The Basics

Philosophy of transcription

Transcribing conversational speech sounds like a simple task – listening to people talk and writing down what you hear – but it is not. Once you have listened to a few audio samples, it will become very clear that people don't talk the way they write essays. Rather than consisting of a well-organized set of clearly defined sentences, conversational speech consists of a haphazard set of linguistic fragments – single words, single phrases, and multiple sentences loosely conjoined together – that only loosely conform to the linguistic concepts of "clause" or "sentence". The conversational speech equivalent of the "sentence" is something called the "utterance" – a loosely defined term that represents a single speech bout usually communicating a single idea bounded by acoustic boundaries like a pause or change in intonation. Add to these fluent but loosely structured forms are real speech disfluencies – hesitations, false starts, restarts and correction, in addition to grammatical speech errors that fall out unnoticed and uncorrected. Add to these forms the complexity of multiple speakers interacting, and a careful dance of give and take of the conversational floor, otherwise known as "turn-taking". This dance is often poorly orchestrated and people can step over each other's turns, resulting in overlapping utterances. In our lab, there are yet two final complications: We are transcribing conversational interactions involving infants and young children whose grasp of their native language is tenuous at best. And we are doing it from recordings in a noisy daycare environment.

Your goal as a transcriber is to convert the auditory stream of speech you hear into a written form. You have two somewhat competing goals: 1. To be as true as possible to what you hear in the auditory recording and 2. To convert this auditory record into something that can be easily analyzed by some simple computer scripts. If you keep these two competing goals in mind, many of the questions you might have about how to transcribe and code something should be resolved. If you are uncertain about the best way to transcribe or code something, it is always a good idea to ask – it's a whole lot easier to clear up misconceptions at the outset than to go back and fix transcripts after the fact (especially if we don't know the problem exists in the first place).

It is important to keep in mind that the conversion of an acoustic speech form to a written form is always an imperfect accomplishment. Two fluent English speakers can in many circumstances hear the same utterances in completely different ways. Your goal as a transcriber is to write down what you hear to the best of your ability. There will be times that you are uncertain what you are hearing. We do have a code (XXX) for speech that is too unclear to be understood. However, please use this code sparingly. If you have a fair degree of confidence in what you are hearing, go ahead and write it down, even if you wouldn't stake your life on it.

One final note: You will be working within a Transcriber file that already contains a variety of codes and other notes generated automatically by the LENA system. Please work around these notes and codes, and do not alter them or use them to make decisions about your own coding, as your work will be used for reliability on the LENA codes. The one exception to this rule is with respect to utterance boundaries: You can assume for the most part that the LENA-defined utterance boundaries are accurate. There are two exceptions to this: 1) they explicitly cut across a sentence (i.e., an utterance as defined by LENA that is not grammatical or fluent, but would be if

you included speech that spills into an adjacent segment) and/or 2) if LENA fails to break up utterances properly over a lengthy stretch – this usually occurs in "OLN" segments (overlapping speech or speech overlapping with other noises).

The nuts and bolts of the task

For a given audio recording you are asked to transcribe:

1. Listen to every segment of the audio recording, regardless of the LENA labels you see.
2. Make sure to listen at the level of the "tree" (each segment on its own), and at the level of the "forest" (across larger stretches of the recording without stopping at segment boundaries).
3. Transcribe what you hear for each segment in normal English words, but omit capital letters and punctuation (with the exception of apostrophes). Only transcribe words – do not transcribe non-verbal noises, physiological noises like sneezing, involuntary laughter (unless the person is saying "haha") or speech produced by a radio or television. However, the notion of "word" is loosely defined – animal sounds and other quasi-words like "oh" can be transcribed if they are intended meaningfully.
4. Please be accurate in your placement of words with a LENA-defined segment. If you select the particular segment to be played, Transcriber will play only the section you have selected. If a word crosses a LENA-defined boundary, you can finish the word in the original segment, but phrases that cross segments should be split across the segment as appropriate. Use the "I" code to designate that the utterance crosses a LENA-defined segment boundary (see below).

Figure 1. selecting segments

5. If you cannot make out what is being said with reasonable assurance, transcribe it as XXX. Babble (word-like sounds that don't mean anything) by a child is transcribed as BBL and an adult transcribed as XXX. Both XXX and BBL should be coded with "U" for codes 3 and 4 (see below). Yet, if you are certain that a segment finishes in the next utterance but you cannot understand what is being said, code 3 can be listed as "F".

Figure 2. BBL utterances

The BBL utterance can be used in many different cases, including in overlapping segments. The codes associated with BBL state that it's the target child producing the speech and all other information is unknown.
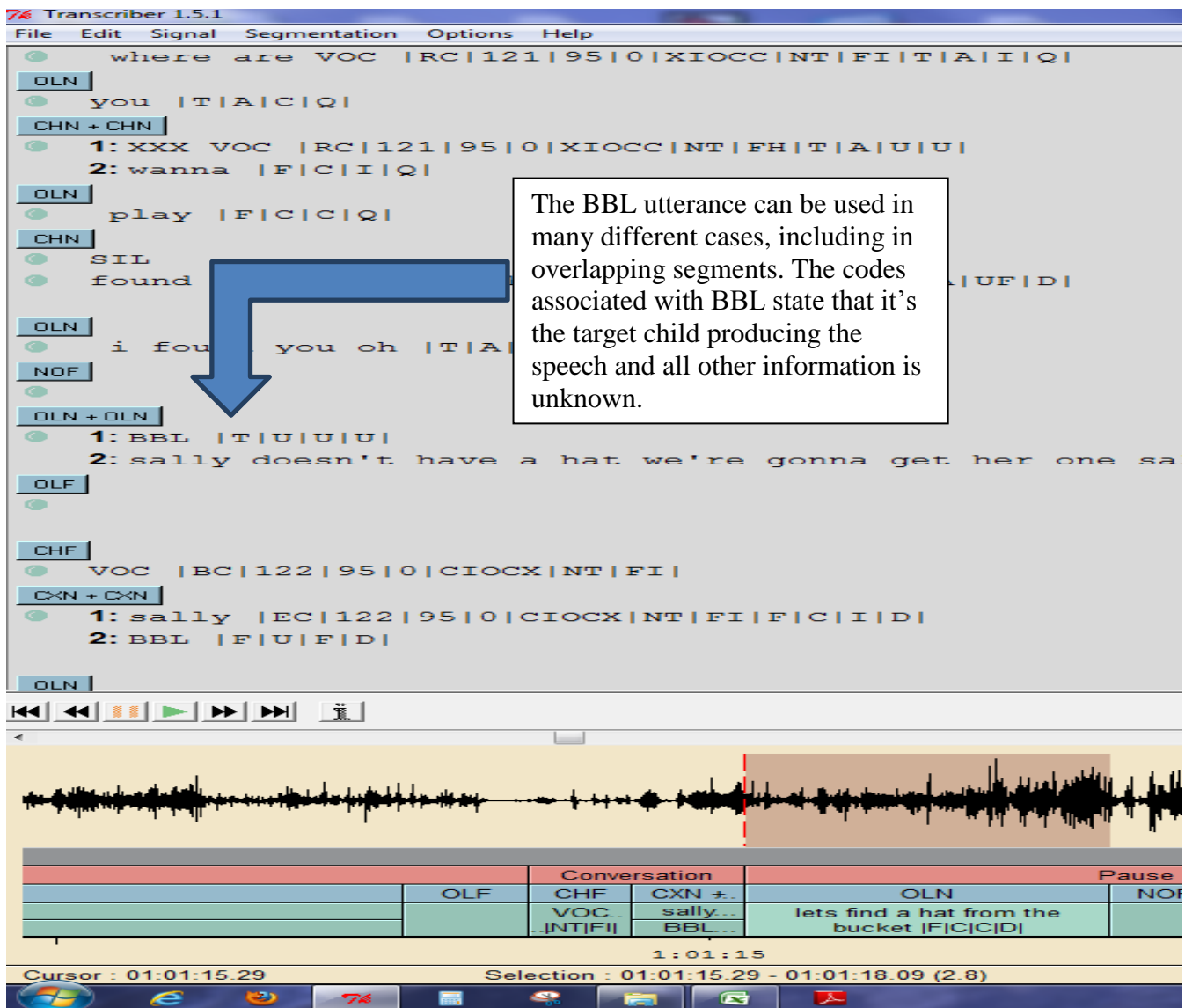
6. Add codes as appropriate. Transcriptions should appear at the beginning of a line of transcription (*before* any LENA "transcription" notes such as "VOC"). Codes should be appended onto the end of the LENA codes (separated by vertical bars), if any, or at the end of the line if there are no LENA codes. In the case of multiple utterances coded within a single LENA segment, place the codes for a given utterance immediately after that utterance, separated by a "." (see below). ***Under no circumstances should you edit any of LENA's codes or notes.***

Figure 3. a one second pause
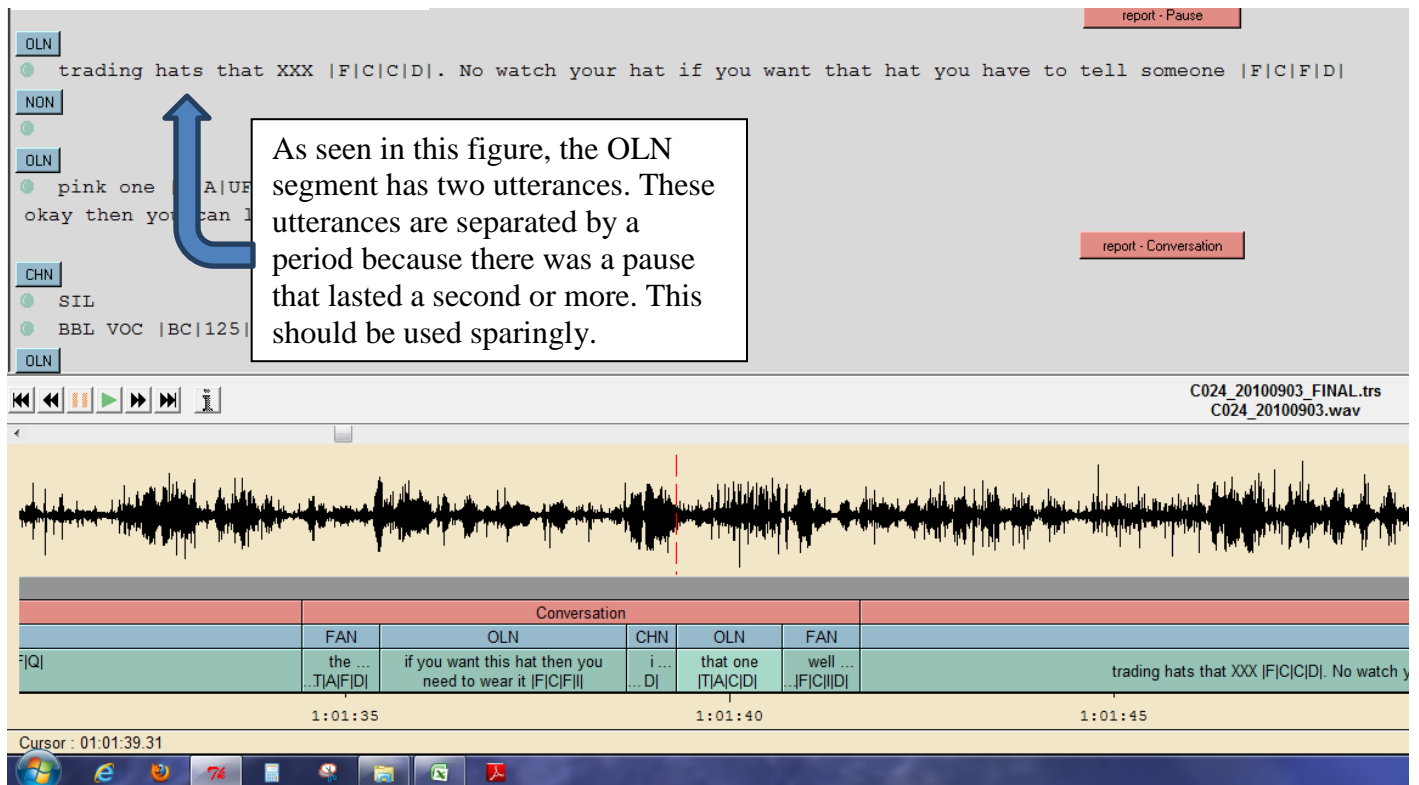
7. Please make note of any difficulties you encounter and bring them to my attention.

The Codes

As noted above, after transcribing the words you hear for each utterance, you should provide a set of four codes indicating who is talking to whom, as well as some basic properties of the utterance. These four codes should be appended on the end of any transcription line for any segment you have transcribed (or, in the case of multiple utterances in a segment, after each utterance, separated by a period). If there are LENA codes, these should be appended to the end of the LENA codes. If there aren't LENA codes, simply put them at the end of the line. They should be separated by vertical bars as follows: |code1|code2|code3|code4|

Code 1: Speaker
      M: Adult male
      F: Adult female
      T: Target child
      O: Other child
      C: Child uncertain
      U: Uncertain

These are fairly self-explanatory. C is reserved for a situation where it is clearly a child speaking, but you are uncertain whether it is the target child or another child. U should be used sparingly.

Code 2: Target listener (based on content and context, NOT tone of voice)
      T: Target child
      O: Other child
      C: Child uncertain
      A: Adult
      U: Uncertain

Again, these codes are fairly self-explanatory. C is reserved for a situation where the target is clearly a child, but it is unclear whether the target child or another child. U should be used sparingly.

Code 3: Utterance grammaticality/completeness

Common codes
F: complete (could be an isolated phrase)
D: disfluency (cut off, false start, etc.)
M: mispronunciation
I: continued in following segment
C: continuation of previous segment
CI: both a continuation and continued
U: uncertain

Uncommon codes
DC: disfluency and continuation
DCI: both a continuation and continued and a disfluency

MC: mispronunciation and continuation
MCI: both a continuation and continued and a mispronunciation
UF: complete and not disfluent but otherwise ungrammatical

Most utterances will be "F". M can be combined with other codes (e.g. "MF" if there is a mispronounced word in an otherwise normal utterance). C and I should only be used when an utterance does not stand on its own. So a segment can be an obvious continuation of another segment and still get an "F", as long as each segment is reasonable on its own. UF is reserved for ungrammatical sentences that are neither a disfluency nor a continuation or continued segment (typically there was some sort of speech error but it was not recognized/corrected). Also note that when combining a disfluency or mispronunciation code with a continuation and/or continued code, the C and/or I should be the last letter in Code 3.

Code 4: Utterance type
Q: question
D: declarative/statement
I: imperative (command)
U: uncertain
R: reading
S: singing

These codes are largely self-explanatory. U should be used sparingly. The code Q for question takes precedence over the "I" and "D" codes. Yet, both the codes "R" for reading and "S" for singing trump the "Q" code.

Figure 4. Examples

In this OLN segment the speaker, a female, has asked a question to a child in the daycare. Questions always trump declarative and imperative sentences!

The utterance in the CHN segment is completed in the OLN segment because of LENA segmenting. Therefore, the transcriber must use the I and C code. These are only used in code 3.

Examples

Target child: can i have an apple mama?
Should be coded as: |T|A|F|Q|

Target child: put the block [finished in the next segment]
Should be coded as: |T|A|I|I|

Adult male: so that you can swim tommy. [completing a previous segment]
Should be coded as: |M|T|C|D|

Adult female: that you don't [both continued and continuing]
Should be coded as: |F|T|CI|D|

Target child: [babble directed conversationally at an adult]
Should be coded as: BBL |T|A|U|U|

Figure 5. Example 1

OLN + OLN

1: BBL |T|U|U|U|

2: sally doesn't have a hat we're gonna get her one sally you going to find a hat from the bucket |F|O|F|Q|

For these OLN segments, they were selected as overlapping speech because both the child and the female were speaking at the exact same moment. How would you code the utterance that the female spoke?
The answer would be |F|O|F|Q|

Figure 6. Example 2

OLN

get your boots on do you have your boots on

What should this OLN segment be coded as? The answer is |F|C|F|Q|. The reason behind the use of these codes are:
- the segment is being spoken by a female speaker.
- we weren't able to identify if it was the target or another child.
- this was a complete utterance and was a question.

Additional Details and Complications

If there are multiple speakers for a given segment:

I. Here are the various ways we deal with multiple speakers in the same LENA-defined segment:

A: Two speakers simultaneously: Use the "overlap" button in Transcriber

1: that's nice |F|U|U|U|
2: yup |F|U|U|U|

indicates that speaker 1 said "that's nice" and speaker 2 said "yup" at (approximately) the same time (overlapping). Select as the speaker for 2 the same speaker code as LENA had for Speaker 1. *Note that occasionally LENA creates "sub-segments" which are multiple bullets within a given segment. When you use the "overlap" button, you will be creating a #2 speaker for each of those subbullets. If the second speaker is only relevant for a sub-set of the subsegments, just leave the unneeded #2 lines blank.*
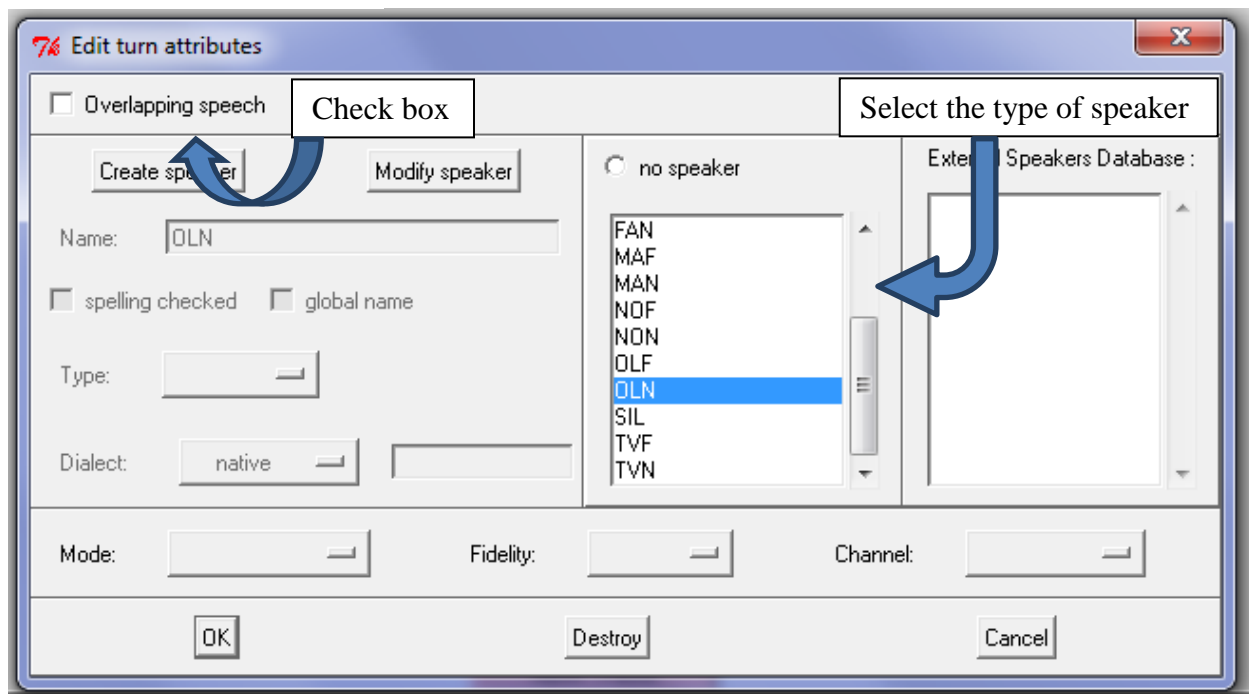
To gain access to the overlap function one must click on the segment button in green. OLN
Once this has been accomplished an editing menu will appear on the screen. You must then check the overlapping speech box and pick the proper type of segment. Please see figure 7 for the layout of the menu.
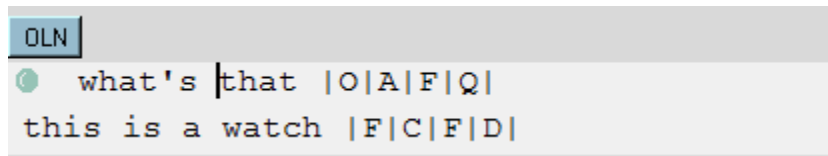
Figure 7. overlap button

B: Multiple speakers one after the other*: Use a new line for each speaker

that's nice |F|U|U|U|
yup |F|U|U|U|
ok |F|U|U|U|

indicates that Speaker 1 said "that's nice", then speaker 2 said "yup", then speaker 3 said "ok". Figure blank is a typical example of two speakers speaking one after the other. To include a second speaker, you must press shift and enter, this will put the cursor beneath the first transcribed segment.

Figure 8. transcriptions of multiple speakers simultaneously

Figure 8. is an example demonstrating how the transcription would with multiple speakers simultaneously.

C: More than two speakers simultaneously and/or combinations of overlap and interspersed speech*: Use the "overlap" button for two of the overlapping speakers, and new lines for the remaining speaker and then angle brackets as needed to indicate timing

1: that's nice <> |F|U|U|U|
<yup> |F|U|U|U|
2: ok |F|U|U|U|

if the "yup" comes after the first speaker or

1: <that's nice> |F|U|U|U|
<yup> |F|U|U|U|
2: ok |F|U|U|U|

if they are simultaneous. (The "2:" speaker is assumed to be simultaneous with the "1:" speaker unless otherwise indicated)

It is not necessary to be absolutely precise in determining what is overlapping with what. You may find that there is some ambiguity in your coding and you will likely have some inaccuracies unless you want to spend 90% of your time figuring out where to put angle brackets. For the overlapping speech, just do the best you can to give a sense of the flow of conversation. It is not nearly as important as determining what goes in what utterance line and what codes to use.

II. Dealing with multiple utterances by the same speaker in a single LENA-defined segment:
For the most part, you should rely on LENA's "judgments". However, occasionally LENA clearly makes mistakes. The two most common cases are when LENA puts a segment boundary in the middle of a sentence due to some kind of pause or hesitation on the part of the speaker, or when LENA concatenates several sentences together in a long string in an "overlapping speech" segment. In the case of LENA breaking up a single utterance, simply use the codes as indicated above. When a transcriber is at all uncertain whether a single segment contains two or more utterances, check to see if there is > 1 second pause between the two speech sections. If so, consider it two utterances. Note that this is ONLY if there is any ambiguity based on the listener's initial judgment. If it is clearly a single sentence, even if there is a pause of more than one second, code it as a single utterance. If it is judged to be multiple utterances by the same speaker AND there is a pause of more than 1 second, code it this way:

that's nice |F|U|U|U| . i like that |F|U|U|U|

NOTE THE PERIOD SEPARATING THE TWO UTTERANCES - THIS PERIOD IS IMPORTANT.

Note also that this formalism can be combined easily with the multiple-speaker options above.

III. C's and I's
If there is any ambiguity about which C's are for which I's, please use numerals, as follows:

1: that's my sister's |F|U|I1|U|
2: it's never in the |F|U|I2|U|

1: blue hat |F|U|C1|U|
2: first place you look |F|U|C2|U|

Miscellaneous Notes:
- Under no circumstances should you edit the speaker ID or LENA codes
- Don't get too carried away with multiple speakers. If eight different people are talking at once, it is unlikely that you can really make out what anyone is saying anyway, and you can just treat this as background noise (i.e., don't bother transcribing/coding it).
- Words such as "gonna" or "wanna" should be transcribed exactly how you hear them. If you hear "gonna", write "gonna". If you hear the more formal "going to", write that.
- If the speaker breaks off a word in the middle or stutters, the ^ symbol can be used to indicate this.
    a. Some examples: "Tommy, p^please drink your milk"; where there is a disfluency in the middle of "please" or "Where's the mi^"; where the person is asking for the milk but doesn't finish the word.
    b. Note: this code should NOT be used to indicate a partial word when the LENA system breaks up a word across segment boundaries, only if it is actually a part-word on the part of the speaker.

Keeping track of, editing, and saving your work

1. For every transcript, indicate in the excel spreadsheet the date you completed *YOUR* version of the transcript. *DO NOT* delete or edit this original version. Save a copy of it with your name on it in the appropriate directory on the LENA computer.
2. Once you and your reliability partner have both transcribed the same section of audio, you will need to get together to compare notes. Create a new version of the transcript, *separate* from either of your originals, and make any changes as needed to create a final version that has corrected any typos and harmonized any discrepancies between the two transcriber-coders. In cases of disagreement, do the best you can do decide which version is the most accurate. In some cases, you may wish to use "XXX" or U codes if you cannot reach an agreement on a transcribed word or code. Remember to consider the "forest" and the "trees" when making edits – sometimes things sound different if you only listen to the segment that is discrepant between the coders.
   If you like, you can use a software program to highlight the changes between your transcription and your partner's. Pease see below for instructions on how to do that.
3. After this process has been completed, there should be four versions of each transcript – the original LENA-generated file, one marked with the name of transcriber A, one marked with the name of transcriber B, and one marked as "FINAL". Please make sure that you indicate the date of completion of all three transcripts in the master excel spreadsheet on Google drive.
4. Once the file is completed it has to be checked using the transcription verifier software. This software checks for any coding errors that must be corrected by transcribers. The shortcut for the software is located on the desktop of the LENA and BOB computer. Please see below for directions on how to use it.

Using the transcription verifier program:
   This section describes how to use a software program to check transcription files for errors, or to highlight the differences between them. There's a manual for this program on the LENA desktop (shortcut to a PDF file just beside the program icon) that contains additional details.

   To start, double-click the transcription verifier shortcut on the LENA or the BOB computer. This brings up the window shown below.
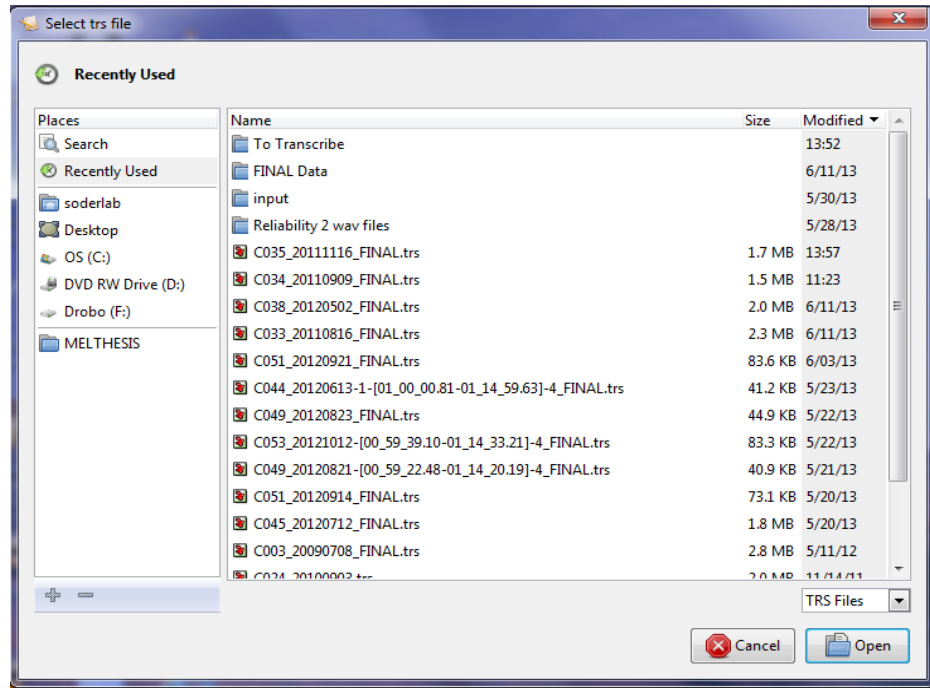
Figure 9. Transcription Verifier
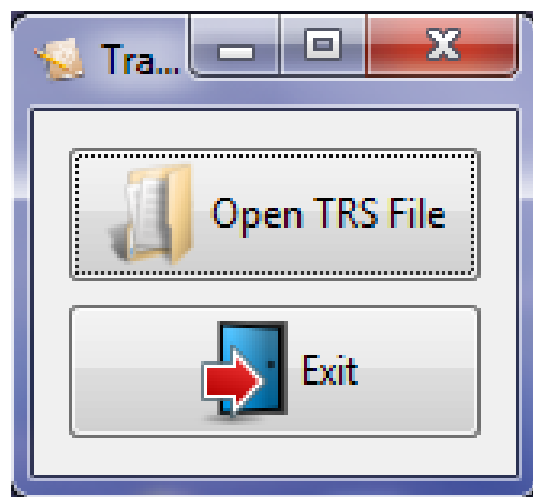
I. Checking files for errors:
   a. Click on the "Open TRS File" button. A screen will appear that will allow you to browse through the files on the computer.
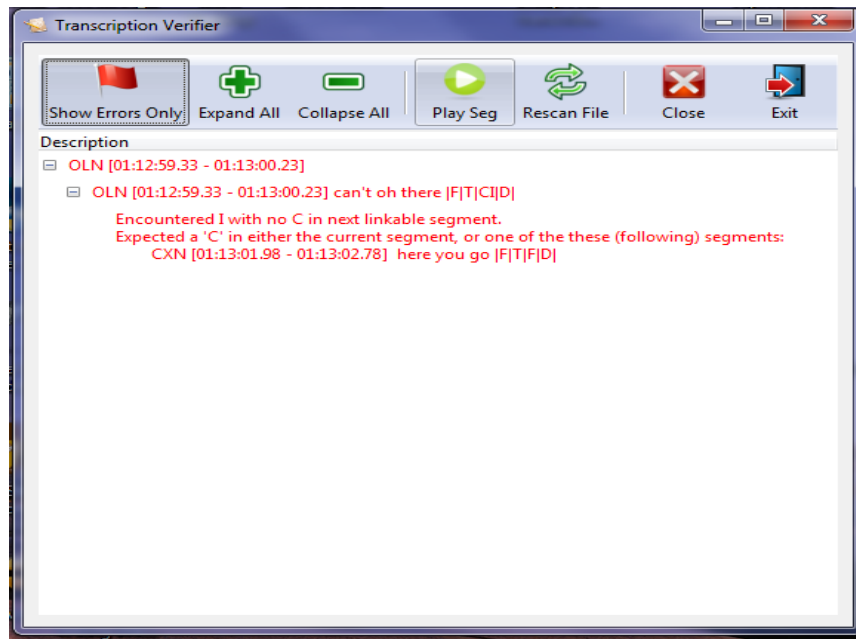
Figure 10. File selection window

   b. We want to select the folder that contains the transcriber files. This folder is currently located on the Drobo external hard drive. To get there, select the "F" drive and then select the "Transcriber Files" folder.

Figure 11. Drive selection
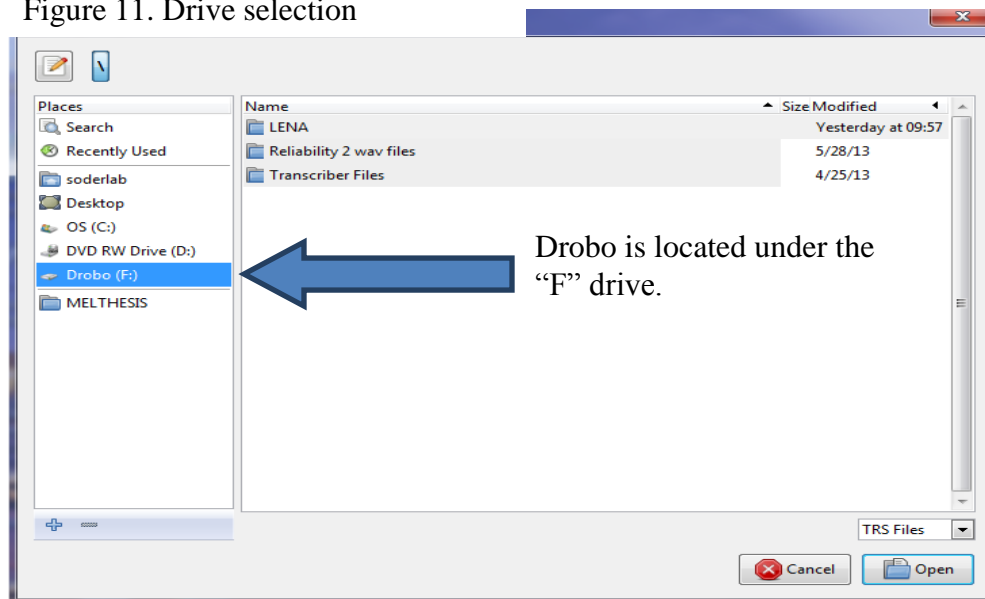
c. Locate the final transcription that both transcribers have helped complete, click it, and hit the "open" button. The program will show a progress bar as it's going through the file. When it completes, you'll see a window with a list of errors that have been found. Each error will appear under one bullet. If there are no errors, the window will be empty.

d. If you like, you can leave the program open while you use Transcriber to fix any errors. When you're done, save the file in Transcriber and then hit the "Rescan File" button in the transcription verifier to refresh the error window.

Figure 12. Error window

II. Comparing files

a. Open the program and click the "Compare Files" button in the main window. You'll see a window that looks like this:

Figure 13. File selection window
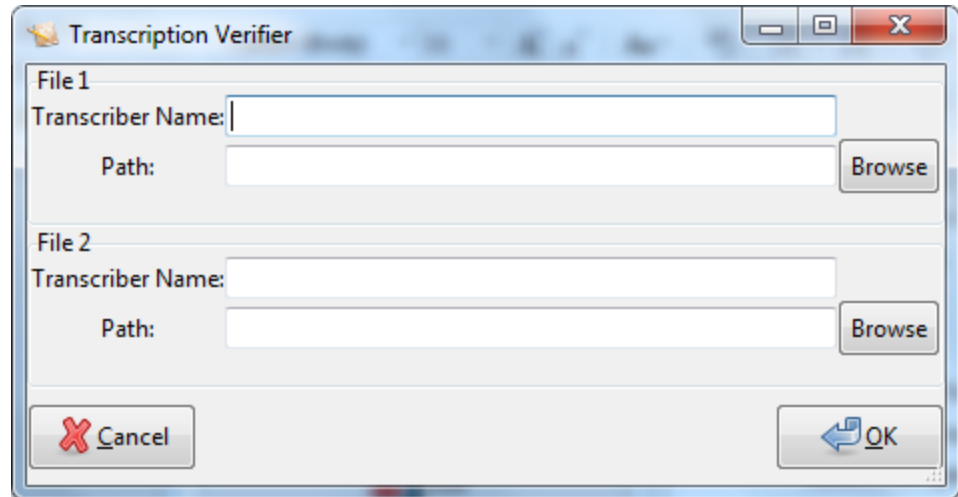
b. Enter the following information for each file:
   - "Transcriber Name" - If you like, you can enter your name in this box – these names will be shown on either side of the split view (shown below) so that you can tell who the work on each side belongs to.
   - "Path" – this is the path to the TRS file you want to compare (you can use the "browse" button to select one).
c. Once you're done, hit "OK". At this point, the program will compare the files and open up a window that shows you the differences, like the one shown on the next page. This may take some time, depending on the length of the file.

Figure 14. Transcription comparison window

The table shows one row for each segment whose transcriptions are different. In this example, the left side shows segments from Horace's file, and the right side from Gertrude's. The highlighted words are the differences. Each entry shows this information:

Figure 15. Segment table description

The highlighting colour shows the type of change that was detected, as described in the legend.

15

A couple of notes: If two segments were joined with the dot (.) symbol, they will appear separately, but with the same start and stop times. If a segment has multiple speakers (like the top one on Harace's side), the table row will contain multiple lines – one for each speaker.
Don't worry about the links and line numbers – they're purpose is to help you find the next segment in the event that you're looking at the whole file (including the segments that match), but we're only looking at the differences.

Booking the LENA and BOB Computers

There are two computers in the transcribing lab in which to transcribe on. Because there are a number of transcribers who will be sharing these computers, it is imperative that you book in advance the time you will be using them on the lab's calendar. Whoever books the computers first, has the priority of working on them. Out of respect for everyone sharing these computers, please make sure to show up during your booked time or delete your booking off of the lab's calendar (at least a day in advance) if you're unable to make your scheduled time.

Laptop Rental System

There are two laptops that can be rented which are only to be used for transcription. They are located in the filing cabinet in the transcribing lab. You must sign-out the laptops on the "Sign-Out" sheet located on the bulletin board in the lab. The laptops can be rented out for the maximum of two weeks and they must be returned by noon of the return date. Laptops should be used for lab business ONLY and may not be connected to the internet in any way. You must use one of the lab's USBs to upload and download transcriber files. You cannot use your own personal USB to do so. Once files have been completed they must be transferred to the LENA computer and saved onto the Drobo. After having transferred files, please do not forget to delete them off the laptop and the USB key.

Pseudonyms

All names in the final transcribed file must be changed to a pseudonym to assure the anonymity of our participants. The name and pseudonyms must all be logged in a spreadsheet on the LENA computer. The spreadsheet is organized by environment and file number.