The SIGNificant Chance Project and the Building of the First Hungarian Sign Language Corpus

Csilla Bartha, Margit Holecz, Szabolcs Varjasi
Research Centre for Multilingualism, Research Institute for Linguistics of the Hungarian Academy of Sciences, 1068 Budapest, Benczúr u. 33.
E-mail: {csillabarthadr, holeczmargit, varjasi.szabolcs}@gmail.com

Abstract

The Act CXXV of 2009 on Hungarian Sign Language and the Use of Hungarian Sign Language recognizes Hungarian Sign Language (HSL) as an independent natural language, moreover it provides the legal framework to introduce bilingual education (HSL-Hungarian) in 2017. In order to establish the linguistic background for bilingual education it was crucial to carry out linguistic research on HSL, which research should be sociolinguistically underpinned and should include corpus-based research. This research also aims to standardize HSL for educational purposes with the highest possible degree of community engagement.
During the SIGNificant Chance project a sign language corpus (approximately 1750 hours) was created. A nation-wide fieldwork was conducted (five regions, nine venues). 147 sociolinguistic interviews and 27 grammatical tests (with 54 participants) were recorded in multiple-camera settings. There were also Hungarian competency tests and narrative interviews conducted with selected participants in order to make the complex description of their different linguistic practices in different discursive contexts possible.
We are using ELAN and three different templates to analyze the collected data for different purposes (sociolinguistic-grammatical template, another for short term project purposes, and one for the dictionary). Some parts of the annotation work has been finished which contributed to the writing of the basic grammar of HSL and the creation of a small corpus-based dictionary of HSL.

Keywords: Hungarian Sign Language, sociolinguistics, SIGNificant Chance Project, corpus building, annotations

1. Introduction

According to some estimates there are 30-40.000, based on other data there are 60.000 D/deaf people living in Hungary which makes them the third largest linguistic minority in Hungary using the Hungarian Sign Language (further on: HSL) as their primary language (Bartha, 2004). The Act CXXV of 2009 on Hungarian Sign Language and the Use of Hungarian Sign Language was an important milestone for the D/deaf community (Bartha et al., 2016). On the one hand, because it recognized HSL as an independent natural language and on the other hand it provides the legal framework for introducing bilingual education in 2017. However, for bilingual education not only theoretical linguistic and applied linguistic researches are necessary but also the standardization of HSL for the purpose of education. We are aware that the notion of standardization (cf. WFD 2014) is quite problematic, however, we conceive standardization as a bottom-up, corpus-based process which is built on data acquired from sociolinguistic sampling. Standardization in our understanding presupposes the widest possible consensus with (and also the involvement of) the signing community and should recognize the diversity of the respective sign language(s). In our bottom-up approach we believe that the standardization process should be based on involving members of the Deaf community. For the linguistic foundation it is essential to have corpus-based, empirical researches applying current sociolinguistic approaches. This is the main goal of TÁMOP 5.4.6/B-13/1-2013-0001 project called Theoretical and practical steps of the standardization of Hungarian Sign Language (SIGNificant Chance).

2. The SIGNificant Chance Project

The corpus created during the project has the following aims: 1. sociolinguistic description of the patterns of language use in the Deaf community 2. analysis of HSL variation and creating an evidence- and corpus-based digital dictionary 3. to provide the grammatical description of the emic categories of HSL 4. measuring competencies of Deaf children in special and mainstream education. Besides providing basic statistic and linguistic (on all linguistic levels) analysis, it also serves educational purposes, like creating educational materials.
The project was materialized between 1st November 2013 and 31st October 2015 involving experts from several fields. Sociolinguists, theoretical linguists, psychologists, sociologist, lawyers and IT professionals helped us among others. Altogether 35 Deaf, hard of hearing and CODA people worked with our colleagues, supporting each other to reach the common goal, under the supervision of Csilla Bartha.

3. The Hungarian Sign Language corpus

3.1 Significance and constitution of the corpus

Recently there are more and more corpus-based sign language researches and descriptions, however, even on international level it is unique to study sign language phenomena using such a vast (more then 1700 hours) corpus which is extremely well documented by sociolinguistic metadata. The participants were selected by applying strict statistical criteria. While there are recent sociolinguistic studies on corpora (see eg. Schembri et al. 2013), in Hungary, there have only been sporadic sign language researches, which were usually based on a small

amount of samples and the analysis was done in a word processor.

In the SIGNificant Chance Project, informants of all the interviews were recorded in the same communicational situation, talking about the same topics; all this provides a unique analysis foundation compared not only to the sign language, but also to the spoken language researches. Since we have comprehensive metadata about the informants, the corpus provides a unique opportunity to conduct qualitative and quantitative analysis as it includes information about gender, age, location, language socialization and many other variable. During our research, each phenomena was analyzed with a bottom-up approach.

For methodological reasons it was necessary to create the sessions without any influence of a hearing person or the Hungarian Spoken Language, therefore no hearing person was allowed in the studio. The fieldworkers of the sociolinguistic interviews, as well as the moderators of the grammatical tests were D/deaf.

During the project it was important to include the Deaf Community: they did not only passively help our hearing colleagues as consultants, they played significant roles as professional supervisors, fieldworkers, annotators, etc.

Together with the Deaf Community, applied linguists, sociolinguists, sign linguists and experts of other fields formed a learning community, all this; spontaneously contributed to raising the prestige of sign language, awakening and increasing the interest towards sign language among the experts involved and the university students. The process of corpus planning and building created a Deaf and Hearing learning community in which our critical and reflexive approach manifested. Furthermore, the level of their metalinguistic and metapragmatic consciousness increased significantly.

The corpus is made of two big parts: sociolinguistic interviews (chapter 3.2) and grammatical tests (chapter 3.3) and the related recordings. During the project recordings were made of each training sessions, including trainings for the fieldworkers and the annotators, workshops and conferences. We used them for documentation and for the purpose of meta analysis. Thus we created a corpus containing a vast amount of material: 1750 hours of recording, which is 6,5 terabyte data.

3.2 Sociolinguistic interviews

In nine venues of 5 regions (Budapest, Szeged, Hódmezővásárhely, Békéscsaba, Debrecen, Kaposvár, Sopron, Győr, Vác) we recorded altogether 147 sociolinguistic interviews (67 metropolitan, 80 rural) which had an average length of 3-4 hours. There were 67 men and 75 women among the informants; the youngest was 21 and the oldest was 82 years old. A statistician helped us with compiling the questionnaire and to choose the right informants in order to acquire data that can be analysed in a quantitative manner. In case of 27 informants both parents, in case of 3 informants only the mother and in case of 1 informant only the father was deaf. The rest was born into a hearing family. During the selection of the participants, the variant of the hearing status of parents was considered to be significant. In order to draw subtle description of the socialization background of the participants, their social networks and the role of languages were mapped. During the fieldwork we recorded sign language users that had different language socialization patterns.

Our questionnaire consisted of 345 questions; we first got some comprehensive data from the results regarding the social situation, language socialization patterns at school and at home, monolingual and bilingual language use and attitude towards the Hungarian Sign Language and different educational programs of Deaf people. Two-third of the questions are closed questions. The rest are narrative, open questions about topics which were mentioned earlier as closed questions at another stage of the interview or topics the informant were keen to speak about. These methods helped in reducing the effect of observer's paradox. The thematic structure of the questionnaire builds from formal to more informal topics so we can also measure the accommodation (cf. Giles et al 1991) between the participants.

During each interview there were 2 fieldworkers and 1 informant. The Fieldworker No.1 followed the questions on a laptop while conducting the interview, thus there was no need of holding the printed version. The Fieldworker No. 2 documented the answers, this way the Fieldworker No. 1 could actively sign and pay attention to the informant. The participants were situated at the three sides of an imaginary square. The informant was on the right side, opposite to the Fieldworker No. 1; the Fieldworker No. 2 documenting the interviews sat a little bit further behind, between the two of them, thus creating a right angle to both participants. We used 3 cameras for recording the interviews: one for the overall picture, one only for the informant and one only for Fieldworker No. 1.

Figure 1: Sociolinguistic interview

Before the live interview we also conducted some pilot tests, which we evaluated. Our aim was not only to avoid technical problems but also to make it sure that the interview is conducted by the fieldworkers in accordance with the guidelines.

In order to prepare the fieldworkers as well as possible and

to get as real data about language use as possible, we organized several trainings for them during which we trained the 16 deaf fieldworkers using different exercises. It was necessary to have 'local' signers conducting the interviews for each venue. Thus made it possible for the informant to clearly understand the fieldworker. On the other hand, dialectal variability was an important analytical aspect in the corpus, and this way we could avoid distortion resulting from speech adaptation (here: signing adaptation). During the fieldwork we asked the fieldworkers to constantly provide feedback for us; we also checked the recordings and we implemented the experiences into the future interviews. Not only the fieldwork but also the whole project was a process of continuous learning for all of us.

For preparing the sociolinguistic interviews we created an online questionnaire; on one hand its purpose was to gain information with the help of the Internet about the most important language use habits of deaf and hard hearing people, on the other hand the data collected contributed significantly to the project as the feedbacks were exceptionally useful in the preparation of the sociolinguistic questionnaire(s). Earlier, there was no research to sum up demographic data (age, gender, professional, education, family relations etc), information about levels of hearing loss, language use habits, attitudes towards Hungarian Languages and Hungarian Sign Language etc. During the trainings – prior to interviews – for the fieldworkers we debated the questionnaire from conceptual, linguistic, wording point of view. The experiences resulting from creating the questionnaire and from the results of the questionnaire provided important research information for the future sociolinguistic field work as well. The questionnaire concluded maximum 66 questions, this number could be smaller depending on the given answers (eg.: question about the child's hearing was asked in case of those who answered to have a child). Each part of the questionnaire (instructions, questions, options etc.) was available in HSL and also in Hungarian. Altogether we received 238 answers, the informants were between the age of 15 and 74; 94 men and 144 women. Based on location there was a big diversity: to the question about the place of residency we received 80 different answers. We succeeded to address the audience based on gender, age, status of hearing etc. It is important to highlight that the online questionnaire was anonymous, it is impossible to identify the informant, therefore we had the chance to reach out to those who would have not answered some questions in real life.

3.2 Grammatical tests

Although previously there were some attempts to describe some parts of its grammar, the first comprehensive, scientific linguistic description of Hungarian Sign Language was created in the Framework of the SIGNificant Chance Project in 2015. It was essential – just like during the whole project– that the grammar should be a result of corpus-based studies using recordings that reflect real language use, it should be based on sign language and avoid applying notions commonly used in spoken languages. Grammar was created by a team consisting of D/deaf colleagues, theoretical linguists and sign linguists. The result of their work was the grammatical test which has eliciting tasks for processing the basic phenomena of sign language. Among others, the following phenomena were tested: WH questions, question words, word order, contrastive topic, quantifiers, negation, etc.

The grammatical test contains 21 exercises. The location of the grammatical fieldwork was Budapest, the informants were always native sign language users; during each session two informants and one deaf fieldworker were present. During the first phase of the grammatical testing we worked with informants from Budapest, 16 recordings were created on 5 cameras. (Besides the 3 cameras used by the sociolinguistic interviews, a bird-eye camera was also used by participant to record the signing from an overhead perspective). During the following phases of the testing we conducted the grammatical tests with rural informants as well. The following criteria played a significant role while choosing the informants: gender, age, education, school type, where they went to school. We conducted the grammatical tests only with such people who have participated in a sociolinguistic interview earlier; this way we received a more complex picture about the language use of each informant adding detailed information about the informants' language socialization and background. At the moment we finished the analysis of the tests recorded with informants from Budapest. Here we recorded 32 informants (15 women and 17 men) in 16 sessions; they were all D/deaf except for one CODA participant. There were four age groups (18-30, 31-45, 46-60 and 61+); it was also necessary to have a proportional distribution not only based on gender and age, but also based on education; and we wanted to have max. 20% professional sign language users (e.g.: sign language teacher). Based on the hearing status of the parents in case of 10 informants both parents were D/deaf and in case of 22 informants both parents were hearing. The whole size of the grammatical corpus is 30 and a half hours.

The fieldworkers during the grammatical testings were deaf people who knew the test well; since they worked on putting them together and since they have some experience in empirical and theoretical linguistics, they could conduct the elicitation exercises in the preferred way, without affecting the natural signing of the informants.

4. Analysis of the corpus

4.1 Preparation and organization of the workflow

We used the ELAN software for analyzing the corpus. On one hand, the results of other international sign language corpus projects proved that this software could help the research aims; on the other hand, looking at the IT competences about operational systems and programs of the researchers and the annotators working with the corpus and the infrastructural conditions, ELAN seemed to be the most appropriate choice.

While elaborating on the annotation methodology of the project, we reviewed the international projects and used

their experiences. We mainly focused on the related works of Johnston (2013) regarding Auslan, but we also gained information from the Dutch (Crasborn et al. 2015) and British (Cormier et al. 2015) practice, furthermore, Ritva Takkinen and her colleagues also supported us. In Hungary we should mention HuComTech's gesture research project where ELAN is used for multimodal analysis (e.g. Abuczki 2013). In Hungary there is no research like this in the field of sign language.

The main aim of the annotation during the SIGNificant Chance Project was to support the sociolinguistic and grammatical work, especially by providing a sample materials and by involving members of the deaf community, because the modern, scientific analysis of sign language is a new experience for most Deaf, therefore it strengthened the positive connection between our colleagues and the community.

Annotators were educated in formal and informal ways. Before the formal education we appointed colleagues with high level of sign language competence, who learnt to use ELAN efficiently in order to support and make the work of the annotators smooth. There were two formal workshops were future annotators learnt the most important aspects of grammatical annotation and the use of the software. In order to adjust to the language skills of the annotators both the trainings and the educational material were available in Hungarian Sign Language and in (written) Hungarian; furthermore we translated the basic ELAN functions into Hungarian.

During the organization we planned and allotted the subtasks based on the individual competencies of the annotators: we prepared a task description, an instruction and we provided constant online (sign language or e-mail) support for the annotators. A work log was written including their notes and impressions.

Most of the annotators did not work in full-time but remotely; however personal meetings were regular in which we documented the feedbacks, we corrected previous works and we coordinated the schedule of the additional work.

4.2 Translating the sociolinguistic interviews

Since several researchers worked on the project who did not sign at all or well enough; besides, our primary aim was not only a grammatical analysis but also the content wise and qualitative analysis of the sociolinguistic interviews, a translation of the interviews was needed.

Our aim was to prepare translations properly segmented in ELAN and to attach them in form of annotations to the videos. However, during the 2 years of the project (including the development of the infrastructure, analysis and preparation of the interviews, development of the dictionary framework etc.) we did not have the chance to fulfill this plan with the limited number of translators, therefore we asked them to insert the translations into a Word chart.

Since we insisted on having CODA or interpreters respected by the members of the deaf community to do the translations to keep the data authentic and accurate, only a small number of translators could work on the recordings during the project, and most of them were from rural areas and they already suffered from a work-overload.

In the future, we plan to integrate these translations into ELAN and to check them.

4.3 Creating the tier structure

Parallel to the fieldworks 3 ELAN templates (sociolinguistic, grammatical and lexical) were created. While establishing these we leaned on the Australian annotation guidelines (Johnston 2013).

We set up 140 tiers for each person (informant and fieldworkers). Besides translation, each linguistic level is represented among them, from phonetics to pragmatics.

The complex tier-structure is a result of constant cooperation of applied linguists, sign linguists and deaf colleagues.

We created controlled vocabularies for certain linguistic types. We defined elements needed to describe handedness, movement (its type, direction and micro movements) and non-manual elements (mouth, eyes, eyebrows, look etc.).

We defined the possible elements based on the results of previous sign language researches and other, non-linguistic, but relevant researches (e.g: emotion and gesture analysis), integrating the feedbacks of the domestic deaf community.

The annotation works started with the lexical (see chapter 5.1) and grammatical (see chapter 5.2) researches; we give a detailed description of the used templates in the relevant chapters. From the conducted sociolinguistic interviews we synchronized 87 and we started the annotation of 76. We started to analyze 15 out of the 16 synchronized grammatical tests.

4.4 Annotation of the corpus

4.4.1 Annotation of the sociolinguistic interviews

Until March 2016, 41 sociolinguistic interviews were fully translated, this means approx. 2500 pages altogether. It will be a very complex task to transfer (and segment) all of them into ELAN and so far we have only been able to check and transfer 5 interviews; we plan to finish the rest in the future.

4.4.2 Processing the grammatical tests

While describing the Hungarian Sign Language grammar, our theoretical linguistic colleagues relied on the sample material created during the so called *focused grammatical research* (annotation), thus while planning the annotation of these parts the main aim was to support the creation of the grammar.

For the annotation of the analyzed phenomena, first the task-based segmentation of full records was needed. This was done by deaf colleagues – who knew the test well enough – on the *interview section* and the *important interview section* tiers; in this case we also used controlled vocabularies for annotation. The next step was the annotation and the segmentation of the elicited phenomena.

The segments of the analyzed phenomena reflect in most cases the borders between sign language utterances, however, we did not systematically checked them so far.

Despite the fact, that due to the data-driven approach we used the right hand - left hand differentiation during the annotation (this way, handedness can be defined only from the data), taking the aims of the project into consideration we used passive and active hand tiers during the *focused grammatical research*. In the future, we plan to transfer and check the created annotations to the tier of right and left hand.

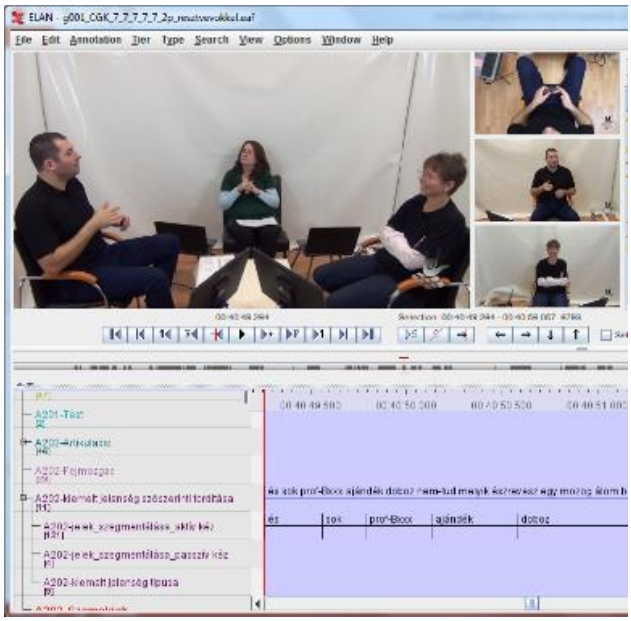

Figure 2: Annotation of the grammatical tests

At the present phase of the analysis we have only worked with actual meanings, however, later we would like to create an ID-gloss database, partly based on the meanings and partly based on the corpus-based online dictionary – to be described in chapter 5.2 – and the previously written Hungarian Sign Language dictionary. Basic grammar mainly focuses on sign language syntax, therefore we did not annotated the grammatical classification of the signs. The reason is that due to the lack of results from proper basic research of Hungarian Sign Language, we wanted to avoid the use of spoken language categories.

Depending on the type of the phenomena we annotated using non-manual components, where we also worked with controlled vocabularies. The annotation and the segmentation was made based on the following tiers: *body movement, head movement, eye-gaze, eye, eyebrow, cheek, chin, articulation, mouth* and *other-non manual*.

As the result of the annotation of the grammatical research, our colleagues created 15363 annotations. They are all approved and checked. Furthermore, during the grammatical annotation of the interviews we segmented each signs occurring in the interval of the *important interview sections*, therefore we created 34440 segments and their filling is going to be an important task in the future. This will serve as the secondary sign material of the ID-gloss database.

4.4.3 Annotation of the dictionary

From the 5 regions we analyzed 6 sociolinguistic interviews from each (altogether 30). Our aim was to have a ratio of 50-50% for women and men in each region, as well as in case of old and young. This criteria was not fulfilled in one region because there we had only a smaller number of interviews thus the proportional selection was not possible.

We annotated 209 pre-defined expressions (their occurrences in the interviews) that are essential in everyday life. Knowing the interview questions it was almost certain that they will be used (e.g. mother, father, and language). We created a separate template for the annotation of the dictionary. The sociolinguistic template would have also been appropriate to gain the needed information for the dictionary, but due to the shortness of the project, and in order to make the annotators' job faster and easier, we used a revised and simplified template. After the segmentation we annotated the following levels: 1. Hungarian translation equivalent. 2. Type of the sign (one handed, two-handed, mirror-symmetrical etc.) 3. Dominant hand, handshape 4. Non-dominant hand, handshape 5. Region, location of signing 6. Type of movement. The above mentioned linguistic information was completed with the code of the informant (enabling us to track other metadata later) and the city.

We did not use controlled vocabularies but we created a virtual keyboard similar to the keyboards on mobile phones and we depicted some elements (handshape, type of movement etc.) by pictograms. By pushing a particular button for a long time, options, such as possible elements of the signing location appeared. We used this method so that the deaf annotators could analyze the material faster, and the pictograms used helped them to rely on their visual competences and not on written language input.

5. Results of the project

5.1 Research on sociology of sign languages and sociolinguistic studies

From the results of the research we first got some comprehensive data about the social status of the Deaf, about language socialization patterns at home and at school, about monolingual and bilingual language use, about their attitude towards the Hungarian language and different educational programs etc. Besides, the recordings also made the corpus-based analysis (qualitative and quantitative) on different levels of sign language use possible. Moreover further researches can be conducted exploiting the database of sociolinguistic metadata.

5.2 Corpus-based dictionary

During the project we created the beta version of the corpus-base dictionary reflecting the dialectal diversity of HSL. We can search with the help of all the annotated characteristics (handshape, location, type of sign, type of movement, direction of movement, sign language – regional – variety); therefore search is not only possible from a spoken language perspective (keyword, topic, grammatical category, first letter) but also from the sign language perspective. Search in sign language was made

easier with the pictogram-based search criteria. There is a definition for each sign and we can search for the English equivalent as well (English translation of all the signs has not been finished yet.).

5.3 Sign language grammar

The first comprehensive, scientific linguistic description of Hungarian Sign Language was created in the framework of the SIGNificant Chance Project in 2015; its script is accessible in the Research Institute for Linguistics of the Hungarian Academy of Sciences. It is outstanding in the sense that it is based on results of analysis of corpora, that reflect real language use, is sign language based it avoids applying notions commonly used in spoken languages to HSL.

5.4 Educational use

Bilingual education is a long-term objective, for which a lot of research is still needed. Since the main aim of SIGNificant Chance Project is to conduct a research that is essential for the establishment of bilingual education, it was very important for us to be able to use it in the field of education. The Hungarian Sign Language corpus by itself can be used as an educational material: it provides an access to authentic texts signed by native users of sign language. Furthermore, ELAN makes possible to subtitle videos faster, example sentences and helping materials can be exported which can be useful in deaf education, education of interpreters and in sign language courses. The corpus can be used as a source at courses focusing on the analysis of grammar phenomena in sign language (see Mesch-Wallin 2008).

6. Further tasks

By having a tier structure, we already have a framework for analyzing most levels of sign language, however, these should be revised from time to time. The next step – based on the pragmatic and discourse analysis experience of Research Center for Multilingualism – will be to work out the tier structure of sign language discourse and pragmatic researches. (The recent structure already includes a rudimentary version).

The creation of an ID-gloss database for corpus analysis (Johnston 2010) is also among our future aims, similarly to the integration of metadata into the corpus. We should regularly use a version-tracking software for documenting the annotations.

It is necessary to deepen the annotation of the corpus, to conduct more corpus-based researches in all levels of the sign language, because researches prior to the SIGNificant Change Project used a non-corpus based approach. Another aim of analyzing the sociolinguistic interviews is to make sign language accessible for those learning the sign language or learning in sign language – regardless of hearing status. In order to have an accurate description of sociolinguistic, dialectal and other levels of Hungarian Sign Language and about the lives of Deaf people, their experiences and language use it is necessary to further annotate the corpus and to publish the materials based on the results for the Deaf community, Sign Language Institutions, for hearing parents of deaf children and for those interested.

7. Acknowledgements

The study could not be realized without the support of SIGNificant Change Project (Támop 5.4.6/B-13/1-2013-0001) and the Methodology Project of the Hungarian Academy of Sciences (SZ-050/2014). We would like to thank the creators of the SIGNificant Project, our deaf and hearing colleagues.

8. Bibliographical References

Abuczki, Á. (2013). A *mondjuk* nem konceptuális használatának vizsgálata multimodális kontextusban. In: Váradi, T. (ed.): Alknyelvdok7. MTA Nyelvtudományi Intézet. Budapest. 3-17.

Bartha, Cs. (2004). Siket közösség, kétnyelvűség és a siket gyermekek kétnyelvű oktatásá-nak lehetőségei. In: Ladányi M., Dér Cs., Hattyár H. (eds.): „…még onnét is eljutni túlra…”. Nyelvészeti és irodalmi tanulmányok Horváth Katalin tiszteletére. Tinta Könyvkiadó, Budapest 313–332

Bartha, Cs., Varjasi, Sz., Holecz, M. (2016). A magyar jelnyelvi korpusz létrehozásának és annotálásának kihívása. In: Tanács, A. Varga, V. Vincze V. (eds.) XII. Magyar Számítógépes Nyelvészeti Konferencia. Szegedi Tudományegyetem, TTIK, Informatikai Intézet, Szeged. 207–219.

Cormier, K., Fenlon, J., Gulamani, S., Smith, S. (2015). BSL Corpus Annotation Conventions (2015) http://www.bslcorpusproject.org/wp-content/uploads/BSLCorpus_AnnotationConventions_v2_-Feb2015.pdf

Crasborn, O., Bank, R., Zwitserlood, I., van der Kooij, E., de Meijer, A., Sáfár, A.: Annotation Conventions for the Corpus NGT. (2015) http://www.bslcorpusproject.org/wp-content/uploads/CorpusNGT_AnnotationConventions_v3_Feb2015.pdf

Giles, Howard – Coupland, Justine – Coupland, Nikolas (eds.) 1991. Contexts of Accomodation. Cambridge University Press. Cambridge–New York–Melbourne.

Johnston, T. (2010). From archive to corpus: transcription and annotation in the creation of signed language corpora. International Journal of Corpus Linguistics, 15(1), 104-129.

Johnston, T. (2013). Auslan Corpus Annotation Guidelines. Online: http://media.auslan.org.au/attachments/AuslanCorpusAnnotationGuidelines_Johnston.pdf

Mesch, J., Wallin, L. (2008). Use of sign language materials in teaching. Online: http://www.ling.su.se/polopoly_fs/1.68035.1325063806!/menu/standard/file/MeschWallin_LREC2008.pdf

WFD (2014). Statement on Standardized Sign Language. http://wfdeaf.org/wp-content/uploads/2014/05/WFD-Statement-on-Standardization-of-Sign-Languages-FINAL-20-MAY-2014.pdf