Conventions for sign and speech transcription of child bimodal bilingual corpora in ELAN

Deborah Chen Pichler, Julie A. Hochgesang, Diane Lillo-Martin and Ronice Müller de Quadros

This article extends current methodologies for the linguistic analysis of sign language acquisition to cases of bimodal bilingual acquisition. Using ELAN, we are transcribing longitudinal spontaneous production data from hearing children of Deaf parents who are learning either American Sign Language (ASL) and American English (AE), or Brazilian Sign Language (Libras, also referred to as Língua de Sinais Brasileira/LSB in some texts) and Brazilian Portuguese (BP). Our goal is to construct corpora that can be mined for a wide range of investigations on various topics in acquisition. Thus, it is important that we maintain consistency in transcription for both signed and spoken languages. This article documents our transcription conventions, including the principles behind our approach. Using this document, other researchers can chose to follow similar conventions or develop new ones using our suggestions as a starting point.

Keywords: bimodal bilingualism, sign language acquisition, child language acquisition, corpus methodology, sign notational conventions, ELAN, transcription

1. Introduction — Building a bi-national bimodal bilingual corpus for L1 acquisition

Bilingualism has long been of interest to those studying language acquisition, as a window into the complex interaction of two languages developing within a single individual. Recent studies have examined bilingual individuals' separation of their two languages, the timing of their acquisition of each language, and the potential influences of the structure of one language on the other, among other topics (see, e.g., Genesee 1989; Meisel 1989; Paradis & Genesee 1996; Nicoladis & Genesee 1997; Bhatia & Ritchie 1999; Paradis & Navarro 2003). The vast majority of these

Language, Interaction and Acquisition 1:1 (2010), **11–40**. DOI 10.1075/lia.1.1.03che
ISSN 1879–7865 / E-ISSN 1879–7873

studies focus on individuals acquiring two spoken languages, and they have greatly increased our understanding of bilingualism in general. However, restricting our investigations to speech-speech bilinguals, or *unimodal bilinguals*, limits our understanding of the bilingual phenomenon. To see the full picture of what bilingualism entails and makes possible, we must also study other categories of bilinguals, such as those learning two sign languages (still very rare), and those learning one sign and one spoken language. This latter category is collectively known as *bimodal bilinguals*, usually comprised of children of Deaf[1] adults, known as *codas*, or Enfants Entendants de Parents Sourds, *EEPS*, in France. Such individuals are very interesting to linguistic inquiry because in addition to alternating between their two languages (*code-switching*), they have the additional potential for producing simultaneous signs and speech, known as *code-blending* (Emmorey et al. 2008). This option is not available to unimodal bilinguals. Whereas the latter type of bilinguals presents a valuable opportunity to study the cross-linguistic influences exerted by their two languages on each other, bimodal bilinguals present an additional opportunity to directly observe the mechanisms and constraints governing the interaction of two languages during bimodal production.

In recent years, a few studies have examined bimodal bilingual production of coda adults (Messing 1999; Berent 2004; Bishop 2006; Emmorey et al. 2008) and children (Prinz & Prinz 1981; Schiff-Myers 1988; Johnson et al. 1992; Singleton & Tittle 2000; Petitto et al. 2001; van den Bogaerde & Baker 2005, 2009). The child studies have examined the development of bimodal bilingualism from the perspective of fluency levels obtained by children in each language, the use of *code-switching* or simultaneous *code-blending*, and the relation between developing bimodal bilingualism and theories of the mechanisms of language development. Our own study seeks to address these questions by examining the development of two signed-spoken language pairs: American Sign Language (ASL) and American English (AE), and Brazilian Sign Language (Libras, also referred to as Língua de Sinais Brasileira/LSB in some texts) and Brazilian Portuguese (BP). Our focus is on understanding the nature of language development more generally, using data from bimodal bilinguals as a potentially extreme case of environmental complexity (see, for example, Chen Pichler et al. 2010; Quadros et al. 2010; Lillo-Martin et al. in press). Our data are being collected and analyzed in three laboratories: two in the U.S. (Gallaudet University and the University of Connecticut), and one in Brazil (Universidade Federal de Santa Catarina).

1. Following conventional practice in the sign language literature, we use the capitalized form *Deaf* to refer to individuals for whom deafness represents a cultural identity rather than a disability, closely associated with the use of a natural sign language such as ASL or Libras.

Critical to achieving our goals is the development of a bimodal bilingual corpus maintained by all three host institutions and utilizing consistent methods for data collection, transcription and analysis. Such practices are, of course, continually evolving as new technologies become available, or new problems come to light. This article summarizes the methodologies and conventions that we have established thus far, building on our previous experience collecting and transcribing longitudinal child sign data (Lillo-Martin & Chen Pichler 2008) and on that of others (MacWhinney 2000; Baker et al. 2009). For those looking for a concise version of our conventions, a summary table of all the notations discussed in this article is provided in Appendix A.

2. Data collection

Our corpus currently consists of longitudinal, spontaneous production data from four bimodal bilingual children in the US and three in Brazil. Longitudinal data have long been a cornerstone of acquisition studies, offering a wealth of information on the processes by which children develop language (see, for example, the over 3000 research articles using the CHILDES data base; bibliography available at http://talkbank.org/usage/). Such data provide several advantages to the researcher: (a) a particular child participant is observed in a natural environment, interacting with people she is very familiar with; (b) the child's development over a period of time is carefully documented; (c) researchers working within a wide variety of theoretical frameworks are able to use such data to address a great range of theoretical issues; (d) researchers can investigate hypotheses about non-target structures used by the child or the relative onset of various target structures; and (e) input provided to the child can also be sampled and studied, when the child is recorded interacting with a parent. Useful resources on the use of spontaneous production data are found in Stromswold (1996) and Snyder (2007), among others.

We are collecting longitudinal data concentrating on the early stages of language development, focusing on bimodal bilingual children between the ages of 18 months and four years of age. All of the children in our project have normal hearing and no known disabilities. Our participants include both girls and boys, as well as children of different racial and ethnic groups. The main inclusion criteria are hearing status and appropriate early linguistic experience in both one signed and one spoken language. In some cases, both parents are Deaf, while in others, only one parent is Deaf and the other is a hearing fluent signer; we require only that the Deaf parent(s) use sign language as the primary means of communication with their children, to ensure that all child participants receive consistent early input in ASL or Libras.

To protect the confidentiality of our participants, both children and adults, we use pseudonyms instead of actual names on all of our tape labels, transcripts and research reports. Following standard policy at American universities, we collect informed consent and video release forms from all parents before they join our project, making sure that they are fully aware of the goals of our research and the ways in which their child's video data will and will not be used. Ideally, we aim to collect approximately three years of longitudinal data from each participant child, but they and their families are, of course, free to withdraw from the study at any time.

Parents provide background information about the estimated percentage of signed *vs*. spoken input their child receives at home and in their educational programs. They also report the number of siblings of the child participants, their hearing status and birth order. Lastly, we ask parents to identify their highest educational level and provide self-ratings of their own experience and fluency in the relevant signed language. We have tried to select a subject sample that is representative of the greater coda population in our respective communities, but given the small size of these communities, it is not always possible to balance across all background factors, nor to apply statistical analyses such as regression or covariance to remove their effects.

Children are video taped in naturalistic sessions of 45 to 60 minutes each on a weekly basis, alternating between sign- and speech-target sessions. With a few exceptions, sign-target sessions involve only Deaf adults (research assistant and/or parent), while speech-target sessions involve only hearing adults (research assistant and/or parent). However, all interlocutors are, in fact, bilingual and have a natural tendency to sometimes code-mix and code-blend while interacting with the child participants. Additionally, the recording environment (usually the child's home or preschool, or for some American participants, Gallaudet University) can be considered strongly bilingual. As a result, some of our sessions are heavily code-blended. We neither actively encourage nor discourage code-blending during filming, preferring to allow the child and interlocutors to interact as naturally as possible. We have found that sessions that are heavily bimodal complement our unimodal sessions well, serving as rich resources for studying aspects of bimodal production.

All research assistants involved in data collection are trained to interact naturally with child participants and are taught strategies for encouraging child production without being overbearing. We also provide parents with a simple training packet including tips on how to optimize filming sessions at home (e.g., turning off noisy fans or televisions, ensuring adequate lighting, choosing activities that encourage the child to stay in one place) and a DVD with short clips illustrating ways in which they can encourage language production (e.g., engaging their child in verbal activities, asking questions that elicit responses beyond simple labels for

objects, and resisting the urge to prod their child to demonstrate every word or sign that he/she knows). Whenever possible, we employ one adult to play with the child while a second monitors the camera and takes notes about the session, noting unusual forms that occur or contextual information that may facilitate the eventual transcription process.

3. Data transcription

Transcription is a technical act that involves an understanding of theory, including transcript design and possible research goals. Ochs (1979: 44) stated, with respect to child language behavior, that "transcriptions are the researcher's data," resulting from a "selective process reflecting theoretical goals and definitions". Unless language behavior is somehow represented, it cannot be analyzed. In this sense, transcription is indispensable. As Ochs observes, the more basic a transcript is, the more accessible it will be to other linguists. The basic transcript is as theory-neutral as possible and includes only the most fundamental descriptive units: utterances[2] and words/signs.

In the interest of making our database compatible with a broad range of potential research foci within a variety of theoretical frameworks, we try to keep our transcripts as basic as possible. Our transcription format includes all the linguistic information identified by Baker et al. (2009) as essential to any transcript of signed language, with the exception of nonmanual signals (see our justification for this exception in Section 3.3.6 of this article). We also strive to minimize interpretation of the data during the transcription process by choosing notational conventions that accurately reflect the form of the sign or speech without imposing any particular analysis. Of course this is not always possible, but our efforts are greatly helped by the use of ELAN (http://www.lat-mpi.eu/tools/elan), a free multi-media annotation program widely used in sign language research. ELAN is ideal for our transcripts because it integrates the transcript and the corresponding digitized video data into a single file, such that each annotation is time-locked with the relevant segment of the video. In this way, researchers can easily locate and view any transcribed utterance in our transcripts, accessing a broad spectrum of nuanced but important information that may be absent from our minimalist transcripts. Furthermore, it permits coding tiers to be added for each new project, so that a

2. In reality, it is impossible to avoid interpretation altogether, since the delimitation of utterances already reflects some interpretation of the data by the transcriber (Hochgesang 2009). For this reason, it is especially important that the criteria used for determining utterance boundaries be explicitly defined.

permanent record can be kept of the coding decisions that are made together with each analyzed utterance. ELAN also offers analysis tools which permit rapid collection of data across any number of existing transcripts (such as word counts and searches). Finally, many other sign language researchers are already familiar with this program, facilitating efficient exchange of transcripts between research teams.

In general, signed data are transcribed by native signing research assistants (either hearing or Deaf), and spoken data is transcribed by hearing research assistants with knowledge of the relevant sign language. Because bimodal bilinguals frequently produce both signed and spoken production within a single filming session, a single video may be transcribed by a number of individuals, calling for a high degree of coordination. To this end, each research lab has designated one research assistant as lab manager, who is responsible for coordinating the transcription process. When a video file is ready for transcription, the lab manager first assigns it to a speech transcriber. We have found it helpful to transcribe speech first, as our child participants frequently produce bimodal utterances in which the spoken element may be more intelligible than the signed element. Once the first transcriber completes the speech transcription, he or she sends the ELAN file to the lab manager, who updates the online transcript log and posts the file to a folder reserved for transcripts awaiting sign transcription. All posted transcripts, as well as logs, current notational conventions and other project-related files, are maintained online in a password-protected account accessible to all members of our respective labs. Thanks to this system, the sign transcriber can then download the ELAN file with completed speech transcription and add the sign transcription. Once a transcript is complete, the lab manager posts it to a folder reserved for transcripts ready for proofing. The online transcription log is updated each time a transcript advances to the next stage, making it possible for all lab members to keep track of which transcripts have been completed, which are currently underway, and which are in line to be transcribed next.

In order to ensure that our transcripts adequately and accurately represent the original data (as judged by native speakers and experienced researchers), the completed transcripts undergo a proofing process through which interrater reliability can be measured. In our laboratories, interrater reliability has typically been measured by assigning sections of the same transcript to different coders (i.e., once for the spoken language, once for the signed language and then twice, specifically for proofing transcription in both languages). We also utilize a feature in ELAN called *compare annotators*, which allows for comparison of like tiers (e.g., when the ASL of one participant has been transcribed separately by two annotators or transcribers). This ELAN feature calculates agreement in terms of segmentation, or how many annotations have been created and how they compare in length of duration (or overlap).

3.1 Tier organization

Data annotation in ELAN occurs on a series of *tiers*, each dedicated to a different aspect of the data, and bearing various hierarchical relationships to each other. The tiers used in our template fall into three broad groups for each participant: (1) translation, (2) signed production (either ASL or Libras), and (3) spoken production (either English or Brazilian Portuguese). Figure 1 shows a screenshot of an ASL/English transcript sample using the ELAN template from our project.

A hierarchical list of all the basic tiers comprising our ELAN transcript template is provided in Appendix B at the end of this article. Although our template

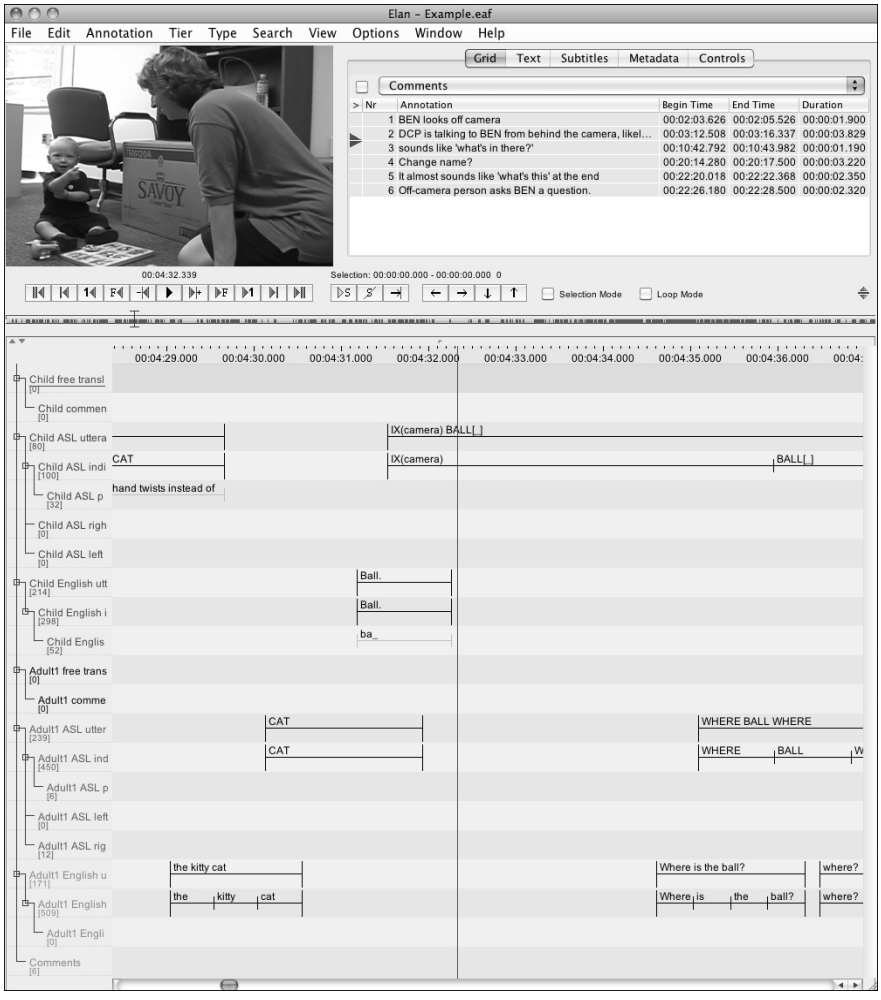


Figure 1. Screenshot of a transcript sample from our project

features distinct groups of tiers for each modality, many of the transcription conventions we use are shared across modalities, as we will discuss in detail later in this section.

3.1.1 *Free translation tiers*

The translation group consists of two tiers, an independent tier labeled *free translation* and a tier labeled (*participant*) *comments* that is dependent on the *free translation* tier. The *free translation* is used for a loose translation of a signed and/or spoken utterance, as interpreted within the given context. The translation may involve interpretation gleaned from context. For example, if a child says/signs, 'Funny!' while looking at his father, the translation might be, 'Daddy's funny!' or 'You're funny!' In cases where participants produce bimodal utterances, perhaps conveying complementary (i.e., non-equivalent) information across the two modalities, the *free translation* tiers convey the global proposition expressed by the speech and sign together. For example, if a participant signs DOG and says *sleeping*, these two utterances are represented as one annotation on the *free translation* tier as: *The dog is sleeping*.

3.1.2 *Tiers related to signed production*

Our transcription template includes five tiers used for signed data: *signed utterance*, *signed individual*, *signed phonology*, *right hand*, and *left hand*. The *signed utterance* tier is the independent tier to which all of the other sign tiers refer and on which they depend hierarchically, as shown in Appendix B. We consider an utterance to be a group of signs delimited by prosodic behavior (e.g., lowering or relaxation of the hands, a longer pause than normal or lengthening of the final sign in the group, and so on). We mark the onset of an utterance at the point when the hand begins to form the hand configuration of the first sign and/or when movement starts. The utterance ends when one or more of the following occurs: the hand changes its configuration, the arm is lowered, signing is paused, or eyegaze shifts to a different location. Utilizing prosodic behavior to delimit utterances is not guided by any specific timing measurements but rather native speaker intuition. It is our intention to capture naturalistic observations in order to discern patterns related to prosodic behavior after we have collected a substantial number of transcripts.

For the *signed individual* tier, the utterance is segmented into individual signs; this is done automatically in ELAN using the tokenization operation. Correct tokenization requires that each gloss on the *signed utterance* tier be separated by a space character, and that multi-word combinations representing a single sign be linked in some way, such as by a hyphen. Annotations for individual signs created via tokenization all have the same time duration and may subsequently be altered

by the transcriber to reflect the onset and offset times of the actual production in the video.

If a transcriber notices any unusual or noteworthy feature related to the form of a given individual sign (e.g., production of the typically two-handed sign CAR as a one-handed sign), this information is noted on the corresponding *signed phonology* tier. The *right hand* and *left hand* tiers allow annotation for the right and left hands separately in cases where a signer switches hand dominance for particular signs. These tiers are also useful when a sign is held on one hand while the other hand moves on to another sign. Such notes can be utilized for further analysis, in which the researcher may be interested in the motivation for such phonological behavior, be it on the phonological, morphological, or syntactic level. One of the advantages of ELAN is that tiers can easily be added (and annotations copied or modified on other tiers) for the purposes of more fine-grained follow-up analysis.

Our transcripts in ELAN employ traditional upper-case English or Portuguese glosses for transcribing signs, a system that offers convenience and high readability, but one with well-known limitations. Due to the lack of standardized writing systems for sign languages, sign glosses are expressed using the written system of a spoken language, typically the dominant language of the community where the sign language is used. These glosses are thus 'twice removed' from the language they are meant to represent (Hochgesang 2009, in progress). Yet such glossing systems are found in almost all published work about sign language linguistics, to the extent that the written glosses themselves have become the actual data used in analyses of signed language structure. Though problems with this state of affairs have been documented (e.g., Johnston 1991; Pizzuto & Pietandrea 2001; Mulrooney 2006; Slobin 2006, 2008), glossing remains a common practice. The question most relevant here is posed by Slobin (2006: 31): "How can [representation of signed languages] be rendered in a format that allows for cross-linguistic and developmental analysis — both between sign languages and in comparison to spoken languages?". Such a concern is especially relevant when both signed and spoken data naturally occur together, as is the case with bimodal-bilingual speakers, and has guided the development of our notational conventions in our transcripts.

To minimize the familiar pitfalls of sign glossing, we have adopted the concept of *ID-glosses* (Johnston 2001) for our corpus. An ID-gloss is a unique label for each lexical sign in the data. It is a lemma, in that the same ID-gloss identifies a set of related signs which may differ in grammatical class, semantic role or inflection (Johnston in preparation). For example, *EAT* is used as a gloss for the uninflected ASL verb *to eat*, aspectually inflected forms such as *to eat regularly* and *to eat continually*, the derived noun *food*, and so on. Our research team maintains running lists of ID-glosses along with sign screenshots for both ASL and Libras online, developed and modified through transcriber discussion, to encourage

consistent choice of glosses across transcribers. The lab managers are responsible for checking all transcripts to ensure interrater reliability with respect to glossing. The multi-file search feature in ELAN (*i.e.* "Search Multiple eaf…") is useful for this purpose.

3.1.3 *Tiers related to spoken production*

The spoken language tiers in our template are organized around the central (independent) *spoken utterance* tier. This tier has two dependent tiers, *spoken individual* and *spoken phonology*, as shown in Appendix B. Parallel to the signed utterance described earlier, a spoken utterance refers to a group of words delimited by prosodic features (e.g., falling tone or a stretch of silence following the last syllable). Spoken utterances are represented using standard orthographic forms on the *spoken utterance* tier, with onset and offset points following the beginning and end of each utterance. For sessions with sufficiently clear audio, the .wav file can be embedded in ELAN, facilitating the visual identification of precise onset and offset points. Annotations on the *spoken utterance* tier are segmented into their component words on the *spoken individual* tier via automatic tokenization, as described earlier for the *signed utterance* tier. Finally, if a word is pronounced unusually (as judged by the native speaker research assistant), it is transcribed using the International Phonetic Alphabet (IPA) on the *spoken phonology* tier.

Many considerations about the transcription of speech are discussed in the manual accompanying the CHILDES database (MacWhinney 2000; see http://childes.psy.cmu.edu/). Although our system is not fully compatible with the CHAT transcription format used in the CHILDES database, we adopt its conventions where possible, because of the excellent foundation it provides for the transcription of speech and its widespread familiarity among language acquisition researchers. By choosing parallel conventions based on CHAT whenever possible for sign and speech, we also hope to facilitate cross-linguistic comparisons and possible full adaptation of our speech transcripts for the CHILDES database in the future, making our data compatible with the CLAN analysis tools (http://childes.psy.cmu.edu/clan/) that are based on the CHAT transcription format.

3.1.4 *Miscellaneous tiers*

In addition to the primary tier categories related to translation, signed production and spoken production, our template includes a number of additional tiers for data that are not particular to any single participant. The *comment* tier is used for miscellaneous information that the transcriber deems important to the readability of the transcript, but which is not represented anywhere else. For example, if the camera sound, for whatever reason, cuts out for a few minutes, or if a participant's utterance is a response to a question signed by a person off-camera, this

information is noted on the *comment* tier. This tier can also be used if a transcriber is unsure of how to transcribe a particular utterance and wants to mark it for a second opinion. Finally, once both spoken and signed data for a video session have been fully annotated, the transcript is proofed by a lab manager or PI in order to ready the transcripts for analysis and to check interrater reliability. Proofing comments or suggestions for changes are recorded on the *feedback* tiers, one for each modality. This allows alternative interpretations of the data to be annotated, while still preserving the annotations that were originally transcribed. Researchers who use the data for various projects can then determine which interpretation is most appropriate for their particular analyses.

3.2 Notational conventions applying to both signed and spoken production

One of the greatest barriers to data sharing is the daunting variability in notational conventions employed by different research teams. In developing notational conventions for our project, we wanted to adopt as far as possible notational conventions that are already familiar and well established for signed and spoken language data. In the field of child language development, by far the most familiar notational conventions are those used in the CHILDES project, as noted above. Although the CHAT transcription format is typically used for spoken language data, a number of the more general conventions can be applied to data in either modality. Our notational conventions for interruptions, unclear utterances and capitalization are all derived from the CHAT conventions, as detailed in the following sub-sections and summarized in the table in Appendix A.

3.2.1 *Interruptions, pauses and trailing off*

We use modified CHAT symbols to indicate when sign or speech is interrupted (either by the signers/speakers themselves or by other participants) or when it "trails off." Note that because tokenization of the *signed utterance* and *spoken utterance* tiers creates individual annotations each time a space is detected, notational symbols must be typed directly adjacent to the preceding word/sign, with no intervening space. Interruptions by others are notated by a single forward slash (/) (e.g., *MOTHER WANT/*), while self-interruptions are notated by two forward slashes (//). Slash notations enclosed within square brackets are also used for *retracing*, or in cases when participants restart an utterance:

[/]	retracing without correction
[//]	retracing with correction
[///]	retracing with reformulation

Pauses within an utterance that are not due to interruption or retracing are notated with the hash symbol (#) directly following the last sign/word preceding the pause, as shown in example (1).

(1) MOTHER LIVE# LONDON

Utterances in which the sign or speech stream trails off are marked with the ellipsis symbol (…) after the last sign/word uttered. For example, if a participant begins to recite the days of the week in spoken English but trails off after the third day, this utterance is transcribed as in example (2).

(2) Sunday# Monday# Tuesday…

3.2.2 *Unclear signs or words*

Spontaneous production, particularly that of young children, frequently yields utterances that are partially obscured or unclear, leading to uncertainty about how best to transcribe them. Following the CHAT format, we use four different notations to distinguish between unclear data: [?], [=?alternative], YYY, and XXX.

In cases where the transcriber is reasonably sure that he or she can identify the unclear sign or word, although some small degree of doubt remains, the notation *[?]* is attached to the end of the unclear word. For example, if a child signs what looks to be *DOG SICK*, but uses a slightly non-targetlike hand configuration in the second sign, this might cast doubt as to the correct interpretation of that sign. The transcriber can encode slight uncertainty by typing DOG SICK[?]. Usually such cases are accompanied by information on the corresponding *phonology* tier specifying the formational features that led the transcriber to employ the *[?]* notation.

In cases where a transcriber is fairly sure of the correct gloss for a word or sign, but can think of an alternative word or sign which might also be feasible, we use the notation *[=?ALTERNATIVE]*. For example, the signs *APPLE* and *ONION* are very similar in ASL, differing only in location: *APPLE* is produced at the cheek, while *ONION* is produced at the temple. If a participant signs what looks to be *GIRL EAT APPLE*, except that the third sign is located somewhere between the temple and cheek, the transcriber may type *GIRL EAT APPLE[=?ONION]*. This notation expresses the transcriber's impression that the third sign was most likely APPLE (perhaps because of contextual information), even though it was signed in a location that makes *ONION* a possible alternative interpretation. As in the case of the *[?]* notation, signs followed by *[=?ALTERNATIVE]* are often accompanied by phonological information on the corresponding *phonology* tier.

The final two notations of uncertainty are similar in that the sign or word is too unclear to be assigned any gloss. If the form of the sign or word under question is clear enough for the transcriber to discern at least some phonetic information

(e.g., sign hand configuration, or the first few phonemes of a spoken word), that sign or word is transcribed as *YYY* or *yyy* and the available phonetic information is noted on the corresponding *phonology* tier. If a sign or word is so unclear that the transcriber is unable to determine any phonetic information at all, it is transcribed as *XXX* or *xxx*.

3.2.3 *Capitalization and punctuation*

For transcription of speech, we follow CHAT conventions in capitalizing only proper names and the English pronoun *I*. First words of sentences are not capitalized. In most cases, this convention does not apply to the transcription of signs, which are glossed entirely in capital letters, following longstanding convention (e.g., *AIRPLANE*). However, as is described in Section 3.3 below, some information on the sign tiers (e.g., referents of points or pronouns) appears in lowercase letters (e.g., an indexical point to Bob's hat is transcribed as *IX(Bob's-hat)*); in these cases, we follow the same capitalization conventions applied to speech transcription. In contrast, our notational system departs from the CHAT conventions in the matter of punctuation: neither signed nor spoken utterances in our transcripts include any punctuation. This decision is consistent with our desire to impose as little analysis as possible regarding the syntactic structure of the utterances recorded in our transcripts.

3.3 Notational conventions specific to the transcription of signed data

As mentioned in the Introduction, we developed our basic glossing conventions for signed data largely from glossing practices adopted in previous work by both our own team and by other researchers in sign linguistics. While there are dedicated transcription systems for representing signed languages at the phonetic level (e.g., *The Movement-Hold Model*, Liddell & Johnson 1989) and the morphological level (*The Berkeley Transcription System*, Slobin et al. 2001), these transcription systems have different purposes from our own. Our goal is to produce a transcript which allows the researcher to readily search the corpus to find utterances of interest for further analysis based on both the transcript itself and the video recording. Consequently, our lab has chosen to use English glosses to represent signed and spoken words. While glossing is admittedly a form of interpretation, it has the advantage of being widely accessible and familiar to a wide range of sign language linguists.

In an effort to minimize variability in glossing across transcribers, we have now begun developing ID-glosses for each sign in our data, as described earlier in Section 3.1.2. While ID-glosses work well for invariable lexical signs or even groups of related signs sharing an identifiable base form (e.g., in the case of verbs inflected to show referents, as discussed below), other signs have interpretations

that are too context-dependent to be assigned a consistent label. Signs falling in this category include pronouns and other index signs, fingerspelled words, name signs, depicting verbs, gestures and emblems. These signs are notated with a label reflecting their category, followed by a context-based description enclosed in parentheses. Labels for categories unique to signed data are represented in capital letters (i.e., *IX* for pronouns or other indices, *FS* for fingerspelled words, *NS* for name signs, and *DV* for depicting verbs), while the labels for categories which can appear as part of either signed or spoken data are represented in lower case letters (i.e., *g* for gestures and *e* for emblems).

3.3.1 *Agreeing or indicating verbs*

Many analyses of sign languages distinguish between *plain* or *lexical* verbs (invariant in form, not directed towards any referent in particular) and *agreeing* or *indicating* verbs (varying in directionality to correspond with a specific referent). In much of the existing sign language literature, the referent of the verbs in the latter category determines the form of the gloss, which explicitly notes the referent of the verb (e.g., *you-GIVE-me* or $_2GIVE_1$ to represent the verb *GIVE* directed from the second person towards the first person). Our notational conventions assign ID-glosses to such verbs, such that the sign verb expressing 'giving' will be glossed simply as *GIVE*, regardless of how it is directed in space. This practice is consistent with both the use of ID-glosses in other sign language corpora (e.g., the Australian Sign Language (AUSLAN) Archive and Corpus: http://www.auslan.org.au/about/corpus/; see also Johnston 1991, 2001) and our intention of creating transcripts that are as theory- and analysis-neutral as possible. Coding of the referents for agreeing or indicating verbs is left to those wishing to conduct analyses in this area.

3.3.2 *Indices and pronouns*

Indices (pointing) directed towards people, objects, or locations are notated with the gloss *IX(referent)*, where the word(s) enclosed in parentheses denote the referent of the index. When the signer points to himself or herself, the referent is labeled as *self* rather than as the signer's pseudonym. This allows us to distinguish between first- and non-first person referents. As mentioned earlier in Section 3.2.3, the referent is typed in lower-case letters unless it is a proper name. If more than one word is required to describe the referent, these words are joined by hyphens. The same general convention is used for other signs directed towards a referent, such as possessive, reflexive and honorific pronouns. These are notated using glosses *POSS(referent)*, *SELF(referent)*, and *HONORIFIC(referent)*, respectively. Thus, if a child points to a Lego castle declaring that it is his, because he built it by himself, such an utterance would be transcribed on the sign utterance tier as shown in example (3).

(3) POSS(self) IX(Lego-castle)# BUILD SELF(self)

3.3.3 *Classifier or depicting verbs*

Perhaps the most challenging sign language verbs for notational systems are *classifier verbs* or *depicting verbs* (Liddell 2003). These signs 'depict' events using a wide combination of locations, motions and conventionalized hand configurations. They are typically very iconically motivated and are interpreted in a highly context-dependent way. Depicting verbs are notated with the label *DV* followed by a description of the depicted event enclosed in parentheses. For example, a verb depicting a vehicle moving down an incline along a straight path is notated as *DV(vehicle-moves-down-straight-path)*. Note that glosses for depicting verbs should include sufficient detail to capture the basic action/event, the position or path of the action, and the entity/entities involved, without being limited to a specific referent. Thus, the depicting verb gloss offered in the previous example is preferable to a gloss such as *DV(Honda-drives-down-Florida-Avenue)*.

3.3.4 *Gestures and emblems*

Gestures and emblems are common in both signed and spoken data and are similar to depicting verbs in that they are often highly iconic (see Goldin-Meadow 2003 and Kendon 2004 for more detailed categorization of gesture in general). For example, gestures can include facial or body mimics (e.g., imitating the angry face of a character in a picture book), common actions (e.g., clapping), reaching with the fully extended arm for a desired object or person (*i.e. gimme*), and locating gestures (e.g., tapping on a chair cushion to invite someone to sit there). In general, these gestures are more conventionalized in form than depicting verbs, and thus occupy a distinct category in our notational conventions. Emblems are even more conventionalized in form than gestures; they are highly recognizable and relatively easy to gloss even for the casual observer (e.g., holding up the index finger to mean *wait-a-minute*). Gestures and emblems are notated with a lowercase *g* or *e*, respectively, followed by a succinct description of the gesture or emblem's meaning enclosed in parentheses (e.g., *g(clapping)*; *g(pouting-face)*; *e(wait-a-minute)*). Many of these gestures and emblems are culture-specific and vary from community to community. In appendix C at the end of this article, we include a list of common gestures and emblems encountered in our data for English and ASL.

3.3.5 *Fingerspelling and namesigns*

Fingerspelled signs, or signs produced using the manual alphabet, are common in some sign languages, particularly ASL, where they occur frequently in the production of even very young signing children. We notate fingerspelled signs with the label *FS* followed by the intended fingerspelled word (rather than an exact letter-by-

letter sequence) in parentheses. For example, if a child intended to fingerspell the word *black* but omitted the letter *c*, such a sign would still be rendered *FS(black)*. In contrast, isolated letters of the manual alphabet are treated as signs and are assigned ID-glosses (e.g., the production of *A* is glossed as *LETTER-A*). Name signs, particularly in ASL, sometimes consist simply of a fingerspelled full or shortened name (e.g., *D-E-B* for one of the co-authors of this article), but this is not normally the case. For the sake of consistency, we notate all name signs using the convention *NS(name)*, where the name of the relevant individual is inserted between the parentheses. Following the capitalization conventions outlined in Section 3.2.3, the first letter of names in these cases is capitalized (e.g., *NS(Debbie)*).

3.3.6 *Nonmanual signals and mouthing*

As rightly pointed out by Baker et al. (2009), nonmanual signals encode essential linguistic information for sign languages and are therefore traditionally included in transcriptions of signed data. Previous versions of our transcription template have included tiers for head movement, body movement, eyebrow position, eyegaze, nose crinkle and mouthing. However, we found that annotating this information for longitudinal, hour-long video sessions was incredibly time-consuming. In our current arrangement, the effort-to-benefit ratio for nonmanual transcription has become even less advantageous, since ELAN allows researchers interested in analyzing nonmanual signals accompanying signed utterances to easily find and view them by clicking on the utterance annotations themselves. Given the importance of nonmanuals to sign language research, we will likely go back and add at least limited information about them to our transcripts in the future, but only once we have made sufficient progress in annotating information relating to the manual component of our signed data.

A notable exception to our current practice of not transcribing nonmanual information relates to the use of mouthed words inserted into a signed utterance, but without any corresponding accompanying sign. In these cases, the mouthed word occupies its own temporal slot, contributing meaning just as a sign would. These cases are glossed using the convention *m(mouthed-word)*. For example, a signer might describe a dress with polkadots by producing the sequence in example (4).

(4) DRESS m(with) DV(polkadots-all-over)

Because the mouthing of the English word 'with' occurs between two signs, we treat it as part of the signed utterance, glossing it on the *signed utterance* tier.

3.3.7 *Reduplicated and held signs*

Reduplication or repetition of sign movement is common in sign languages, and is particularly frequent in child and child-directed signing (Maestas y Moores 1980;

Holzrichter & Meier 2000; Meier et al. 2008). Cases in which a sign is produced with more than the lexically specified number of movement cycles are marked with the symbol [+]. This notation is also employed whenever lexically reduplicated signs are produced with *fewer* than the usual number of movement cycles. For instance, the ASL sign *MOTHER* is normally signed with two movements to the chin; if a child signs it with four movements to the chin, then signs it again with only a single movement to the chin, both of these cases are notated as *MOTHER[+]*. Thus the [+] symbol only indicates that a sign is unusual in some way with respect to cyclicity; it does not specify the number of cycles produced (in contrast to the very similar notation for repetition proposed by Baker et al. 2009). Such details can be separately noted on the corresponding *phonology* tier if desired.

Another modification of form that results in an extended sign duration is perseveration, or the holding of a sign in one location. Returning to the ASL sign *MOTHER* used above, a signer may hold the hand in static contact with the chin at the end of this sign, rather than lowering the hand or moving it into position for the next sign, as is customary. In our notational conventions, held signs are marked by a single underscore enclosed in brackets. Thus, the above example is glossed as *MOTHER[_]*. This notation is also used when signs with repeated hand-internal movement are held in place (with continued hand-internal movement). For example, the ASL sign for *MOM* is produced by contacting the thumb of the hand to the chin and wiggling the four outstretched fingers. If the articulation of this sign is extended longer than usual, based on native speaker intuition, it is glossed as *MOM[_]*.

3.3.8 *Order of notational conventions*

During the course of transcription, a transcriber may find the need to employ more than one of the notational conventions described above to a single gloss. For instance, an unclear sign that repeats will require two notational conventions, [?] and [+]. Since the notation [?] applies to the form of the sign while the [+] notation applies to the production of that sign (regardless of its form), this sign will be transcribed as SIGN[?][+]. In general, the notation conventions that apply to the internal form of a sign or word itself will appear first, followed by the notational conventions that apply to the production of the sign or word.

Sometimes more than one notational convention regarding the production of a sign is required. In these cases, order of notation is determined by order of occurrence. For example, if the sign MOTHER is produced repeatedly and then held for a marked duration (e.g., at the end of the utterance or to maintain the floor while another person interrupts), this is notated as MOTHER[+][_].

3.4 Notational conventions specific to the transcription of speech data

As stated earlier, our transcription conventions for speech are largely adapted from those presented in the CHAT manual developed for the CHILDES project. Our discussion here focuses on just a select few special cases that either occur with high frequency in our data, or that are unique to speech within a Deaf context. For more information, we refer the reader to the CHAT manual, freely downloadable at http://childes.psy.cmu.edu/.

3.4.1 *Standard notations*

In general, speech is represented by standard orthography, regardless of variations in speaker dialect or accent. This is also true of child forms that are noticeably non-target; in these cases, the transcriber enters the standard orthographic form of the word on the *spoken utterance* tier, and then transcribes the form that was actually pronounced, using the IPA, on the corresponding *phonology* tier.

3.4.2 *Special notations: Shortenings and assimilations*

In colloquial speech, speakers often omit certain sounds from words (e.g., pronouncing the English word *because* as simply *'cause*). The CHAT manual refers to these forms as *shortenings*, using parentheses to enclose the part(s) of the word left off by the speaker (e.g., *(a)bout, (be)cause*). For a table with numerous examples of English shortenings, we refer the reader to the CHAT manual (Table 3, Section 6.6.6, p. 47–48).

Table 1. Common English assimilations

Nonstandard	Standard	Nonstandard	Standard
coulda(ve)	could have	mighta	might have
dunno	don't know	need(t)a	need to
dyou	do you	oughta	ought to
gimme	give me	posta	supposed to
gonna	going to	shoulda(ve)	should have
gotta	got to	sorta	sort of
hadta	had to	sorta	sort of
hasta	has to	wanna	want to
hafta	have to	wassup	what's up
kinda	kind of	whaddya	what did you
lemme	let me	whyntcha	why didn't you
lotsa	lots of		

Shortenings are distinct from *assimilations*, another common category of specially notated words. The CHAT manual considers some assimilations (e.g., English *gimme* and *gonna*) common enough that they should be transcribed as pronounced and provides a list of recommended spellings for them (Table 4, Section 6.6.7, p. 48), reproduced in Table 1.

3.4.3 *Sound effects and imitations*

Sound effects are noted using the symbol *&=* followed by a label representing the relevant sound. The CHAT manual provides a table of the most common sound effects for English (p. 59), some of which are reproduced in Table 2 below.

Table 2. Common sound effects

&=cries	&=gasps	&=groans	&=growls	&=grunts	&=hums
&=laughs	&=moans	&=mumbles	&=roars	&=sighs	&=sings
&=squeals	&=vocalizes	&=whimpers	&=whines	&=whistles	&=yells

When a participant produces a sound imitating that made by another person, animal, or machine, we use the label *imit* inserted between the *&=* notation and the label for the imitated entity, as illustrated in Table 3 below.

Table 3. Imitation sound effects

&=imit:baby	&=imit:dog	&=imit:lion	&=imit:motor	&=imit:plane

3.4.4 *Speech without full voice*

We frequently encounter speech that is not fully voiced in our video sessions, particularly during code-blended utterances. We transcribe whispered speech on the *spoken utterance* tier, but enter the label *whispered* on the phonology tier for that utterance. Similarly, completely voiceless speech (i.e., mouthed only), unaccompanied by signs, is also transcribed on the *spoken utterance* tier, accompanied by the label *–voice* on the *spoken phonology* tier for the relevant utterance. These instances are distinct from the mouthing behaviors described in Section 3.3.6 in that the latter occur as insertions within clearly signed utterances, while the former do not. For example, if a child hands his mother a book and signs a request that she read it to him, she might mouth 'okay' as she opens the book to begin reading. We recognize that these cases and those described in Section 3.3.6 are fairly simple examples of mouthing behavior that occur in signed or code-blended discourse. More complicated cases may require modification to our current notational conventions.

Occasionally, parents or children employ what is known as *Deaf voice*, a specific speech quality associated with Deaf speakers. Deaf voice ranges widely in

intelligibility, depending on both the speaker and the listener. Utterances in Deaf voice that are intelligible to the transcriber are simply transcribed as regular speech on the *spoken utterance* tier, with the optional accompanying label *Deaf voice* on the *spoken phonology* tier. Utterances in Deaf voice that are not intelligible to the transcriber are transcribed as *&=vocalize* on the *spoken utterance* tier, borrowing the notational convention employed for sound effects.

4. Coding and searching in ELAN

All our data analyses are conducted using the ELAN transcripts so that researchers can view the signed/spoken utterances on video while coding for structures of interest. Typically, we add a tier or multiple tiers, sometimes using the Controlled Vocabulary feature offered by ELAN, coding for such attributes as blend type, word order, target cyclicity, etc. Because these codes are kept on separate tiers, it is possible to re-visit coding when necessary and review the signed utterance on which a particular code was based. It is also possible to use the same transcript for multiple analyses, viewing only the tiers relevant for each analysis.

ELAN offers complex search tools, allowing searches and changes across multiple transcripts. ELAN also provides detailed information on annotation frequency, duration, etc., through the feature *annotation statistics* as well as other viewers on the ELAN window. The full functionality of ELAN facilitates a wide range of potential analyses. In addition, the transcript tiers can be exported into a variety of other programs, such as CLAN and Microsoft Excel, for further classification and quantification.

4.1 Identifying bimodal utterances

In our analyses, we sometimes focus on examining the structure of signed utterances, while at other times we focus on examining the structure of spoken utterances. Quite often, however, we are interested in examining utterances that are code-blended, expressed by both sign and speech. Sometimes the sign and speech convey the same information (*congruent blends*); occasionally the sign and speech each contribute a different component of the message (*complementary blends*) as also observed by Emmorey et al. (2008) and van den Bogaerde & Baker (2009). Our transcripts do not directly notate blends. The signed portions are transcribed separately from the spoken portions. However, blends can be identified by examining whether both signed and spoken utterances overlap within a single annotation on the free translation tier. If a researcher has a criterion of overlap between sign and speech to identify blends, these overlaps can be easily determined using

the ELAN program, provided the extent of the signed and spoken utterances is accurately annotated, as discussed previously.

5. Comparison with other systems

We chose to use ELAN for our transcription tool, as noted already, mainly because of the facility ELAN provides for examining both video and multiple tiers of transcript simultaneously. This is very useful for sign language research in general, but especially so for our work on bimodal bilinguals, because it allows a clear graphic display of the timing of signed and spoken utterances, as well as other overlapping features such as comments and analysis codes.

We chose to maintain a number of our well-known conventions for sign transcribing, because of their usefulness and familiarity both for our own work and for the broader community. However, we have also taken a major step in the adoption of ID-glosses and associated transcription philosophy. We do not expect a reader to be able to reproduce exactly the signs that were made based on the written transcript. This is not necessary, because the video is available; nor is it desirable, because linguistic research on sign languages is not yet at the point where analysis can be conducted on the basis of transcript alone. The desire to combine familiar notational conventions with consistency and searchability drives our approach and our choices.

Due to these choices, our notation is unfortunately not currently fully compatible with the CHAT format used by CLAN programs in CHILDES, as previously noted. We adopted CHAT conventions where possible, but when the CHAT conventions contradicted our other needs, the latter prevailed. For example, all symbols modifying basic glosses must be typed adjacent to the previous gloss, without a space, in order for the ELAN tokenization process to correctly segment utterances into words. In some cases, conflicts were found between our established sign language conventions and those of CHAT. For example, the use of the + symbol for repetition is long-standing in sign linguistics, and we chose to keep this use of this symbol for both signed words and spoken words, to maintain consistency.

Each of these decisions is made within the context of our overall transcription philosophy and guiding principles. While other researchers may have different priorities, we hope that articulating our decisions as we have here will allow for productive comparison and discussion, and clarity for readers of our research reports.

6. Concluding remarks

The goal of this article has been to provide an overview of the methodological aspects of our project on bimodal bilingual language acquisition. Given that this project spans three research labs on two continents, one of our major priorities is to establish consistent and efficient practices for data collection and transcription. An additional goal is that our transcripts capture sufficient detail to be useful for a broad range of research questions while avoiding notational conventions that inherently interpret the data within a specific theoretical framework. Although the practices outlined in this article were originally developed to address research questions related to bimodal bilingual development, we hope they will be of use to researchers studying sign language acquisition in other contexts as well.

Acknowledgments

We warmly thank the bimodal bilingual children and their families who are participating in our project, as well as the research assistants who tirelessly film and transcribe the data for our research. The project described in this article is supported by Award Number R01DC009263 from the National Institute on Deafness and Other Communication Disorders (United States). 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 (United States). Additional support was provided by a Gallaudet priority grant, and by CNPQ Grant (Brazil) #200031/2009–0 and #470111/2007–0.

References

Baker, A., Bogaerde, B. van den & Woll, B. (2009). Methods and procedures in sign language acquisition studies. In A. Baker & B. Woll (Eds.), *Sign language acquisition*, 1–49. Amsterdam: John Benjamins.

Baker, A. & Woll, B. (Eds.) (2009). *Sign language acquisition*. Amsterdam: John Benjamins.

Berent, G.P. (2004). Sign language-spoken language bilingualism: Code mixing and mode mixing by ASL-English bilinguals. In. T.K. Bhatia & W.C. Ritchie (Eds.), *The handbook of bilingualism*, 312–335. Malden, MA: Blackwell.

Bhatia, T.K. & Ritchie, W.C. (1999). Language mixing and second language acquisition: Some issues and perspectives. In E.C. Klein & G. Martohardjono (Eds.), *The development of second language grammars: A generative approach*, 241–267. Amsterdam: John Benjamins.

Bishop, M. (2006). *Bimodal bilingualism in hearing, native users of American Sign Language*. Unpublished doctoral dissertation. Washington, DC: Gallaudet University.

Bogaerde, B. van den & Baker, A.E. (2005). Code-mixing in mother-child interaction in deaf families. *Sign language & linguistics* n° 8(1–2), 151–174.

Bogaerde, B. van den & Baker, A.E. (2009). Bimodal language acquisition in Kodas (kids of deaf adults). In M. Bishop & S.L. Hicks (Eds.), *Hearing, mother-father Deaf: Hearing people in Deaf families*, 99–131. Washington, DC: Gallaudet University Press.

Chen Pichler, D., Quadros, R.M. & Lillo-Martin, D. (2010). Effects of bimodal production on multi-cyclicity in early ASL and LSB. Supplement to the proceedings of the 34th Boston University conference on language developement (BUCL). Casadilla press.

Emmorey, K., Borinstein, H.B., Thompson, R. & Golan, T.H. (2008). Bimodal bilingualism. *Bilingualism: Language and cognition* n° 11(1), 43–61.

Genesee, F. (1989). Early bilingual development: One language or two? *Journal of child language* n° 16, 161–79.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Hochgesang, J.A. (in progress). *Is That a Y or … ? Representation of hand configuration data in different notation systems for child acquisition of ASL*. Unpublished doctoral dissertation. Washington, DC: Gallaudet University.

Hochgesang, J.A. (2009). Is there a sentence in ASL? Insights on segmenting data in ASL. Presented at the *Sign Language Corpora: Linguistic Issues* workshop at University College London, London, England, July 24–25, 2009.

Holzrichter, A.S. & Meier, R.P. (2000). Child-directed signing in American Sign Language. In C. Chamberlain, J.P. Morford & R. Mayberry (Eds.), *Language acquisition by eye*, 25–40. Mahwah, NJ: Lawrence Erlbaum Associates.

Johnson, J. M., Watkins, R. V. & Rice, M.L. (1992). Bimodal bilingual language development in a hearing child of deaf parents. *Applied psycholinguistics* n° 13(1), 31–52.

Johnston, T. (in preparation). *From archive to corpus: Transcription and annotation in the creation of signed language corpora*, manuscript. Department of Linguistics, Macquarie University, Australia.

Johnston, T. (2001). The lexical database of Auslan (Australian Sign Language). *Sign language and linguistics* n° 4(1/2), 145–169.

Johnston, T. (1991). Transcription and glossing of sign language texts: Examples from Auslan (Australian Sign Language). *International journal of sign linguistics* n° 2(1), 3–28.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Liddell, S.K. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge: Cambridge University Press.

Liddell, S.K. & Johnson, R. (1989). American Sign Language: The phonological base. *Sign Language Studies* n° 64, 195–277.

Lillo-Martin, D. & Chen Pichler, D. (2008). Development of sign language acquisition corpora. *Proceedings of the 6th International Conference on Language Resources and Evaluation, LREC 2006, Marrakech, Morocco*. Paris: ELRA.

Lillo-Martin, D., Quadros, R.M., Koulidobrova, H. & Chen Pichler, D. (in press). Bimodal bilingual cross-language influence in unexpected domains. *Proceedings of the 2009 Conference on Generative Approaches to Language Acquisition (GALA)*. Cambridge Scholars Press.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Maestas y Moores, J. (1980). Early linguistic environment: Interaction of deaf parents with their infants. *Sign language studies* n° 26, 1–13.

Meier, R., Mauk, C., Cheek, A. & Moreland, C. (2008). The form of children's early signs: Iconic or motoric determinants? *Language learning and development* n° 4(1), 63–98.

Meisel, J. M. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the life span: In health and in pathology*, 13–40. Cambridge: Cambridge University Press.

Messing, L. (1999). Two modes — two languages? In L.S. Messing & R. Campbell (Eds.) *Gesture, speech and sign*, 183–200. Oxford: Oxford University Press.

Mulrooney, K. (2006). *The structure of personal narratives in American Sign Language*. Unpublished doctoral dissertation. Gallaudet University, Washington DC.

Nicoladis, E. & Genesee, F. (1997). Language development in preschool bilingual children. *Journal of speech-language pathology and audiology* n° 21(4), 258–70.

Ochs, E. (1979). Transcription as theory. In E, Ochs & B. Schieffelin (Eds.), *Developmental pragmatics*, 42–72. New York: Academic Press.

Paradis, J. & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in second language acquisition* n° 18(1), 1–25.

Paradis, J. & Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of input? *Journal of child language* n° 30, 371–393.

Petitto, L. A., Katerelos, M., Levi, B., Gauna, K., Tetrault, K. & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of child language* n° 28(2), 453–496.

Pizzuto, E. & Pietrandrea, P. (2001). The notation of signed texts: Open questions and indications for further research. *Sign language and linguistics* n° 4(1–2), 29–45.

Prinz, P. M. & Prinz, E.A. (1981). Simultaneous acquisition of ASL and spoken English, phase II. *Sign language studies* n° 30, 8–88.

Quadros, R.M., Lillo-Martin, D. & Chen Pichler, D. (2010). Two languages but one computation: Code-blending in bimodal bilingual development. To be presented at the conference on Theoretical Issues in Sign Language Research; Purdue University, Indiana; October 2010.

Schiff-Myers, N. (1988). Hearing children of deaf parents. In D. Bishop & K. Mogford (Eds.), *Language development in exceptional circumstances*, 47–61. New York: Churchill Livingstone.

Singleton, J. L. & Tittle, M.D. (2000). Deaf parents and their hearing children. *Journal of Deaf studies and Deaf education* n° 5(3), 221–36.

Slobin, D. (2008). Breaking the molds: Signed languages and the nature of human language. *Sign language studies* n° 8(2), 114–130.

Slobin, D. (2006). Issues of linguistic typology in the study of sign language development of deaf children. In B. Schick, M. Marschark & P.E. Spencer (Eds.), *Advances in the sign language development of Deaf children,* 21–45. New York: Oxford University Press.

Slobin, D. I., Hoiting, N., Anthony, M., Biederman, Y., Kuntze, M., Lindert, R., Pyers, J., Thumann, H. & Weinberg, A. (2001). Sign language transcription at the level of meaning components: The Berkeley Transcription System (BTS). *Sign Language & Linguistics* n° 4, 63–96.

Snyder, W. (2007). *Child language: The parametric approach*. Oxford: Oxford University Press.

Stromswold, K. (1996). Analyzing children's spontaneous speech. In D. McDaniel, C. McKee & H.S. Cairns (Eds.), *Methods for assessing children's syntax*, 23–54. Cambridge, MA: MIT Press.

Résumé

Cet article présente l'adaptation des outils d'analyse linguistique existant pour les langues des signes à la description de données bilingues bimodales. Nous annotons dans ELAN des données écologiques recueillies dans le cadre d'études longitudinales d'enfants entendants de parents sourds. Ces enfants sont en train d'acquérir soit la Langue des Signes Américaine (ASL) et l'anglais américain (AE), soit la Langue des Signes Brésilienne (Libras) et le portugais brésilien (BP). Dans le but de créer un corpus utilisable pour des analyses variées sur le développement linguistique, nous visons à établir un processus systématique pour la transcription des énoncés en langues signées et en langues orales. Ce chapitre décrit nos conventions de transcription et le raisonnement sur lequel elles sont fondées. Nous espérons que ce document servira à d'autres chercheurs qui souhaiteraient adopter les conventions présentées ici ou les modifier selon leurs besoins.

Appendix A: Summary of notational conventions

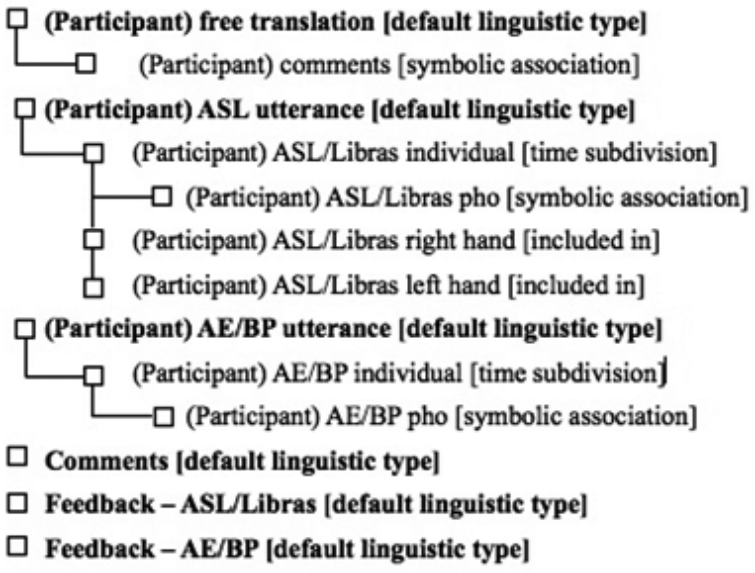
Item	Convention	Example
Capitalization and punctuation	Only proper names are capitalized; no end-of-sentence punctuation	*doggie go away*
Sign language glosses	All capital letters; multi-word glosses must be linked with hyphens	*RABBIT* *GIVE-UP*
RH/LH hand	When the two hands articulate different signs overlapping each other, use these tiers	
Pointing to people	Use the label IX followed by referent in lowercase letters (except for proper names), enclosed in parentheses	*IX(self)* *IX(mother)*
Pointing to objects	Use the label IX followed by referent in lowercase letters, enclosed in parentheses (hyphens between words).	*IX(dog)* *IX(pig-puzzle-piece)*
Pointing to locations	IX followed by location in lowercase letters, enclosed in parentheses (hyphens between words)	*IX(outside)* *IX(inside-refrigerator)*
Possessives	Use the label POSS followed by referent in lowercase letters (except for proper names), enclosed in parentheses (hyphens between words)	*POSS(self)* *POSS(Maria)*
Reflexives	SELF followed by referent in lowercase letters (except for proper names), enclosed in parentheses (hyphens between words)	*SELF(self)* *SELF(mother)*
Indicating verbs	Provide the ID-gloss for the sign only; do not add information about referents	*GIVE* *GO*

Depicting verbs	Gloss with label *DV* followed by description in parentheses (hyphens between words)	*DV(bird-sits-on-tree)*
Fingerspelled words	Gloss with label *FS* followed by the unhyphenated word in parentheses	*FS(Nokia)* *FS(apple)*
Name signs	Gloss with label *NS* followed by name in parentheses. Codenames are used to protect children's privacy	*NS(Debbie)* *NS(BEN)*
Repeated signs	Add [+] (the plus symbol enclosed in brackets) to end of gloss	*MOTHER[+]*
Held signs	Add [_] (underscore enclosed in brackets) to end of gloss	*MOTHER[_]*
Pause within utterance	Represent pauses with # (a single hatch mark) attached to previous gloss	*IX(self) CHOOSE# RED*
Interruption	Add / (slash) to end of last gloss before interruption	*WANT/*
Self-interruption	Add // (double slash) to end of last gloss before interruption	*WANT//*
Retracing without correction	Add [/] (slash enclosed in brackets) to end of last gloss before retracing	*IX(self) WANT[/] IX(self) WANT IX(toy)*
Retracing with correction	Add [//] (double slash enclosed in brackets) to end of last gloss before retracing	*IX(self) WANT[//] IX(self) DON'T-WANT IX(toy)*
Retracing with reformulation	Add [///] (triple slash enclosed in brackets) to end of last gloss before retracing.	*IX(self) WANT[///] IX(toy)* *IX(self) WANT*
Trailing off	Add … (ellipsis) to end of last gloss before trailing off	*WANT…*
Gesture	Gloss with label *g* followed by concise meaning in parentheses	*g(angry-face)*
Emblem	Gloss with label *e* followed by name of emblem in parentheses	*e(come-here)*
Showing	Gloss with label *show* followed by name of object shown in parentheses	*show(toy)*
Mouthing	Gloss with *m* followed by word mouthed	*m(okay)*
Word is not clear (meaning)	Add [?] (question mark in brackets) to end of unclear gloss; add description of phonetic form on phonological tier if necessary	*WANT APPLE[?] PLEASE*
Word is not clear (alternative)	Type best guess first as gloss, followed by [=?ALTERNATIVE] (equal sign followed by question mark and alternative gloss in brackets)	*WANT APPLE[=?ONION]*

Word is not clear (form)	Gloss each unclear word as YYY (there may be more than one). Add description on phonological tier of each YYY gloss	*WANT YYY PLEASE*
Word is not clear	Gloss each unclear word as XXX (there may be more than one)	*WANT XXX PLEASE*
Shortenings	Put the unpronounced part of a word in parentheses	*(be)cause*
Sound effects	Use &= (ampersand and equal sign) before the sound (such as cries, laughter, and whistles)	*&=cries* *&=laughs*
Imitations	Use &=imit: (ampersand, equal sign, *imit*, and colon) before the sound imitation (such as sounds imitating another person, animal or machine)	*&=imit:baby* *&=imit:plane*

Appendix B : Hierarchical organization of our Bimodal Bilingual template

Parent (independent) tiers are shown in bold type (insert participant pseudonym in parentheses), while child (dependent) tiers are marked by lines indicating their relationships. There are some tiers dependent on dependent tiers (e.g., the phonological tiers are dependent on the individual tiers which, in turn, are dependent on the utterance tiers). The relationship between each child tier and its parent tier is specified in square brackets (ELAN linguistic stereotype). With the exception of the *Comment*, *Feedback-ASL/Libras* and *Feedback-AE/BP tiers*, the tier hierarchy shown here is repeated for each participant.

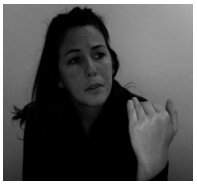


Appendix C: Common gestures and emblems

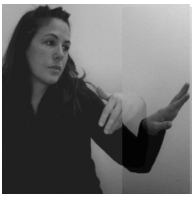
***Yes* and *no* gestures**	Headnods and shakes used in isolation (without accompanying signs) are transcribed as gestures, *i.e. g(yes)* or *g(no)*.
Showings	A demonstrating gesture with object in hand can be transcribed with *show(object)*.

show(cookie)

Emblems Verb-like, highly conventionalized emblems are transcribed as *e(meaning-of-emblem)*.



e(come-here)
palm toward signer, bending finger(s) in

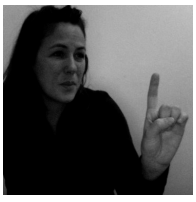


e(go-away)
flat hand flicks away from signer

e(no)
index finger wags back and forth

e(shh)
index finger to lips

e(stop-it)
full open palm pushed forward sharply once



e(wait-a-minute)
index finger held up toward addressee

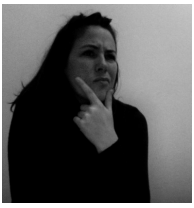
e(hold-on)
full open palm repeatedly pushed forward in small movements

e(stay-there)
full open palm pushed forward in one movement

e(cut-it-out)
open hand flicks palm toward ground

Other gestures Gestures are also conventionalized, but not to the same extent as emblems. These are transcribed as *g(meaning-of-gesture)*.

g(clapping)

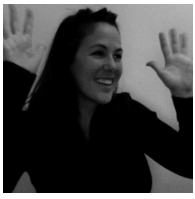


g(hmm)
L hand on mouth with finger tap

g(huh)
two flat hands palm up

g(oops)
flat hand to mouth

g(reaching)
arm fully extended



g(yea)
Deaf applause (hand oscillation)

g(hey)
hailing gesture

g(well)
two flat hands flip to palm up

g(here)
tapping a location with hand

g(ow)
flat hand, bent wrist, shakes as if in pain

g(voilà)
open hand indicates object

g(pshaw)
flat hand faps downward dismissively

g(see-what-I-mean)
hands on hips, facial expression with raised brows

Author's addresses

Deborah Chen Pichler
Department of Linguistics
Gallaudet University
800 Florida Ave NE
Washington DC 20002
USA

deborah.chen.pichler@gallaudet.edu

Julie A. Hochgesang
Department of Linguistics
Gallaudet University
800 Florida Ave NE
Washington DC 20002
USA

julie.hochgesang@gallaudet.edu

Diane Lillo-Martin
University of Connecticut
Department of Linguistics
337 Mansfield Road Unit 1145
Storrs CT, 06269-1145
USA

lillo.martin@uconn.edu

Ronice Müller de Quadros
Universidade Federal de Santa Catarina
Centro de Comunicação e Expressão
Campus Reitor João David Ferreira Lima
Bairro Trindade
Florianópolis
Santa Catarina
Brasil
CEP 88040-970

ronice@cce.ufsc.br