**Coding of Syntactic Unit (SU) and Mean Length of Utterance (MLU) in ASL**

This document describes ASL MLU coding methods currently in use at the University of Connecticut in the Sign Linguistics and Language Acquisition (SLLA) Lab (PI: Prof. Diane Lillo-Martin). This coding guide is still being fine-tuned as we develop ASL MLU as a measure of syntactic complexity in ASL. This document also contains the current guidelines for transcribing and segmenting syntactic units/utterances in ASL for use with ASL MLU, ASL-IPSyn, and other measures of linguistic development and complexity. Please contact the lab manager for the latest version of this coding guide. Contact information and project information are available on the SLLA Lab website at <https://slla.lab.uconn.edu>. Many ASL glosses used in this guide are from ASL SignBank. More information about ASL SignBank, including a live link to ASL SignBank, is also available on the SLLA Lab website. More detailed information about annotation conventions used with this project is also available at this [Dropbox Showcase link](https://showcase.dropbox.com/s/SLAAASh-data-annotation-protocols-RBYwuaOZame6b82tzEVk7).

**1. What counts as a syntactic unit/utterance (SU)?**

* 1. Use the Syntactic Unit tier to transcribe an utterance to be analyzed
  2. Use a preponderance of evidence from three areas to help determine what is a single syntactic unit:
     1. Syntax – “The largest linguistic unit that is held together by rigid grammatical rules” (Track, 1999:273).
     2. Semantic – “A combination of words expressing a complete thought and making complete sense” (Gartside, 1981:239)
     3. Prosodic – Sentences breaks can be indicated by a single, or some combination of the following prosodic cues: lowered hands, a pause, lengthening of a sign, a hold, a blink, a nod, a change in eyebrow height, and/or a shift in head or body position.
     4. See Fenlon et al. (2007) & Crasborn (2007) for more discussion.
  3. One syntactic unit consists of a single main clause, with any number of embedded clauses, and up to one conjoined clause.
  4. When there is no manual conjunction sign, only conjoined clauses without long prosodic breaks and in which there is some meaning connection between the two clauses should count as a single syntactic unit. This is to help avoid run-on sentences that artificially inflate MLU.

Consider examples (i) and (ii) below. In (i), there is no clear prosodic break (indicated by the lack of a comma), and there is a relationship between the action of going to the store and buying the bread (at the store). In example (ii), there is a prosodic break (indicated by the presence of the comma) and no clear relationship between the two actions, besides a possible temporal relationship. Under our guidelines (i) is considered one syntactic unit all together while (ii) is considered two separate syntactic units.

* + 1. IX\_1 GOix SHOP BUY BREAD
    2. IX\_1 GOix SHOP, WATCH MOVIE
  1. From the CHAT manual section 9 (pp 57-60): In instances where a child repeats a word/sign such as “milk milk milk milk”, it is misleading to compute this as one utterance with a MLUw of 4 words. Instead, a better approach is to count this as 4 separate syntactic units, each with a MLUw of 1 word. Additionally, when a child is listing words without any evidence of syntactic relationship, such as “red, white, blue, green”, the same procedure of counting each word as a separate utterance should be followed to avoid inflation of MLU.
     1. For instances of counting or saying the ABCs, the whole episode should be grouped into one syntactic unit such as “ONE TWO THREE FOUR” as opposed to “ONE” “TWO” “THREE” “FOUR”. This one syntactic unit is excluded from MLU as a routine under our guidelines. Our annotation method also clarifies the exact extent of the routine being produced.
     2. For instances of listing numbers, letters, items, or any other non-routine sequence, each item should be separated into its own syntactic unit to allow for potential inclusion in analyses. For example, each of these items produced consecutively (with no evidence of syntactic relationship between them) would be divided into a 1-word syntactic unit: “HORSE” “BIRD” “MONKEY” “LETTER-R” “SIX” “FOUR” “SEVEN”.
  2. Do not count actions towards a syntactic unit: leave them off of the tier entirely.
     1. Exclude actions transcribed typically as &=reaches, &=jumps, and so on as well as actions, including attention-getting actions which may be transcribed as actions beginning with “&=”, or as one of two specific ASL SignBank glosses i(hey) and E(here).

1. **How many utterances are analyzed for MLU?**
   1. Up to 100 MLU-analyzable syntactic units (fewer are used if there are not 100 such units in the session)
   2. If the child is interacting with a less familiar interlocutor or in a less familiar situation (location, equipment, etc.), the first 10-25 utterances are skipped over. This is not necessary for familiar situations.
2. **What types of utterances are excluded (and included) for MLU?**
   1. Utterances containing XXX or YYY
      1. **Include**utterances with only one XXX/YYY. Count this XXX/YYY as one word and one morpheme.
      2. **Exclude** utterances with more than one XXX/YYY.
      3. **Exclude** completely unintelligible utterances containing only XXX/YYY.
   2. Imitations of an interlocutor’s utterance
      1. **Exclude exact** imitations of an interlocutor’s utterance.
      2. **Include partial** imitations of an interlocutor’s utterance.
      3. **Include partial** imitations of an interlocutor’s utterance that add additional signs.
      4. Note: For un-transcribed interlocutors, be sure to pay attention to the seconds preceding each utterance of the target participant in order to identify potential exact imitations as opposed to partial imitations.
   3. **Exclude** exact repetitions of a signer’s own utterance
   4. Utterances containing IX**.** *Some IXes may be produced with a looser handshape than an adult-like production. These should be treated the same as any other IX produced by the signer.*
      1. **Exclude** utterances consisting of IX produced in isolation
      2. **Include** utterances with an IX produced alone that is modified to indicate more detailed location information (e.g., around the corner)
      3. **Include** utterances with a series of IX
      4. **Include** utterances with an IX and other lexical signs
      5. Continue to **include** IX when apparently being used as a filler similar to English “like”, if the utterance would have been included
         1. e.g. MOTHERstr ANDasym IX\_1 GOix IX\_1 STORE IX\_1
   5. **Exclude** routines, interjections and actions
      1. See appendix for list of routines
      2. Interjections are transcribed with i(xxx) according to SLAASh conventions. See exact glosses in ASL SignBank.
      3. **Includeinterjections** transcribed as i(xxx) and **routines** when used in combination with other linguistic elements.
         1. **Exception**: Always **exclude** i(hey), E(here), and other attention-getting measures whether combined with other linguistic elements or produced alone.
      4. Actions are transcribed with &=xxx according to SLAASh conventions
         1. Some actions may be transcribed improperly and should be **excluded** regardless of how they were transcribed, for example, E(give-me) instead of &=reaches
         2. As discussed above (section 1f), **exclude actions** even when used in combinations with other linguistic elements.Do not transcribe actions on the Syntactic Unit tier at all.
   6. **Include** emblems transcribed as E(xxx); they are treated as regular signs and glossed exactly as shown in ASL SignBank.
   7. **Exclude** incomplete utterances
      1. Utterances that are partially off-camera and may have had additional elements before or after what the camera has captured.
         1. Use e.g. “…SIGN” or “SIGN…” to mark the incompleteness
      2. Trailing off marked by …
      3. Interruption marked by /
      4. Self-interruption marked by //
   8. **Include** utterances with retracing or reformulation marked by [/], [//] or [///]
      1. Only count the final version of the retraced or reformulated elements of the utterance. (Elements which have been retraced or reformulated over may be enclosed in <angle brackets> for ease of reading, though this convention is not uniformly followed.) For example, these example syntactic units/utterances would each have a MLUw of 3:
         1. <IX(MOT) WANT>[//] IX\_1 WANT EAT
         2. IX(FAT)[///] EAT IX(FAT) WANT
         3. IX\_1[/] IX\_1 LIKE FROG
3. **What counts as a morpheme in a lexical sign?**
   1. **Do not count** nonmanual markers. *Nonmanuals such as i(yes) and i(no) can be included in the Syntactic Unit tier, but they aren’t required to be included in the syntactic unit, and they will not count toward MLU in any case. In our corpus, nonmanuals are typically not visible due to the camera angles and film quality. Therefore, we do not consistently transcribe or count nonmanual markers.*
   2. **Do not count** two-handed signs twice. *They should be counted once overall.*
      1. BALL signed on both hands counts as 1 word & 1 morpheme
      2. LOOK signed on both hands counts as 1 word & 1 morpheme, though with additional contextual meaning this could be coded as 2 depicting signs indicating two sets of eyes looking. Without evidence of this reading, code similar cases conservatively as 1 word & 1 morpheme.
      3. Two simultaneous but unique signs should be **counted separately**.
   3. Each individual sign that makes up a compound sign counts as a morpheme in the one sign. This means a compound sign will often be 1 word and 2 morphemes.
      1. DRIVER (CAR+PERSONb)
      2. PARENTS (MOTHER+FATHER)
      3. SNOWMAN (SNOW+MAN)
      4. **Do not count** additional morphemes in signs of potentially compound origin that are not transparently a compound and are not used productively, such as ONE-MOREix
   4. Verbal Agreement. *Only count agreement if it is clearly visually different from the appearance of the unmarked no-agreement form in the person’s signing.*
      1. Location modification to indicate source/goal. In the example below, *a* and *b* below are used to indicate different locations
         1. SCHOOL-a CHURCH-b a-GOix-b
      2. A verb produced in the location where an action should be performed. Sometimes similar items will be coded as depicting signs (DS), but the decision of verb in location vs. DS will be determined on a case-by-case basis. If an item is coded as a DS then instead follow the guidelines for how depicting signs are coded (see below, pt. 5).
         1. LEAVE-ALONE [towards item]
         2. JUMP [produced above child’s bed]
      3. Person agreement. *Add 1 morpheme for each agreed location.*
         1. TELL-you (+1)
         2. you-GIVEo-me (+2)
         3. EXPLANATION [towards camera] (+1)
         4. BRUSH-HAIR [on child’s head] (+1)
      4. Plural agreement. *Judge plurality based on context and prosody, not by counting a specific number of repetitions.*
         1. GIVEx[+] towards multiple 3rd-person entities gains an additional plural morpheme as well as regular person agreement
   5. Numerical Incorporation. Add 1 morpheme.
      1. TWO-WEEKS
      2. FOUR-DAYS
      3. SIX-YEARS-OLD
   6. Negative Incorporation. Add 1 morpheme.
      1. DON’T-KNOW
      2. DON’T-LIKE
      3. DON’T-WANT
      4. NOT-NEED
   7. Pronominal Modification
      1. Modifications of IX to add plurality or other aspects of meaning. *Judge possible IX plurality by using prosody.*
         1. IXarc
         2. IX(around-the-corner)
         3. IX(balls)[+]
      2. Spatial modification of pronouns. *Count 1 additional morpheme for each point of reference past the 1st point (as in a regular IX sign).*
         1. TWO-OF-US (+1)
         2. THREE-OF-US (+2)
      3. Agreement of THAT. *Add +1 additional morpheme for the referent.*
         1. THAT [in neutral space] (no added morpheme)
         2. THAT [towards book] (+1)
   8. Repetition for aspect or emphasis or plural marking. *When uncertain about the difference between a verb with plurality (GROW[+]) as opposed to a noun (PLANT), use context and your intuition to determine how to code the sign.*
      1. HUG[+] “I hug everyone.” (+1)
      2. GROW[+] “The trees grow.” (+1)
   9. Other agreement. *Count agreement of other types of sign based on number of added meaning units*.
      1. SAME-AS generally counts as +2, 1 for each endpoint. However, SAME-AS moved in a triangle pattern to show 3 agreed things would get +3. SAME-AS moved in an arc around an area would only count as +1.
      2. Signs such as AREA produced over a specific region of space as opposed to in neutral area gains +1 morpheme for the spatial agreement.
4. **How are depicting signs (DS) coded?**
   1. **DS(ca) - constructed actions**
      1. Count these toward MLU if there is a manual element to the constructed action. In this case, transcribe the sign on the Syntactic Unit tier as DS\_handshape(description) as opposed to DS(ca) with no elaboration.
      2. Do not count these toward MLU if there is no manual element to the constructed action. Exclude the DS(ca) from the utterance, but count the other signs as usual. *As mentioned above,* *nonmanuals are typically not visible in our corpus. Therefore, we do not consistently transcribe nonmanual markers, nor count nonmanual markers toward MLU.*
   2. **What counts as a word when coding depicting signs (DSes)?**
      1. Each handshape that represents an object, entity, or location.
         1. DS\_b2(animal-runs)
         2. DS\_f(curved-thin-pipe)
         3. Both active and passive handshapes each count as a word, e.g. DS\_b2(animal-runs) and DS\_b(slanted-surface).
      2. Two handshapes used together to indicate the size and shape of an object count as one word in total for the two handshapes together. *This is still true whether or not the two handshapes make contact with each other.*
         1. DS\_i(thin-reed)
         2. DS\_c(broad-vertical-pillar)
      3. Two handshapes articulated either symmetrically or alternatingly count as two separate words, each containing their own set of morphemes. *This is still true whether or not the two handshapes make contact with each other.*
         1. DS\_b5(light-flashes) DS\_b5(light-flashes-alternatingly)
         2. DS\_1(antenna-wiggles) DS\_1(antenna-wiggles)
         3. DS\_s(rowing-strenuously) DS\_s(rowing-strenuously)
      4. Two unique handshapes produced concurrently count as two separate words, each containing their own set of morphemes. Also see point 5bi. This includes both active and passive handshapes – movement is not required.
         1. DS\_5(tree) DS\_2(person-approaches-tree)
         2. DS\_b2(person-sits) DS\_1(entity-zooms-by-person)
      5. If a depicting sign is held and referenced over multiple syntactic units, count that sign (and all of its morphemes) as part of all syntactic units in which it participates. The held DS is italicized for emphasis in the following example.

1st SU: TREE *DS\_5(tree)* BIRD DS\_b2(bird-sits-in-tree)

2nd SU: *DS\_5(tree)* FATHERstr DS\_b2(person-approaches-tree)

* 1. **What counts as a morpheme in a size and shape DS (SASS)?**
     1. A meaningful **handshape**
        1. This credit is given whether or not a child’s handshape is fully adult-like. You can think of this as a basic morpheme, similar to the morpheme that each non-depicting sign receives automatically.
        2. Two handshapes used together symmetrically count as one word together (see above section) and also as one morpheme together.
     2. A **path** of movement that indicates the extent of the described entity
     3. A **manner** of production giving additional information
        1. e.g. repeated DS\_f to show multiple scattered small objects
     4. The **spatial relationship** involved. *This can be the relationship between two DSes, or between the DS and a real-world referent. A spatial relationship is often involved when there is no path movement.*
        1. e.g. DS\_f placed at particular locations to show button placement
  2. **What counts as a morpheme in a handling or entity DS?**
     1. A meaningful **handshape**
        1. This credit is given whether or not a child’s handshape is fully adult-like. You can think of this as a basic morpheme, similar to the morpheme that each non-depicting sign receives automatically.
        2. Handshape may be the only morpheme credited in a DS. This occurrence, however, is extremely rare.
     2. A **path** of movement that indicates the path of movement of an object
        1. e.g. zigzag, uphill, downhill, to the side, straight forward
        2. If the path is more than a straight progression between the source and goal, location can be credited as well as path (see iii.)
     3. The meaningful **location(s)** involved
        1. Can credit a production in a particular location. *One morpheme per meaningful location of the DS. Meaningful locations may be real-world locations or those pictured e.g. in a book.*
           1. DS\_bo(eyedrop-in-left-eye) [+1]
           2. DS\_bo(eyedrops-in-both-eyes) [+2]
           3. In a book: DS\_2(person-sits-on-log) [+1]
        2. Can credit both source and goal if both are separately meaningful.
           1. Consider cases such as a DS depicting the act of putting a pen cap on a pen. The goal location is meaningful, being on the pen, but the source location is not meaningful outside of “not on the pen” being in contrast to the goal location.
           2. DS\_fo(put-cap-onto-pen) [+1]
           3. DS\_fo(move-cap-from-table-to-pen) [+2]
        3. Often, locations are not meaningful other than as the endpoints of the (meaningful) path and will not count as additional morphemes.
        4. Rarely, a location can count as indicating existence of an entity.
           1. e.g. DS\_3 produced with a very brief downward movement to indicate that a car or train exists in the proposition
     4. A **manner** of production giving additional information
        1. Multiple manner morphemes could co-occur within a single sign (e.g., quickly and repeated)
        2. Hand-internal movement may be a clue towards meaningful manner
           1. e.g. the wiggling of fingers in DS\_5(bubbles-move-upwards) could indicate the manner “fizzily”
           2. e.g. handshape change from 2 to b2 to indicate bending the legs at the peak of each jump
        3. Orientation change may be a clue towards meaningful manner
           1. e.g. a twist of the wrist in DS\_3(truck-drives-bouncily) to indicate the bouncy manner of the truck’s travel
        4. Speeding up or slowing down of a sign may be a clue towards meaningful manner. However, be very conservative when crediting modifications in movement speed. The modification should be obvious/easy to see, and for a clear meaningful reason.
        5. Possible manners include jumping, limping, wobbling, bubbling, quickly, slowly, repeated
        6. For example: A DS showing something bouncing forward would have +1 path morpheme (straight forward) and +1 manner morpheme (bouncing)
     5. A meaningful **spatial relationship** between two DS handshapes
        1. Count this morpheme only once for the pair of DSes involved.
           1. DS\_b2(chair) DS\_b2(chair-across-from-other-chair) [+1]
        2. Multiple spatial relationships (beginning & ending) may be counted in rare circumstances when these aspects of the meaning are not already covered by other already-credited morphemes
        3. Often used when there is no movement involved
     6. Non-manual signals (**NMS**) are not included at all in our MLU coding, in part due to video visibility constraints in our corpus.
  3. **What should you do if you disagree with a transcription?**
     1. If a sign was not transcribed at all by the transcriber, go ahead and transcribe it as if you were the ASL transcriber, including Free Translation. You may then use it in your syntactic unit. If you’re not confident doing so, then just add the missing sign to the ASL feedback tier and do not include it in the Syntactic Unit tier until it can be checked.
     2. If you are very confident that a sign was transcribed incorrectly, use your interpretation in the Syntactic Unit tier. Do not change the transcription, but add comments to the ASL feedback tier for future checking.
     3. If you think a sign is a lexical sign but it was glossed as a DS or you think a sign is a DS but it was glossed as a lexical sign or it was glossed both ways at different points in the file: These instances of the sign in question should all be treated and transcribed the same, either all as a lexical sign or all as a DS. Perhaps SignBank did not include the relevant sign when the transcript was first transcribed. Or perhaps the DS is a lexical sign within the family (in other words, a home sign). Use your best judgment, and if you aren’t certain then go with what was transcribed originally, and leave your questions on the ASL feedback tier for future checking.
     4. If the free translation is longer than the syntactic unit (many utterances included in one free translation) then split up the SU as you think it should be split up and leave a note on the ASL feedback tier for the free translation(s) to be checked for that time duration.
     5. In general, for disagreements, always leave a comment or question on the ASL feedback tier to enable future checking. Then either code using your interpretation when you are very confident, or stay with the ASL transcriber’s interpretation in less certain cases.

**Appendix 1**

**List of Routines**

BYE-BYE

COOL

E(perfect)

FS(ok)

HELLO

NO

PEEK-A-BOO

PLEASE

SORRY

SWELL

THANK-YOU/GOOD

WOW

YAY

YEAH/OIC

YES

Letters of the alphabet

Slow fingerspelling emphasizing each letter

Counting