Advancing the Utility of the Transcript: A Computer-Enhanced Methodology*
Methods XII, Moncton, New Brunswick
August 2nd, 2005

Tyler Kendall <tsk3@duke.edu>
Duke University & NC State University

Project Overview
• Building researcher tool(s) to store, analyze, interact with, and present linguistic data
• Attempting to treat natural speech data like any other type of data.
• Focused initially (version 1) on applications in discourse analysis (transcript-centric) but expanding to be a complete digital audio library and analysis suite (version 2).

Transcript Features
• User control of the format (e.g., vertical  columnar  paragraph  tablature (Figure 1))
• User control over the amount and kind of data displayed (e.g., translations for multilingual transcripts, syntactic information, nonlinguistic information, time information, pause information...)
• Direct link to Praat phonetic software
  o Utterance-level coding (instrumentally-determined utterances – uploadable from Praat TextGrid files)
  o Active link to Praat allows for on-the-fly phonetic analysis and audio clip extraction (Figure 2)

Toward a Complete Digital Archive
• Through a partnership with the NC State University Libraries
• Working toward a complete digital archive of the North Carolina Language and Life Project’s sociolinguistic interviews (> 1,500 tapes)
• Users can access, download, and extract from the audio, listen to the online, and can annotate sound files (Figure 3)
• Searchable (e.g., by research site, by demographic information, by transcript text, by annotations on the audio files, etc.)
• Queryable (can generate summary statistics from transcript/annotation information)

Selected References
• Edwards, Jane and Martin Lampert, eds. (1993), Talking Data: Transcription and Coding in Discourse Research, Hillsdale, NJ: Lawrence Erlbaum Assoc, Inc.

* I would like to gratefully acknowledge the financial and infrastructural support of the NC State University Libraries for this project. I would also like to thank Walt Wolfram, Erik Thomas, Amanda French, Phillip Carter, Christine Mallinson, and Drew Grimes for the use of their data and/or for their help and feedback on various stages of this project.

Figures
(1) e.g., vertical format  columnar format
(2) pitch analysis example
(3) listen & annotate example