## Transcription systems used

System A is Santa Barbara DT2 used in transcripts made by the second author and Extract 5.1.

Du Bois, J. W. (2006). *Representing discourse (a work in progress)*. Retrieved 24 July 2020 from <http://www.linguistics.ucsb.edu/projects/transcription/representing>

System B is common conventions used in Conversation Analysis (CA) also used by others beyond this (often less detailed if not CA).

System C is the VOICE corpus system, also used in the Asian Corpus of English (ACE).

VOICE Project. (2007). *VOICE Transcription Conventions* [2.1]. Retrieved 20 October 2020 <http://www.univie.ac.at/voice/voice.php?page=transcription_general_information>

## In more detail

### Santa Barbara DT2

This transcription system is not linked to one particular way of thinking about language but attends to language at the micro level.

The main website for relevant materials is the one maintained by Mary Bucholtz and John W. Du Bois found at <http://transcription.projects.linguistics.ucsb.edu/index.html>. Within this site there is a lot of information to explore, including a detailed bibliography. It differs from Conversation Analysis in not respelling words and considered computer-based searching, so that, for instance, symbols have distinct meanings (this is why a semi colon is used after the speaker attribution, so it is not the same as the lengthening symbol of the colon). On the *Representing Discourse* page, there are a number of appendices for the manual (which needs to be ordered – this will not be necessary for your class). The *Comparison* document shows the differences between the basic symbols used in these systems, an older version of this system and Conversation Analysis, which is helpful if you are accessing files from the Santa Barbara Corpus of American English.

### Conversation Analysis (CA)

This system is linked to the type of analyses done in CA, which look at structures of conversation and grew out of sociological interests in this. As mentioned above, these symbols are also used widely by others (with some adaptations).

Often authors use what is call the Jefferson system (developed by Gail Jefferson). A discussion of its ideas and symbols can be found in Hepburn and Bolden (2012) (open access versions are available, find via a Google Scholar search). Part of the EMCA wiki found here <http://emcawiki.net/Transcription_Resources>, provides links to multiple online training resources for those wanting to learn more about how to do this kind of transcription. In addition to special symbols, CA uses text styling based on literary conventions (underlining, bolding, capitalisation) and respells words to better represent how they are said (although, this is not without introducing new problems into transcriptions).

### VOICE transcription system

This is a system developed specifically for the ELF corpus of Vienna-Oxford International Corpus of English (VOICE). There is a focus in the design on ease of searchability and ELF features of interest. It is quite different to the other two systems used in its appearance but does capture the same sorts of information.

Within this system, spellings are carefully standardised to support corpus-based analysis (i.e., if you wanted to research *probably*, you would just need to type this in, not a range of spelling such as *probly* etc. that aim to represent specific productions, unlike CA transcriptions). The VOICE website contains both the corpus and details on the transcription system (see lefthand menu): <https://www.univie.ac.at/voice/page/index.php>.

### References

Hepburn, A., & Bolden, G. B. (2012). The conversation analytic approach to transcription. In J. Sidnell (Ed.), *The handbook of Conversation Analysis* (pp. 57–76). Wiley. <https://doi.org/https://doi.org/10.1002/9781118325001.ch4>