# SALT Transcripts

"Systematic Analysis of Language Transcripts" software (SALT) is a system designed, at the University of Wisconsin-Madison primarily by Jon F. Miller, for speech/language therapists to help them perform Language Sample Analysis (LSA).

SALT transcripts are essentially plain-text documents that use predefined transcription conventions that cover specification meta-data, as well as textual annotations, as seen in Figure 1.

LaBB-CAT recognises and imports SALT transcripts (files ending in .slt). The transcript conventions cover a closed set of meta-data and annotation types, and many of the defined annotations can be imported into LaBB-CAT, as long as attributes and layers are pre-defined in LaBB-CAT in order to receive the annotations. If layers are not specifed in LaBB-CAT, transcripts can still be imported, but annotation data is necessarily lost during import.

Links between SALT annotations and LaBB-CAT layers must be explicitly configured in LaBB-CAT i.e.

1. You must create appropriate word/phrase layers and participant/transcript attributes to hold the SALT annotations you want to parse out of the transcripts.
2. Then you must select the *converters* menu option, and press the configuration button for the *SALT transcript* converter and specify/save the settings, as shown in Figure 2. (There will be two SALT options in the list, one for importing SALT transcripts, and one for exporting them, but they both use the same settings, so pressing the configuration button for either will have the same effect.)

## Meta-data header

The following SALT headers can be imported as participant or transcript attributes into LaBB-CAT:

- ParticipantId
- Language
- Gender
- Dob - date of birth
- Doe - date of sample
- Ethnicity
- Context - sampling context
- Subgroup - Subgroup/story
- Collect - collection point number
- Location - place of sample

```
$ Child, Examiner
+ Language: English
+ Name: EXAMPLE
+ Gender: F
+ Dob: 1/11/2016
+ Doe: 1/12/2021
+ Ca: 5;1
+ Ethnicity: Māori
+ Context: Nar
+ Subgroup: TATP
+ Collect: 4
- 0:00
E Mm.
- 0:01
C Tama say goodbye.
- 0:04
E Yeah.
- 0:04
C (That and) and his mum was busy cook/ing.
- 0:10
C But he[EP:she] was ring/ing on the phone.
- 0:12
C But he really want/*ed to play with him[EP:her].
```

Figure 1: Example of SALT transcript, including a meta-data header, Child/Examiner speaker turns, and morpheme/error annotations. (Times are typically in whole seconds, but it's recommended to have more granular synchronization, as seen in Figure 4)

| | |
|---|---|
| **C-Unit Layer** | cunit |
| **Target Participant Layer** | main_participant |
| **Comment Layer** | comment |
| **Parenthetical Layer** | parenthetical |
| **Proper Name Layer** | entity |
| **Repetitions Layer** | repetition |
| **Root Layer** | root |
| **Error Layer** | error |
| **Sound Effects Layer** | sound_effect |
| **Pause Layer** | silence |
| **Bound Morpheme Layer** | bound_morpheme |
| **Maze Layer** | maze |
| **Partial Word Layer** | partial_word |
| **Omission Layer** | omission |
| **Code Layer** | code |
| **Language Attribute** | transcript_language |
| **Participant ID Attribute** | participant_id |
| **Gender Attribute** | participant_gender |
| **Date of Birth Attribute** | participant_date_of_birth |
| **Date of Sample Attribute** | transcript_recording_date |
| **Current Age Attribute** | transcript_ca |
| **Ethnicity Attribute** | participant_ethnicity |
| **Context Attribute** | transcript_context |
| **Sub-Group Attribute** | transcript_subgroup |
| **Collection Point Attribute** | [none] |
| **Location Attribute** | transcript_location |
| **Date format** | d/M/yyyy |
| **Parse Inline Conventions** | ☑ |

Figure 2: Example of a SALT converter configuration in LaBB-CAT

The assumed format for date values (`Dob`/`Doe`) – i.e. whether month or day comes first, and whether the year has two or four digits – is specified by the SALT converter's *Date format* setting. However, dates are stored by LaBB-CAT in ISO format - i.e. `yyyy-MM-dd`

## Word tag annotations

The following annotations on words can be extracted into *word layers* if desired (when setting up each word layer to receive the tags, ensure with 'Alignment' set to *None*):

- Root - word stems marked with |, e.g. `wented|went`
- Bound Morpheme - affixes marked with / e.g. `ring/ing`, `wante/*ed`
- Partial Word - cutoffs suffixed with * e.g. `ever*` `everybody`

## Other in-situ annotations

The following annotations, which can span multiple words, or do not apply directly to words, can be extracted into *phrase layers* if desired:

- C-Unit - i.e. divisions into utterance based on where full-stops/periods and other specific punctuation appear in the transcript
- Parenthetical - enclosed in `((` and `))`
  e.g. `the boy ((I can/'t remember his name)) left`
- Proper Name - multi-word names delimited by `_`
  e.g. `Mr_Jones`
  – these are split into individual tokens where the `_` delimiters occur, which are tagged on the phrase layer specified by the *Proper Name Layer* setting, if specified.
- Repetition - repetitions delimited by `_`
  e.g. `heaps_and_heaps|heaps`
  – similarly these are tokenised on `_` and the multiple tokens are tagged on the selected *phrase layer*.
- Error - tags in square brackets, where the tag starts `E`
  e.g. `he[EP:she] wented|went[EO:went]`
  – these can encompass multiple words, e.g utterance errors `[EU]`
- Code - tags in square brackets, where the tag doesn't start with `E`
  e.g. `it's your turn to tell the story X[REDACTED]`
- Sound effect - prefixed by `%`
  e.g. `%yip_yip went Schnitzel_von_Crumm`
- Pause - prefixed by `:`
  e.g. `um :02 because`
- Maze - false starts, repetitions, etc. enclosed in `(` and `)`
  e.g. `(That and) and his mum`

- Omission - missing words prefixed by *
  `call/ing *on the phone`
- Comment - enclosed in { and }
  e.g. `{pretends to eat}`
  There is a system *span layer* that can be used to extract comments.
- background noises, etc. - comments that start with a single word followed by a colon can be extracted into their own *span layer* e.g. comments like `{NOISE:phone ringing}` can be mapped at upload time to LaBB-CAT's *noise* layer.

## Other Settings

Each SALT transcript starts by declaring the IDs of the speakers, the first speaker being the 'target speaker'. This translates to LaBB-CAT's concept of a 'main participant', and there is a setting called *Target Participant Layer* making this link explicit.

As previously mentioned, the expected format of dates is specified by the *Date format* setting. Options are:

- `M/d/yyyy` - month first, four digit year
- `d/M/yyyy` - day first, four digit year
- `M/d/yy` - month first, two digit year
- `d/M/yy` - daye first, two digit year

Also, all of the above annotation parsing can be disabled by un-ticking the *Parse Inline Conventions* check box. If this is done, then no annotation parsing is done by LaBB-CAT, and all words, codes, pauses, etc. will appear as-is in the word layer, no annotations or tags will be added, and the text on the word layer will not be normalized.

## Orthography Normalization

Unless parsing of inline conventions is disabled, the word token spellings are normalized during input, so that the resulting word orthographies are the normal spellings that might be found in dictionaries, recognised by syntactic parsers, etc.

This is to ensure that SALT transcripts can be automatically annotated in LaBB-CAT in the same way that other formats can be.

For example, an utterance in a SALT transcript might look like this:

`But he[EP:she] was ring/ing on the phone.`

In LaBB-CAT, after upload, this utterance will look like this:

```
But he was ringing on the phone.
```

i.e. the slashes marking morphem boundaries, and any comments, codes, and any other annotations are stripped out, and where possible the spelling are normalised (e.g. `smile/ing` will become `smiling`).

If layers have been configured to receive the annotations, then no information is lost; it's parsed out into separate layers in LaBB-CAT, as seen in Figure 3.



Figure 3: Error and bound morpheme tags have been extracted out on to corresponding annotation layers in LaBB-CAT

## Synchronization

It is imperative that each utterance is bounded by a synchronization time mark - i.e. a line starting with : which specifies the time the utterance occured in the recording.

LaBB-CAT needs utterance start/end times in order to synchronized playback, and also perform automatic annotations that relate the transcript to the audio, like forced alignment.

The SALT transcript conventions generally include examples of time synchronization to the nearest second. However, this is often not accurate enough for some LaBB-CAT processing, and specifying times accurate to the millisecond are recommended, as shown in Figure 4.

```
-  0:00.34
E  Mm.
-  0:01.26
-  0:02.06
C  Tama say goodbye.
-  0:04.06
-  0:04.14
E  Yeah.
-  0:04.99
-  0:05.77
C  (That and) and his mum was busy cook/ing.
-  0:10.16
-  0:10.21
C  But he[EP:she] was ring/ing on the phone.
-  0:12.25
```

Figure 4: Utterance synchronization to the milliscond is recommended