

## Transcription Presentation Format

### Objective

The purpose of this document is to describe the format to be used for producing and grading transcriptions in this course.

### Conventions

The main transcription system we will use in the course is described in *Discourse Transcription* (Du Bois et al., 1992), as updated in our class handouts. Note that, although the theory, discussion, and examples of transcription *categories* as presented in the earlier documents generally remain valid, the transcription *symbols* have undergone a few revisions in the time since the earlier publications appeared. For all the transcribing we do in this course, it is important to use the most recent updated symbol conventions, as indicated in class handouts.

### Font

Please use a Unicode font. This will give you the widest range of options for using special symbols, including characters capable of writing any of the languages of the world in their traditional standard orthography, plus the International Phonetic Alphabet, and many other specialized symbols. Fixed-width fonts can help with aligning overlap brackets, but are not necessary.

### Tabs

Tabs are used in transcribing to help organize the presentation of transcription information, displaying it in a usefully iconic way. For our purposes, tabs are to be used for one purpose only: to distinguish between separate “fields” of data in each line. For a basic transcription, there are two fields. The first field in each line indicates who the speaker is (via a speaker label written in capital letters, followed by a semi-colon). The second field represents the actual utterance, or transcribed speech.

Insert exactly one tab character in each transcription line:

- if there is a speaker label in the line, the tab is inserted immediately following it
- if there is no speaker label in the line, the tab is the first character of the line

Note that you should NOT use tabs for other purposes, such as inserting blank space in a line in order to align overlaps vertically (just insert spaces instead).

When a line of transcription goes beyond the right margin, let the text wrap to the next line without inserting tabs.

### Line spacing

Do not use double-spacing (only use single spacing!), and do not leave blank lines between turns.

### Timestamp

Starting with Transcription Homework #3, occasionally (about every 30 seconds) insert a time stamp into your transcription, indicating the location of the corresponding audio in your digitized audio file. For example, to add a timestamp at 60.6 seconds into the computer audio file, insert the following notation in

its own line: <T= 60.6>. This notation allows you and your instructor to quickly find the relevant place in the audio file. You may wish to add additional time stamps at various points in your transcription, for convenience in locating specific portions of the transcription, for checking, feedback/consulting, and so on.

### **Index (line numbers)**

It is useful to index the transcription to make it easier to refer to specific portions of the data. Each line (or other unit) in the transcription is given its own unique index. There are at least two methods which are particularly useful for this: numbering each line, and/or timestamping each line.

**NOTE: DO NOT ADD LINE NUMBERS OR INDEXES UNTIL THE FINAL PROJECT.**

### **Anonymity**

Names of participants used in the transcripts (both in speaker labels, and when one discourse participant utters another participant's name) should not be the speakers' real names, unless you have received written consent from the participant to use their first name. Otherwise, make up an appropriate pseudonym – one which is comparable sociolinguistically and prosodically – and use it consistently. Any pseudonym (or other words you have modified because they could compromise anonymity, such as an utterance of a participant's street address or phone number) should be indicated with the pseudograph symbol (i.e. tilde ~). (Note, however, that the tilde is used only for uttered words on the recording, not for speaker attribution labels, even if these are pseudonyms.)

For speaker attribution labels, give the complete first name, all in capital letters, not just an initial (e.g. use JILL rather than J).

Before disseminating your recording via the web or CD-ROM (or for your midterm project), you will need to make sure that any anonymity-sensitive words have been “bleeped” or otherwise obscured. You will learn simple techniques for doing this in this course.

### **Grading**

I will count off points for errors in the transcription conventions. I will not count off points for the parts of the process that require your judgment as researchers, including the length of IUs and their final intonation. I will, however, make a note when I disagree with you—but this doesn't mean your judgment is wrong.