Collecting an American Sign Language Corpus through the Participation of Native Signers

Pengfei Lu[1] and Matt Huenerfauth[2]

[1] The City University of New York (CUNY)
Computer Science Doctoral Program, Graduate Center
365 Fifth Ave, New York, NY 10016, USA
[2] The City University of New York (CUNY)
Computer Science Department, Queens College
65-30 Kissena Blvd, Flushing, NY 11367, USA
pengfei.lu@qc.cuny.edu, matt@cs.qc.cuny.edu

Abstract. Animations of American Sign Language (ASL) can make more information, websites, and services accessible for the significant number of deaf people in the United States with lower levels of written language literacy – ultimately leading to fuller social inclusion for these users. We are collecting and analyzing an ASL motion-capture corpus of multi-sentential discourse to seek computational models of various aspects of ASL linguistics to enable us to produce more accurate and understandable ASL animations. In this paper, we will describe our motion-capture studio configuration, our data collection procedure, and the linguistic annotations being added by our research team of native ASL signers. This paper will identify the most effective prompts we have developed for collecting non-scripted ASL passages in which signers use particular linguistic constructions that we wish to study. This paper also describes the educational outreach and social inclusion aspects of our project – the participation of many deaf participants, researchers, and students.

Keywords: American Sign Language, animation, accessibility technology for people who are deaf, data collection, social inclusion, motion capture.

1 Introduction

American Sign Language (ASL) is the primary means of communication for one-half million deaf people in the U.S. [1], and ASL and English have distinct vocabulary and word-order. Due to educational and language exposure reasons, many deaf adults in the U.S. have relatively low levels of written English literacy [2], making it difficult to read English text on websites or other sources. Animations of American Sign Language (ASL) make information and services accessible for these individuals.

We are studying how to create high-quality ASL animations by analyzing recordings of people. We are in the middle of a 5-year study to collect, annotate, and analyze an ASL motion-capture corpus of multi-sentential discourse. Our goal is to create an annotated collection of ASL movement data using video camcorders and

C. Stephanidis (Ed.): Universal Access in HCI, Part IV, HCII 2011, LNCS 6768, pp. 81–90, 2011.

motion capture equipment, and we are seeking computational models of various aspects of ASL linguistics. These models will help us to synthesize more understandable ASL animations, enabling their use in applications for deaf individuals with low literacy. For our research, we want the stories we collect to contain certain linguistic phenomena but not others. This paper explores different ways we have prompted people to perform ASL stories in order to optimize this.

Section 2 describes the corpus collection goals of our study and other related corpus building projects. Section 3 summarizes our previous work on eliciting and collecting our corpus using alternative prompting strategies and the main results. Section 4 presents our corpora collection, annotation procedure, and the experiments in the second year. Section 5 contains the results, conclusions, and future research plans. Section 6 discusses the participation of people who are deaf in this research.

2 Our Corpus Collection Goals

For our corpus, we record native ASL signers performing short stories or passages in our laboratory while being videotaped and while wearing motion-capture gloves, an eye-tracker, a head-tracker, and a set of sensors on a special bodysuit. This equipment records the signer's handshape; hand location; palm orientation; eye-gaze vector; and joint angles for the wrists, elbows, shoulders, clavicle, neck, and waist. Three video cameras record front, side, and face-close-up views. Details of our studio configuration appear in [3]. Another native signer (called the "prompter") sits behind the front-view camera to converse with the performer being recorded. Signers tend to perform more natural ASL performances in this type of conversational setting [4]. Because we want to record natural ASL performances (to use as a basis for our research), we do not pre-script the passages to be performed; however, it is necessary to give the performer prompts to encourage them to sign about a particular topic for some period of time. As discussed below, we also want to encourage the performer to use particular linguistic constructions (and not others) in their ASL signing (without giving them a specific script to perform). Therefore, we have experimented with various forms of prompting to elicit ASL passages that are optimally suited to our research needs. Identifying a set of such prompts is the focus of this paper.

After we record a human signer performing a multi-sentence passage in ASL, then our team of ASL linguistics experts watch the recording and create a timeline of the signs performed. The experts also analyze the performance to note various linguistic constructions and other details about the performance, and this information is also added to the timeline for each story. Details are described in [4]. This linguistic annotation of each story facilitates our later analysis and use of the data in the corpus.

Several researchers have collected video-based corpora of sign language, e.g. [5][6][7], or short sign language recordings via motion-capture, e.g. [8][9]. However, our project is the first to record a large corpus of sign language passages while using motion-capture equipment. Previous researchers have also designed schemes for annotating the referential use of signing space [10] on a timeline, but our project is the first to analyze this linguistic use of signing space in a motion-capture corpus.

2.1 Spatial Reference Points in Our ASL Corpus

Signers associate people, things, and concepts under discussion with 3D locations around them in space [5][11][12]. The typical way in which a signer establishes a spatial reference point (SRP) is by pointing to a location around them in space immediately before or after mentioning someone or something for the first time. Signers may set up several of these SRPs around them in space during a conversation. After being established, these points are later used by the signer in the following ways:

- The signer will point at the location to refer to the person or thing.
- Some signs change how they are performed to aim at these SRPs (e.g., the motion path of some verb signs goes from their subject toward their object).
- Signers may aim their eyes or head at these SRPs for linguistic reasons.

SRPs are frequently used and are essential to *human* ASL performances, and they are also important for producing good quality ASL *animations*. Huenerfauth [13] found that native signers' comprehension of ASL animations improved when the animations included association of entities with locations in the space and the use of verbs whose motion paths were modified based on these locations. Our research focuses on adding these capabilities to ASL animation synthesis technologies; we believe that mathematical functions of verbs' motion paths can be induced from 3D motion data we are collecting from human signers performing those verbs. We are also analyzing the ASL passages we are collecting to learn when to associate entities with 3D locations, where to place them, and how these locations affect sign movements. The models of ASL spatial use we learn will be embedded into ASL animation software to produce more natural looking and understandable animations. Therefore, it is desirable that the stories and passages we collect in our corpus contain many examples of signers setting up SRPs and using them in a story or passage.

Fig. 1 shows a sample of some of the linguistic information for one story in our corpus; it is a timeline of an ASL passage discussing a girl using her computer. For our project, we add various linguistic information to a timeline corresponding to the video and motion-capture recording of each story collected. Fig. 1 shows only a subset of that information: the sequence of signs, the establishment of SRPs (indicated by a line on the "SRP#1 Establishment" row of the timeline), and the references made to those SRPs during the passage (indicated on the "SRP#1 Reference" row of the timeline). In this case, the first time that the signer points to a location in 3D space around his body, he establishes an SRP at that location to represent the girl being discussed; this SRP is referred to again later in the passage when the signer performs another "POINT" sign. A loose translation of the passage in Fig. 1 would be: "Wow. There was this 12-year-old girl, and she was on the Internet typing…"

Fig. 1. Example of a timeline from a story from our corpus that contains an SRP

2.2 Classifier Predicates in Our ASL Corpus

Classifier predicates (CPs) are a linguistic construction in ASL that also uses the space around the signer's body – but in a different way than SRPs. CPs are complex signs in which the signer creates movement for the hands (or sometimes the body) to indicate the spatial arrangement, size, shape, or movement of people/objects in a 3D scene being described [14]. During CPs, entities under discussion are associated with locations in space around the signer, but unlike SRPs, during CPs, the arrangement reflects a real-world 3D configuration or arrangement of objects. CPs are not our current research focus, and because they lead signers to use space around their bodies in a different way than SRPs, we don't want to record stories that contain a lot of CPs in our corpus. Building this corpus (recording people performing ASL and then linguistically analyzing the recordings) is very time-consuming, we therefore want to optimize the stories that we collect so that they contain primarily SRPs and not CPs.

2.3 How We Evaluate Our Prompting Strategies

Research projects collecting video recordings of sign language for linguistic study have used scripts or various prompting strategies to encourage signers to perform stories or sentences that contain specific linguistic phenomena of interest to the researchers [5][6][7][15]. We have adopted some of these prompts for our project and have invented others; sections 3 and 4 discuss how we evaluate the success of our prompting strategies. For our research, an ideal ASL passage to be collected would:

- *Be long enough to allow for establishment of SRPs.* If a story is too short, then the signer might not set up many SRPs or refer to them in the story. So, we will count the length of the stories we collect – as measured in seconds of time or in the total number of signs performed. By measuring the length of the stories collected using each of our prompting strategies (details in sections 3 and 4), we will be able to determine which prompting strategy is most effective.
- *Contain several SRPs established by the signer.* Collecting stories in which signers establish large numbers of SRPs around them in space can sometimes be difficult; so, we will count the number of SRPs established during each story we collect to measure the effectiveness of our different prompting strategies.
- *Contain many pointing signs or verbs that refer to SRPs.* With many examples of these spatial references (SRs), we will be able to study diverse forms of spatial use and reference in ASL signing. [1] Some linguists regard a signer pointing to himself (to say "I" or "me") as a form of spatial reference; so, we'll also count the number of these "first person references" in our data as a separate total.
- *Contain as few CPs as possible.* So, we will count the number of CPs that occur during the stories we collect; unlike the other items we discussed above, we would prefer to see a small number of CPs in the stories collected in the corpus.

[1] When a signer points to a location in space the first time, this establishes an SRP for some entity under discussion; we refer to this as an *SRP establishment*. When the signer points to the SRP again later, then this is a *spatial reference* to a previously established SRP. We also count the first pointing gesture that established the SRP as a spatial reference.

3 Summary of Our Previous Study: Year 1 of Our Project

During the first year of our project, we recorded and annotated 58 ASL passages from 6 signers (approximately 40 minutes of data). For this data collection, the prompter behind the camera used 9 different prompting strategies to elicit ASL passages:

- Tell a story: Invent a story using this topic: "If I had a genie that could grant three wishes, I'd…"
- Children's book: Read this short children's book, and then explain the story as you remember it.
- Repeat Conversation: Watch this 3-minute video of an ASL conversation or of a captioned English conversation, and then explain what you saw.
- Wikipedia Article: Read this 300-word Wikipedia article on "The History of Racial Segregation in the United States," and now explain/recount the article.
- Recount Movie/Book: Tell me about your favorite movie or your favorite book.
- Compare (not people): Compare two things: e.g. Mac vs. PC, Democrats vs. Republicans, high school vs. college, Gallaudet University vs. NTID, travelling by plane vs. travelling by car, etc.
- Personal Intro/Info: Introduce yourself, describe some of your background, hobbies, family, education, etc.
- Hypothetical Scenario: What would you do if: You were raising a deaf child? You could have dinner with any two famous or historical figures?
- Compare (people): Compare two people you know: your parents, some friends, family members, etc.

In a prior paper [4], we presented the characteristics of the stories that signers performed in response to each of these different prompts. Our results from analyzing the stories collected during year 1 are summarized here: The "tell a story," "children's book," and "repeat conversation" prompts elicited ASL stories with high CP/SR ratios (undesirable). These prompts that related to spatially/visually descriptive topics led to many CPs performed by signers. The "wikipedia article" and "recount movie/book" prompts yielded long story lengths, high number of SRPs established, and modest CP/SR ratios (desirable). While they elicited shorter passages, the "compare" and "personal intro/info" prompts also yielded stories with low CP/SR ratios (desirable). The "personal intro/info," "tell a story," and "hypothetical scenario" prompts led to many first-person references; the signer often discussed himself in these passages.

4 Our New Study: Year 2 of Our Project

The analysis of the different prompting strategies in year 1 of our project (discussed above) guided our data collection procedure in year 2. This section presents a new study we conducted to analyze our revised set of prompts used during this second year of the project. Specifically, we stopped using the "tell a story," "children's book," and "repeat conversation" prompts from year 1, and we continued to use the "wikipedia article," "recount movie/book," "compare," and "personal intro/info." Each of these is explained in Table 1. In an effort to encourage signers to tell even

longer stories, use more SRPs, and use fewer CPs, we tried several new prompting strategies during year 2 (which are also listed in Table 1). As we gained additional experience at recording signers using motion-capture equipment and analyzing stories, we were able to collect a larger set of stories in this second year of the project. The number of stories collecting using each type of prompt is also listed in Table 1. One of the new prompting approaches ("photo page") involved showing a page of images to a participant to encourage him to tell a story; an example of a page of images similar to those used in our study is shown in Fig. 2. Typically, the photographs included popular celebrities, athletes, or politicians who were currently in the news.

Table 1. Types of prompts used in Year 2 and number of stories of each type collected (N)

Type of Prompt	N	Description of This Prompting Strategy	Year This Was Used
News Story	12	Please read this brief news article (about a funny or memorable occurrence) and recount the article.	Year 2
Opinion / Explain Topic	5	Please explain your opinion on this topic (given) or explain the concept as if you were teaching it to someone.	Year 2
Compare (not people)	19	Compare two things: e.g. Mac vs. PC, Democrats vs. Republicans, high school vs. college, Gallaudet University vs. NTID, travelling by plane vs. travelling by car, etc.	Year 1 & 2
Compare (people)	2	Compare two people you know: your parents, some friends, family members, etc.	Year 1 & 2
Personal Intro/Info	8	Introduce yourself, describe some of your background, hobbies, family, education, etc.	Year 1 & 2
Personal Narrative	3	Please tell a story about an experience that you had personally.	Year 2
Photo Page	5	Look at this page of photos (of people who are in the news recently) and then explain what is going on with them.	Year 2
Recount Movie Book	9	Recall a book you're read recently or a movie you saw, and then explain the story as you remember it.	Year 2
Wikipedia Article	3	Read a brief Wikipedia article on some topic and then explain/recount the information from the article.	Year 1 & 2

Fig. 2. Example of what the page of photos looked like for the "photo page" prompts

In the second year of our study, we recorded and annotated 66 ASL passages from 3 signers (approximately 75 minutes of data). As was done in year 1 of the project, we used a set of prompts to elicit unscripted multi-sentential single-signer passages. The prompts used in year 2 are listed in Table 1. As was done in year 1 of the project, our team of ASL linguistics experts analyzed the stores collected to produce a timeline of each performance that includes the sequence of signs, the establishment of SRPs, the references to SRPs, the use of CPs, and other linguistic phenomena of interest to our research. In order to evaluate the set of prompts used during year 2 of our project, we calculated the average passage length (measured in the number of signs performed or the number of seconds); the results are shown in Fig. 3. We would prefer longer stories in our corpus because this increases the opportunity for the signer to establish several SRPs and to refer to them again. Further, we have found it easy to record very short stories from signers during our recording sessions; finding prompts that encourage a signer to perform a longer multi-sentence passage are therefore valuable to identify. Note: Error bars in Figures 3, 4, 5, and 6 indicate the standard error of the mean for each value.

Fig. 4 displays the average number of classifier predicates per second and the average number of spatial references per second in each type of ASL story. As discussed in section 2, we would prefer stories in our corpus with a low number of CPs and a high number of SRPs. The "opinion explain topic," "news story," and "compare people" prompts all led to high CP/SR ratios and long story lengths.

Fig. 5 displays the average number of the spatial reference points established per story – for each prompting strategy. For example, if a signer sets up three different points in space around their body to represent three entities under discussion in a passage, then we would say that such a story would have 3 SRPs *established*. If the signer continues to refer to these entities multiple times throughout the story (i.e., pointing to these locations in space again and again during the story), then the number of *spatial references* in the story would be much higher.

Fig. 6 displays the average number of 1st-person references per second in the stories collected using each prompting strategy. As discussed in section 2, we present the results for first-person references separately because some ASL linguists would consider this a form of spatial reference and some would not. We would say that a first-person reference occurs whenever the signer is pointing to himself. The "personal narrative" and "personal" prompts led to a high number of 1st-person references; this is not surprising since these prompts led signers to discuss themselves.

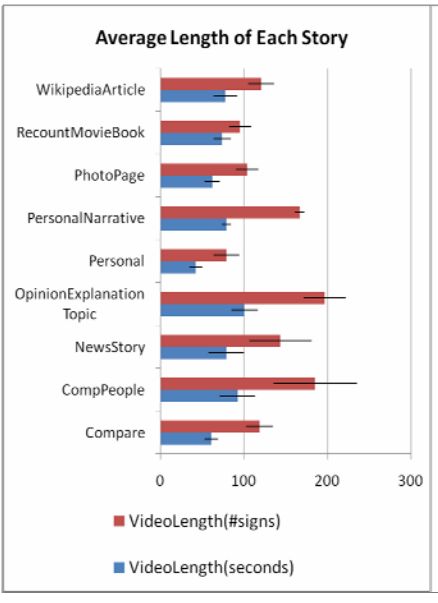

Fig. 3. Length of the ASL stories collected

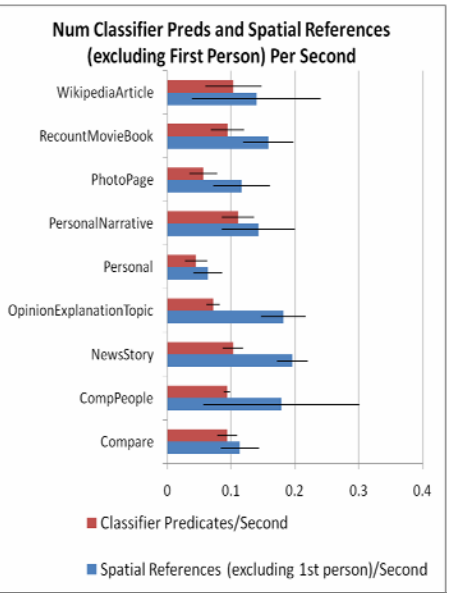

Fig. 4. Number of classifier predicates and spatial references per second in each story

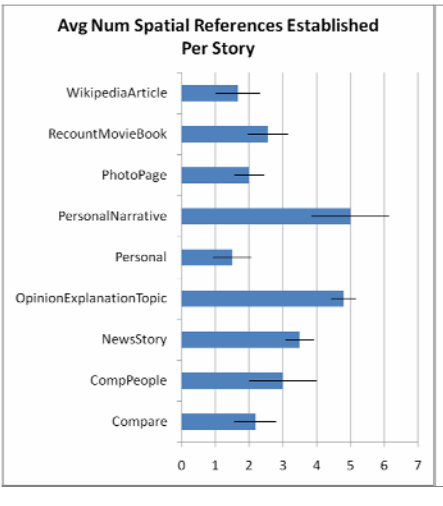

Fig. 5. Number of spatial reference points established in each story

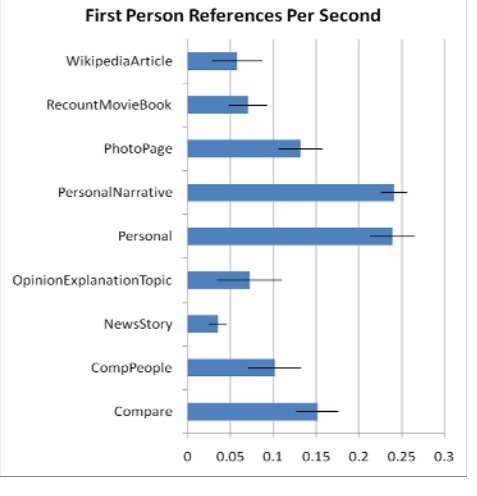

Fig. 6. Number of first-person references in each ASL story

5 Discussion and Future Work

We were pleased with the results of the "opinion/explanation topic," "compare people," "news story," and "recount movie/book" prompts, which led to good story

lengths and high SR/CP ratios. The "opinion/explanation topic" and "news story" prompts led to large numbers of SRPs established in the signing space. In year 3 of the project, we intend to use a larger proportion of these prompts during our data collection. While the results of our study have immediate benefits for our own research project, other linguistics and computer science researchers who are conducting sign language data collection will benefit from our comparison of various forms of prompting. This research therefore has benefits for sign language animation synthesis and sign language recognition research, which has accessibility benefits for people who are deaf. We are not aware of other systematic analysis of the benefits of various forms of linguistic prompting used in a sign language data collection study.

As we gather and analyze our corpus of ASL performances, we are beginning to analyze this data to study how human ASL signers set up SRPs in space, how they perform ASL verbs whose motion paths change based on SRP arrangements, and other details of sign language performance (e.g., the timing and speed of signs). We intend to incorporate these findings into our ASL animation synthesis technology to produce more natural and understandable animations of ASL [3][4][13].

6 Inclusion of Deaf Participants, Students, and Researchers

As discussed in section 1, ASL animation technology has the potential to make more information accessible to people who are deaf that have lower levels of English literacy – ultimately leading to fuller social inclusion of these individuals. Our research project also has a more immediate impact in this regard due to the extensive participation of people who are deaf in the various stages of the research process. Based on New York City, we advertise through Deaf community websites and other local resources to identify participants to come to our lab to be recorded for the corpus and to participate in experiments evaluating the quality of ASL animations.

Several deaf and signing researchers also participate in our project (including deaf high school and undergraduate students). Each summer, two or three students from local deaf schools or mainstream programs in the greater New York City area participate in three-month research experiences at our lab. In addition, we have also hosted a deaf undergraduate student visiting from Gallaudet University and another undergraduate student majoring in linguistics with excellent fluency in ASL for summer research experiences at the lab. Further, a graduate-level deaf research assistant also coordinates the projects at the lab throughout the year. Further details of the participation of deaf students and researchers in our project are described in [16]. Our goal is for these students to gain knowledge and practical experience working in a research laboratory and insight into the process of applying for and succeeding in future study and careers in scientific research – ultimately leading to fuller inclusion of people with disabilities in the fields of computer science and accessibility.

Acknowledgments. This research was supported by the National Science Foundation (award #0746556), Siemens (Go PLM Academic Grant), and Visage Technologies AB (academic software license). W. Clarke, K. Gallagher, A. Krieger, J. Lamberton, A. Pagan, J. Penzellna, and M. Turtletaub assisted with data collection/annotation.

References

1. Mitchell, R., Young, T., Bachleda, B., Karchmer, M.: How Many People Use ASL in the United States? Sign Language Studies 6(3), 306–335 (2006)
2. Traxler, C.: The Stanford achievement test, ninth edition: national norming and performance standards for deaf and hard-of-hearing students. J. Deaf Studies & Deaf Education 5(4), 337–348 (2000)
3. Lu, P., Huenerfauth, M.: Collecting a Motion-Capture Corpus of American Sign Language for Data-Driven Generation Research. In: Workshop on Speech and Language Processing for Assistive Technologies (SLPAT), Human Language Technologies / North American Association for Computational Linguistics Conference (HLT-NAACL), pp. 89–97 (2010)
4. Huenerfauth, M., Lu, P.: Eliciting spatial reference for a motion-capture corpus of American Sign Language discourse. In: Workshop on the Representation and Processing of Signed Languages, LREC, pp. 121–124 (2010)
5. Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., Lee, R.: The syntax of ASL: functional categories and hierarchical structure. MIT Press, Cambridge (2000)
6. Bungeroth, J., Stein, D., Dreuw, P., Zahedi, M., Ney, H.: A German sign language corpus of the domain weather report. In: Vettori, C. (ed.) Workshop on the Representation and Processing of Sign Languages, LREC, pp. 2000–2003 (2006)
7. Efthimiou, E., Fotinea, S.E.: GSLC: Creation and Annotation of a Greek Sign Language Corpus for HCI. In: Stephanidis, C. (ed.) HCI 2007. LNCS, vol. 4554, pp. 657–666. Springer, Heidelberg (2007)
8. Brashear, H., Starner, T., Lukowicz, P., Junker, H.: Using multiple sensors for mobile sign language recognition. In: IEEE International Symposium on Wearable Computers, p. 45. IEEE Press, New York (2003)
9. Cox, S., Lincoln, M., Tryggvason, J., Nakisa, M., Wells, M., Tutt, M., Abbott, S.: Tessa, a system to aid communication with deaf people. In: ACM SIGACCESS Conference on Computers and Accessibility, pp. 205–212. ACM Press, New York (2002)
10. Lenseigne, B., Dalle, P.: A Tool for Sign Language Analysis through Signing Space Representation. In: Sign Language Linguistics and Application of Information Technology to Sign Languages Conference, Milan, Italy (2005)
11. Liddell, S.: Grammar Gesture and Meaning in American Sign Language. Cambridge University Press, Cambridge (2003)
12. Meier, R.: Person Deixis in American Sign Language. In: Fischer, S., Siple, P. (eds.) Theoretical Issues in Sign Language Research. Linguistics, vol. 1, pp. 175–190. University of Chicago Press, Chicago (1990)
13. Huenerfauth, M.: Improving spatial reference in American Sign Language animation through data collection from native ASL signers. In: Stephanidis, C. (ed.) UAHCI 2009. LNCS, vol. 5616, pp. 530–539. Springer, Heidelberg (2009)
14. Supalla, T.: Morphology of Verbs of Motion and Location. In: Caccamise, F., Hicks, D. (eds.) 2nd Nat'l Symposium on Sign Language Research and Teaching, pp. 27–45 (1978)
15. Nishio, R., Hong, S., Konig, S., Konrad, R., Langer, G., Hanke, T., Rathmann, C.: Elicitation methods in the DGS (German Sign Language) corpus project. In: Workshop on the Representation and Processing of Signed Languages, LREC, pp. 178–185 (2010)
16. Huenerfauth, M.: Participation of high school and undergraduate students who are deaf in research on American Sign Language animation. In: ACM SIGACCESS Accessibility and Computing, vol. (97), pp. 14–24. ACM Press, New York (2010)