including one state channel, cater for the needs of Polish children with impairments. In the context of facilities for the disabled, special attention is paid to the programmes of the Minimini plus channel. The broadcaster's website promises the availability of both subtitles and audio description for all broadcast programmes (cf. Minimini plus, n.d.).

Despite the authors' consultations with the technical department of cable and satellite TV, these options are not available. Neither the broadcaster nor the cable operator was able to explain this situation. The state TV channel ABC has prepared its range of programmes for children with disabilities. The offer includes one or two programmes with audio description every day. At the time of writing this article, there are a Czech-Chinese cartoon called "Krecik i Panda" (from 1 to 3 episodes of the series per day) and a production of TVP ABC, "Nela - a small reporter" (1st episode of the series). Moreover, on Saturdays, the "Supełkowe ABC"

programme is broadcast. This is a programme that meets educational goals, the content of which is entirely translated into sign language (cf. Telewizja Polska TVP ABC, n.d.).

The choice of content for which services are provided is controversial, despite the undeniable educational advantages. Services for people with disabilities are intended to counteract exclusion, so creating programmes designed exclusively for people with disabilities, instead of providing more transmissions with audio description or subtitles of children programmes, is not conducive to inclusion. Programmes that feature collections of school accessories, clothing and toys are not available for children who have to deal with communication barriers in Poland.

5 Services in the public space for people with hearing impairments and poor eyesight

It is difficult to identify effective systemic solutions in Polish TV to disseminate methods that allow access to the media for people with hearing or sight impairments. The standards set by Polish TV differ from good examples known in Europe. However, a change of perspective reveals that there are many good practices taking place in Poland in other areas, such as the design of public spaces and the ways in which communication barriers are being overcome.

Let us look at one example, from the third largest city in Poland, Łódź, with approximately 700,000 residents. The history of the development of this city with industrial roots is only 180 years old, but due to scarce investment in the second half of the twentieth century, the public space in the city has grown obsolete. It is only in recent years that a fresh breeze has been introduced to this 'city of four cultures' (at the turn of the 19th and 20th century, Łódź Metropolis was co-created by its Polish, German, Jewish and Russian populations), its textile factories and the Łódź Secession, the city's Art Nouveau heritage. At long-standing tourist attractions, it is difficult to find convincing forms of adjustment to the needs of deaf and blind people. The management of each institution decides about initiatives individually and it is difficult to talk about a systematic policy implemented in this field. One significant exception is the Museum of History of the City of Łódź, where young artists have constructed a scale model of the largest hall of the Poznański Palace. This enables blind and visually impaired people to become acquainted with the architectural details of the palace by touch.

However, new projects and investments in the city do have elements meeting the best European standards and adhering to directives prescribing accessibility for people with disabilities. For the implementation of the EC1 project, a new culture centre in the city, the management has followed the guidelines of the National Centre for Film Culture, which lays down the use of audio description and Braille description. In addition, people implementing the project were required to be advised on the planned facilities by entities specialized in audio description. It is a pity that these common-sense guidelines are relativized by the phrase "whenever possible"[45]. The implementation of modern and very ambitious projects and the infinite potential for opportunities and amenities in the context of equal opportunities for citizens should be self-evident.

Most of the manifestations of openness to the needs of disabled people in Poland are the result of efforts by specialized non-governmental organizations. As can be seen from the examples mentioned above, these have become indispensable in the drive to meet the needs of disabled people. The measurable effects are numerous services in the field of audiovisual media. Actions promoting audio description of films and TV programmes in Poland are the focus of many R&D groups and individual researchers involved in audiovisual translation, such as Fundacja Audiodeskrypcja, Fundacja Widzialni, Agnieszka Szarkowska, Izabela Künstler-

[45] "[…] Whenever possible, institutions should provide the blind with the opportunity to experience an exhibition – in spite of its strongly audiovisual component. Therefore, solutions using audio description and descriptions in Braille are preferred, especially when the message is as compatible as possible with the main idea. Entities (associations and foundations) specialized in audio description technology should be involved in the implementation of such projects [...]".

Zawisza, Agnieszka Chmiel, Iwona Mazur, Barbara Szymańska or Tomasz Strzymiński, who strive to provide equal opportunities for access.

The successful initiatives of non-governmental organizations are also visible in the field of website accessibility. In 2017, the number of websites published by disabled-friendly public institutions increased by 25 % compared to the previous year. This is testimony to the difficult work in this field done by the Widzialni Foundation and organizations with a similar profile of activity. The high quality of Internet access shows the possibility of adapting the Polish media to the desired standards and of applying good practices even with little governmental involvement (cf. Fundacja Katarynka, n.d.).

6 Future outlook and conclusions

What can be done to transfer the good practices of NGOs in the field of website accessibility to the TV market in Poland? According to the forecasts of the National Council of Radio and TV, 50 % of programmes in Poland will be provided with services for the disabled by 2022. The planned pace of development is such that in 2018 the level of services will increase to 15 %, in 2019 to 25 %, and between 2020 and 2021 to 35 %. Such a slow implementation of long overdue changes is obviously a consequence of the aforementioned economic factors, but it is also caused by the structure of the audiovisual translation market in Poland. Decades of reliance on voice-over translation have resulted in media companies getting used to low expenditure on translations of their programmes. Moreover, the mechanisms that make up the dubbing industry have yet to be built, and a lack of trained dubbing actors means that film and theatre actors must be employed instead, enabling only small-scale production. In addition, audiovisual translation methods are taught in only a few academic centres, and do not cover the full range of options. It will also be difficult to introduce new methods, as viewers are used to the traditional division of tasks in audiovisual translation. Finally, given that the public media are the role model for the entire media market in the implementation of solutions relevant to society, actions and initiatives will be necessary to prevent broadcasters acting illegally.

In short, allowing the blind and visually impaired and the deaf and hearing-impaired to have full access to the media is a self-evident responsibility of countries that want to be regarded as modern and fully democratic. These tasks currently rest with the public TV and its financial structure, making it unlikely that the situation on the Polish TV market will improve quickly without further bottom-up initiatives involving private sector, R&D and non-governmental organizations.

Sources

Literature

Bogucki, Łukasz (2010): “The demise of voice-over? Audiovisual translation in Poland in the 21st century.” Barbara Lewandowska-Tomaszczyk; Marcel Thelen (eds): *Meaning in translation*. Frankfurt am Main: Peter Lang, 415-424

Chmiel, Agnieszka; Iwona Mazur (2012): “AD reception research: some methodological considerations.” Elisa Perego (ed): *Emerging Topics in Translation: Audio Description*. Trieste: EUT Edizioni Università di Trieste, 57-80

Chmiel, Agnieszka; Iwona Mazur (2014): *Audiodeskrypcja*. Poznań: Pracownia Wydawnicza WA UAM

Ciborowski, Mateusz (2010): “Pod lupą: Chodź, opowiedz mi świat.” *Pochodnia 4* [871].

ITC Guidance on Standards for Audio Description (2000). http://www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/audio_description/index.asp.html (30 April 2018)

Jankowska, Anna (2013a): Tłumaczenie skryptów audiodeskrypcji z języka angielskiego jako alternatywna metoda tworzenia skryptów audiodeskrypcji. Unpublished doctoral dissertation. Kraków: Uniwersytet Jagielloński

Jankowska, Anna; Martyna Mentel; Agnieszka Szarkowska (2015): "Why big Fish isn't a FAT cat? Adapting voice-over and subtitles for audio description in foreign films." Łukasz Bogucki; Mikołaj Deckert (eds.): *Accessing Audiovisual Translation*. Frankfurt am Main: Peter Lang, 137-148

Korycińska-Wegner, Małgorzata (2011): *Übersetzer der bewegten Bilder. Audiovisuelle Übersetzung – ein neuer Ansatz*. Frankfurt am Main: Peter Lang

Künstler, Izabela; Urszula Butkiewicz; Robert Więckowski (2012): "Audiodeskrypcja – standardy tworzenia." http://dzieciom.pl/wp-ontent/uploads/2012/09/Audiodeskrypcjazasady-tworzenia.pdf (30 April 2018)

Michoń, Marcin (2016): "Kommunikative und technische Nezessität als prägender Faktor in der audiovisuellen Übersetzung." Łukasz Bogucki; Barbara Lewandowska-Tomaszczyk; Marcel Thelen (eds.): *Translation and Meaning. New Series, Vol. 2, Pt. 2*. Frankfurt am Main: Peter Lang, 77-91

Szarkowska, Agnieszka; Monika Laskowska (2015): "Poland – a voice-over country no more? A report on an on-line survey on subtitling preferences among polish hearing and hearing-impaired viewers." Łukasz Bogucki; Mikołaj Deckert (eds.): *Accessing Audiovisual Translation*. Frankfurt am Main: Peter Lang, 179-197

Szymańska, Barbara; Tomasz Strzymiński (2010): "Standardy tworzenia audiodeskrypcji do produkcji audiowizualnych." http://www.audiodeskrypcja.org.pl/index.php/standardy-tworzenia-audiodeskrypcji/do-produkcji-audiowizualnych (30 April 2018)

Tomaszkiewicz, Teresa (2006): *Przekład audiowizualny*. Warszawa: Wydawnictwo Naukowe PWN.

Internet sources and documents

Business Insider (2017): https://businessinsider.com.pl/media/tv-radio/ogladalnosc-telewizji-w-pazdzierniku-2017-roku/n5yypms (15 September 2017)

Fundacja Audiodeskrypcja (n.d.): www.audiodeskrypcja.org.pl (15 September 2017)

Fundacja Katarynka (n.d.): www.fundacjakatarynka.pl (15 September 2017)

Fundacja Widzialni (n.d.): http://widzialni.org/wyniki-raportu-dostepnosci-2017 (15 September 2017)

Krajowa Rada Radiofonii i Telewizji (2013): http://www.krrit.gov.pl/Data/Files/_public/PortaPDFdocumentls/0/Nadawcy/aktualnosci/porozumienie_nadawcow.pdf (15 September 2017)

Minimini Plus (n.d.): https://www.miniminiplus.pl/program (15 September 2017)

Sejm Rzeczpospolitej Polskiej (2011): http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WDU20110850459/T/D20110459L.pdf (15 September 2017)

Telewizja Polska TVP (2018a): http://www.tvp.pl/dostepnosc/napisy-dla-nieslyszacych (15 September 2017)

Telewizja Polska TVP (2018b): http://www.tvp.pl/dostepnosc/programy-z-audiodeskrypcja-w-tvp/5267057 (15 September 2017)

Telewizja Polska TVP (2009): www.tvp.pl/dostepnosc/telewizja-bez-barier (15 September 2017)

Telewizja Polska TVP ABC (n.d.): http://abc.tvp.pl/18490761/supelkowe-abc (15 September 2017)

TV Puls (2018): http://tvpuls.pl/udogodnienia (15 September 2017)

Authors

Agnieszka Stawikowska-Marcinkowska is Junior Professor at the Institute of German Philology of the University of Łódź, Poland. Her research focuses on two areas: the contribution of corpus linguistics to the study of legal and business language and its phraseology, and contrastive discourse analysis. She is currently working on the specialist language of communication at the stock exchange. E-mail: agnieszka.stawikowska@uni.lodz.pl

Marcin Michoń is Junior Professor at the Institute of German Philology of the University of Łódź, Poland. His research focus is on the areas of text linguistics, translation process research, interpreting, audiovisual translation and mass media communication. E-mail: marcin.michon@uni.lodz.pl