

Transcribing Procedures

1. Introduction

This document gives a broad overview of transcribing procedures. It focuses on procedures, methods, and practical mechanisms that we will use to organize the practice of transcription in this course. A more fine-grained treatment of the details of the transcription process itself can be found in the main textbook for the course (*Discourse Transcription*).

In the approach we take in this course, transcribing a recording entails two main tasks: transcription and checking. During much of the course, you will transcribe one minute of spoken discourse per week. Each week you will also check your partner's minute of transcription from the week before. Each of these tasks involves a variety of subtasks, to be described below.

2. Recordings

The recordings you will work with in this course will be digitized on a computer. This facilitates collaboration and exchange of data between course participants. It also makes the transcribing easier and more effective, enabling use of computer-assisted transcription techniques (see below).

It is important to use consistent conventions for creating and naming the computer files for recordings and transcriptions. Consult the relevant Appendices (Ethnographic Recording and Transcription Format) for details.

The recording file that we will be using for most of our work is the Research Segment file (20-40 minutes long). Within this file, we will be mostly working on the Transcription Excerpt, a 4 minute long subset of the Research Segment. It is very important for you to mark the exact start and end of the Transcription Excerpt within the Research Segment, using your sound editor (see the Appendix on "Digitizing and Editing Recordings"). It is the Research Segment file that you will use for time-aligning your transcription with the corresponding audio (see "Alignment" below).

3. Delicacy

In the initial stages of learning about transcription, the transcriptions you will be asked to produce are relatively "broad," as opposed to "narrow". But a broad transcription does not mean a loose one. It can be as rigorous in its own terms as a narrow one is, for what it seeks to represent. As the course goes along, the transcriptions will become more detailed, especially in the area of prosody. For a given week's assignment, the expected level of delicacy will be specified in the assignment schedule handout. See the handout on "Transcription Delicacy Hierarchy" for specific information on which transcriptional features are to be included at each level of delicacy.

4. Partners

To gain the benefits of working and talking with your peers about transcribing issues, you will be paired with a partner. All the transcribing and checking work that you do in this course will be done with the support of your partner (and vice versa). Normally you and your partner will be working on different conversations, so that you will gain broader experience through working on two different conversations.

5. Checking

In addition to transcribing your own conversation, each participant will check their partner's transcription, on a weekly basis. Because you are transcribing in a team, you will gain the benefit of a second opinion on how to represent what is going on in your conversation. When you check your partner's conversation, you should offer your insights and alternative interpretations. You should expect to receive the same careful attention in return from your partner.

It is important to take the task of checking seriously. When done well, checking requires about the same time, effort, and attention as does transcribing in the first place. When you turn in your checking of your partner's transcription to the instructor (e.g. using the "Record Changes" or "Track Changes" function of your word processor), it is expected that there will be lots of marks indicating corrections of your partner's work, reflecting the fact that you are doing a careful job of it.

Note that you are only responsible for checking those features that your partner has made a serious attempt to transcribe. For example, if you notice that your partner's transcription contains no pauses at all (even though pauses were supposed to be transcribed in the assigned delicacy level), you are not responsible for identifying all the pauses for your partner. Instead, simply indicate that the necessary feature was entirely omitted, and leave it to your partner to find them in the revision stage. On the other hand, if your partner's transcription does have pauses marked, and thus displays a good-faith effort to identify the pauses in the recording (with a plausible level of accuracy and completeness), you are then responsible for checking and identifying any pauses your partner may have missed.

6. Revision

As soon as you have received your partner's checking of your transcription, you should go back and review your recording in order to update and improve your transcription. Of course, the suggested corrections which are marked in your transcript are themselves interpretations, and hence are not necessarily "correct". You have to review the recording and evaluate the suggestions for yourself. Consider the suggestions of your partner and your instructor carefully. Think about what the different transcriptions mean, and what justifications exist for your original transcription versus the proposed revisions. If a proposed change seems valid to you, incorporate it in your revision.

7. Reconciliation

On points where you disagree with your partner, it is very valuable to have a "reconciliation" meeting to go over the corrected transcription together. You can learn a lot by trying to argue for or against a particular transcription decision with your partner—that is, give specific reasons for your position, naming any relevant auditory cues, and so on. Try to set up a regular weekly time to meet for this purpose. Even if you agree with all of your partner's suggestions or corrections, a reconciliation meeting to discuss the ideas as they apply to your transcription can be valuable.

8. Updating

You may choose to incorporate changes into your transcription as you go along, so that when you turn in your transcription file for the following week, any earlier portions of the transcriptions will be revised, up-to-date, and brought up to the standard of your current knowledge about transcription. Or, you may prefer to simply incorporate corrections of errors each week, and wait on upgrading the delicacy level of your previous minutes of transcription until you can revise a larger portion at one time.

In any case, when the time comes for turning in the *revision* of Minutes 1-4 of your transcription (i.e. normally around the 6th or 7th week of the course), you will need to fully revise and correct the entire four-minute version. In addition, you will need to bring the whole transcription up to a consistently high level of delicacy.

9. Explication

A number of class sessions will be devoted to transcription sessions, in which we will not only transcribe, but also discuss the transcription decisions you make, and your rationale for them. Of course, identifying the right transcription category for a phenomenon you observe to be present in the recording data is important, and this decision is overtly expressed by selecting the right symbol for what you hear or see. But what is more important is to be able to understand, and to articulate, the reasons for invoking one particular transcriptional analysis over another. Because a transcription is always an interpretation of reality, it is important to be able to explicate the representations you create. To explain and justify your analysis to other session participants, you need to understand the concept behind each transcriptional category, as well as the specific cues which signal it. You need to be able to motivate and explicate your interpretation.

So that we can make the most effective use of our limited time for in-class transcription sessions, you are expected to be well prepared, understanding the assigned readings, as well as all updated material from the handouts and course packets in a way that lets you explain your decisions. You should try to develop a command of all the transcribing categories, cues, conventions, symbols, and practices described in the materials assigned up to a given date, and be able to explicate and motivate them as we discuss specific transcription problems in class.

10. Consultation

Because of the number of students in the class and the large quantities of transcription that are collectively generated during the course, it is not possible for the instructor to provide feedback on every word of every student's transcription. However, the instructor will check selected portions of your work at intervals during the course, and offer oral feedback, in class and in private consultation sessions.

To get the most effective input, you (along with your transcription partner) should present your transcription for oral evaluation several times during the quarter. For each consultation session, you are expected to present a portion of the transcription you have recently completed. We will go over it together to talk about relevant transcription issues. Half of the time of each session will be devoted to your transcription, and half to your partner's. (If the number of students in the course is unusually large, we may need to do some of these instructor consultation sessions in larger groups.)

For each transcription consultation session, you should come prepared with the following:

- Your recording should be available for playback on the instructor's computer, in the form of a digitized computer file (audio should be in .WAV format), accessed either over the network (preferable), on CD-R, or both. Make sure you locate and *test* your computer file to be sure it works before you come to the feedback session, so that we do not waste time trying to figure out how to access your file and get it to work.

Du Bois: Transcribing Procedures

- Pick a *specific section* of your recording that you found challenging or interesting to transcribe, and which you would like to talk about. The amount of material you select should be about half a page (or a page at most), which generally provides more than enough to talk about.
- You should know *exactly* where in the computer file to find each item you want to talk about. Identify the exact time location of the relevant section, and insert a timestamp in your transcription (e.g. <T=127.3>).
- You should bring a paper printout of your selected page of transcription (no more than *one* page), with one copy each for yourself, your partner, and your instructor.

Note that the instructor consultation sessions are an integral and essential part of the course. You are responsible for reserving time for these sessions with the instructor.

11. Questions

In the consultation session, you will get “live” feedback on your transcribing. In addition, you are likely to be asked impromptu questions about your transcription by your instructor. (You can also raise questions of your own about how to transcribe the specific phenomena in your conversational excerpt, but theoretical discussion of concepts is not the main focus of these transcription consultation sessions.)

Examples of typical questions you may be asked include:

- Is there is an intonation unit boundary between these two words? Why or why not?
- What prosodic cues can you identify, which you could use to argue for the boundary, or against it?
- What intonation unit cues are missing in this case? Does that change your interpretation?
- Who does this pause belong to? Why?
- Where are the primary and secondary accents in this line? Why?
- and so on...

Of course no one is expected to know how to answer such questions at the beginning of the course—that is what the course is for. All of the above are questions we will talk about together in class, so you will have a chance to learn how to master this kind of analysis. When the time comes in the course that we have arrived at this level of transcription, you will be expected to come prepared to explain your transcription decisions, along the lines of the above questions.

12. Computer-Assisted Transcription (CAT)

As much as possible we will make use of techniques of computer-assisted transcription during this course, learning how to use software such as Sound Forge, VoiceWalker, and/or PRAAT. This means that all recordings will need to be digitized for use on a computer. For this reason, you should:

- create a CD-ROM disk (that is, “burn a CD”) which contains a copy of your digitized computer files, along with those of your partner(s)
- bring your CD containing the relevant recordings with you to class sessions, team meetings, feedback sessions, etc.
- also make sure that you always copy the latest up-to-date versions of your transcription and recording files to the appropriate folder in the Linguistics Lab network.

Several software tools that we will use for computer-assisted transcription are available free to be downloaded from the web, while others may need to be purchased. Students should refer to the relevant course materials for information on how to obtain these items for use on your own computer.

13.Alignment

Using a tool like SoundForge, with its Regions function (or Transcriber, Transana, EXMARaLDA, etc.), you should begin to align your transcription with the corresponding recording as soon as possible during the transcription process. This means that for each intonation unit you have defined, you should identify the corresponding “Region” of the recording. This will make it possible to play back the exact portion of the recording which corresponds to each of your intonation units. This makes it much easier for you, your partner, and your instructor to check the accuracy of your transcription. It also makes it easier to gradually refine the granularity of your transcription, as you listen again and add more detail.

For your alignment work, you should use your Research Segment file. (If, after this course is over, you decide you want to expand your work with this Research Segment for your own future research projects, you will be glad you can simply add to all the work you’ve already done on the Transcription Excerpt by extending your transcription and alignment to additional portions of the Research Segment.)

14.Bleeping

As you time-align your transcription with the corresponding portions of the recording, you should also use the Regions function to mark the location on your recording of any names or other words that could compromise anonymity (see the section on “Anonymity” in the “Transcription Format” Appendix). Before publishing the data, it will be necessary to bleep or otherwise mask these words, so they cannot be recognized.

15.Language

For doing the transcriptions in this course, you are welcome to work on any language you know very well. For purposes of (unassisted) transcription, this means you should know the language natively, near-natively, or at least fluently.

16.Schedule

Consult the syllabus and the assignment schedule for the dates when transcriptions and checkings are due. Transcriptions and checkings are always due weekly, on the assigned date. Please note that just before class is probably not a good time to try to email or print out work due on that day, as the computer or printer you want may be in use, or otherwise unavailable. In any case, do not miss class in order to “finish up” a late assignment—the class sessions are too important to miss.

17.Timeliness

It is important that you do your transcription on time, in order to keep up with transcription activities and discussions in class, and because your partner and instructor need to receive your work on time in order to be able to provide you with feedback. The checking also needs to be done on schedule, in order to give your partner feedback on their transcriptions in time for them to incorporate improvements into their next revision. For these reasons, TRANSCRIPTION AND CHECKING ASSIGNMENTS TURNED IN LATE WILL RECEIVE REDUCED CREDIT.

18.Exchange

For exchanging and submitting transcriptions in this course, we will use electronic means as much as possible. When the time comes each week to turn in an additional increment of transcription, please do the following BEFORE CLASS on the date due:

- Place a copy of all files relevant to the current assignment in the designated directory for this course. For example, using a computer on the Linguistics Lab network, find the “Everyone” drive, and copy your files into the Ling212 directory, under the specific subheading for your class and year (e.g. “Class05”).
- Include the following items:
 - your updated transcription file. The required format may be (1) SoundForge .wav or .sfx or similar file; (2) OpenOffice.org Writer .odt file, and/or (3) Microsoft Word .doc file.
 - your updated audio file (.wav format) if appropriate.

By keeping your transcription and audio files up-to-date and accessible, you make it possible for your partner(s), classmates, and instructor to check your transcription, give feedback, and so on.

You MAY need to do the following, as an alternative and/or backup means of exchanging data:

- Email your *transcription* file as an attachment (format may be SoundForge .sfx, OpenOffice.org Writer .odt, or Microsoft Word .doc) to two people: (1) one copy to your partner, to be checked, and (2) one copy cc’ed to your instructor, for evaluation.
- Email your *checking* file (containing your mark-up of your partner’s transcription from the previous week) as an attachment (using the computer format in which you received it) to two people: (1) one copy to your partner, for revision and reconciliation, and (2) one copy cc’ed to your instructor.

19.Format

An important part of transcribing is to use a consistent format for presenting and exchanging information. See the handout on “Transcription Format” for a detailed discussion of this topic, including important conventions for naming your transcription and audio files.

Finally, transcribers want to know: What is...

20.The Point

Why transcribe? Transcription is more than just manual labor—though it surely is hard work. Nor is it merely a means to an end: It’s not just a way to get discourse data to use in your research. Rather, the very practice of transcribing has value in and of itself. Many transcribers find that if you put effort and insight into your transcription practice, keeping your eyes (and ears) open and thinking about what you are doing, the transcribing process will teach you more about discourse than any book, article, or lecture ever could.

[rev. 11-Oct-2005]