SALT Transcription Rules

**Contents**

[A. C-unit utterance segmentation 2](#_Toc505088732)

[B. Transcript format 2](#_Toc505088733)

[C. Bound morphemes 3](#_Toc505088734)

[D. Spelling conventions 5](#_Toc505088735)

[E. Parenthetical Remarks 6](#_Toc505088736)

[F. Unintelligibility 6](#_Toc505088737)

[G. Mazes (filled pauses, repetitions, revisions) 7](#_Toc505088738)

[H. Overlapping speech and interjections 7](#_Toc505088739)

[I. Pauses (unfilled) 8](#_Toc505088740)

[J. Omissions and errors 8](#_Toc505088741)

**Transcription Resources**

Reference documents found at <http://saltsoftware.com/resources/tranaids>

* Summary of Transcription Conventions
* Summary of C-unit Segmentation Rules

Free online classes found at <http://saltsoftware.com/training/self-paced-online-training>

* 1300: Transcription Quick Start
* 1301: Transcription – Getting Started
* 1302: Transcription – Transcript Format
* 1303: Transcription – Utterance Segmentation
* 1304 - 1306: Transcription Conventions
* 1308: Transcription – Practice Samples



Help built into SALT software

* F1 key for context-specific help (*provides transcription conventions while in editor*)
* Resources accessible from the Help menu

1. **Communication Unit (C-Unit) Utterance Segmentation**

Communication units are defined as “an independent clause and its modifiers”. In other words, a C-unit includes one main clause with all subordinate clauses attached to it. A communication unit is an utterance that cannot be further divided without the disappearance of its essential meaning. Subordinate clauses depend on the main clause to make sense. They cannot stand alone or be separated from the main clause.

*Helpful Tips*

Conjunctions: Below is a list of common coordinating and subordinating conjunctions that may help in understanding C-units. Coordinating conjunctions alert us that the utterance can stand alone as a separate C-unit. However, if an utterance contains a subordinating conjunction, the utterance should not be separated into separate C-units because a) the essential meaning would change, and b) dependent clauses cannot stand alone.

* Coordinating Conjunctions (signals to segment into new C-unit)

*and, but, or, and then, then, so* (not “*so that*”)

* Subordinating Conjunctions (do not segment into new C-unit)

Early development: *because, that* (“*that*” may also be an implied subordinator), *when, who*

Later development: *after, before, so that, which, although, if, unless, while, as, how, until, as\_\_as , like, where*

Pauses and intonation: Do not ignore pauses and intonation when segmenting utterances but, whenever reasonable, segment utterances based on grammar rules.

*Rules and Examples*

For comprehensive rules with examples, download the “Summary of C-unit Segmentation Rules” at <http://saltsoftware.com/resources/tranaids>

|  |
| --- |
| Practice #1: C-unit Segmentation |

1. **Transcript Format**

Line identifiers. Each entry must begin with one of the following identifiers:

|  |  |
| --- | --- |
| $ | Identifies speakers |
| + | Header information, e.g., demographics |
| C or E or ..... other | Speaker identifier |
| = | Transcriber comment |
| ; or : | Between-utterance pause |
| - | Timing line |

Ending punctuation. All utterances must end with one of the following punctuation marks:

|  |  |
| --- | --- |
| . ! ? | Statements, exclamations, questions |
| > | Abandoned utterances |
| ^ | Interrupted utterances |
| ~ | Intonation prompts |

Timing Lines.

Mark beginning and ending time. SALT uses timing lines to calculate elapsed time and speaking rate. Begin timing lines with a hyphen followed by the elapsed time in :sec, min:sec, or hours:min:sec.

Transcriber Comments

= Comment lines begin with an equal sign

{ } Comments within utterances are enclosed within braces

*Example of transcript format:*

$ Child, Examiner

+ Name: Chance

+ Gender: M

+ CA: 8;1

+ Context: Nar

+ Subgroup: SSS

- 0:00

C (Um) he/'s the guy that (make/3s) make/3s hay out of (go\*) gold.

C And (he he he) he use/3s (um) something to spin it around with because (um) the king only (was was was) was greedy.

C So he said this will (m\*) make the hay out of gold.

:03

E Can you tell me what happen/ed first in the story?

C (Um :02) well there was something XX {whispers}.

= lots of background noise

C And he jump/ed up and down and say[EW:said] "no, no, that was/n't his name".

C But at the end (he) she said "(is her na\*) is his name (Rump\*) Rumpelstiltskin"?

.....

C And that/'s all.

- 3:12

**C. Bound Morphemes**

*Format: use the root form of the word followed by the bound morpheme, i.e., root/bm.*

**1. Plurals and Possessives**

* /S for plurals, e.g., *dog/s, story/s*
  + **Do not mark plurals if there isn’t a singular form, e.g., *news, tidings***
  + **Do not mark irregular forms, e.g., *mice*, *wolves, deer, geese***
* /Z for possessive inflection, e.g., *Mary/z car*
  + Do not mark possessive pronouns, e.g.,*mine,* ***his, hers, its, ours, yours, theirs***
* /S/Z for plural and possessive, e.g., *baby/s/z names*

|  |
| --- |
| Practice #2: Bound Morphemes – Plurals and Possessives |

**2. Verb inflections**

* **/ING for present progressive** 
  + As a rule of thumb... present progressive verbs, which are marked as bound morphemes, typically follow an auxiliary "be" verb (is, are, am, was, were, be, been). Use the root form of the verb followed by the bound morpheme “/ing”.

Examples:

He is swim/ing at the lake.

Are they swim/ing now?

They were make/ing a cake.

They/’re leave/ing soon.

* Gerunds, which end in -ing, should not be marked as a bound morpheme. Gerunds function as nouns, i.e., gerunds function as subjects, direct objects, indirect objects, and objects of prepositions.

Examples:

Dave enjoys **swimming** more than any other sport.

They went **swimming**.

I dream about **winning**.

I look forward to **going** home.

Hint: If an “-ing” word follows a preposition, it is a gerund.

|  |
| --- |
| Practice #3: Bound Morphemes – Present possessive & Gerunds |

* **/ED for regular past tense**, e.g., *look/ed, stop/ed, try/ed, believe/ed*
  + Do not mark irregular forms, e.g., *were, saw, went, had, made*
  + Past participles, passive forms, and predicate adjectives

Although sometimes confused with past tense verbs, these forms are not verb forms and the -ed at the end of these words are NOT marked as bound morphemes. For example,

Past participles: Her hair was well **brushed**. She was **tired**. The porcupine was **named** Fluffy.

Passive forms: The store was **robbed** last night. The vase is **smashed**.

Predicate adjectives: The grape/s became **shriveled** in the sun. I like car/s **produced** in Japan.

**Hint: regular past tense /ED verbs never follow a BE or GET verb.**

|  |
| --- |
| Practice #4: Bound Morphemes – Regular past tense |

* /3S for 3rd person singular verb forms, e.g., *she run/3s*
  + **Do not mark irregular forms, e.g., *has, was***
  + **Do not mark forms if the sound of the root changes, e.g., *do 🡪 does***
* **/EN for past participle**, e.g., *take/en, eat/en, prove/en*
  + Do not mark irregular forms, e.g., *gotten, spoken, seen, been*
  + **Do not mark forms if the sound of the root changes, e.g., *write 🡪 written***

**Hint: regular form is present tense + EN as a separate syllable. /EN verbs always follow a HAVE verb.**

|  |
| --- |
| Practice #5: Bound Morphemes – 3rd person singular (/3s) & Past participle (/en) |

1. **Contractions**

* /’T, /N’T for negatives, e.g., *can/’t, did/n’t, is/n’t*
  + **Do not mark forms if the sound of the root changes, e.g., *do 🡪 don’t, will 🡪 won’t***
  + **Do not mark the word *ain’t* as there isn’t a non-negative form**
* /’LL, /’M, /’D, /’RE, /’S, /’VE for contracted “will”, “am”, “would”, “are”, “is”, “have”

Examples: *they/’ll, I/’m, she/’d, we/’re, he/’s, we/’ve*

* /H’S, /H’D, /D’S, /D’D, /’US for contracted “has”, “had”, “does”, “did”, “us”

Examples: *He/h’s been sick.*

*We/h’d better go.*

*What/d’s he do for a living?*

*Why/d’d he leave?*

*Let/’us go.*

|  |
| --- |
| Practice #6: Bound Morphemes - Contractions |

**D. Spelling conventions**

Since each word with even a slightly different spelling appears as a different word in SALT analysis, it is important to follow standard spelling conventions to ensure consistency within and between transcripts.

* Abbreviations should either be spelled out or left as an abbreviation but WITHOUT the period, e.g., *Mr, Mon, Monday*. Periods may only be used to mark the end of an utterance.
* Yes words: *ok, aha, mhm, uhhuh, yeah, yep, yes*
* No words: *no, ahah, mhmh, uhuh, nah, nope*
* Hyphenated words follow standard spelling conventions, e.g., *mother-in-law, pick-me-up*.
* Numbers and dates: You should transcribe numbers and dates as a single word in either written form or as digits, e.g.,*21 or twenty-one*, *April\_13\_2018* or *4-13-2018*. Choose a format and be consistent.
* Clock time: Do not use colons when transcribing clock time because colons will be interpreted as pause times. Type out the words connected with an underscore, e.g., *eight\_thirty, five\_oclock*
* Counting and “spelled” words use the underscore to connect the numbers or letters,   
  e.g., 1\_2\_3\_4\_5, C\_A\_T
* Concatenatives: Words, such as gonna, wanna, and hafta, occur when the syllables of common words are blended together. Transcribe concatenatives as a single word.

|  |  |  |  |
| --- | --- | --- | --- |
| *betcha* (bet you)  *coulda* (could have)  *gonna* (go/ing to) *gotta* (got to) *hafta* (have to) | *liketa* (like to)  *lookit* (look at it)  *musta* (must have)  *oughta* (ought to) | *outta* (out of)  *shoulda* (should have)  *sposta* (suppose to) *trynta* (trying to) | *useta* (used to) *wanna* (want to) *whatcha* (what are you)  *woulda* (would have) |

* Other Shortened Words: If the speaker says *'cuz* instead of *because*, either type out the full word *because* or use the vertical bar to identify the word. In the following example, the speaker said “cuz”. No matter which form you use, the word root in this example is analyzed as “because”.

Example: *He was sad because they left.*

*He was sad cuz|because they left.*

* Proper Names and Titles: Enter proper names or titles as a single linked word, joined together with the underscore character, e.g., "Little Red Riding Hood" should be transcribed as *Little\_Red\_Riding\_Hood* and "Grandma Jones" should be *Grandma\_Jones*. When words are linked, they are counted as a single unit (word) in SALT analyses.
* Idiosyncratic Forms: Young children often produce speech that differs from the adult version. Although these idiosyncratic forms are not adult-like productions, they are stable productions by the child rather than babbling, mispronunciations, or word errors. When transcribing idiosyncratic forms, enter a percent sign (%) at the beginning of the idiosyncratic form. For example,

C My %vroom {car}.

E Would you like a cookie?   
C  %coopa {cookie}.

* Sound Effects: Sound effects are non-word vocalizations which represent specific sounds such as those made by an animal, e.g., *meow*, or an object, e.g., *vroom*.
  + If the sound effect is essential to the meaning or structure of the utterance, it should be entered as any other word except it should begin with a percent sign (%).

C The dog went %grr.   
C Then the boy heard %ribbit\_ribbit.

* + If the sound effect is not essential to the meaning or structure of the utterance, note its occurrence with a comment using braces { } or a separate line beginning with an equal sign =. For example,

C The dog growl/ed {child makes growling sounds} at them.   
C They fell down.

= child makes falling down sounds

* Sounds with specific meanings: *hmm, huh* (question or affirmation), *IDK* (intoned *I don’t know*), *uhoh* (something is wrong), *shh* or *%shh* (be quiet), *psst* or *%psst* (to get someone’s attention)
* Other spellings: *ain’t, alot, atta, no one, oh, ooh, oop, opps, oopsy*

**E. Parenthetical Remarks**

A parenthetical remark is a word or clause that has been added by the speaker as an explanation, comment, or request for help. To code parenthetical remarks, enclose them in double parentheses. Parenthetical remarks should be coded if 1) they occur in the middle of an utterance and do not contribute to the meaning of the utterance, or 2) they occur as stand-alone utterances but do not contribute to the content of the sample.

Parenthetical remarks which occur in the middle of an utterance

When a parenthetical remark occurs in the middle of the utterance does not contribute to the rest of the utterance, enclose it in double parentheses. For example,

C The boy ((I can/'t remember his name)) left the house.   
C Then the ((what/'s that call/ed)) < > gopher bit him on the nose.   
E <gopher>.

Parenthetical remarks which occur as stand alone utterances

When a parenthetical remark occurs as a stand-alone utterance, you need to decide whether or not you want that utterance included when analyzing the sample.

For conversational samples, you may want these utterances included as they occur naturally in conversation. For example,

C Last weekend I went to my grandma/z house.

C  I forgot what I was go/ing to say.   
E What did you do at your grandma/z?

For story retell and expository samples, however, you may not want them included unless they contribute to the content of the narrative. For example,

C  ((I skip/ed a page)).  
C  (((Where) where was I))?

C ((Can I stop now))?

**F. Unintelligibility**

Unintelligibility can be the result of many things including the speaker turning away from the microphone or speaking too softly, equipment failure, or background noise. It can also be caused by the speaker's phonological difficulties or due to an unfamiliar listener’s perception. As a general rule, if you cannot understand all or part of an utterance after listening three times, it should be marked as unintelligible.

Marking Unintelligible Segments: Use an X to mark an unintelligible word. Use XXX when the entire utterance is unintelligible. And use XX to mark a segment which is anywhere between a single word and the entire utterance. Note that it is often difficult to distinguish whether the unintelligible segment consists of a single word, multiple words, or the entire utterance (after all, it's unintelligible).

Example *1*: unintelligible word

C He X away yesterday.

Example *2*: unintelligible segment, somewhere between a single word and the entire utterance.

C He XX yesterday.

Example *3*: unintelligible utterance

C  XXX.

**G. Mazes (filled pauses, repetitions, revisions)**

Filled pauses**.** Words or vocalizations that fill in pauses should be placed in parentheses. Standard fillers include: *AH, EH, ER, HM, UH, UM*. Nonstandard fillers should be coded as [FP], e.g., *like[FP], you\_know[FP]*.

Part-words. Use an asterisk to replace the portion of the word that is missing when a speaker fails to complete a word, e.g., *(b\* b\*)*. This form is used to mark stuttering as well as part-word revisions. Unless part-words occur at the end of an interrupted or abandoned utterance, they are treated as maze components and are parenthesized. To mark stuttering in the middle of a word, separate the parts of the word, before and after the stuttering, with underscore characters, e.g., *hap\_ (p\* p\*) \_py*.

Repetitions or revisions. The speaker’s final expressive attempt is considered the resolved utterance, and the earlier attempts are mazed. All repetitions and revisions should resolve themselves so do not end an utterance with a repetition or revision unless the utterance is abandoned or interrupted.

Adjacent mazes. Combine adjacent mazes into a single maze.

Examples:

C And then (um he went) he went to the park.

C And (when the b\*) if the batter hit/3s the ball, he run/3s (to um) to first base.

|  |
| --- |
| Practice #7: Mazes (filled pauses, repetitions, revisions) |

**H. Overlapping speech and Interjections**

Transcription is linear but sometimes the language sample is not. Use angle brackets < like these > to mark the overlapping speech in both speakers’ utterances.

Example: E Which book do you want <to read>?

C <This one>.

Use empty angle brackets < > for interjections that occur within the other speaker's utterance but do not overlap the speech.

Example: C I don’t want it < > anymore.

E <Ok>.

|  |
| --- |
| Practice #8: Overlapping speech & Interjections |

**I. Pauses (unfilled)**

Enter the length of each significant pause (2 seconds or longer) at the point where it occurs in the transcript. Pause length is measured to the nearest second.

Pauses within utterances

Enter the length of the pause in seconds at the position in the utterance where it occurs. Pause time should be separated with a blank space from any adjacent words. If the pause occurs immediately before or after a maze, include it within the maze. For example,

C The movie :02 was (um :05) Batman.

Pauses between utterances

Format pause in seconds or minutes:seconds. If the pause occurs between utterances of different speakers, begin the pause line with a colon. If the pause occurs between utterances of same speaker, you almost always begin the pause line with a semicolon. You would only begin a pause line with a colon if you want to force a turn change.

Pause time may be entered as:

:sec or ;sec        { *this simplified format was introduced with SALT 18* }

:  :sec or ;  :sec

:  min:sec or ;  min:sec

|  |
| --- |
| Practice #9: Pauses (unfilled) |

**J. Omissions and Errors**

Omissions

An omission occurs when one or more words or bound morphemes that are obligatory for grammatical correctness are absent. For example, if the speaker says, "His name Larry", the copula "is" should be marked as an omission. On the other hand, if the examiner asks, "What are you doing?" and the speaker responds "Eating", there isn't an omission because no obligatory content was left out.

* *Omitted Words*

The asterisk symbol (\*) is used to indicate an omitted word. At the point in the transcript where the word was omitted, type an asterisk followed by the omitted word. There should be no blank spaces between the asterisk and the omitted word. For example,

*This \*is a cookie.*

* *Omitted Bound Morphemes*

A slash followed by an asterisk is used to indicate the omission of a bound morpheme in obligatory context. Type the slash, the asterisk, and the missing bound morphemes at the point in the transcript where they've been omitted. For example,

*The car go/\*3s here.*

* *Omitted Contractions: Words vs Bound Morphemes*

Omitted contractions may be transcribed in two ways, as an omission of a word or as an omission of a bound morpheme. It doesn't make much difference which way you mark this type of omission since omissions are not included in any of the calculations based on words or morphemes. The important thing is to mark that the omission occurred. For example,

*She \*is leave/ing now.* OR *She/\*’s leave/ing now.*

Errors

SALT recognizes a special category of codes called "error codes". These codes are only used to mark errors in syntax and morphology. They are not used to mark errors in semantics or pragmatics. SALT contains a list of default error codes which can be edited to suit your purposes. They include:

[EO:=] overgeneralization error (all codes that begin with “EO:”)

[EP:=] pronoun error (all codes that begin with “EP:”)

[EW] extraneous word

[EW:=] other word-level error (all codes that begin with “EW:”)

[EU] utterance-level error

* *Overgeneralization errors, the [EO:\_\_\_] code*

Consider the following examples of utterances containing overgeneralization errors:

C He **felled**.   
C There were **deers** in the woods.   
C That/'s **hises**.

To mark overgeneralization errors, first identify these words as instances of the root form, e.g., "*fall*", "*deer*", and "*his*". To identify the root form, follow each overgeneralized word with a vertical bar '|' and the root form. Then insert the code [EO] at the end of the overgeneralized word. Put the target word inside the brackets following a colon:

C He felled|fall**[EO:fell]**.   
C There were deers|deer**[EO:deer]** in the woods.   
C That/'s hises|his**[EO:his]**.

C He lookeded|look/ed**[EO:looked]** in a tree.

Notice, in the first three examples, that the bound morphemes in the overgeneralized words are not marked because we don't want to give the speaker credit for a bound morpheme on a word that is not a real word. In the last example, however, the speaker is given credit for one of the past tense bound morpheme.

* *Pronoun errors, the [EP: \_\_\_] code*

Use this code to mark errors involving personal, possessive, or reflexive pronouns. The correct pronoun is put inside the brackets following a colon. For example,

C  Him**[EP:he]** was sick.   
C So them**[EP:they]** look/ed for the frog.

Notice that each error code is attached to the specific pronoun that is used incorrectly.

* *Other word-level errors, the [EW] and [EW:\_\_\_\_] codes*

Use the code [EW] to mark other words that are used incorrectly. The correct word, if known, is put inside the brackets following a colon. For example,

C The big frog were**[EW:was]** mad.

C  They look/ed on**[EW:in]** a hole.   
C And then the turtle was of**[EW]** better.

Notice that each error code is attached to the specific word that is used incorrectly. In the last example, the word "of" is an extraneous word; there is no correct substitution.

* *Utterance-level errors, the [EU] code*

Use the utterance-level error code [EU] to mark the following types of errors:

- Errors that cannot be associated with a specific word

- Utterances which contain more than two omissions and/or word errors; to avoid over-correcting them

Insert the [EU] code between the last word and the ending punctuation mark. For example,

C He book [EU].   
C And they came stop/ed [EU].

C He was have/ing a frog [EU]?

C The kid tell everything to him happen [EU].

When NOT to mark omissions and errors:

Don’t hold the speaker to “grammatically perfect” language. Instead, consider whether or not the language is appropriate for the speaker’s age and the context in which it’s used. Does the utterance sound ok or does it “sting your ear”? Also, do not mark utterances which contain unintelligible segments since it’s possible that the unintelligible segments might have made the utterances acceptable.

|  |
| --- |
| Practice #10: Omissions and Errors |

|  |
| --- |
| Practice #11: Putting it all together – symbols and codes |
| Practice #12: Putting it all together – sample transcript |