Starting point in ELAN:

Phonetic utterance

(For target kids only: Target utterance)

Translation

XDS – X[person]-directed speech

Timepoints

Participant

Recording (including all of the metadata captured by the recording/EAF filename)

Goal endpoint:

Phonetic utterance

(For target kids only: Target utterance)

Morphological parse of the phonetic utterance

(For target kids only: Morphological parse of the target – needs to be associated with morphological parse of the phonetic form)

Translation

XDS

Timepoints

Participant

Recording

How to organize an EAF for this endpoint:

Phonetic utterance – Time-aligned root tier

(For target kids only: Target utterance, as a symbolic association to the phonetic)

Morphological parse of the phonetic utterance – Symbolic subdivision of the phonetic utterance, with the following structure:

Utterance is divided into words (word = symbolic subdivision)

Words are divided into morphemes (morpheme = symbolic subdivision)

Morphemes are associated with a UR and a gloss (symbolic associations)

(For target kids only: Morphological parse of the target – Symbolic subdivision of the target utterance, with the same structure as the phonetic)

Translation – Symbolic association to the root

XDS - Symbolic association to the root

Timepoints – Feature of the root

Participant – Feature of the root

Recording – Feature of the file

Repeated Y/N – Added during analysis (does not need to be in EAF)

FLEx takes in:

Phonetic utterance

Target utterance

Translation

(optionally other info)

{addressee}

FLEx outputs:

Phonetic utterance divided into words

Words associated with word-level morphological analysis

Words divided into morphemes

Morphemes associated with UR and gloss

Workflow for script:

Take ELAN file

Export all phonetic utterances and give each utterance a unique ID (unique across the entire dataset – can create ID as filename plus participant plus timepoint)

Export translations and give each translation the unique ID of its phonetic utterance

Export target utterances and give each target the unique ID of its phonetic utterance

Note: Do not import with the native EAF import function of FLEx. It assigns unique IDs but they do not make sense.

Script will output an XML file that will look like this:

<node id=a\_x\_4 type=utterance participant=x> utterance

<node> translation </node>

</node>

<node id=a\_x\_4 type=target participant=x> target

<node> translation </node>

</node>

You may need to edit this output in order to import to FLEx.

When imported to FLEx the XML should look like this as a text:

ngE27ma4 na4ngE27ma4

translation: xxx

note: a\_x\_4

note: phonetic

note: x

ngE5ma2 na4ngE27ma4

note: a\_x\_4

note: target

note: x

Parse this text in FLEx

Export the text from FLEx using the FLExText function

<line> ngE27ma4 na4ngE27ma4 </line>

<word> ngE27ma4 </word>

<morpheme> ngE27ma4 </morpheme>

<morpheme-gloss> there:anaphoric </morpheme-gloss>

…

<word> na4ngE27ma4 </word>

…

<note: target>

<note: a\_x\_4>

<note: x>

Use a script to bring this information back into the EAF

Workflow for this script:

For every line exported from FLEx, look up the annotation in ELAN that has the same unique ID

For this annotation, create as many child annotations as there are words in the FLEx export

For each word, create as many child annotations as there are morphemes in the FLEx export

For each morpheme, create annotations on separate symbolic association tiers with the UR and gloss

Overall steps are:

1. Save EAF and back it up

2. Run script to export EAF to simpler XML format (possibly FLExText format)

3. Import the output of 2 to FLEx using native import function

4. Parse in FLEx

5. Export the output of 4 to FLExText using native export function

6. Run script to turn FLExText from 5 into additional annotations on the EAF