### Annotation of Child Language Corpora:
A comparison of two methods with special emphasis on bilingual data

**Diane Lillo-Martin & Debbie Chen Pichler**  
University of Connecticut (USA) & Gallaudet University (USA)

Collaborators: Ronice Müller de Quadros & Julie Hochgesang  
Universidade Federal de Santa Catarina (Brazil) & Gallaudet University (USA)

**Sign Linguistics Corpora Network**  
Workshop 3: Annotation  
Stockholm, Sweden 14-16 June 2010

#### Appendix A: Summary of notational conventions

<table>
<thead>
<tr>
<th>Item</th>
<th>Convention</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Capitalization and punctuation</strong></td>
<td>Only proper names are capitalized, no end-of-sentence punctuation</td>
<td>(gaggle go away)</td>
</tr>
<tr>
<td><strong>Sign language glosses</strong></td>
<td>All capital letters; multi-word glosses must be linked with hyphens</td>
<td>EABBET GIVE-UP</td>
</tr>
<tr>
<td><strong>RH/LH hand</strong></td>
<td>When the two hands articulate different signs overlapping each other, use these signs</td>
<td>X(self) (mother)</td>
</tr>
<tr>
<td><strong>Pointing to people</strong></td>
<td>Use the label IX followed by referent in lowercase letters (except for proper names), enclosed in parentheses</td>
<td>X(dog) X(puzzle-piece)</td>
</tr>
<tr>
<td><strong>Pointing to objects</strong></td>
<td>Use the label IX followed by referent in lowercase letters, enclosed in parentheses (hyphens between words),</td>
<td>X(inside) X(refrigerator)</td>
</tr>
<tr>
<td><strong>Possessives</strong></td>
<td>Use the label POSSS followed by referent in lowercase letters (except for proper names), enclosed in parentheses (hyphens between words)</td>
<td>POSSS(self) POSSS(Maria)</td>
</tr>
<tr>
<td><strong>Reflexives</strong></td>
<td>SELF followed by referent in lowercase letters (except for proper names), enclosed in parentheses (hyphens between words)</td>
<td>SELF(self) SELF(mother)</td>
</tr>
<tr>
<td><strong>Indicating verbs</strong></td>
<td>Provide the ID-gloss for the sign only, do not add information about referents</td>
<td>GIVE GO</td>
</tr>
<tr>
<td><strong>Depicting verbs</strong></td>
<td>Gloss with label JV followed by description in parentheses (hyphens between words)</td>
<td>OR(bird-sits-on-tree)</td>
</tr>
<tr>
<td><strong>Fingerspelled words</strong></td>
<td>Gloss with label FS followed by the unhyphenated word in parentheses</td>
<td>P=Nokia N=Apple</td>
</tr>
<tr>
<td><strong>Name signs</strong></td>
<td>Gloss with label NN followed by name in parentheses. Codenames are used to protect children’s privacy</td>
<td>X=Debbie X=N(BEN)</td>
</tr>
</tbody>
</table>

---

#### Annotation of Child Language Corpora

### Repeated signs
Add ```[*]``` (the plus symbol enclosed in brackets) to end of gloss

### Held signs
Add ```[~]``` (underbrace enclosed in brackets) to end of gloss

### Pause within utterance
Represent pauses with ```//``` (a single hatch mark) attached to previous gloss

### Interruption
Add ```[/]``` (slash enclosed in brackets) to end of last gloss before interruption

### Self-interruption
Add ```//``` (double slash enclosed in brackets) to end of last gloss before interruption

### Retracing without correction
Add ```///``` (triple slash enclosed in brackets) to end of last gloss before retracing

### Retracing with correction
Add ```[/]``` (slash enclosed in brackets) to end of last gloss before retracing

### Retracing with reformulation
Add ```///``` (triple slash enclosed in brackets) to end of last gloss before retracing

### Trailing off
Add … (ellipsis) to end of last gloss before trailing off

### Texture
Gloss with label g followed by concise meaning in parentheses

### Emblem
Gloss with label e followed by name of word as problem in parentheses

### Slowing
Gloss with label s followed by name of object shown in parentheses

### Mouthing
Gloss with m followed by word mouthed

### 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

### 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)

### Word is not clear (form)
Gloss each unclear word as YYY (there may be more than one). Add description of phonological tier of each YYY gloss

### Word is not clear
Gloss each unclear word as XXX (there may be more than one). Add description of phonological tier of each XXX gloss

### Shortenings
Put the unpronounced part of a word in parentheses

### Sound effects
Use &```(````)``` (ampersand and equal sign) before the sound (such as cries, laughter, and whooshes)

### Imitations
Use &```(````)``` (ampersand, equal sign, initial and colon) before the sound imitations such as sounds imitating another person, animal or machine
Appendix B: Hierarchical organization of our Bimodal Bilingual template

Parent (independent) tiers are shown marked for [default linguistic type] (insert participant 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-AEBP tiers, the tier hierarchy shown here is repeated for each participant.

<table>
<thead>
<tr>
<th>(Participant) free translation [default linguistic type]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Participant) comments [default linguistic type]</td>
</tr>
<tr>
<td>(Participant) ASL/Libras Utterance [default linguistic type]</td>
</tr>
<tr>
<td>(Participant) ASL/Libras Individual [time subdivision]</td>
</tr>
<tr>
<td>(Participant) ASL/Libras phrase [symbolic association]</td>
</tr>
<tr>
<td>(Participant) ASL/Libras right hand [included]</td>
</tr>
<tr>
<td>(Participant) ASL/Libras left hand [included]</td>
</tr>
<tr>
<td>(Participant) AEBP Utterance [default linguistic type]</td>
</tr>
<tr>
<td>(Participant) AEBP individual [time subdivision]</td>
</tr>
<tr>
<td>(Participant) AEBP phrase [symbolic association]</td>
</tr>
<tr>
<td>Comments [default linguistic type]</td>
</tr>
<tr>
<td>Feedback - ASL/Libras [default linguistic type]</td>
</tr>
<tr>
<td>Feedback - AEBP [default linguistic type]</td>
</tr>
</tbody>
</table>

These appendices are from Chen Pichler, Hochgesang, Lillo-Martin & Quadros (2010).

Acknowledgments

- Warm thanks to the participating bimodal bilingual children and their families; and to our hard-working research assistants.
- Financial support comes from: Award #R01DC009263 from the National Institute on Deafness and Other Communication Disorders; the Gallaudet Research Institute; and CNPq (Brazilian National Council of Technological and Scientific Development) Grant #200031/2009-0 and #470111/2007-0.

Research Employing the Bibibi data:


Works Cited


