Linking an ID-gloss database of ASL with child language corpora

Julia Fanghella[1], Leah Geer[2], Jonathan Henner[3], Julie Hochgesang[4], Diane Lillo-Martin[1], Gaurav Mathur[4], Gene Mirus[4], Pedro Pascual-Villanueva[1]

[1]University of Connecticut, [2]University of Texas at Austin, [3]Boston University, [4] Gallaudet University
{julia.fanghella, diane.lillo-martin}@uconn.edu,{julie.hochgesang, gaurav.mathur, gene.mirus}@galluadet.edu,
leah.geer@utexas.edu, jhenner@bu.edu, pedrodpedro@gmail.com

Abstract

We describe an on-going project to develop a lexical database of American Sign Language (ASL) as a tool for annotating ASL corpora collected in the United States. Labs within our team complete locally chosen fields using their notation system of choice, and pick from globally available, agreed-upon fields, which are then merged into the global database. Here, we compare glosses in the database to annotations of spontaneous child data from the BiBiBi project (Chen Pichler et al., 2010). These comparisons validate our need to develop a digital link between the database and corpus. This link will help ensure that annotators use the appropriate ID-glosses and allow needed glosses to be readily detected (Johnston, 2011b; Hanke and Storz, 2008). An ID-gloss database is essential for consistent, systematic annotation of sign language corpora, as (Johnston, 2011b) has pointed out. Next steps in expanding and strengthening our database's connection to ASL corpora include (i) looking more carefully at the source of data (e.g. who is signing, language background, age, region, etc.), (ii) taking into account signing genre (e.g. presentation, informal conversation, child-directed etc), and (iii) confronting the matter of deixis, gesture, depicting verbs and other constructions that depend on signing space.

Keywords: ID-gloss, ASL corpora, lemmatization

1. Introduction

A lexical database that lists a unique gloss, also known as ID-gloss, for each sign is indispensable for annotating corpora consistently (Johnston, 2010). The human transcriber, when left to rely on their own memory for retrieval of unique glosses, is more likely to produce errors in the transcript. Continuous use of the database during the transcription process allows the human transcriber a more efficient retrieval system that will reduce the amount of errors in the transcript. As corpora grow, they feed the lexical database in turn, providing tokens of signs that need unique glosses. This paper reports on an on-going project, the ID-Gloss Project as reported in Alkoby et al. (2010), to develop a lexical database of American Sign Language (ASL) as a tool for annotating sign language corpora collected in the United States.

1.1. Database design

The design of our ID-gloss database (Alkoby et al., 2010) is unique, and reflects the current scholarly approach to sign linguistics in the United States. Several different research labs work on sign languages, but no set of systematic, consistent, nationally accepted glosses exists. For this reason, our database was developed so as to permit different research groups to provide site-specific information corresponding to a common set of lexical signs. It is not that we wish for a set of standard glosses to be used in ASL research but rather we wish to facilitate cross-lab data comparison, which is the aim of this project. Each lab completes locally chosen fields (gloss, phonological information, word class, etc.) using their annotation system of choice, and picking from a global template of available fields that was designed in collaborative meetings. The fields are then merged into the global database, thereby providing complementary information for each sign. Figure 1 provides a schematic diagram of our system. The diagram illustrates that the global site houses all of the media files, which are linked to each of the local databases and the local database is linked with the respective local template (a subset of the global template).

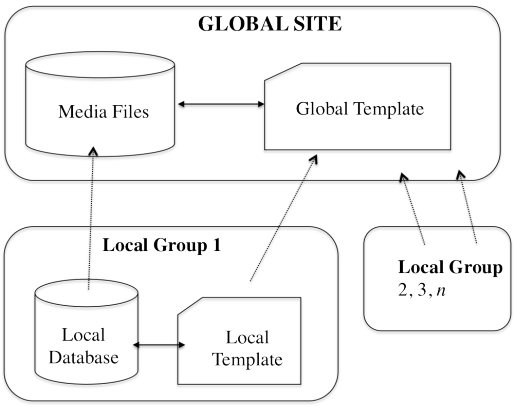


Figure 1: Schematic diagram of the databasae structure.

1.2. Approach: methods of database construction

In this section, we will discuss how our database is built. We begin with a description of what types of annotations each group contributes to the database and focus on the phonological and phonetic notes and descriptions we make for each sign. Other annotations include morphosyntactic information like part of speech and sociolingusitic information like regions or age-groups that might use a particular sign, among other things.

1.2.1. Research groups

Our approach currently capitalizes on the annotation perspectives of three different research groups. The first group

comprises researchers at Gallaudet University and the University of Connecticut (for convenience, this group is referred to as G/UC). This group follows transcription conventions developed by Chen Pichler et al. (2010). Phonological information about signs is entered using Stokoe Notation (Stokoe et al., 1965) as well as phonetic information about hand configuration using Sign Language Phonetics and Architecture – SLPA – (Johnson and Liddell, 2011). The second group, referred to as BTS henceforth, includes researchers from Boston University and Gallaudet University using the Berkeley Transcription System, or BTS (Slobin and Hoiting, 2002). The third group includes researchers at the University of Texas (UTX). This group also uses SLPA for each of the formational parameters of signs.

1.2.2. Data annotation

The three notation systems mentioned in §1.2. have individual strengths that contribute to the structural integrity of the ID-gloss database. Stokoe's notation system is the oldest and most well known by many sign language linguists. The notation system is based on three major sign parameters, namely the *dez* (handshape), *tab* (location) and *sig* (movement). The handshape in a sign can be described by 19 possible labels and additional diacritics. The location of a sign can be represented by symbols that represent certain areas of the body, from the face to the hips. The possible placements can be more specific (forehead, mid-face, chin, or cheek/temple) or more general (trunk). The movement can be represented by 24 symbols. These include directionality (upward or downward movement), internal movement such as wiggling, and movements to contact and grasping movements.

BTS notation was developed in order to be compatible with the CHAT transcription system (MacWhinney, 2000) used by the CLAN analysis programs in Child Language Data Exchange System (CHILDES). The full BTS system allows for annotation of polycomponential signs, such as classifiers and depicting signs. With respect to handshape notation, BTS allows for more fine-grained distinctions between than Stokoe (e.g., the ASL 'A', 'S', and 'T' handshapes all receive different notations in BTS but receive the same label in Stokoe Notation). The BTS uses 68 distinct subhandshapes, elaborated from 10 more abstract handshape categories.

Johnson and Liddell (2011) propose a segmental approach, the SLPA, to the phonetic notation of signs wherein each segment is notated with information about the handshape, movement, placement (like Stokoe's *location*), contact, and orientation of the hand(s). The use of Johnson and Liddell's notation system is time consuming, but it is useful for gathering phonetic-level detail about the production of each sign in the database, something neither Stokoe nor BTS notations provide. Additionally, while Stokoe Notation provides a single label for each parameter of the sign (e.g., [A] for the handshape pictured in Figure 2), SLPA provides a componential notation for the behavior of separate elements in each part of the sign. For example, the hand configuration in SLPA is represented by a series of symbols that describe the joint behavior of each finger and thumb as well as the arrangement (relationship between fingers, such as crossing) and contact, if any, between the fingers and thumb. Thus, the [A] in Stokoe notation is phonetically annotated as [LEE<1FF=2FF=3FFe=4FFe] in SLPA.

Figure 2: [A] handshape.

The choice of each notation system depends on a research lab's theoretical orientation and research goals. The database allows for each group to use their preferred notation system. Another advantage of the database system is that other research labs can then see what notations are used by other research labs.

1.3. Linking the database with an ASL corpus

One advantage to our overall design is flexibility in assigning glosses to signs, since it allows local groups to gloss the same sign in potentially different ways. Furthermore, the different groups may provide different sets of additional information for each sign. While each group gives at least information in the fields *gloss* and *alternative gloss*, the groups provide different phonological information, and choose between optional fields for morphosyntactic, sociolinguistic, and other types of information. Each lab has access to the information entered by other labs, so the work is mutually beneficial. Moreover, the flexibility afforded by the system may lead to eventual convergence on glosses, a desirable outcome. In our project, the annotation of corpora and the building of a lexicon have so far been independent processes, but here we evaluate our progress to date, and discuss concerns for the continued development of these resources.

2. Methods

For the present paper, we are comparing the glosses for signs in our ID-gloss database to annotations of spontaneous child ASL data from the BiBiBi project (Chen Pichler et al., 2010). The child data were annotated by coders trained to use consistent glosses, but they worked without access to the developing database. The first set of signs selected for inclusion in the database were expected to be ones that would likely occur in the child language corpus. Thus, this makes a good test case for examining the efficacy of a cyclical approach to simultaneously building a lexicon and corpus. Specifically, as we start comparing the corpus annotations to the signs in the database, we notice that some of our predictions are borne out (e.g., signs we predicted would be used were indeed used), but other signs are missing from the database. For example, consider the table in Table 1.

Related Signs in Corpus	ID-gloss	Missing (or inconsistently glossed) signs
EAT/FOOD	EAT	
PICK/FIND	PICK	
SEARCH/LOOK-FOR	SEARCH	check consistency
MY/MINE	POSS(self)	check consistency
SAME/SAME-AS	SAME	SAME-TIME

Table 1: BiBiBi glossing conventions: need verification with ID-gloss database

We can see the potential challenges in glossing: it is tempting to use context to distinguish 'search' from 'look-for', for example, but doing so is incompatible with the goal of maximizing searchability by having a unique gloss per sign type. This process helps to fuel the cyclical process to adding signs to the lexical database then returning to transcription with the expanded list. This report includes comparison of approximately 650 of the 1000 signs currently in our database, which have been assigned ID-glosses by the G/UC group.

Five sessions of spontaneous ASL data from one child in this corpus were selected for analysis. These sessions were collected when the child, Ben, was age 1;07 (years;months), 1;10, 2;01, 2;04, and 2;07. The total number of individual child productions at each session is given in the second column of Table 1. Out of these productions, the following were eliminated: uninterpretable productions (coded as YYY or XXX according to conventions), gestures, mouthing (in the absence of a manual sign), fingerspelling (coded as FS), pointing (coded as IX or POSS), depicting (coded as DV). The remaining items are lexical tokens, the number of which is given in the third column of the table. Finally, repeated tokens of the same type within a session were reduced, providing the number of lexical types, given in the fourth column.

Age	Total child utts	Lexical Tokens	Lexical Types
1:07	459	105	42
1:10	854	340	77
2;01	445	175	60
2;04	625	275	95
2;07	454	213	81

Table 2: Data set used for analysis

Each of the lexical types produced by Ben was compared against the set of ID-glosses entered by the G/UC team. We calculated the proportion of Ben's lexical types that were shared with the database and contrasted them with the those that were not yet in our database (unshared types).

3. Results

The results of our analysis are presented in Figure 3.

Overall, 63% of Ben's lexical types are included in the database. As the figure illustrates, the proportion of shared types decreases over the year's worth of data, from 76% to 66%.

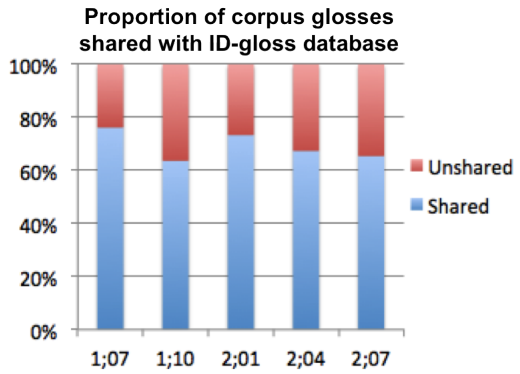


Figure 3: Proportion of glosses shared with database.

4. Discussion

In this section, we will discuss what we better understand about the process of ID-glossing based on the comparison we have just outlined in §2. and §3. We particularly focus upon the challenges that are presented by the missing glosses (particularly phonological variants) and challenges that are involved with growing the database (e.g., how can we involve the community, a question that we have been concerned with from the onset of our project). But first, a quick discussion of the results presented in §3.

4.1. Interpretation of the results

As we saw from Figure 3, the proportion of lexical types produced by Ben that were also in the database decreased. One possible interpretation of this is that Ben's lexicon grew, as one might expect, from 1;07 to 2;07, but the database had not yet been updated with examples of his newer lexical items. This reinforces the circular process we have been describing, wherein the database encourages consistent transcription, but corpus transcription encourages expansion of the database when new items are encountered.

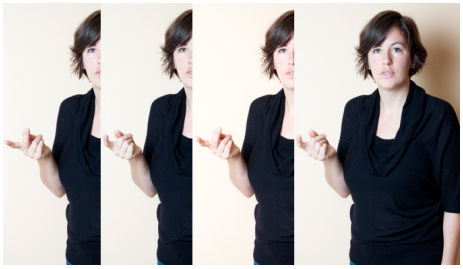
4.2. Missing items & glossing challenges

Many of the glosses that do not appear in the database currently are signs that will be added in the near future (e.g., CHICKEN, FALL-DOWN, GRANDMOTHER, WATERMELON). Others are numbers; they are annotated conventionally, but not included in the database. Perhaps a future step regarding numbers might be to include signs with numeral incorporation, as these do vary, and we would want to be able to capture any variation with an appropriate gloss in our database. A few glosses indicate that the annotators did not follow the annotation conventions consistently (e.g., MINE was used although the conventions called for POSS(self)). Despite the overall utility of the database, which reaffirmed the need for ID-glossing, several problem cases were revealed that deserve attention. The issue that will be addressed here is, how many ID-glosses should be assigned for signs that resemble one another and represent the same concept, or the issue of phonological variation.

4.2.1. Treating phonological variation

Consider, for example, the word 'dog': in order to account for different versions of the ASL sign for 'dog', the database includes unique glosses for two distinct phonological forms for the same concept: DOG$_{snap}$ and DOG$_{slaphip}$. The annotators of the child corpus used three different glosses, however: DOG$_{snap}$, DOG$_{slapsnap}$, and DOG. The use of the gloss DOG was the result of human transcriber error. The transcriber should have appended 'snap' to the gloss. This error could have been avoided if the transcriber had access to the database. For the other variant, DOG$_{slapsnap}$, it would be possible to add to the database, and include the phonological information distinguishing the three variants in every annotation in the child database, eliminating the use of the underspecified gloss DOG. An alternative option would be to have one unique gloss DOG and leave the identification of phonological variants to secondary tagging. The question of how many distinct glosses are needed for the different forms of 'dog' – and other items that have multiple phonological forms – is one that needs to be addressed more thoroughly. Johnston (2010) discusses this very issue and suggests using a separate ID-gloss for each phonological variant, and to tag phonological information (such as handshape or movement) onto the gloss. This was the approach taken here with the different versions of 'dog'. An approach similar in principle – provide phonological information to distinguish sign variants – is used by the researchers who provide glosses using BTS. With that system, all phonological variants with different initial handshapes will necessarily be distinctive. DOG$_{snap}$ is glossed [KT]DOG and DOG$_{slaphip}$ is labeled [BU]DOG, where [KT] and [BU] are the names of the initial handshapes used to produce these signs. However, DOG$_{slapsnap}$ would have the same representation as DOG$_{slaphip}$, since they both use the same initial handshape. To the extent that these variants should be differentially glossed, the technique of using handshape information to distinguish varying phonological forms for the same concept is appealing, although it is not sufficient to distinguish all cases.

Another method, which is now the standard for the Auslan corpus (Johnston, 2011a), is to use a single ID-gloss for all *minor* phonological variants, and then specify the phonological information through secondary tagging (e.g. on a separate tier in ELAN). What is unclear, however, is what is meant by "minor" in terms of phonological variants. Does this mean one quantifiable difference, or more? In our view, the changes in the the phonological forms of DOG$_{slaphip}$, DOG$_{snap}$ and DOG$_{slapsnap}$ are not minor. There are several changes form to form, which can be seen in Figure 4. Specifically, DOG$_{snap}$ is produced in neutral space and requires the index finger to flex and make contact with the thumb in a snapping motion which is repeated (Figure 4a). DOG$_{slapsnap}$, pictured in Figure 4b, requires the hand, with all fingers extended to make contact with the thigh then move to neutral space where the fingers change their configuration and snap once. Finally, DOG$_{slaphip}$ does not have the snapping motion of DOG$_{snap}$ or DOG$_{slapsnap}$ but has the patting motion from DOG$_{slapsnap}$ - that is, DOG$_{slaphip}$ is when the hand needs to move to the thigh and make contact twice or so (Figure 4c).



(a) DOG$_{snap}$

(b) DOG$_{slapsnap}$

(c) DOG$_{slaphip}$

Figure 4: Three different tokens of 'dog' in ASL

There are pros and cons to each approach. Under the first approach, the transcriber would be burdened with generating (or requesting) a new ID-gloss for any new phonological variant produced in a text and adding it to the database. However, this would be a burden only for the first time a particular variant is encountered. If the same variant is used again, the transcriber need only select the correct ID-gloss from the database. On the other hand, if we follow Johnston's approach, the transcriber would not be as burdened the first time. Yet, the burden would be delayed until time came for secondary tagging.

The transcriber would have to add the same phonological information on a secondary tier every time the variant appears, leading to redundancy within the transcription.

4.2.2. What is the purpose of the analysis?

While there are clear benefits and challenges with either approach, the following question should be considered carefully before deciding which approach to follow: What is the goal of the analysis? If, for instance, the analysis is more focused on the meaning (syntactic / morphological structure of the child's utterances), and less on phonological variation, then Johnston's approach should suffice. On the other hand, if the focus of the analysis is on the frequency

and environment of the phonological variants themselves, then Johnston argues that it will be sufficient to conduct a search on the secondary tagging.

What is clear from this discussion is that comparisons between the annotated data and the glosses in the database validate our need to develop an electronic link between the lexical database and the corpus. Such a link (currently under development) will serve two important purposes: it will help to ensure that annotators use the appropriate ID-glosses, and it will allow new glosses needed to be readily detected (Johnston, 2010; Hanke and Storz, 2008). As the database grows, this strict process will also allow researchers to make accurate estimates of lexical frequency, something that is lacking in most sign language research to date and something which has implications for further work in various subfields of linguistics.

4.3. Challenges in growing our database

While modest evidence indicates that the ID-gloss database will be useful for annotating the child language corpus, König et al. (2010) identified several issues that accompany the development of a gloss database. Three of these issues will be discussed here: a) ethical issues related to assigning glosses to signs, b) reconciling the variety of gloss notation systems used by various labs, and c) challenges in glossing signs that do not have one-to-one translation equivalents.

4.3.1. Glossing & community involvement

First, who gets to decide the glosses for the signs, the researcher(s) or the language community? Hochgesang et al. (2010) offer a framework for addressing this issue (but see also Dudis et al. (2009) for another perspective). In particular, it is important to be transparent throughout the research process and involve the community of language users at each juncture. To give an example, consider the ASL signs EAT and FOOD. It may not be clear at this point if there is any systematic way in which the form of these signs differs when they appear in citation form (though it is certainly clear with context), the community of ASL users may have opinions about which English word is a better unique identifier, and it is these intuitions researchers should be attentive to and consider when assigning glosses to lexical items. Discussion on how to best allow community participation is on-going.

4.3.2. Different glosses, different labs

The second issue concerns consistency in glossing. We will discuss this issue as it relates to both within-lab and across-lab concerns. The difficulties of utilizing a glossing system for a signed language are well established (Pizzuto and Pietrandrea (2001) and see also our discussion in §4.2.). One way to maintain gloss consistency within a group is by using an available dictionary for as many signs as possible. The BTS group uses The American Sign Language Handshape Dictionary (Tennant and Brown, 1998) for glosses where possible, supplemented with an explicit handshape symbol at the beginning of each sign, as we discussed in §4.2.1.. As we have already mentioned, the development of an electronic link between corpus and database will even more greatly facilitate consistency within each group. In the United States, the issue of cross-lab inconsistency is a major concern. The diversity of sign language research labs, and the lack of national glossing standards, results in differences in glossing at the lexical level. Our database was specifically designed to allow for these differences and nevertheless permit cross-lab comparisons and eventual cross-corpus searches. In addition to the lab's primary gloss for a sign, each lab completes information about alternate glosses in a separate field in the database. If the database is queried, the displayed results will match either the main gloss or the alternate. This allows more flexibility in use as well as maximizing the ability to search the database. To continue the example from above, consider the sign for 'eat': One lab used the gloss EAT with the alternative gloss FOOD, while another lab did the opposite. In a cross-lab search, all of the relevant information is still retrieved with a single query.

4.3.3. Signs without English translation equivalents

The third, and final, issue we will discuss here is the difficulty in assigning glosses to signs that do not have suitable translation equivalents between ASL and English. Thus far, our database does not include, classifiers , depicting signs (except for one) and other polycomponential forms, but in order for the database to reflect language use in ASL corpora, these are forms that will need to be included. Johnston (2011b) offers possible solutions, based on his work in developing an ID-glossed corpus for Auslan. He makes distinctions between signs that are fully lexical, those that are partially lexical and those that are non-lexical (e.g., gesture and emblems). For partially lexical items, or signs that are regular in form but for which the meaning is conditioned by the context of utterance, Johnston (2011b) suggests using some sort of indication of what type of partly lexical item it is (classifier, depicting signs, etc.), what handshape is used, and what does it mean in a particular context. To use Johnston's example, a partly lexical sign might be glossed DSH(F):describe-as-appropriate, where "DSH" indicates that it is a depicting sign with an "F" handshape. Another similar type of example from our database is pictured in Figure 5. Here, the UTX coder uses "DS" to indicate that the form is a type of depicting sign, "F" to indicate the handshape, and *"long-skinny"* to reflect the type of noun being classified. When this form is used in naturalistic signing, as opposed to citation form, the *"long-skinny"* designation can be replaced with whatever is indicated by the given context.

Figure 5: DS(F):long-skinny.

We will have to decide lab-internally, or as a group, what

sorts of glosses we find most appropriate and fitting for this category of signs. One approach would be to adopt the glossing techniques of (Johnston, 2011b) wherein we distinguish between fully, partly, and non-lexical items. This would allow us, crucially, to capture enough about the form of the sign to encourage consistent application of a unique ID-gloss, but allows for flexible additions to the semantic content

5. Conclusions

Here we have shown that, despite challenges that persist in developing an ASL lexical database, the linking of transcripts within a corpus to such a database will aid in understanding crucial facts about the language. An ID-gloss database is essential for consistent and systematic annotation of sign language corpora, as Johnston (2010) has pointed out and as we have attempted to demonstrate. We provided preliminary results from a comparison between our database and an annotated corpus that did not have the benefit of an ID-gloss database. There are several logical next steps to consider in expanding the ID-gloss database and in strengthening its connection to ASL corpora. In closing, we will mention three of these steps. First, we should look more carefully at the source of the data and document the metadata. It is important to know who is signing, what is their language background, how old are they, where did they grow up and where do they live now. All of this information will help generate a more complete picture about how ASL is used and what differences exist between groups (e.g., regional groups, age groups etc). Second, we should take into account the genre of signing. Was a particular text from a presentation or an informal conversation? Was an adult directing signing at an infant or child? This will contribute valuable information that can lead to descriptions of distinct linguistic registers in ASL. Lastly, as we mentioned briefly in §4.3.3., we need to confront the matter of deixis, gesture, depicting verbs and other constructions that depend on signing space. Each of these questions will help grow our database, as well as allow for more accurately annotated corpora and thus strengthen the link between the the two.

6. Acknowledgements

The project described in this article was supported by Award Number R01DC009263 from the National Institute on Deafness and Other Communication Disorders. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Deafness and Other Communication Disorders or the National Institutes of Health.

7. References

K. Alkoby, J.A. Hochgesang, G. Mirus, and P. Pascual Villanueva. 2010. Construction of an ID Gloss database. In *Proceedings from the* 10^{th} *meeting of the Theoretical Issues in Sign Language Research Conference*, West Lafayette, IN.

D. Chen Pichler, J.A. Hochgesang, D. Lillo-Martin, and R. Müller de Quadros. 2010. Conventions for sign and speech transcription of child bimodal bilingual corpora in ELAN. *Language, Interaction and Acquisition/Langage, Interaction et Acquisition*, 1(1):11–40.

P. Dudis, G. Mathur, and G. Mirus. 2009. Bringing a corpus in line with deaf communities. In *Proceedings from the Sign Language Corpora: Linguistic Issues workshop*, University College of London.

T. Hanke and J. Storz. 2008. iLex – a database tool for integrating sign langauge corpus lingustics and sign language lexicography. In O. Crasborn, E. Efthimiou, T. Hanke, E.D. Thoutenhoofd, and I. Zwisterlood, editors, *Proceedings of the Third Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora at the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, pages 64–67, Marrakech, Morocco.

J.A. Hochgesang, P. Pascual Villanueva, G. Mathur, and D. Lillo-Martin. 2010. Building a database while considering research ethics in sign language communities. In *Proceedings of the* 4^{th} *Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies at Language Resources and Evaluation (LREC'10)*, pages 112–116, Valletta, Malta, May.

R. E. Johnson and S. K. Liddell. 2011. A segmental framework for representing signs phonetically. *Sign Language Studies*, 11(3):408–463.

T. Johnston. 2010. From archive to corpus: Transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics*, 15:104–129.

T. Johnston, 2011a. *Auslan Corpus Annotation Guidelines*. Centre for Language Sciences, Department of Linguistics, Macquarie University, Sydney, Australia.

T. Johnston. 2011b. Lexical frequency in sign languages. *Journal of Deaf Studies and Deaf Education*, 36:1–31.

S. König, R. Konrad, G. Langer, and R. Nishio. 2010. How much top-down and bottom-up do we need to build a lemmatized corpus? In *Proceedings from the* 10^{th} *meeting of the Theoretical Issues in Sign Language Research Conference*, West Lafayette, IN.

B. MacWhinney. 2000. *The* CHILDES *Project: Tools for Analyzing Talk.* Lawrence Erlbaum Associates, Mahwah, NJ, 3^{rd} edition.

E. Pizzuto and P. Pietrandrea. 2001. The notation of signed texts: Open questions and indications for futher research. *Sign Language & Linguistics*, 4(1):29–45.

D. I. Slobin and N. Hoiting. 2002. The Berkely Transcription System for sign language research (BTS). In Gary Morgan and Bencie Woll, editors, *Directions in sign language acquisition*. John Benjamins, Amsterdam.

W. C. Stokoe, D. C. Casterline, and C. G. Croneberg. 1965. *A dictionary of American Sign Language on linguistic principles*. Linstok Press, Silver Spring, MD.

R. A. Tennant and M. G. Brown. 1998. *The American Sign Language Handshape Dictionary*. Gallaudet University Press, Washington.