SignStream®3 Collection File XML Documentation

Last Updated: May 2025

Author: Gregory Dimitriadis (Rutgers University)

Distributed through the American Sign Language Linguistic Research Project (ASLLRP)
by Carol Neidle (carol@bu.edu) from http://www.bu.edu/asllrp/SS-XML-doc.pdf

For further information about SignStream, see http://www.bu.edu/asllrp/SignStream/3

Abstract: This document describes the structure and outline of a SignStream®3 collection XML file, and how it relates to the application's front-facing features from the perspective of a general user. While working with SignStream, one may wish to inspect the actual data underlying an entity being modified, or attempt to debug an error that is occurring. The collection file contains most of the information necessary to do this, and this document also provides guidance on where to find other files and resources, both within the SignStream®3 installation package and potentially externally on the Internet. With knowledge about how the data is organized in the XML (and note that multiple aspects of a single entity may be stored separately across interrelated components), it may be possible to engage in basic automated parsing and interpretation of collection files. However, we also expect to release in coming months a self-standing application that will help in extracting information directly from the XML files.

Quick Facts

- A SignStream®3 collection file has a ".ssc" extension, but the file type is technically a Zip Archive.
- Within this ".ssc" zipfile is an xml file, typically named "macrounit1.3.xml".
- Data within the xml file are separated into "TRACKS" which correspond one-to-one to Signstream®3 fields in the application window.
    - Within each "TRACK" is a list of "ENTITY" objects, each of which represents just a portion of a single event/entity (nonmanual, gloss, gesture, etc.).
- The file also contains a hierarchical listing of Temporal Partitions, Segment Tiers, Utterances, and Statement Fields (which correspond to Track/Field objects).
- Start/End values within the XML file are "timecodes," which work out to the time value displayed in green text in a video window of the SignStream application multiplied by the frame rate of the video.

Brief Outline of Collection XML Contents

The XML file summarized below is found within a *.ssc file, representing a SignStream®3 collection. It is essentially a standard zipfile, just with a different file extension, and so can be unzipped with any utility program. Inside is a single XML file, containing the below structure. Many of the elements are only linked by various ID attributes, so many times the best way to navigate between different parts of the structure is to work with a particular ID at a time, to locate it across the various elements.

Key Elements:

- MEDIA-FILES: the videos associated with this collection
    - PACKET_DURATION: units between each frame of the videos in this collection – units vary based on video type, frame rate, codec, etc.
    - MOVIE_DURATION: total units in the videos for this collection – units vary based on video type, frame rate, codec, etc.
    - MOVIE_TIMESCALE: this value divided by the packet duration will get the frame rate for the video(s) used in the collection.
- LOCAL-CODING-SCHEMA defines either overridden or custom (user-created) nonmanual field definitions
    - If the ID of the FIELD is an existing SignStream field ID, then the one within this coding schema element is overriding the SignStream default.
    - If the ID is new to SignStream, it is a user-created field.

- The VALUE objects within define the possible input values for the nonmanual field.

- GLOSS-WINDOW represents the “Database View” layer which contains the following:
  - TEMPORAL-PARTITION objects representing containers for segment tiers
    - SEGMENT-TIER objects are inside, each holding a list of utterances and the associated field and utterance data.
- STATEMENT-FIELD (inside SEGMENT-TIER objects) represents the entire data for a particular field (nonmanual/manual/etc)
  - The FIELD-ID here denotes which field this is.
    - Each Field ID corresponds to a particular field of a given name – and specified in the defCodingScheme.xml .
    - e.g. "eye brows, "eye aperture", etc.
  - This FIELD-ID corresponds 1-to-1 with TRACK attribute FID (under DATABASE).
  - ENTITY-IDS is a comma-separated list of ID's, each of which references an ENTITY object found under a TRACK.
- TRACK elements hold all the entities in the collection for a particular field across all segment tiers.
  - Attribute FID corresponds to STATEMENT-FIELD attribute FIELD-ID.
  - Each ENTITY under a TRACK represents one entity/event in the collection, which may be linked to other entities in the SignStream application. (E.g. a regular manual gloss is composed of multiple ENTITY elements – a dominant gloss entity, a nondominant gloss entity, a handshape entity, etc.)
  - Not all the attributes of an ENTITY will be set for every type (field type) of entity.
    - An entity type is determined by the TRACK (and its FID) in which it is contained.
  - All ENTITY objects should have an ITEM within which has the start ("S") and end ("E") time and the value (the textual version) of the entity.
    - Often there are multiple ITEM elements under an ENTITY. This is particularly the case with chained entities – such as a nonmanual entity with an onset, the main value of some string (e.g., “eyebrows raised”) and its offset: 3 distinct ITEM elements stored under one ENTITY
    - ITEM elements can be any of the following:
      - Nonmanual main event (specific field values for a nonmanual field)
        - e.g. values "raised" or "lowered" for the field "eye brows"
      - Manual main event (any text values for a manual field)
      - Onset
      - Offset
      - Initial Hold
      - Final Hold
    - Not all ENTITY objects will have a PROPERTIES object inside; it is valid only for gloss entity objects.
      - The PROPERTIES attributes correspond to the settings of the Morph-Phon window in the SignStream application.
    - Every ENTITY will also contain a SEGMENT-TIER-UTTERANCE-REFS element.
      - Within this, there can be any number of REF elements, which contain a SEGMENT-TIER-UTTERANCE_ID which refers to an utterance which intersects with or contains the entity.

<u>Complete XML Structure:</u>

- SIGNSTREAM-DATABASE
  - Attributes:
    - DATABASE-VERSION
    - EXCERPT
    - EXCERPT_NUMBER
    - PARTICIPANT
    - SIGNSTREAM-VERSION
    - SOURCE
    - S

- E
- UTILS
  - Attributes:
    - TP_MOVIE_TIME
    - TP_ONSCREEN_LOCATION
    - TP_PIXEL_IN_TIME_FRAME
- REFERENCES
  - CODING-SCHEMES
    - List: REF
      - Attributes:
        - ID
        - FILE_NAME
        - SIGNATURE
  - PARTICIPANTS
    - List: REF
      - Attributes:
        - COMMENTS
        - DOMINANT_HAND
        - FULL_NAME
        - GENDER
        - ID
        - LABEL
        - NATIVE_LANGUAGE
        - PARENT_INFO
        - YEAR_OF_BIRTH
  - ANNOTATORS
    - List: REF
      - Attributes:
        - ID
        - LABEL
- MEDIA-FILES
  - Attributes:
    - MOVIE_DURATION
    - PACKET_DURATION
  - List: MEDIA-FILE
    - Attributes:
      - ID
      - PATH
- GRAPH-FILES
  - List: GRAPH-FILE
    - Attributes:
      - ID
      - NAME
      - PATH
- MEDIA_FILES_ON_SCREEN_LOCATION
  - List: MEDIA-FILE
    - Attributes:
      - FILE_NAME
      - LOCATION_X

- LOCATION_Y

- LOCAL-CODING-SCHEME
  - List: FIELD
    - Attributes:
      - ID
      - List: FIELD-VALUES
        - List: FIELD-VALUE
          - Attributes:
            - ID
            - List: FIELD-VALUE-ATTRIBUTES
              - List: FIELD-VALUE-Attribute
              - Attributes can include: NAME, LABEL, PREFIX, TYPE, GROUP, COLOR
- DATABASE
  - TRACKS
    - List: TRACK
      - Attributes:
        - FID
      - List: ENTITY
        - Attributes (some of):
          - COMP_END_ENTITY_ID
          - COMP_START_ENTITY_ID
          - CORRESPONDING_GLOSS_ENTITY_ID
          - DEPENDENT_GLOSS_ID
          - DEPENDENT_HAND_ID
          - DEPENDENT_HAND_SHAPE_ID
          - DEPENDENT_LOCATION_ID
          - DEPENDENT_SIGN_TYPE_ID
          - ID
          - LOCKED_DOM_NDOM_ALIGNMENT
          - LOCKED_GLOSS_PHON_ALIGNMENT
          - MARKED_NUMBER_OF_HANDS
          - REF_ENTITY_ID
        - List: ITEM
          - Attributes:
            - VID
            - S
            - E
            - ORDER
          - Value: field value name
        - PROPERTIES
          - Attributes:
            - DOM_HAND_START_END_HANDSHAPE
            - NDOM_HAND_START_END_HANDSHAPE
            - NAME_SIGN
            - NO_OF_HANDS
            - PASSIVE_BASE_ARM
            - SIGN
            - START_END_HANDSHAPES
            - TWOHANDED_HANDSHAPES
        - SEGMENT-TIER-UTTERANCE-REFS
          - List: REF
            - Attributes:

- SEGMENT-TIER-UTTERANCE_ID

- NOTES
  - List: NOTE
    - Attributes:
      - ID
      - TIME_CREATED
      - TIME_UPDATED
      - SUBJECT
      - BODY
- WINDOW
  - DISPLAY-TIME-PERIOD
  - VISIBLE-FIELD-REFS
    - List: REF
      - Attributes:
        - FIELD_ID
  - TEMPORAL-PARTITIONS
    - List: TEMPORAL-PARTITION
      - Attributes:
        - ID
        - LABEL
      - SEGMENT-TIER
        - Attributes:
          - ID
          - INDEX_FIELD_ID
          - IS_DEFAULT
          - LABEL
          - PARTICIPANT_ID
        - VISIBLE-FIELD-REFS
          - List: REF
            - Attributes:
              - FIELD_ID
        - VISIBLE-PANEL-REFS
          - List: REF
            - Attributes:
              - ORDER
              - PANEL_ID
              - PANEL_PERCENTAGE
        - List: STATEMENT-FIELD
          - Attributes:
            - ENTITY-IDS
            - FIELD-ID
        - List: SEGMENT-TIER-UTTERANCE
          - Attributes:
            - ID
            - LABEL
            - S
            - E
          - VISIBLE-FIELD-REFS
            - List: REF
            - Attributes:
              - FIELD-ID

Frame Calculations

To get the frame rate value for the video(s) associated with the collection:

> ffprobe -v error -select_streams v -of default=noprint_wrappers=1:nokey=1 -show_entries stream=r_frame_rate <FILENAME>

You can download the ffprobe executable here: https://evermeet.cx/ffmpeg/

Frame/Timecode Calculation

(Time Value) = (Collection Timecode Value) / (Frame Rate)

or

(Collection Timecode Value) = (Time Value) * (Frame Rate)

Important Field Identifiers:

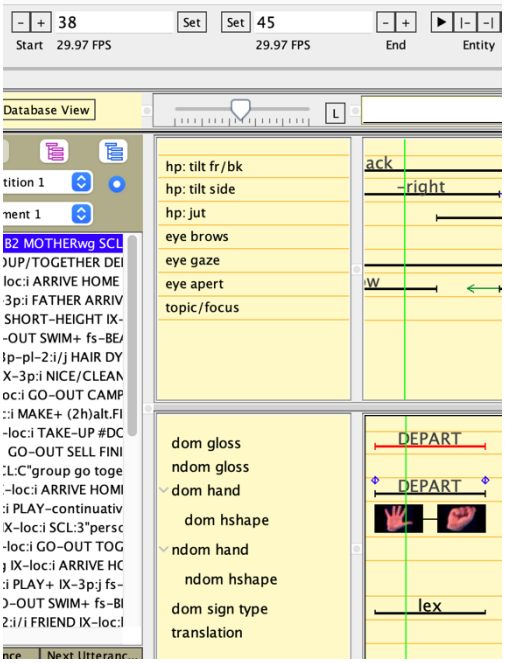
SignStream Field Label	XML Field ID
dom gloss	10000
ndom gloss	10001
dom hand	10501
dom handshape	1000
ndom hand	10502
ndom handshape	2000
dom sign type	26000
ndom sign type	26001
translation	30000

Location of Coding Scheme file (defCodingScheme.xml)

- From your default SignStream installation directory (call it SIGNSTREAM_DIR) <SIGNSTREAM_DIR>/newXMLfiles/defCodignScheme.xml
- This is the default coding scheme, which can be overridden locally per collection.
- The ID's and other items listed below are defined in this coding scheme file.

XML Data: One-Handed Gloss

- Retrieving information for a gloss by its name and start/end time:
- Say the gloss is a one-handed dominant hand gloss named "DEPART" and it starts at timecode 38 and ends at timecode 45



- In the XML, first look for the TRACK identified by the field ID for “dom gloss”

```
<TRACK FID="10000">
```

- Then within that track look for “DEPART”- there may be many if there are multiple glosses over the entire collection with the same name, so look for those with the values S="38" and E="45"

```
<ENTITY COMP_END_ENTITY_ID="" COMP_START_ENTITY_ID="" CORRESPONDING_GLOSS_ENTITY_ID="" DEPENDENT_GLOSS_ID="" DEPENDENT_HAND_ID="285874" DEPENDENT_HAND_SHAPE_ID="407791" DEPENDENT_LOCATION_ID="" DEPENDENT_SIGN_TYPE_ID="160417970058943608" ID="135887861686253" LOCKED_DOM_NDOM_ALIGNMENT="F" LOCKED_GLOSS_PHON_ALIGNMENT="T" MARKED_NUMBER_OF_HANDS="F" REF_ENTITY_ID="">
    <PROPERTIES DOM_HAND_START_END_HANDSHAPE="" NAME_SIGN="F" NDOM_HAND_START_END_HANDSHAPE="" NO_OF_HANDS="1" PASSIVE_BASE_ARM="F" SIGN="LEXICAL" START_END_HANDSHAPES="DIFFERENT START/END HANDSHAPE" TWOHANDED_HANDSHAPES=""/>
    <ITEM E="45" IID="" ORDER="1" S="38" VID="135887861686252">DEPART</ITEM>
    <SEGMENT-TIER-UTTERANCE-REFS>
        <REF SEGMENT-TIER-UTTERANCE_ID="771167"/>
    </SEGMENT-TIER-UTTERANCE-REFS>
</ENTITY>
```

- This element – the entire ENTITY - is a top-level representation of the DEPART gloss we are looking for. From here, we can inspect all the other sub-components of the gloss: the dominant hand portion, the dominant handshapes, etc.
- The first thing to take note of here is the ID value of this ENTITY. If you search for this ID you will find it in a few places within the XML.
    - ID="135887861686253"
    - If you search for this value, you will see it also as a REF_ENTITY_ID attribute on other ENTITY elements. Those ENTITY elements are the sub-components of the gloss, and they refer back to the main gloss portion.
    - You will also see it as part of a comma-separated list attribute called ENTITY-IDS on a STATEMENT-FIELD element.
- Looking at the linked Hand sub-component (DEPENDENT_HAND_ID="285874")
    - We can see here it has the same start/end times as the main ENTITY:

```
<ENTITY COMP_END_ENTITY_ID="" COMP_START_ENTITY_ID="" ID="285874" LOCKED_GLOSS_PHON_ALIGNMENT="T" REF_ENTITY_ID="135887861686253">
    <ITEM E="45" IID="" ORDER="1" S="38" VID="5145114">DEPART</ITEM>
    <SEGMENT-TIER-UTTERANCE-REFS>
        <REF SEGMENT-TIER-UTTERANCE_ID="771167"/>
    </SEGMENT-TIER-UTTERANCE-REFS>
</ENTITY>
```

- Looking at the linked Handshape sub-component (DEPENDENT_HAND_SHAPE_ID="407791")
  - We can see that it contains two ITEM elements within this one ENTITY
  - Each ITEM here represents one handshape. In this case, the start handshape (represented by HSH_004) shares the start timecode of 38, and the end handshape (represented by HSH_082) shares the end timecode of 45.
  - HSH_004 and HSH_082 refer to handshape image codes defined in the SignStream defCodingScheme.xml file.
  - (Note: if there were more handshapes in between the start and end handshapes – e.g. as would occur with a fingerspelled sign – those additional handshapes' timecodes would be spread out evenly distributed between the start and end times.)

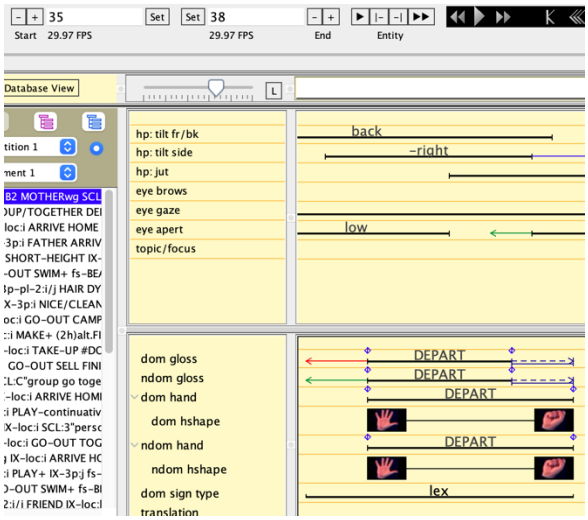
```
<ENTITY COMP_END_ENTITY_ID="" COMP_START_ENTITY_ID="" ID="407791" LOCKED_GLOSS_PHON_ALIGNMENT="T" REF_ENTITY_ID="285874">
    <ITEM E="38" IID="HSH_004" ORDER="1" S="38" VID="4"/>
    <ITEM E="45" IID="HSH_082" ORDER="2" S="45" VID="82"/>
    <SEGMENT-TIER-UTTERANCE-REFS>
        <REF SEGMENT-TIER-UTTERANCE_ID="771167"/>
    </SEGMENT-TIER-UTTERANCE-REFS>
</ENTITY>
```

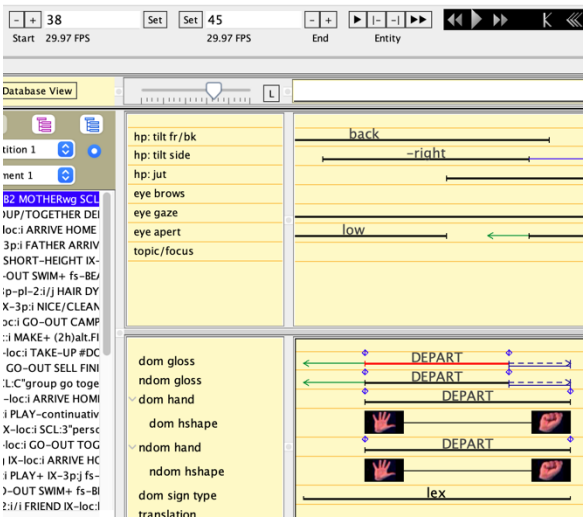
XML Data: Two-Handed Gloss, Changing Properties

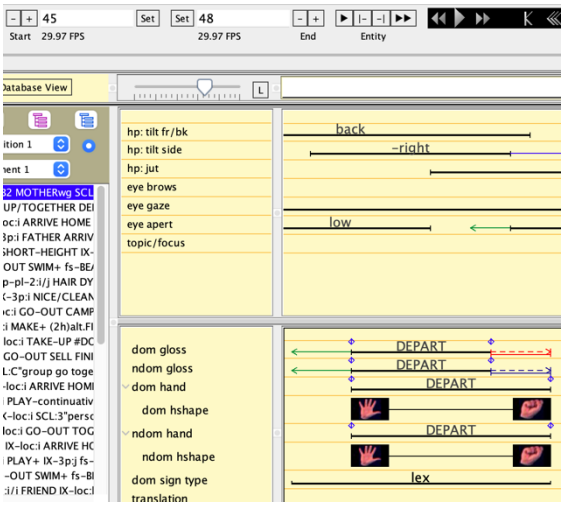
Let's complicate our existing one-handed gloss slightly to demonstrate some additional points about the elements reviewed above.

First, we will make this a two-handed sign, so there is now a complimentary and parallel nondominant portion of the gloss, in addition to the dominant portion. Second, we will add a few "associated elements" as the SignStream application call them: an "Onset" to the beginning, and a "Final Hold" to the ending.

- The new gloss looks like so – note the 3 distinct images for each of the elements:







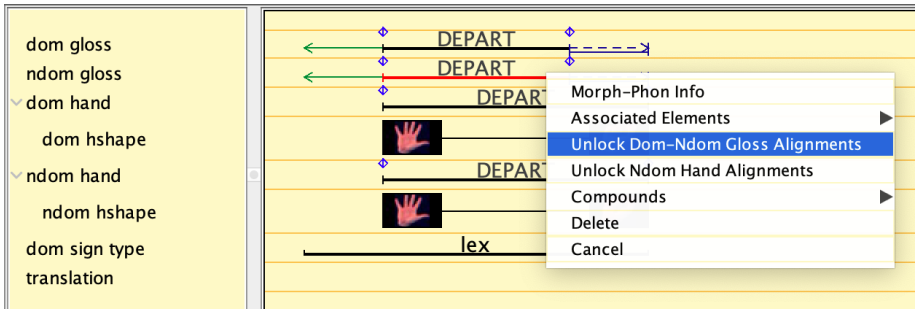
- If we look at the XML for the dominant gloss ENTITY for DEPART, we can see what has changed:

```
<ENTITY COMP_END_ENTITY_ID="" COMP_START_ENTITY_ID="" CORRESPONDING_GLOSS_ENTITY_ID="17455247271642037" DEPENDENT_GLOSS_ID="17455247271642037" DEPENDENT_HAN
D_ID="285874" DEPENDENT_HAND_SHAPE_ID="407791" DEPENDENT_LOCATION_ID="" DEPENDENT_SIGN_TYPE_ID="16041797005894360 8" ID="135887861686253" LOCKED_DOM_NDOM_ALIGNMENT="T" LOCKE
D_GLOSS_PHON_ALIGNMENT="T" MARKED_NUMBER_OF_HANDS="F" REF_ENTITY_ID="">
    <PROPERTIES DOM_HAND_START_END_HANDSHAPE="" NAME_SIGN="F" NDOM_HAND_START_END_HANDSHAPE="" NO_OF_HANDS="2" PASSIVE_BASE_ARM="F" SIGN="LEXICAL" START_END
_HANDSHAPES="SAME START/END HANDSHAPE" TWOHANDED_HANDSHAPES="SAME HANDSHAPES DOM NDOM"/>
    <ITEM E="38" IID="" ORDER="1" S="35" VID="600000">Onset</ITEM>
    <ITEM E="45" IID="" ORDER="2" S="38" VID="135887861686252">DEPART</ITEM>
    <ITEM E="48" IID="" ORDER="3" S="45" VID="600002">Final Hold</ITEM>
    <SEGMENT-TIER-UTTERANCE-REFS>
        <REF SEGMENT-TIER-UTTERANCE_ID="771167"/>
    </SEGMENT-TIER-UTTERANCE-REFS>
</ENTITY>
```

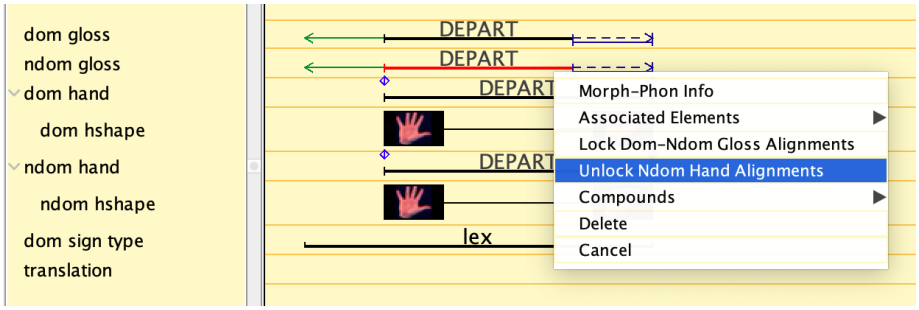
- There is now a value to represent the link to the nondominant portion of the gloss (DEPENDENT_GLOSS_ID="17455247271642037")
- There are now multiple ITEM elements whereas there was only one previously. The new ITEM elements represent the Onset and Final Hold on the application.
    - The VID values for these can be mostly ignored. They each correspond to the “type” of associations which can be applied onto a gloss, and there are no references to them elsewhere.
- Note that the value of the ITEM containing the main portion of the gloss – the DEPART – is unchanged in its start and end times. But now the Onset ITEM has an earlier start time, and the Final Hold ITEM has a later end time. So, collectively, this ENTITY has a start time of 35 and end time of 48.
- Let’s look at the nondominant portion of the gloss next:

```
<ENTITY COMP_END_ENTITY_ID="" COMP_START_ENTITY_ID="" CORRESPONDING_GLOSS_ENTITY_ID="135887861686253" DEPENDENT_GLOSS_ID="135887861686253" DEPENDENT_HAND_ID
="17455247271652039" DEPENDENT_HAND_SHAPE_ID="1745524727166204O" DEPENDENT_LOCATION_ID="" DEPENDENT_SIGN_TYPE_ID="" ID="17455247271642037" LOCKED_DOM_NDOM_ALIGNMENT="T" LOC
KED_GLOSS_PHON_ALIGNMENT="T" MARKED_NUMBER_OF_HANDS="F" REF_ENTITY_ID="">
    <PROPERTIES DOM_HAND_START_END_HANDSHAPE="" NAME_SIGN="F" NDOM_HAND_START_END_HANDSHAPE="" NO_OF_HANDS="2" PASSIVE_BASE_ARM="F" SIGN="LEXICAL" START_END
_HANDSHAPES="SAME START/END HANDSHAPE" TWOHANDED_HANDSHAPES="SAME HANDSHAPES DOM NDOM"/>
    <ITEM E="38" IID="" ORDER="1" S="35" VID="600000">Onset</ITEM>
    <ITEM E="45" IID="" ORDER="2" S="38" VID="17455260392552044">DEPART</ITEM>
    <ITEM E="48" IID="" ORDER="3" S="45" VID="600002">Final Hold</ITEM>
    <SEGMENT-TIER-UTTERANCE-REFS>
        <REF SEGMENT-TIER-UTTERANCE_ID="771167"/>
    </SEGMENT-TIER-UTTERANCE-REFS>
</ENTITY>
```

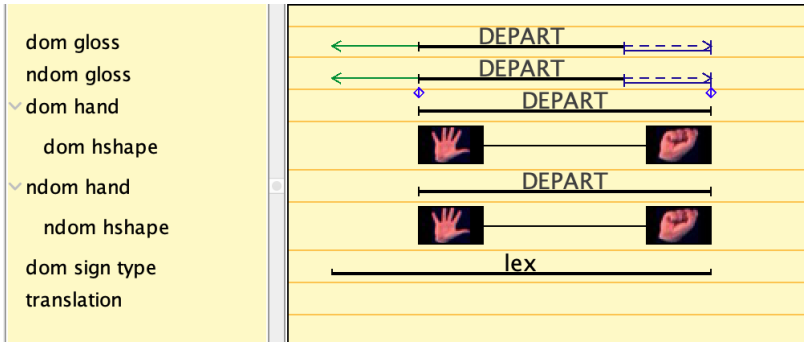

- This nondominant portion has the same amount of linkages as the dominant portion did, meaning that it refers to a dependent hand, dependent handshape, etc. These are the nondominant versions of those same entities.
- Currently the dominant and nondominant portions of the gloss are “locked” both in terms of to each other, as well as each gloss being locked to their hand alignments. We can observe what happens when we unlock both of those. We’ll do it here for this nondominant portion:



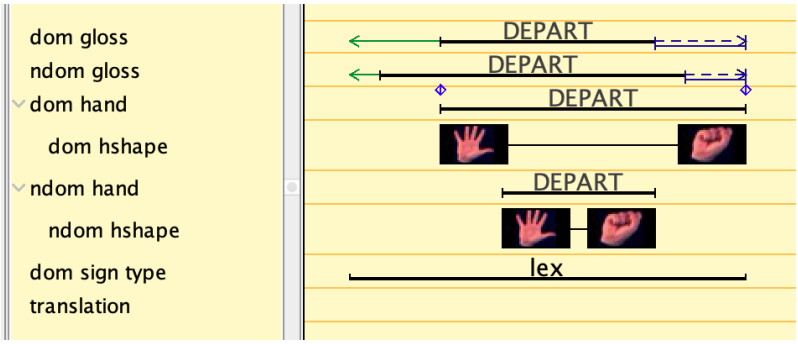
- Note the lack of blue diamonds between the dominant and non-dominant portion of the gloss, after unlocking them:



- Then note the lack of blue diamonds over the “ndom hand” line:



- Now with these unlocks, it is possible to set the start and end times of the various components separately:



- The XML representation of the nondominant gloss now looks like this:

```
<ENTITY COMP_END_ENTITY_ID="" COMP_START_ENTITY_ID="" CORRESPONDING_GLOSS_ENTITY_ID="135887861686253" DEPENDENT_GLOSS_ID="135887861686253" DEPENDENT_HAND_
ID="17455247271652039" DEPENDENT_HAND_SHAPE_ID="17455247271662040" DEPENDENT_LOCATION_ID="" DEPENDENT_SIGN_TYPE_ID="" ID="17455247271642037" LOCKED_DOM_NDOM_ALIGNMENT="F"
 LOCKED_GLOSS_PHON_ALIGNMENT="F" MARKED_NUMBER_OF_HANDS="F" REF_ENTITY_ID="">
                <PROPERTIES DOM_HAND_START_END_HANDSHAPE="" NAME_SIGN="F" NDOM_HAND_START_END_HANDSHAPE="" NO_OF_HANDS="2" PASSIVE_BASE_ARM="F" SIGN="LEXICAL" START_E
ND_HANDSHAPES="SAME START/END HANDSHAPE" TWOHANDED_HANDSHAPES="SAME HANDSHAPES DOM NDOM"/>
                <ITEM E="36" IID="" ORDER="1" S="35" VID="600000">Onset</ITEM>
                <ITEM E="46" IID="" ORDER="2" S="36" VID="17455260392552044">DEPART</ITEM>
                <ITEM E="48" IID="" ORDER="3" S="46" VID="600002">Final Hold</ITEM>
                <SEGMENT-TIER-UTTERANCE-REFS>
                    <REF SEGMENT-TIER-UTTERANCE_ID="771167"/>
                </SEGMENT-TIER-UTTERANCE-REFS>
            </ENTITY>
```

XML Data: Segment Tiers and Utterances

- Note in the previous image, the element SEGMENT-TIER-UTTERANCE-REFS. This contains potentially multiple references to Utterance ID's which either contain or intersect with this ENTITY. Searching for that Utterance ID toward the end of the XML you will find its definition:

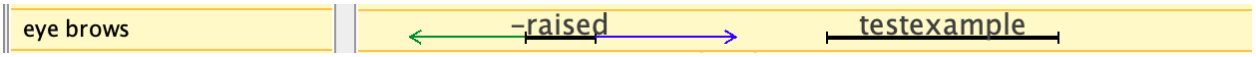
```
<SEGMENT-TIER-UTTERANCE E="96" ID="771167" LABEL="U1" S="0">
```

- Looking higher in the hierarchy above this SEGMENT-TIER-UTTERANCE, you will also find the SEGMENT-TIER and TEMPORAL-PARTITION of which it is a part:

```
<TEMPORAL-PARTITIONS>
    <TEMPORAL-PARTITION ID="55829211" LABEL="Temporal Partition 1">
        <SEGMENT-TIER ID="55827552" INDEX_FIELD_ID="" IS_DEFAULT="T" LABEL="Segment 1" PARTICIPANT_ID="1">
```

XML Data: Field Values

- As shown above for manual glosses, the value of the entity is the gloss name which is given to it; in the case of glosses, it is the one ITEM element that is not an associated element of the gloss, that contains that gloss name.
- Nonmanual entities are slightly different, in that their value can either be chosen from a predefined list, or alternatively it can be freeform text. As an example of the former:



- Here, "-raised" is a default value of the defined nonmanual field "eye brows" - in the XML it appears as so:

```
<ENTITY COMP_END_ENTITY_ID="" COMP_START_ENTITY_ID="" ID="1374250585858453" REF_ENTITY_ID="">
    <ITEM E="18" IID="" ORDER="1" S="13" VID="500000">Onset</ITEM>
    <ITEM E="21" IID="" ORDER="2" S="18" VID="5">-raised</ITEM>
    <ITEM E="27" IID="" ORDER="3" S="21" VID="500003">Offset</ITEM>
    <SEGMENT-TIER-UTTERANCE-REFS>
        <REF SEGMENT-TIER-UTTERANCE_ID="771167"/>
    </SEGMENT-TIER-UTTERANCE-REFS>
</ENTITY>
```

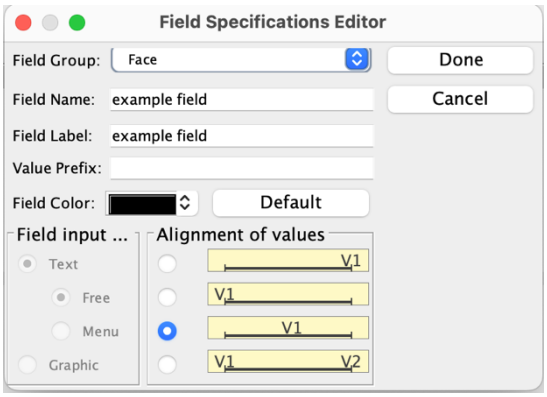
- The VID of “5” here corresponds to the identifier given for this value under field “eye brows” in the defCodingScheme.xml file.
- The second entity “testexample” is a value which a user of SignStream added as a custom value to the “eye brows” field. It looks similar in structure:

```
<ENTITY COMP_END_ENTITY_ID="" COMP_START_ENTITY_ID="" ID="17465458957802041" REF_ENTITY_ID="">
    <ITEM E="41" IID="" ORDER="1" S="31" VID="17465458957802040">testexample</ITEM>
    <SEGMENT-TIER-UTTERANCE-REFS>
        <REF SEGMENT-TIER-UTTERANCE_ID="771167"/>
    </SEGMENT-TIER-UTTERANCE-REFS>
</ENTITY>
```

- The VID here (“17465458957802040”) is not found in the defCodingSchem.xml file, but instead is a generated value created specifically for this SignStream collection. At the top of the collection XML you will see a LOCAL-CODING-SCHEME element, within which is a collection of new and overridden definitions of fields and values, specific to this collection. Here is our added value:

```
<FIELD ID="10">
    <FIELD-VALUES>
        <FIELD-VALUE ID="17465458957802040">
            <FIELD-VALUE-ATTRIBUTES>
                <FIELD-VALUE-ATTRIBUTE NAME="testexample"/>
                <FIELD-VALUE-ATTRIBUTE LABEL="testexample"/>
            </FIELD-VALUE-ATTRIBUTES>
        </FIELD-VALUE>
    </FIELD-VALUES>
</FIELD>
```

- This indicates that for the “eye brows” field (having ID=“10”) there is a new FEILD-VALUE with the NAME and LABEL of “testexample”
- We can also define a completely new field for this collection:



- This field is configured as a “free” type of nonmanual field, meaning that the user can enter any value for the value of a nonmanual entity, instead of being constrained by the predefined values. In the XML it appears like so:

```
<FIELD ID="17465464984712478">
    <FIELD-ATTRIBUTES>
        <FIELD-ATTRIBUTE NAME="example field"/>
        <FIELD-ATTRIBUTE LABEL="example field"/>
        <FIELD-ATTRIBUTE TYPE="NMNFRT"/>
        <FIELD-ATTRIBUTE ALIGNMENT="CENTER"/>
        <FIELD-ATTRIBUTE GROUP="FACE"/>
        <FIELD-ATTRIBUTE ORDER="11"/>
    </FIELD-ATTRIBUTES>
</FIELD>
```

- And looking at that generated FIELD ID="17465464984712478" we find a TRACK element with the same identifier and all the entities it contains:

```
<TRACK FID="17465464984712478">
    <ENTITY COMP_END_ENTITY_ID="" COMP_START_ENTITY_ID="" ID="17465465161322479" REF_ENTITY_ID="">
        <ITEM E="34" IID="" ORDER="1" S="24" VID="17465465197312698">freeformtext</ITEM>
        <SEGMENT-TIER-UTTERANCE-REFS>
            <REF SEGMENT-TIER-UTTERANCE_ID="771167"/>
        </SEGMENT-TIER-UTTERANCE-REFS>
    </ENTITY>
</TRACK>
```

Final Thoughts: Other XML Data

The previous examples involved a gloss entity, as that is the most complicated form of entity found within a SignStream collection. There are other, lesser complicated, forms of entities as well, which nonetheless follow a similar structure. Most of the same standards and rules apply to them, so the same lessons can be applied to those entities across other fields as well. In looking to find entities from a specific field, it will probably be necessary to consult the defCodingScheme.xml file within the SignStream®3 distribution folder to find the ID value of the field, and then use that find to the relevant entities in the collection file for that field.

For additional SignStream information and reference materials, see https://www.bu.edu/asllrp/SignStream/3/ .