



Summary of SALT Transcription Conventions

1. Transcript Format. Each entry begins with one of the following symbols. If an entry is longer than one line, continue it on the next line.

- \$ Identifies the speakers in the transcript; always the first line of the transcript. Example: \$ Child, Examiner
- C Child/Client utterance. The actual character used depends on the \$ speaker line.
- E Examiner utterance. The actual character used depends on the \$ speaker line.
- + Header information such as name, age, context, codes used in sample (e.g., + CA: 5;7). Also used for inserting test scores.
- Time marker. Example of two-minute marker: - 2:00
- : Pause between utterances of different speakers. Example of five-second pause: : :05 or :05
- ; Pause between utterances of same speaker. Example of three-second pause: ; :03 or ;03
- = Comment line. This information is used for transcriber comments and is not analyzed in any way.

2. End of Utterance Punctuation. Every utterance must end with one of these six punctuation symbols.

- | | | | |
|---|--|---|--|
| . | Statement, comment. Do not use a period for abbreviations. | ^ | Interrupted utterance. The speaker is interrupted and does not complete the thought/utterance. |
| ! | Surprise, exclamation. | > | Abandoned utterance. The speaker does not complete the thought/utterance but has not been interrupted. |
| ? | Question. | | |
| ~ | Intonation prompt. | | |
- Example: E And then you have to~

3. { } Comments within an utterance.

Example: C Lookit {C points to box}.

Nonverbal utterances with communicative intent are placed in braces. Example: C {nods}.

4. Unintelligible Segments. X is used to mark unintelligible sections of an utterance. Use X for an unintelligible word, XX for an unintelligible segment of unspecified length, and XXX for an unintelligible utterance.

Example 1: C They went X X park. Example 2: C He XX today. Example 3: C XXX.

5. Bound Morphemes. Words which contain a slash "/" indicate that the word is inflected or contracted in a regular manner. The root word is entered in its base spelling followed by a slash "/" and then the bound morpheme.

English and Spanish

/S Plural. Examples: kitten/s, baby/s, rana/s, flor/s.

Do not mark words that end in "s" but represent one entity (e.g., pants, binoculars).

Do not mark irregular forms (e.g., mice, geese, deer) or when the sound of the root changes (e.g., leaves, wolves).

English only

/Z Possessive inflection. Examples: dad/z, Mary/z

Do not mark any possessive pronouns (e.g., mine, his, hers, ours, yours, its, theirs).

/S/Z Plural and Possessive. Example: baby/s/z

/3S 3rd Person Singular verb form. Examples: go/3s, tell/3s, try/3s

Do not mark irregular forms (e.g., has, was) or when the sound of the root changes (e.g., do→does, say→says).

/ED Past tense. Examples: love/ed, die/ed

Do not mark irregular forms (e.g., did, grew, had, sank) or predicate adjectives (e.g., was tired, are bored, got fixed).

Transcription hint: regular past tense /ED verbs never follow a BE or GET verb.

/EN Past participle. Examples: take/en, eat/en, prove/en

Do not mark irregular forms (e.g., gotten, spoken, seen, been) or adjectives (e.g., was eaten, were taken). Also, do not slash verbs when the sound of the root changes (e.g., write→written).

Transcription hint: Regular form: present tense + EN as separate syllable. /EN verbs always follow HAVE, HAS, or HAD.

/ING Progressive verb form. Examples: go/ing, run/ing, bike/ing

Do not mark the gerund use of the verb form (e.g., went swimming, reading is fun).

/N'T, /'T Negative contractions. Examples: can/'t, does/'n't

Do not mark irregular forms (e.g., won't) or when the sound of the root changes (e.g., do→don't).

/'LL, /'M, /'D, /'RE, /'S, /'VE Contracted → WILL, AM, WOULD, ARE, IS, HAVE

Examples: I/'ll, I/'m, I/'d, we/'re, he/'s, we/'ve

/H'S, /H'D, /D'S, /D'D, /'US 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 the boy look there? Let/'us go.

6. Bound Pronominal Clitics (Spanish only). Pronominal clitics may be either bound or unbound. When bound, they are preceded by a plus sign. Examples: gritando+le, deja+lo, dá+me+lo

7. Mazes. Filled pauses, false starts, repetitions, and reformulations.

() Surrounds the words/part-words that fall into these categories. Combine adjacent mazes.

Example: C And (then um) then (h*) he left.

8. Omissions. Partial words, omitted words, omitted bound morphemes, and omitted pronominal clitics are denoted by an asterisk (*).

* Following one or more letters indicates that a word was started but left unfinished.

Example: C I (w* w*) want it.

* Preceding a word indicates that an obligatory word was omitted.

Example: C Give it *to me.

/* Following a slash the * is followed by the obligatory bound morpheme which was omitted.

Example: C The car go/*3s fast.

+* Following a plus sign the * is followed by the obligatory Spanish pronominal clitic which was omitted.

Example: C Él está gritando+*le a la rana.

9. Overlapping Speech. When two speakers are speaking at the same time, the words which occur at the same time are surrounded by angle brackets <>. Example: C Can I have that <one>?

E <Uhhuh>.

When one speaker interjects in the middle of another speaker's utterance, use empty angle brackets to indicate the position of the interjection. Example: C I want you to do it <> for me.

E <Ok>.

10. Linked words. The underscore “_” is used to link multiple words so they are treated as a single word. Uses include:

Titles of movies and books.

Examples: Beauty_and_the_Beast, Frog_Where_Are_You

Proper names.

Examples: Mrs_Jones, Sr_Rojas

Words or phrases repeated multiple times.

Example: He ran ran_ran|ran as fast as he could.

11. Root identification. The vertical bar “|” is used to identify the root word.

Overgeneralization errors.

Example: C He goed|go[EO:went] to the park by himself.

Linked words repeated for emphasis.

Examples: C The boy ran very very_very|very fast. C Dijeron rana rana_rana|rana dónde estás.

Spanish only: Inflected word forms.

Example: C Había|haber una vez un niño que tenía|tener una rana.

Spanish only: Diminutives.

Example: C El perrito|perro tumbó|tumbar las abeja/s.

English only: Root form of irregular verbs.

Note that the root forms of irregular verbs are not identified in any of the SALT reference databases.

Example: C The bird flew|fly/ed away.

12. Sound Effects and Idiosyncratic Forms %. The percent sign is used to identify sound effects which are essential to the meaning or structure of the utterance. Non-essential sound effects are entered as comments. Strings of the same sound are linked together.

Example 1: C The dog went %woof_woof.

Example 2: C The dog barked {woof woof}.

The percent sign is also used to identify idiosyncratic forms. These are immature productions which are consistent in reference to an object, person, or situation.

Example 1: C See %vroom {car}.

Example 2: C My %coopa {cookie}.

13. Spelling Conventions.

- Abbreviations: Periods are not legal word characters. Abbreviated words should either be written out or left as an abbreviation but without the period. Examples: Mr, Mister, Mrs, Dr, Sra, Señora
- Filled pause words: AH, EH, ER, HM, HMM, UH, UM, MM, and any word with the code [FP]
- Yes words: OK, AHA, MHM, UHHUH (*English & Spanish*)
YEAH, YEP, YES (*English only*)
SÍ (*Spanish only*)
- No words: NO, AHAH, MHMH, UHUH (*English & Spanish*)
NAH, NOPE (*English only*)
- Hyphenated words follow standard spelling conventions. Examples: mother-in-law, pick-me-up, good-hearted, twenty-five
- Numbers (*examples*): twenty-one or 21; dieci-siete or 17 (*choose a format and be consistent*)
- Clock time: Do *not* use colons when typing clock time because it will be interpreted as a pause. Type out the words connected with an underscore character. Examples: eight_thirty, la_una_treinta
- Counting and “spelled” words use the underscore character to connect the numbers or letters.
Examples: 1_2_3_4_5, C_A_T
- Shortened words: Either use the vertical bar to identify the word or use the full word. In the following example, the speaker said “cuz”. No matter which format you use, the word root in this example is analyzed as “because”.
Example: C He was sad cuz|because they left.
C He was sad because they left.
- English concatenatives:

BETCHA (<i>bet you</i>)	LIKETA (<i>like to</i>)	OUTTA (<i>out of</i>)	USETA (<i>used to</i>)
COULDA (<i>could have</i>)	LOOKIT (<i>look at it</i>)	SHOULDA (<i>should have</i>)	WANNA (<i>want to</i>)
GONNA (<i>going to</i>)	MUSTA (<i>must have</i>)	SPOSTA (<i>supposed to</i>)	WHATCHA (<i>what are you</i>)
GOTTA (<i>got to</i>)	OUGHTA (<i>ought to</i>)	TRYNTA (<i>trying to</i>)	WOULDA (<i>would have</i>)
HAFTA (<i>have to</i>)			
- 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, %PSST (to get someone's attention)
- Other English spellings:
AIN'T, ALOT, ATTA, NOONE, OH, OOH, OOP, OOPS, OOPSY
- Reflexive vs Non-reflexive pronouns (Spanish only):
The following pronouns can be used both reflexively and non-reflexively: ME, TE, SE, OS, NOS. Attach the code [X] when used reflexively. Examples: C El niño se[X] fue con el perro.
C El perro me ayudó a conseguir la rana.

14. [] Codes. Codes are used to mark words or utterances. Codes are placed in brackets [] and cannot contain blank spaces. Codes used to mark words are inserted at the end of a word with no spaces between the code and the word. Codes used to mark utterances are inserted at the end of the utterance or at the position of the feature you wish to mark, separated by at least one space from adjacent word(s).

a) Codes used to mark errors in the reference database samples:

[EO: __] used to mark overgeneralization errors.	C He falled fall[EO:fell].
[EP: __] used to mark pronoun errors.	C And them[EP:they] found the frog.
[EW: __] used to mark other word-level errors.	C He were[EW:was] look/ing.
[EW] used to mark extraneous words.	C And then the boy is a[EW] sleep/ing.
[EU] used to mark utterance-level errors.	C And they came to stop/ed [EU].
[FP] used to mark non-standard filled pause words.	C The dog (um like[FP]) fell down.

b) Other codes used in the Bilingual Spanish/English reference database samples:

[F] used to mark fragments due to utterance segmentation in modified communication units.	C The gopher look/ed out of the hole. C and bit the boy [F].
[CS] used to mark code-switched words.	C The dog fell from la[CS] ventana[CS].
[WO] used to mark utterances with non-standard word order.	C And then fell down the dog and the boy [WO].
[I] used to mark vocabulary provided by the examiner (imitated word).	C And then the :05 <> owl[I] scare/ed him. E <Owl>.
[X] used to mark Spanish reflexive pronouns.	C El niño se[X] fue con el perro.