

DOCUMENT RESUME

ED 138 082

FL 007 994

AUTHOR Skopek, Lucienne; And Others  
 TITLE Methods for Cataloguing, Storing, and Retrieving Large Volumes of Tape-Recorded Conversations. Working Papers in Sociolinguistics, No. 32.  
 INSTITUTION Southwest Educational Development Lab., Austin, Tex.  
 SPONS AGENCY Social Science Research Council, New York, N.Y. Committee on Sociolinguistics.  
 PUB DATE May 76  
 NOTE 26p.  
 EDRS PRICE MF-\$0.83 HC-\$2.06 Plus Postage.  
 DESCRIPTORS Cataloging; Connected Discourse; Discourse Analysis; \*Information Retrieval; \*Information Storage; \*Language Research; Language Role; Language Usage; Medical Vocabulary; Patients (Persons); Physicians; \*Research Methodology; \*Sociolinguistics; \*Tape Recordings

ABSTRACT

A major problem in doing linguistic research from tape-recorded material is finding specific tape content for later, detailed analysis of data. A project on use of language in medicine being carried out at the Cornell University Medical College has developed a method of cataloguing taped material that eliminates the need for transcriptions and permits rapid locating of a specific tape segment. The project studies taped conversations between doctors and patients to observe how speakers hear and understand each other in natural conversation and to use findings to teach medical students the uses and functions of language in medicine. Contents of tapes are organized into eight categories: address, or time of utterance; speaker; main conversational division; specific speech act; lexical content; miscellaneous linguistic and non-linguistic information; attitude displayed by the speaker, and the conceptual category - an abstract interpretation of the subject discussed. Categories are explained here with examples, and an appendix furnishes a transcript of recorded conversation and its corresponding cataloguing sample.  
 (CHK)

\*\*\*\*\*  
 \* Documents acquired by ERIC include many informal unpublished \*  
 \* materials not available from other sources. ERIC makes every effort \*  
 \* to obtain the best copy available. Nevertheless, items of marginal \*  
 \* reproducibility are often encountered and this affects the quality \*  
 \* of the microfiche and hardcopy reproductions ERIC makes available \*  
 \* via the ERIC Document Reproduction Service (EDRS). EDRS is not \*  
 \* responsible for the quality of the original document. Reproductions \*  
 \* supplied by EDRS are the best that can be made from the original. \*  
 \*\*\*\*\*

ED:36032

WORKING PAPERS IN SOCIOLINGUISTICS

Number 32

May, 1976

*Bruce Fraser*

U.S. DEPARTMENT OF HEALTH  
EDUCATION & WELFARE  
NATIONAL INSTITUTE OF  
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

*FLCO 79974*

Methods for Cataloguing, Storing, and Retrieving  
Large Volumes of Tape-Recorded Conversations

by

Lucienne Skopek  
Eric J. Cassell  
Cornell University Medical College  
New York, New York

and

Bruce Fraser  
Boston University  
Boston, Massachusetts

Sponsored by the Social Science Research Council  
Committee on Sociolinguistics  
Prepared and distributed at

Southwest Educational Development Laboratory  
211 East Seventh Street  
Austin, Texas 78701

Methods for Cataloguing, Storing, and  
Retrieving Large Volumes of Tape-Recorded Conversations<sup>1</sup>

One of the major problems encountered in doing linguistic research from tape recorded material is the following: how to find tape content for later, more detailed analysis of data? To deal with this problem in research on the use of language in the medical setting, we have devised a system which allows rapid access to any segment of recorded conversation. We present here a description of our methodology in the hope that it will not only serve as a guide for investigators engaged in similar research with recorded data but also to elicit comments and suggestions.

The method commonly employed in analyzing taped conversations is to first transcribe (typescript) the recording and then code the relevant material from transcriptions. Transcription of conversation is a laborious process requiring from six to eight hours of typing for each hour of tape and the search for specific data even from typed material remains difficult and time-consuming. Even so, because working directly from tapes is a laborious and tedious task, investigators tend to use typescripts, often without realizing that, no matter what mode of transcription has been used (phonetic or orthographic), they are no longer truly studying the utterances but only the coding of them. Since we thoroughly believe that the study of natural language must be based on sources as close as possible to the actual spoken word, we have developed a method of cataloguing tape recorded material which enables us

---

<sup>1</sup>This work was supported by a grant from the Robert Wood Johnson Foundation.

to eliminate the need for transcriptions and to rapidly find a segment of interest on the original tape.

Our project on the use of language in medicine is being carried out at the Cornell University Medical College and involves the study of conversations between patients and doctors recorded in natural medical settings. Our long-term goal is to study verbal interaction in a medical setting, to draw conclusions on how speakers hear and understand each other in natural conversation, and to use this material to teach medical students the uses and functions of language as a tool in medicine. To this end we have collected data which represent over 900 hours of recorded material involving more than 2000 patient interviews and 800 patients. The largest segment represents recordings of patients in private practice, while the remainder involves hospital in-patients.<sup>2</sup>

#### Recording

In order to use this recorded material as a teaching tool, the project has developed recording methods which approximate studio quality stereophonic reproduction. Several recording systems are employed. The private practice method uses a separate Sennheiser wireless transmitting microphone for the doctor and patient. The two receivers are connected to a 1/4 inch stereo Revox A-77 tape recorder located in a separate room; each receiver records on a separate track. Doctor-patient interactions with hospitalized patients are also recorded stereophonically employing a Superscope portable stereophonic cassette recorder and separate high

---

<sup>2</sup>Informed consent has been obtained for each recorded patient, and all tapes and individual patients have been assigned random numbers.

