Hrvatska revija za rehabilitacijska istraživanja 2022, Vol 58, (Special Issue) Sign Language, Deaf Culture, and Bilingual Education str. 109-120

OPEN SCIENCE VERSUS DATA PROTECTION – CHALLENGES AND SOLUTIONS IN SIGN LANGUAGE ACQUISITION STUDIES

VERA KOLBE

Humboldt-Universität zu Berlin, Faculty of Cultural, Social & Educational Sciences, Institute for Rehabilitation Sciences, Department of Sign Language Pedagogy & Audio Pedagogy, Ziegelstr. 4, Berlin, Germany

contact: vera.kolbe@hu-berlin.de

Received: 15.09.2021.
Accepted: 31.05.2022.

PRELIMINARY REPORT
UDK: 81’221.24:001.891
81’221.24:303
doi: https://doi.org/10.31299/hrri.58.si.6

Abstract: *Research within Deaf communities needs to adhere to the ethical requests of the partner communities involved. These ethical requests can be met via open science practices that are implemented in the project strategies of the European Commission. Open science refers to transparent, collaborative, and accessible research including citizen science. However, researchers studying the acquisition of sign languages are challenged by the General Data Protection Regulation implemented by the European Parliament. Researchers who study sign languages frequently handle personal data, i.e. video data. Such data cannot be fully pseudonymised since facial expressions contain relevant linguistic inputs and cannot be blurred. Hence, strict data protection measures are necessary, but these measures are contradictory to open science practices. How can we meet the demands of the data protection regulations, as well as enable open science practices. This article identifies strategies used in a sign language acquisition study involving more than 100 sign language videos of children’s narrations, which was conducted at the Department of Sign Language Pedagogy and Audio Pedagogy, Humboldt-Universität zu Berlin, Germany.*

Keywords: *open science, data protection, sign language acquisition studies, research data processing, sign language research ethics*

1. INTRODUCTION

Researchers studying sign language acquisition need to be aware of the research ethical requests of Deaf communities. These requests are similar to those of diverse linguistic and cultural minorities (Harris, Holmes, & Mertens, 2009, p. 112), as well as ethical principles for research with human subjects (Pollard, 2002, p. 162). In this article open science practices are described with respect to the research ethical requests of Deaf communities.

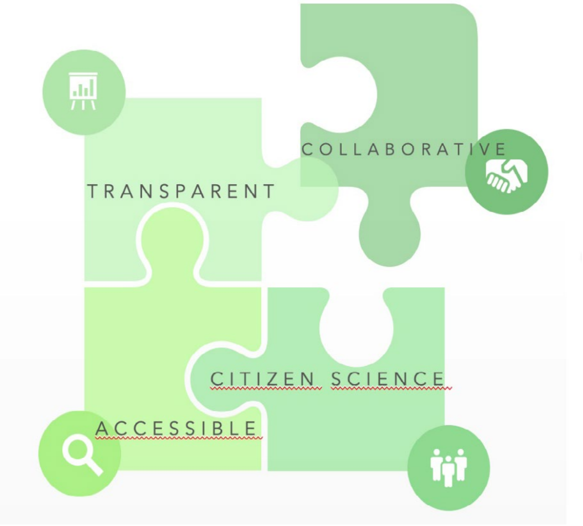
The General Data Protection Regulation (GDPR) of the European Parliament clearly defines the legal demands associated with data protection while processing or storing data. In addition, the European Commission has requested the implementation of open science practices in its project strategies (e.g., Horizon 2020). Data protection and open science practices are two conflicting requirements that pose challenges for researchers, especially those handling video data involving children. Acquisition studies in sign languages are based on language samples from children, i.e. video sequences of children signing. Sign languages are visual languages that are expressed not only through the hands, but also through body posture and facial expressions. Therefore, these video clips cannot be fully pseudonymised. As the faces of individuals in these videos cannot be blurred or covered with a black bar since this would eliminate important linguistic information. Hence, video clips used in sign language research studies are considered personal data under the data protection regulations of the European Union (EU) and the processing of these videos falls under restrictive legislation. Since children are specially protected, the regulations are even more restrictive when it comes to processing video clips of children in sign language acquisition studies.

This paper presents experiences made during a large sign language acquisition study that was based on more than 100 videos of children´s narrations in German sign language. This paper is structured as follows: After a short introduction to the research ethical requests of Deaf communities presented in Section 1, the open science requirements of the EU are outlined in Section 2, while Section 3 provides implementation examples from our acquisition study in German sign language. Section 4 describes the requirements of the GDPR when processing video data of children, and Section 5 presents solution strategies used for the handling of the video data of children in this study. Finally, Section 6 describes the limitations as well as future possibilities with respect to increased exchangeability and more open science in acquisition studies of sign languages.

1.1 Research ethical requests of Deaf communities

Sign languages are the natural languages (Sandler & Lillo-Martin, 2017, p. 371) of Deaf communities around the world. “*Deaf communities form cultural groups with practices and values that are in some cases distinct from those of non-Deaf communities*” (Hill, Lillo-Martin, & Wood, 2018, p. 2). Research ethical requests of Deaf communities are often similar to the requests of many linguistic and cultural minorities: they demand culturally sensitive research designs (Harris et al., 2009, p. 112) and question whether the Deaf community´s interests, values, and norms are represented in research design and content (Pollard, 2002, p. 6). These requests demand beneficence of research leading to more social justice for the Deaf communities, that are involved as partner communities in research projects (Harris et al., 2009, p. 109). The Sign Language Linguistic Society (2016) has provided specific directions that take into account responsibilities towards Deaf individuals and Deaf communities, researchers and the public, that need to be considered when conducting sign language research. Ideally, studies use a community-based participatory research design. Section 2 provides further information on research ethical requests of Deaf communities along with implementation examples in Section 3.

2. OPEN SCIENCE



Graphic 1. *Illustration depicting the principles of open science (European Commission, 2017, p. 5): designed by V. Kolbe based on Graphic Node (2019)*

The European Commission is highly engaged in fostering open science practices and implementing them in the requirements of publicly funded European projects. They define open science as having to be transparent, collaborative, accessible, and fostering citizen science (European Commission, 2017, p. 5). The need for open science to be transparent at all stages of research has been put into place in order to enable all stakeholders to connect, participate, and review research processes, as well as research results. Transparent research design and data sharing makes research reproducible and aims at enhancing sound scientific conduct. It also enables collaborative research practices, where different researchers can effectively access data from previous studies. Science should be accessible to stakeholders and the public, especially to the communities involved in and affected by a research topic. Citizen science focuses on the participation of the wider public, e.g., non-scholars, in research projects, thus widening the social impact of research and bringing in new perspectives. The means for open science are open scholarly communication, open access publications, and openly accessible research data (European Commission, 2017, p. 6).

Open science practices will reform how we conduct research projects, as well as our attitude towards sound and rigorous science. Open science practices aim to enhance research quality through networking and data sharing. They also aim to increase social impact through higher visibility of research results (European Commission, 2017, p. 5; 2018, p. 4). The above-mentioned principles of open science aim for further research integrity, quality, and social impact. At the same time the critical claims of Deaf communities as communities affected by minority language research effects can be subsumed within these categories.

2.1 “Science needs to be transparent.”

One of the major prerequisites of transparent research is the “conflict of interest” statement that is required by journals in order to clarify the connections between funding and research content. In the case of research with partner communities, it is important to be transparent about the possible influence of our research and the corresponding results for the partner community. A transformative research-philosophical approach demands that research should contribute to improved social justice for all involved communities:

> *“Beneficence is defined in terms of the promotion of human rights and increased social justice. An explicit connection is made between the process and outcomes of research and furtherance of a social justice agenda.” (Harris et al., 2009, S. 109)*

We need to ask ourselves whether our research is beneficent and whether our research results will contribute to this social justice agenda or whether there will be conflicts of interest. The beneficence cannot be judged only by the researchers, but needs to be recognised and valued in the Deaf communities themselves (Pollard, 2002, p. 6).

2.2 “Science needs to be collaborative.”

The need for collaboration must be highlighted between different research teams and disciplines, as well as in the research team itself. In its ethics statement, the Sign Language Linguistic Society requests increasing participation of Deaf scholars (Sign Language Linguistic Society, 2016) . Regarding the collaboration between Deaf and hearing researchers, we need to be aware of the potential power issues that can arise.

> *“Sign language users and communities have been traditionally marginalized and researchers must always be aware that this might result in power inequalities between sign language consultants and researchers.” (Sign Language Linguistic Society, 2016)*

These power issues might also exist between hearing and Deaf scholars within the research team itself. We need to be aware of how the languages used in such mixed teams are organised because the selection of the main language for communication might create an imbalance towards users of sign languages. For example, because of the time lag that occurs with an interpretation or sign language interpreters who are not familiar with the research topic (Singleton, Martin, & Morgan, 2015, p. 11).

2.3 “Science needs to be accessible”

Science is made accessible to the general public through open access publications. In our case, we need to ensure accessibility especially for Deaf communities. It is important to take efforts to make all content (written or spoken) accessible in the sign language of the Deaf community involved in the research project. This can be achieved by producing websites that present content not only in written and spoken language, but also in the sign languages studied, e.g., NaKom DGS project website (https://nakom.hu-berlin.de/de/en/nakom/nakom.html), or by organising lectures bilingually (https://nakom.hu-berlin.de/de/forschungsergebnisse).

2.4 “Science needs to be citizen science”

Citizen science invites non-scholars to get involved in research projects. It also questions the ownership of research since it demands further participation and involvement of citizens (in our case, members of the Deaf communities) in research design and review processes.

> *“To conduct sign language research, scholars must first and foremost respect the wishes of the signing community, and be careful to involve*

> *members of the Deaf community in appropriate ways." (Sign Language Linguistic Society, 2016).*

This demand is met by a community-engaged research approach (Harris et al., 2009; Ross et al., 2010) that involves members of the partner community, e.g., Deaf communities, as equal members of the research team during all phases of the research project. As Deaf communities are close-knit communities, researchers need to carefully evaluate the impact of their research on all members of the partner community.

3. IMPLEMENTATION EXAMPLE IN A SIGN LANGUAGE ACQUISITION STUDY

The NaKom DGS-Test project conducted at Humboldt-Universität zu Berlin, Germany, adapted the "British Sign Language Production Test - Narrative Skills" (BSL PT) by Herman et al. (2004) to DGS (Kolbe, 2021). This test uses sign language narrations from children to assess their narrative competences. In the present study, DGS narratives of 103 children, aged 4-11 years, were collected in order to gather information about the course of sign language acquisition in DGS in the tested categories, as well as to develop reference measures for the test. In the sample of 103 children who provided DGS narratives, 72 children are native DGS signers, born into a family with Deaf parents who communicate in DGS. The logo of the NaKom DGS study is shown in Figure 1. We used the logo in the project flyer (link provided in Section 5), as well as in all communications with the associations of the Deaf, parents, and schools.

The following is an overview of the ethical considerations and open science practices adopted during the different stages of our research project. Members of the Deaf community have identified the need for sign language assessment and acquisition studies to provide necessary insights for fostering children´s sign language competences, and for reassurance about the development of a child´s competences in DGS. Deaf and hearing educators are requesting research about sign language acquisition in DGS, because DGS is taught in schools, yet very little research has been carried out so far.

Figure 1. *Logo of the NaKom DGS-Test sign language acquisition study focusing on the assessment of narrative production in DGS by children aged 4-11 years old, n = 103 children (Logo design: Department of Sign Language Pedagogy and Audio Pedagogy, Humboldt-Universität zu Berlin, Germany)*

Our research team consisted of Deaf and hearing researchers. In the research team, we communicated in DGS since all researchers had sign language competency. As part of the research design, we put forth several possibilities for collaboration with and involvement of members of the Deaf community as well as for discussion of research content in many phases of the study. Through their participation in the research team, Deaf scholars are becoming experts in the field of sign language acquisition and will hopefully continue to do research in this field, especially since there is a lack of research on this subject in Germany.

Given that our research involves the Deaf community, the potential impact of this research should be made **transparent**. The NaKom DGS study is establishing the first sign language production test for narrative competences in DGS and its results will have a high impact on educators, parents, members of the Deaf community and scholars. On the one hand, it provided the first set of results for many areas covered by the test instrument, thus raising awareness and providing reference measures for German sign language development of children. On the other hand, the NaKom DGS-Test is considered to be the first published test instrument for DGS, thus these results cannot be compared to the results of other test instruments. To avoid negative effects of this study on individuals or partner communities and to focus our awareness of the effects in respect to the social justice agenda, it is necessary to define these in the study design phase. Following Ross et al. (2010, p. 12), this should be done using an advantage-risk analysis (Table 1). In addition to the advantage-risk-analysis, Table 1 also contains lines showing the measures taken to prevent or minimise the risks added by Kolbe (2021, p. 89). It includes the impacts of the study on participating members of the partner community, as well as the partner community itself before, during, and after the research process. The impact on their agency was also analysed. Instead of providing details on the experiences during the research process, the first part of Table 1 lists the general questions that should be asked with respect to the different fields. The effects of the research results and the effects on agency are highlighted through examples of the analysis conducted as part of our study.

Table 1. *Impact of research analysis on partners, participants, and Sign Language Community, including measures taken (Kolbe, 2021, p.89), adaptation based on Ross et al. (2010), and examples from NaKom DGS study (Kolbe, 2021, p.89)*

	Partners and participants (PP)	**Partners and participants as members of the Deaf community**	**Deaf community**
Experience during research process (General questions provided)			
Risks	What are the risks that may affect PP personally?	What are the risks that may affect PP as members of the Deaf community?	What are the risks that may affect the Deaf community?
Advantages	What are the advantages that PP may benefit from personally?	What are the advantages that PP as members of the Deaf community might benefit from?	What are the advantages for the Deaf community?
Measures taken	What measures are taken to prevent or counter the risks for PP?	What measures are taken to prevent or counter the risks for PP as members of the Deaf community?	What measures are taken to prevent or counter the risks for the Deaf community?
Effects of research results (Examples from study provided)			
Risks	The test may not be able to assess competences of children correctly.	Children may be recognised in language samples and pictures; the language competences of children may be criticised	Only few acquisition studies in DGS conducted so far; NaKom DGS provides first reference measures based on results of “only” 72 native signing children with Deaf parents; no other assessment instrument for DGS acquisition has been published so far
Advantages	Competences of child can be compared to the development of reference groups - educators can adapt teaching accordingly	Direct awareness towards own sign language competency	Large acquisition study in DGS (n = 103); first reference measures for language acquisition in DGS in the areas covered by the test instrument
Measures taken	Collect a large sample (as large as possible) of native signing children with Deaf parents from all over Germany for establishing reference measures; conduct statistical analyses to establish validity; international comparison of test results with BSL PT (Herman et al., 2004) and ASL Expressive Skills Test (Enns, Zimmer, Broszeit, & Rabu, 2019); in cooperation with educators, test results should be used for defining areas of high and low competences for didactical planning, thus expanding sample for reference measures	Informed consent of parents and children divided into the following: - consent to participate in study (anonymous participation, except for video analysis by research team) - consent to publish parts of video clip or video frames as part of a scientific publication (child may be recognised) - consent to use language samples in tester training (child may be recognised) Decision of research team: minimal publication of language samples of children, instead examples from adult signing are published whenever possible.	Discussion of intermediate and final results with Deaf experts; bilingual publication of test results on project homepage (DGS + German; English + International Sign); presentation and discussion of results also in DGS, not only in scientific community, but in Deaf community as well

Effects on agency			
Risks	Child may feel pushed to participate in study; awareness of own sign language competences	Publication of video clips and frames - children may be individually approached about content	Research team was led by a hearing researcher
Advantages	Positive experience of own competences in DGS and its value	Sign language competences of children are acknowledged and valued	Research team consisted of Deaf scholars; new knowledge in Deaf community about DGS development of children
Measures taken	Emphasised that test participation is voluntary; asked for consent of child before test was administered	Substitute video clips and frames of children with adults	DGS was used as the language of communication within the research team and in discussions with Deaf experts for equal accessibility

If new risks or advantages were identified during the research process, they should be added to the advantage-risk-analysis. The effect of measures taken should be assessed repetitively and adaptations documented.

Based on the details provided in Table 1, it is clear that the research study was designed to be collaborative. In many stages of the research process, experts from the Deaf community were included in discussions of preliminary results and questions. Deaf and hearing scholars in the research team communicated in DGS to ensure equal accessibility for all members. The NaKom DGS-Test is a test adaptation of the BSL PT (Herman et al., 2004), which has also been adapted to ASL (Enns et al., 2019). Thus, any problems that arose during the study were discussed in detail with the authors of the original test and the author of the ASL adaptation.

Extensive efforts were taken in order to make the results **accessible**, especially to Deaf communities: the results were published on the multilingual project homepage in German and DGS, as well as in English and International Sign. All letters to parents and children, as well as consent information and the project flyer were translated into DGS. The DGS video clips are accessible on the project homepage (link to project homepage NaKom DGS). The link and the QR-code leading to the website content should be included in the printed documents. All associated journal articles or publications were made open access or made accessible after the legal time limit. To enable equal accessibility of the sign language acquisition results, video lectures were produced bilingually in German and DGS and published on the project homepage. In a step towards **citizen science**, these results were made accessible not only for interested scholars, but also for the interested public, especially the Deaf community.

4. DATA PROTECTION REGULATIONS

The beginning of the NaKom DGS-Test project coincided with the implementation of the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and the free movement of such data, as well as the repealing of Directive 95/46/EC. The General Data Protection Regulation (GDPR) entered into force in May 2018 (GDPR, 2016/679). It is an important regulation that meets the transparency demands of open science, but tends to be very challenging when it comes to collaboration or accessibility.

The GDPR applies to all data subjects located in the EU, even if the data is processed outside the EU, as well as to all data processed in the EU, even if the data was collected elsewhere (Mondschein & Monda, 2019, p. 61). As researchers, we have the roles of data controllers and data processors. Sign language data is typically in the form of video clips that cannot be fully pseudonymised, because facial expressions carry relevant linguistic information for sign languages. This means that our research material is considered personal data. Since the data cannot be made anonymous when reproduced in publications or shown as parts of language samples in conferences (i.e., video clips), the person signing might be recognised. In acquisition studies of sign languages, the signing children might be recognised, because Deaf communities are tight-knit communities, where connections and acquaintances are manyfold.

The fundamental right to the protection of personal data demands that all data be processed in a lawful and transparent manner. When handling personal data, the controller and/or processors (i.e., the researchers) are obliged to keep a track record of all processing under the person´s responsibility ((82), Article 24 GDPR, 2016/679). A record of processing activities has to be maintained (Article 30 GDPR, 2016/679) and needs to be presented upon request.

Additional information about the data subject must be kept separate from the pseudonymised data to ensure that the information cannot be attributed to an identifiable natural person (Article 4 (5) GDPR, 2016/679). Because sign language data cannot be fully pseudonymised, we need to be even more careful when storing the additional data, e.g., contact addresses or collected background information.

The GDPR describes the legal framework under which personal data can be processed in a lawful manner (Article 6 GDPR, 2016/679): the most frequently used legal basis in the context of research is consent (Mondschein & Monda, 2019, p. 63). The controller and processors need to be able to prove that valid consent was acquired in a lawful manner. The data subjects possess the right to withdraw consent at any time. Since the German Research Foundation (Deutsche Forschungsgemeinschaft, 2013) recommends that research data should be stored for ten years, a study participant might decide after nine years that he or she wants the sign language video to be deleted. Therefore, efficient and solid long-term consent management and data management is necessary (Mondschein & Monda, 2019, p. 63).

Consent must be easily accessible, formulated in clear and plain language (Article 7 GDPR, 2016/679), and provided in a distinguished act. Consent can be given in writing, electronically, or orally ((32) GDPR, 2016/679). In the case of research in sign languages, consent can also be signed. In sign language acquisition studies, we often need to get informed consent from parents who are themselves members of the Deaf communities, which means that we need to provide all information in sign language, as well as in written form. Unfortunately, this is not yet provided in all studies involving members of Deaf communities (Singleton et al., 2015, p. 10). However, it can be easily achieved nowadays by producing video clips in sign languages, hosting them on a website, and making them accessible via website links or QR-Code printed on associated written information.

The consideration 33 to the GDPR (2016/679) acknowledges that often it is not possible in scientific research to fully identify the purpose of the personal data at the time of data collection. In this case, data subjects need to be able to give consent for areas of research or parts of research projects.

Article 89 of the GDPR (2016/679) outlines several safeguards and derogations for data processing in the scientific context, provided that technical and organisational measures, such as data minimisation and pseudonymisation, are fulfilled.

Children receive assitional protection since they may not be fully aware of the risks and consequences concerning the processing of their personal data ((33) GDPR, 2016/679). Therefore, the risk of discrimination as a result of the processing of personal data is explicitly mentioned. When processing video clips of children, we need to be aware of how differently they can be perceived and whether future teenagers would like to have videos from their childhood publicly available.

All information and communication with a child about the processing of their personal data must be conducted in an easily understandable language ((58) GDPR, 2016/679). Associations or other institutions, such as universities, are supposed to develop a code of conduct on how to ensure that children receive special protection and how to obtain parental consent (Article 40 2.(g) GDPR, 2016/679).

The GDPR is organised in rights and principles. We need to respect the rights of the data subjects, e.g., right to access one´s personal data, right to rectification, erasure and restriction of processing, while following the principles, e.g., accountability, accuracy, data minimisation, and storage limitation. The principle of accountability defines that we as data processors are responsible for data protection and the reporting of data breaches. Since the data subjects need to be informed about possible data breaches, this principle emphasises again the necessity of a solid long-term data management strategy. Another principle that needs to be provided by the controller is accuracy of the data. Data minimisation is an important principle that must be considered when designing the background questionnaires since the amount of data stored should be as minimal as possible: what information do we really need for our analysis? What type of information is only interesting at first glance, but will not be used as the study progresses? Storage limitation is a painful request when we think about the amount of work spent on data collection, because it demands that the researchers must specify a time frame for the deletion of data (Mondschein & Monda, 2019, p. 62).

5. SOLUTIONS WHEN HANDLING VIDEO DATA OF CHILDREN

In order to ensure that we fulfil all the requirements of the GDPR, we are currently making compromises regarding open science possibilities. However, we hope that future technical solutions might lead to more open exchange possibilities (discussed in Section 6).

Informed consent was obtained using a flyer that was produced in clear and easily understandable language and contained basic information about the purpose and scope of the study, data protection measures, all consent information, as well as consent withdrawal and data deletion information along with contact details (see flyer on the NaKom DGS project homepage). The contact address for the research team is very important. The address needs to be stable and accessible online for the declared duration of time corresponding to data storage. In our case, this corresponded to a ten year storage duration for primary data (Deutsche Forschungsgemeinschaft, 2013, p. 21). The flyer was designed to be visually appealing and understandable, and therefore contained many pictures and graphics. The QR code, as well as a link to the DGS translation was printed on all letters and the project flyer.

In the NaKom DGS project, two strategies were used to get in touch with parents and their DGS signing children in Germany. On the one hand, contact was established via the Ministry of Education in some federal states of Germany, as well as through school administration and school headmistresses and headmasters. Another strategy was to contact associations of the Deaf, Deaf community-based media, and other private contacts.

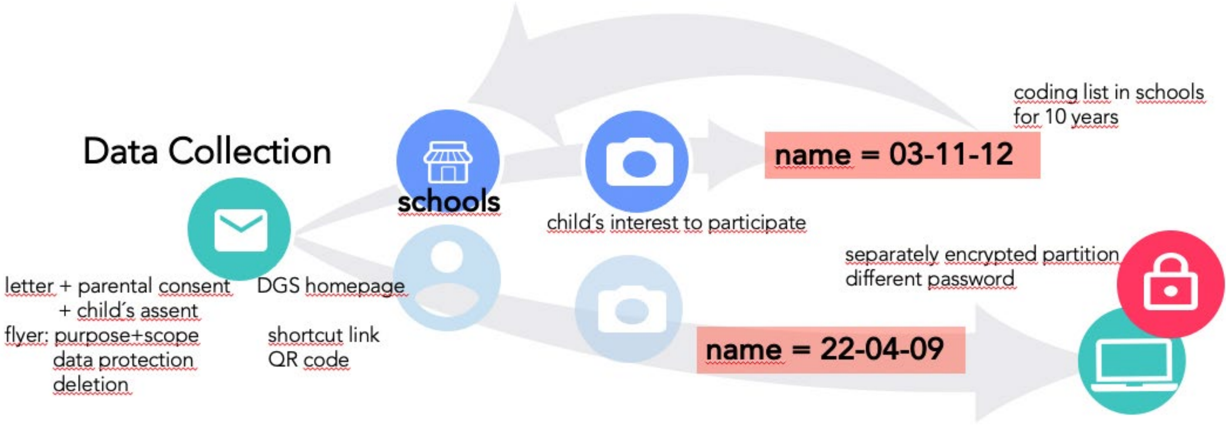


Figure 2. *Measures for data separation and data protection used during the data collection phase categorised into the steps used in the two contact strategies: via schools (mid-blue symbols) and via the Deaf community (light blue symbols) (designed by V. Kolbe based on Graphic Node (2019))*

The contact with ministries and school administration was established in Spring 2018, which was a time of insecurity due to the upcoming GDPR (2016/679). In the process, the content of the project flyer had to be modified due to requests from school authorities, resulting in the use of language that was not easily understandable in certain federal states.

For the purpose of our study, consent was established in two steps. First, we obtained consent for participation in the study. Next, after the data collection process, we obtained consent for the publication of parts of the language sample in scientific publications and congresses. Prior to data collection, the parents were contacted in order to obtain parental consent regarding their child's interest in participating in the study. When we visited the schools, the children were asked again for their consent before the beginning of the data collection process[1].

To ensure data protection, the children's narratives were not filed under their real names, instead a coding strategy was used with numbers. The coding lists linking the real children´s names to the codes remained in the schools and were filed for the duration of the data storage period, i.e., ten years.

When contact was established via the Deaf community, contact addresses also had to be managed. Parents and children were contacted via an article in the German newspaper for the Deaf (https://gehoerlosenzeitung.de), through associations for the Deaf, associations for parents of Deaf children, as well as private contacts. The contact addresses and coding lists were secured on a computer in a separately encrypted partition of the hard drive using a different password, which is symbolised in Figure 2 with the red lock.

All sign language videos were stored in a special restricted access server at the Department of Sign Language Pedagogy and Audio Pedagogy at Humboldt-Universität zu Berlin (HU), Germany. All additional information from the background questionnaires was pseudonymised and stored on this server.

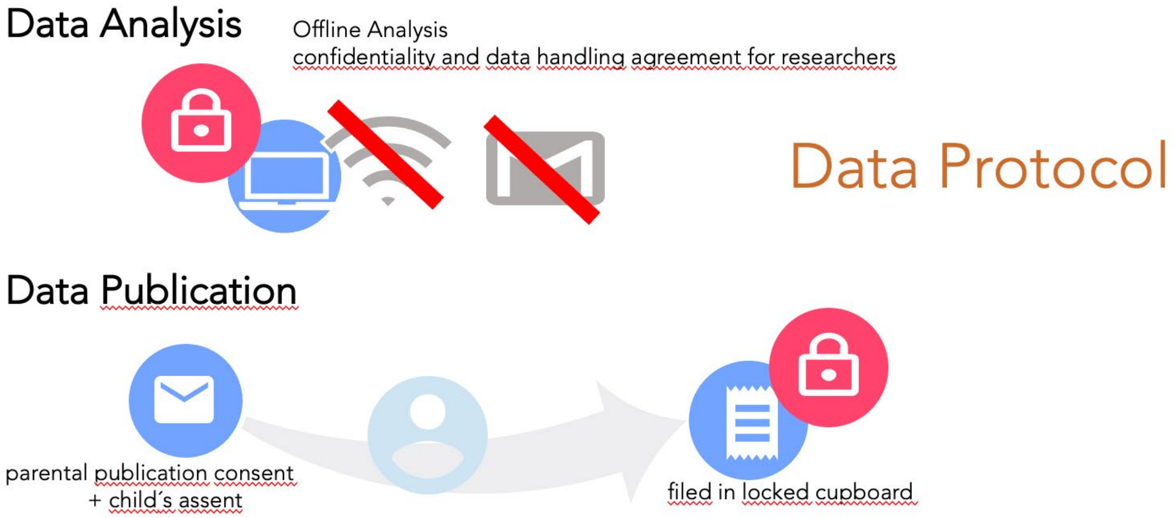


Figure 3. *Measures for data protection during data analysis phase prior to data publication (designed by Kolbe based on Graphic Node (2019))*

---

[1] One of the reviewers suggested that consent could be obtained from children using a comic version of the consent form where the information is provided as pictures or drawings. Then the child could "sign" with their name or draw a small picture of themselves.

For the purpose of data analysis, the necessary videos were downloaded via a secure network at HU and stored in a separate encrypted partition of the hard drive using a separate password (Fig. 3). This password was not included in the automatic password system of the computer, and had to be filled in manually each time in order to open the partition. The open source encryption software, VeraCrypt (IDRIX, 2017), was used to encrypt the data since it is compatible with different operating systems: Windows, Linux, MacOS. Data analysis was conducted offline, i.e., the computer was disconnected from the LAN or WLAN. All videos and the partitioned section were closed before going online and connecting to the internet, since the data is encrypted only when the partitioned section is closed. The video material could not be sent by email, unless the file itself was encrypted prior to being sent. For discussion and communication within the team, we used the cloud service of HU, but researchers were only allowed to access it via the VPN-tunnel program acknowledged through HU.

All participating researchers had to sign a data protection and confidentiality agreement, which included specific information concerning data protection, data processing, and the non-restorable deletion of videos at the end of their participation in the research team or at the end of the project.

As required by GDPR (2016/679), we wrote a data protocol to track all data processing and storage.

Before data publication, a second letter asking for consent was sent to the parents. In this letter, parents and children were asked if they would consent to the publication of sequences of the children´s narrations or snapshots from the video to be used in scientific publications or lectures. If contact was made via the Deaf community, a hard copy of these letters had to be filed in a locked cupboard since they contained the real names of the parents and children.

As specified earlier, the agreement was that all data collected had to be deleted after a duration of ten years after collection. We plan to contact the children who participated in the study and their parents after a duration of nine years in order to offer them a chance to view the stored data and request them for consent for a prolonged data storage. This might be beneficial since the children will be older and more capable of understanding the scope of this demand. Legally, many will be old enough to provide consent for themselves. However, we are aware that this might not be successful in all cases and could lead to a reduction in the number of language samples.

6. LIMITATIONS AND FUTURE SOLUTIONS

In Section 5, we presented the solutions that we used for data protection. To fulfil the requirements of the GDPR (2016/679), we went to great lengths to ensure that all the data collected remains protected and all additional data was separated and pseudonymised, as seen in Figures 2 and 3 (red lock symbols). Therefore, given these measures, our research design could be considered as resembling “locked” science, rather than open science. However, we would like to design studies that ensure both data protection and open science.

Our research should be accessible in order to foster collaborations with other research teams and citizens, both internationally and nationally. For some research questions, it would be very interesting to exchange language samples, but we cannot do this with not pseudonymised video data. Unlike spoken language samples (https://childes.talkbank.org), we cannot collect sign language samples in a database. This is a problem that we still need to find a solution for.

Solution strategies that can be realised with the currently available technological possibilities are notation systems, ELAN (2019) annotations, or shadow signing (Fries 2020, p. 222). There are a number of possibilities available for transcribing or notating sign languages, e.g., sign writing (https://www.signwriting.org) or the Hamburg Sign Language Notation System (https://www.sign-lang.uni-hamburg.de/dgs-korpus/index.php/hamnosys-97.html). However, many people find it difficult to read those notations and depending on the research question being considered not all necessary information may be provided in the notations.

Using the freely accessible ELAN (2019) software, it is possible to annotate videos in sign languages and compare data. Since annotations can focus on various aspects, a shared annotation file might not include specific aspect needed for another research question. In order to enable data sharing, the face of the children could be blurred with a grey field across the eye area, although this leads to the loss of important sign language information. Details on the missing facial information (e.g., eye brow position, direction of eye gaze, eye closure) could be provided in the annotations. However, the facial information deleted must be annotated by the research team who collected the data prior to exchange, which will be time consuming.

Another possibility could be shadow signing (Fries, 2020, p. 222). Shadow signing indicates that another person is filmed copying the communication of a signer. Shadow signer clips of the children´s narrations could be produced and thus pseudonymised, making the data exchangeable. It is questionable whether an adult signer could shadow sign for a child and whether shadow signing can be used not only on content level, but also on a linguistic level.

We could also use technically advanced solutions, such as avatar programs or sign recognition via motor analysing sensors. But we would need avatar programs that include manual aspects of sign languages, as well as facial expression and body posture, thus providing all the necessary linguistic information. Motor analysing sensor techniques require specialised equipment and can be used only in a lab setting. We found that filming the children for the purpose of the study could create an artificial/inauthentic setting that could influence the children´s narratives. Therefore, a special lab setting may influence the narrations produced by children even more since they are not familiar with such lab settings.

Hence, we continue to look for solutions to make sign language data more accessible while meeting the regulations of the European GDPR. We aim for our research not only to fulfil the requirements of the GDPR, but also those of open science:

We aim for our research to be transparent, as well as accessible, collaborative, and include citizen science practices.

Declaration of interest statement

The author declares no potential conflicts of interest.

Acknowledgements

I would like to thank all the children and their parents for their consent and interest in participating in this study. I would also like to thank the teachers and school administration, as well as associations and members of the Sign Language Community for their time and support during the NaKom DGS-Test sign language acquisition study.

REFERENCES

Deutsche Forschungsgemeinschaft. (2013). Vorschläge zur Sicherung guter wissenschaftlicher Praxis. Retrieved from http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf

ELAN. (2019). Computer software. Nijmegen, NE: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan

Enns, C., Zimmer, K., Broszeit, C., & Rabu, S. (2019). *American Sign Language: Expressive Skills Test*. Winnipeg: Northern Signs Research, Inc.

European Commission. (2017). Providing researchers with the skills and competencies they need to practise Open Science. Retrieved from https://cdn1.euraxess.org/sites/default/files/policy_library/ec-rtd_os_skills_report_final_complete_2207_1.pdf

European Commission. (2018). Open Science Policy Platform Recommendations. Retrieved from https://ec.europa.eu/research/openscience/pdf/integrated_advice_opspp_recommendations.pdf

Fries, S. (2020). *Gewalterfahrungen Gehörloser Frauen: Risikofaktoren, Ressourcen und Gesundheitliche Folgen*. Wiesbaden: Springer Fachmedien Wiesbaden GmbH.

GDPR. (2016/679). Retrieved from https://eur-lex.europa.eu/eli/reg/2016/679/oj

Graphic Node. (2019). Infographics Prime - Templates. Apple App Store. Retrieved from https://www.graphicnode.com

Harris, R., Holmes, H. M., & Mertens, D. M. (2009). Research Ethics in Sign Language Communities. *Sign Language Studies, 9*(2), 104. doi:10.1353/sls.0.0011

Herman, R., Grove, N., Holmes, S., Morgan, G., Sutherland, H., & Woll, B. (2004). *Assessing British Sign Language Development: BSL Production Test (Narrative Skills)*. London: City University.

Hill, J. C., Lillo-Martin, D., & Wood, S. K. (2018). *Sign Languages*.

IDRIX. (2017). VeraCrypt (Version 1.21). Retrieved from https://www.veracrypt.fr/en/Home.html

Mondschein, C., & Monda, C. (2019). The EU's General Data Protection Regulation (GDPR) in a Research Context. In P. Knubben, M. Dumontier, & A. Dekker (Eds.), *Fundamentals of Clinical Data Science* (pp. 55-71). doi:10.1007/978-3-319-99713-1_5

Pollard, R. Q. (2002). Ethical Conduct in Research Involving Deaf People. In V. Gutman (Ed.), *Ethics in mental health and deafness* (pp. 162-178). Washington, D.C: Gallaudet Univ. Press.

Ross, L. F., Loup, A., Nelson, R. M., Botkin, J. R., Kost, R., Smith, G. R., Jr., & Gehlert, S. (2010). Human subjects protections in community-engaged research: a research ethics framework. *Journal of empirical research on human research ethics, 5*(1), 5-17. doi:10.1525/jer.2010.5.1.5

Sandler, W., & Lillo-Martin, D. (2017). Sign Languages. In M. Aronoff & J. Rees-Miller (Eds.), *The Handbook of Linguistics* (pp. 371-396). Chichester, UK: John Wiley & Sons Ltd.

Sign Language Linguistic Society. (2016). Ethics Statement for Sign Language Research. Retrieved from https://slls.eu/slls-ethics-statement/

Singleton, J. L., Martin, A. J., & Morgan, G. (2015). Ethics, Deaf-Friendly Research, and Good Practice When Studying Sign Languages. In O. Eleni (Ed.), *Research methods in sign language studies* (pp. 7-20). Chichester: Wiley Blackwell.