

A SYSTEM FOR COMPUTER AIDED PROCESSING OF PHONETIC MEASUREMENTS¹

Peter Holtse and Jens H. Stellingner

Abstract: A set of computer programs is described which will sample and average e.g. intonation, air pressure, or EMG activity. Data processing is essentially user controlled, but trivial decisions may be left to the computer if this is desired. The programs are accessible to users with little experience with computers.

1. Introduction

During the last two years a system of computer programs has been developed which aims at providing the phonetician with a processing tool for the kind of DC-signals normally recorded on e.g. Mingograms. The programs (called the Multi-Channel Processing system, or MCP) should enable the researcher to process statistically whole patterns of simultaneous events as well as single, selected measuring points.

It has always been a problem, especially when dealing with physiological recordings, that for practical reasons the investigation is limited to selected points in time which the researcher hopes will clarify his theory. It has been found, most notably in the field of EMG, that single points are in most cases only poor descriptions of the complex behaviour of the speech organs during the articulation of a particular speech utterance. Intonation is another good example of the importance of the pattern rather than a point.

1) The time code and the timing program was designed by Mogens Møller, who has also been of invaluable help with suggestions and discussions, especially during the early planning stages.

The programs described here is an attempt to provide a means of describing such complex measurements combined with the usual description of points. And it is hoped that they will be found useful in dealing with a variety of different measurements, such as for instance air pressure, air flow, EMG activity, intonation, and combinations between these and the audio signal.

Although the original reason for building the system was a project in EMG (see Fischer-Jørgensen 1974), the system is a general one, not written with only one kind of signal in mind. This means that in many ways it is incomplete, since obviously it cannot have every conceivable function ready made. However, the system is modular, so that many functions not thought of during the first planning stages can be implemented by changing only small sections of the complex of programs.

It was intended that the programs should be fairly easy to operate for the user with little or no knowledge of computer programming. This requirement is of course in direct conflict with the design features mentioned above. However, we believe that the original goal has been partially reached, at least for some of the more obvious processing functions.

2. System survey

The complete MCP system consists of four main parts:

1. The analogue recording process, in which the material to be investigated is recorded on a frequency modulating tape recorder.

2. Data editing and preparation. During this part of the process the FM-recordings are transferred to Mingograms and inspected visually. All information describing contents, the quality of the recorded material, where it is to be found, etc. is noted and transferred to punch tape.

3. The actual sampling process, in which the recorded material is transferred from analogue tape to digital representation and stored in digital form with its appropriate description of contents as it was specified during stage two.

4. The processing stage, during which the experimenter causes the digitized, stored data to be averaged and displayed in various ways.

3. Hardware set up

The MCP system is run on the lab.'s PDP-8/E computer with 32k of memory. It relies heavily on a 1.6M disk and uses a dual DEC-tape drive as additional storage. Data is sampled through an eight channel multiplexed analog-to-digital converter controlled by a real-time-clock. Data, results, etc. are manipulated and displayed via a Tektronix 4014-1 graphic terminal and a DEC-writer LA30.

4. Description of the system components

4.1 Analogue recording process

The MCP system assumes that the material to be processed is recorded on an eight channel FM-tape recorder. Of these eight channels one is used by the system for a special time code which uniquely identifies any position on each FM-tape. The remaining seven channels are available to the user. The time code is generated by a computer program and may be recorded on channel eight during the actual experimental recording, or it may be inserted some other time. Fig. 1 shows a system flow chart of the analogue recording process.

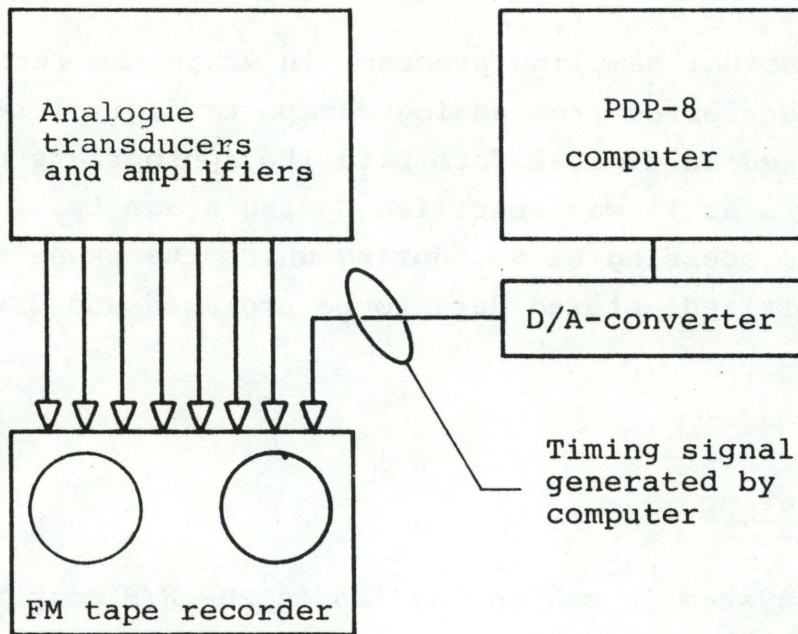


Figure 1

System flow chart of the analogue recording process. Seven data channels are recorded on the analogue tape recorder together with a computer generated timing signal.

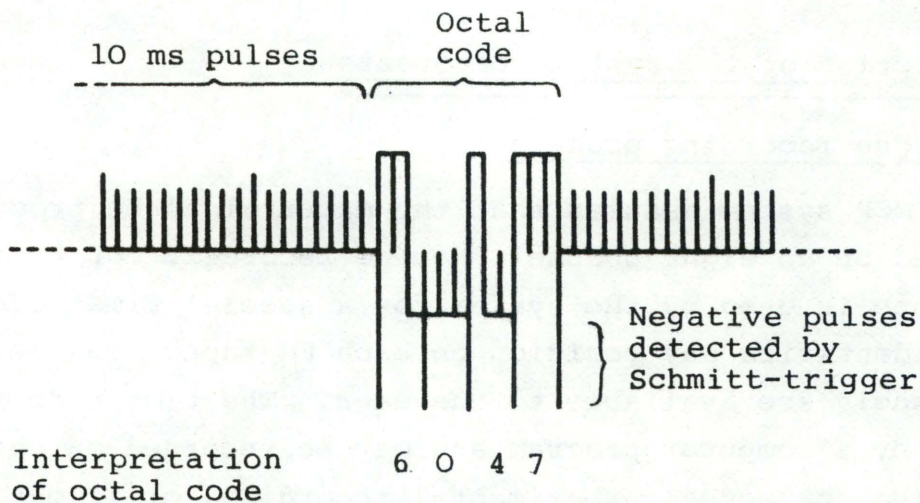


Figure 2

Time signal and twelve bit binary code displayed on Mingograms. Upward shift of zero line is interpreted as a binary one. Downward shift as a binary zero.

4.1.1 The time code

The time code is generated by a special computer program (MCTIME) and recorded on the analogue tape via the computer's analog-to-digital converter. The code consists of a short positive pulse every 1/100 second. Every two seconds the zero line is shifted up or down during twelve cycles to form a twelve bit binary pattern as exemplified in fig. 2. As may be seen from fig. 2, the basic 10 ms marking is recognizable even during the shifting of the zero line. The twelve bits are combined by threes and interpreted as a four digit number to the base of eight. Hence the code is spoken of as the Octal Code.

In order to facilitate recognition of the code signal during replay into the computer and to ease visual reading of the code, the beginning of every octal code and the end of every three-bit octal digit is signalled by a short, but large, negative pulse. During replay the time signal is led to both the analog-to-digital converter of the computer and to a schmitt-trigger entrance. The schmitt-trigger detects the negative pulses at the beginning of the code and starts the A/D-converter sampling the base line during the following twelve centiseconds. These samples are interpreted as explained above.

4.2 Off-line preparation of data

The contents of the FM-tapes, complete with time codes, is printed on Mingograms and inspected visually. This phase is of paramount importance, since the system is not an automatic data evaluator. Determining the quality of the recordings and where the interesting periods are located is left entirely to the experience of the researcher, who must later pass on his judgements to the MCP system. The results of the researcher's evaluation are collected in special files, known as Sampling Control files (or S-files). For each combination of speaker and utterance type one separate file must be collected. These files will later control the sampling and further data processing.

A Sampling Control file must contain the following information: The first record of the file identifies the speaker and utterance type. The second record describes the contents of the seven data channels to be sampled.¹ One record describes where every repetition of this utterance is located, i.e. the number of the analogue tape on which the individual item is recorded, its nearest preceding octal code, and the distance in centiseconds from the octal code to the beginning of the sampling window. The sampling window is a 1.25 second wide window placed round the item to be sampled. Only information within the sampling window will be stored in digital form by the computer.

Another record describes a number of time references within each sampling window. As many as seven time references may be associated with each item. One or more time references are used as line-up-points during the later averaging procedure, but the researcher is free to use the time references for his own purposes as well, such as describing the time distance between various events within the sampling window.

Finally, the S-file contains a record describing the quality of each channel of each sampled item. Thus each item/channel combination is assigned a number in the range zero to seven, where zero is normally assumed to describe an empty channel. Otherwise there is no inherent hierarchy of the numbers of the quality description, and the researcher may use them at will to describe the quality of his recordings.

Special coding forms are available for preparing the Sampling Control information. This information is later punched on paper tape. The punch tape files are read into a computer program, known as the MCP compiler, which checks the tapes for spelling mistakes, missing information, etc. From each input

1) If an experiment needs less than seven channels the remaining channels are marked as empty. They are sampled and stored but will not appear on any display.

tape the compiler produces a proofread S-file containing a complete description of all the repetitions of the utterance in question. This file is stored on magnetic tape (DEC tape) for easy reference later on. Furthermore, the compiler produces a printout of the complete S-file. The prints of the S-files provide the researcher with a useful survey of his sampled material.

The MCP compiler is equipped with an extensive default system, which allows the user to enter only really necessary information from punch tape, while information which is in any way predictable is supplied automatically by the default system. For instance, describing the contents of a channel will make the compiler assume that all repetitions of this utterance as spoken by that particular speaker are of acceptable quality. Therefore, the quality of the recordings need only be described explicitly in the (hopefully rare) cases where the recordings for some reason or other are not up to expectations.

Fig. 3 shows a listing of the Sampling Control information for one utterance.

4.3 Digital sampling process

The sampling process is controlled from the program MCSAMP. First the user enters a list defining the speakers and utterance types he wants sampled during the particular run of the program. MCSAMP then reads the appropriate Sampling Control information from the S-files created during stage two and merges the control information into one file. This file is sorted so as to minimize tape movement on the analogue tape recorder.

When sorting is completed the program makes its first request to the operator for a specific analogue tape to be mounted and starts reading the time code. Every time the program encounters a code combination defined in the control file it will sample a 1.25 second wide window from all seven channels with a sampling frequency of 200 Hz and store this record on the disk.

#####

208

IPUC DATALAB.

S S I - D A T A

DATE: 76/ 9/ 8

#####

MCP PROJECT NAME: # EMG-74

FILENAME: # DEMONZ. JS

SUBJECT: # JS

RESEARCH ITEM: # DEMONSTRATION

CONTENTS OF CHANNEL-NR. : # 1 BICP
2 TEMP
3 FRON
4 CT
5 VOC
6 F0
7 I

NUMBER OF RECORDS
IN DATAFILE: # 6

ANALOGUE TAPES
AND SAMPLING-POINTS: # REP. NR. AT-NR. OCT-CODE DEC-CODE
1 67 316 10
2 67 325 32
3 67 404 41
4 67 413 126
5 67 477 39
6 68 051 142

TIME-REFERENCES: # REP TA TB TC TD TE TF TG
1 40 79 24 53 70 93 112
2 40 80 23 51 70 93 112
3 40 80 24 54 72 93 110
4 40 82 15 52 73 97 114
5 40 83 22 53 75 98 118
6 40 85 21 57 77 98 118

ACCEPTABILITY LEVELS
IN CHANNEL NR. # 1 2 3 4 5 6 7
1 4 6 2 3 5 4 1
2 1 2 3 4 5 6 0
3 0 1 2 3 4 5 6
4 5 5 5 5 5 5
5 7 7 7 7 7 7
6 7 7 7 7 7 7

#####

Figure 3

Listing of Sampling Control information

When necessary the program asks the operator to rewind the analogue tape or mount a new tape.

The relatively low sampling frequency implies that most signals need to be hardware preprocessed before they can be sampled into the computer. Thus EMG activity must, for instance, be hardware filtered and integrated. And intonation is investigated by sampling the output of a pitch meter. However, due to the modular design of the system this limitation can, if desired, be circumvented by altering only the sampling program (MCSAMP). These alterations need not affect the controlling and sorting programs at all.

When all requested items have been sampled a special sorting program (MCSORT) is called. This program reads the sampled records and sorts them into separate files (called Z-files) for each speaker/utterance type. At the head of each Z-file is inserted the Sampling Control information from that speaker/utterance type combination. Thus all sampled items for a given utterance and descriptions of contents and quality for that utterance are stored in one file. Finally, the Z-files are transferred to DEC-tape for long term storage.

Fig. 4 shows a system flow chart of the complete data-entry process as described in sections 4.2 - 4.3.

4.4 Processing of data

During the design phase of the MCP system it was realized that it is impossible to foresee in all detail what the processing requirements of the future users of the system will be. It was therefore decided to let the user control the procedures directly in some detail rather than making a complicated automatic system. This means, of course, that users must spend some time typing out orders to the system. But the interactive nature of the system will, on the other hand, allow the user to see the results of his orders almost immediately, if he wishes.

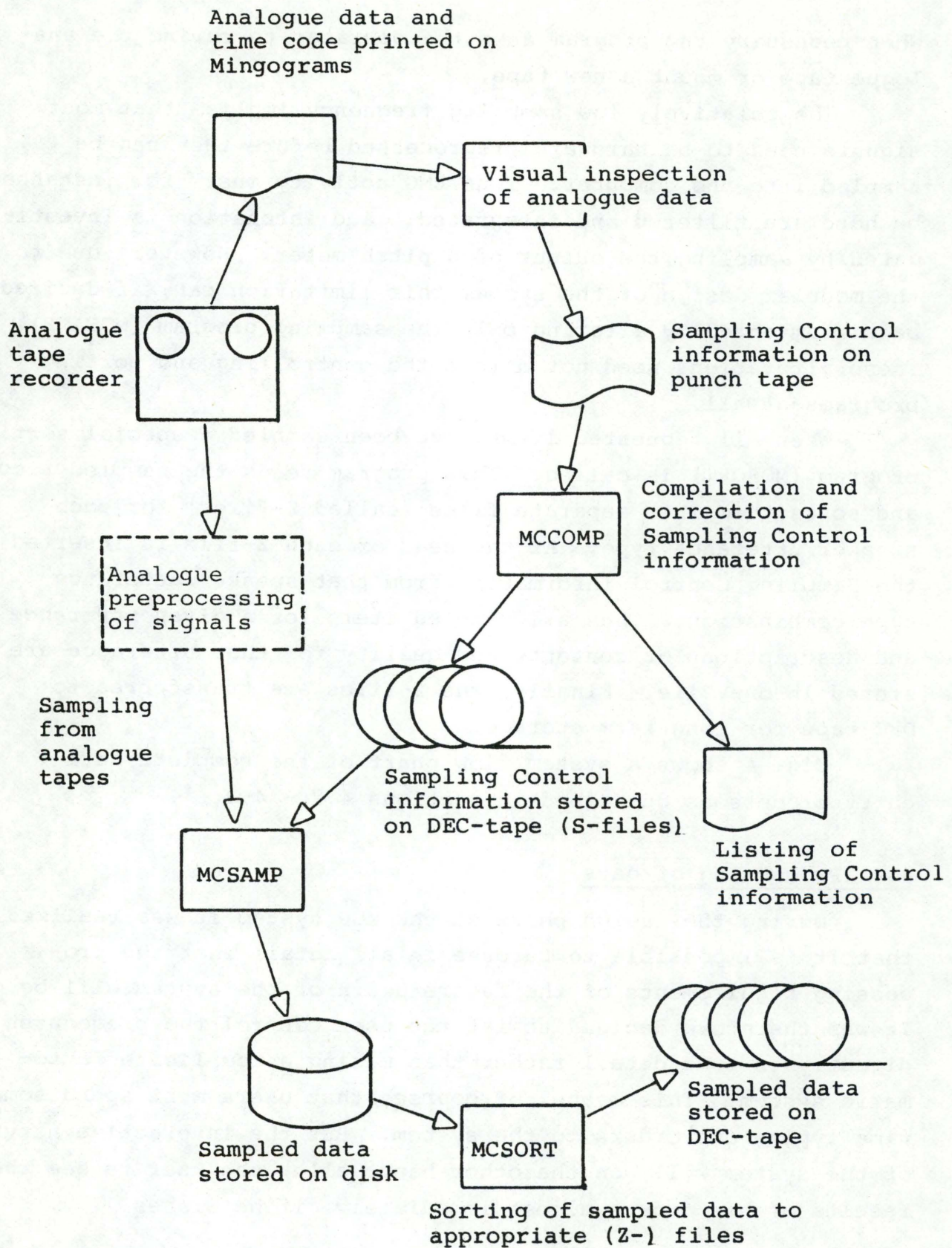


Figure 4

System flow chart of the MCP data entry process.

At the moment the processing procedures consist of the following five programs:

MCMODF - A data modification program which works directly on the Z-files. The contents of one specified record of the file is displayed on the terminal screen, and the user may enter various modification into this record and its corresponding description in the file head. Erroneous data values may, for instance, be deleted or time references may be altered. It is possible to delete single measurements in the middle of a specified channel without discarding the whole channel.

MCAVEX - Calculates mean values, standard deviations and number of tokens from one or more Z-files. The results are stored on the disk in a special X-file. For each input file (i.e. Z-file) to the program the user must define which of the time references should be used as line-up-point for the averaging process. Averaging is carried out with the input records aligned to these points, and the results are a 250 points wide mean window.

This possibility of defining a separate line-up-point for each channel is an important feature of the system. It is possible, for instance, to average the contents of one channel round a line-up-point at, say the explosion of a stop consonant, while another signal in the same recording is averaged round a line-up-point located at the beginning of a neighbouring /s/.

Channels to be averaged are defined by their contents - not by number. Thus it is possible to ask for, for instance, the EMG activity of the Vocalis muscle to be averaged without worrying about whether this was stored in channel two or three. This feature is particularly useful if somebody wants to average a given type of recording from two different recording sessions where the contents of the channels for some reason have not been ordered in the same way.

MCAVEM - Is an averaging program similar to MCAVEX, but calculates median values and quartile distances instead of means and standard deviations. The output is known as an M-file.

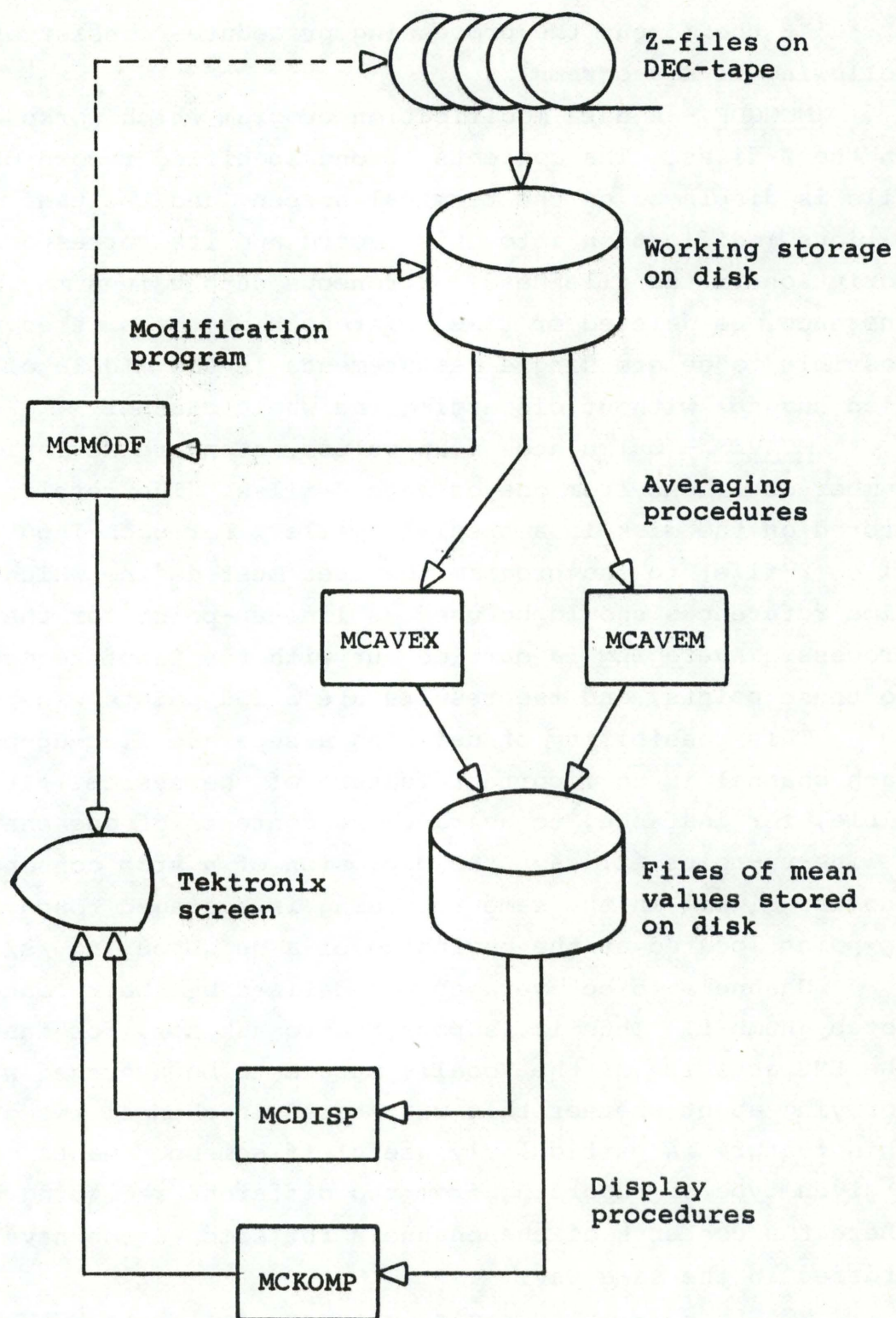


Figure 5

System flow chart of the processing part of the MCP system.

MCDISP - Displays the contents of all the channels of one average file (X- or M-) on the graphic screen.

MCKOMP - Displays several specified channels from specified average files for comparison on the graphic screen.

The logical interaction of the processing programs is shown in fig. 5.

Great pains has been taken to make communication with the processing procedures as simple as possible to the user not familiar with the working of computers. This problem has been approached by providing a set of supervising programs which allow the user to control the processing with approximations to normal words. Furthermore, a default system has been incorporated so that the programs will assume certain standard values for the input parameters if they are not explicitly defined by the user. This feature will in most cases cut down appreciably the amount of typing to be done by the user.

Hard copies of the displayed data and results are at the moment obtained by photographing the display screen. Figs. 6, 7, 8, 9 show examples of displays.

5. System capacity

The MCP system will handle almost any number of different utterance types and speakers, although the number of DEC-tapes available will usually put a practical limit to the amount of material that can be stored at one time.

The maximal length of one Z-file (containing raw, sampled data) is 17 repetitions. Since, however, the averaging programs will combine a maximum of 49 different input files into one mean file, utterances repeated more than 17 times per speaker are simply stored in two or more files and combined during the averaging process.

During one run the sampling and sorting programs will sample from a maximum of ten different analogue tapes, sort the sampled records into a maximum of 63 Z-files (one for each speaker and utterance type), or sample and sort a maximum of 450 items - whichever of these three limits is reached first. Experiments exceeding the limits must run the sampling and sorting programs several times.

References

- | | |
|---|---|
| Beauchamp, K.G. 1973: | <u>Signal processing: Using analog and digital techniques</u> (London) |
| Fischer-Jørgensen, E. 1974: | "Electromyographic investigation of Danish consonants, stress, and stød", <u>ARIPUC</u> 8, p. 203-206 |
| Knuth, D.E. 1973: | <u>The art of computer programming</u>
Vol. 3: Sorting and searching
(London) |
| Lorin, H. 1975: | <u>Sorting and sort systems</u> (London) |
| Madnick, S.E. and
J.J. Donovan 1974: | <u>Operating systems</u> (New York) |
| Port, D.K. 1971: | "The EMG data system", <u>Haskins SR</u>
25/26, p. 67-72 |

Legend to figs. 6,7,8,9 (pp. 216-219):

Figure 6

Display of sampled data from utterance: 'Det er wienerbrødet de siger' [de:'vi:ʔnʌbrœð?ddi'si(:ʌ)] as spoken once by speaker 'BM'.

Time references associated with this item are 'TA' at the boundary between [v] and [i:] and 'TB' at the explosion of [b].

The contents of each data channel is printed to the left of the display. The two top channels are EMG recordings. The two bottom channels are fundamental frequency and intensity. (The figures to the right of the display are internal scaling factors used by the display routines.)

The display was obtained by photographing the Tektronix-screen, which is slightly convex. This reproduction technique unfortunately introduces a slight distortion to the scale, especially towards the edges of the screen.

Figure 7

Results of averaging all repetitions of the utterance by speaker 'BM' which is exemplified in fig. 6.

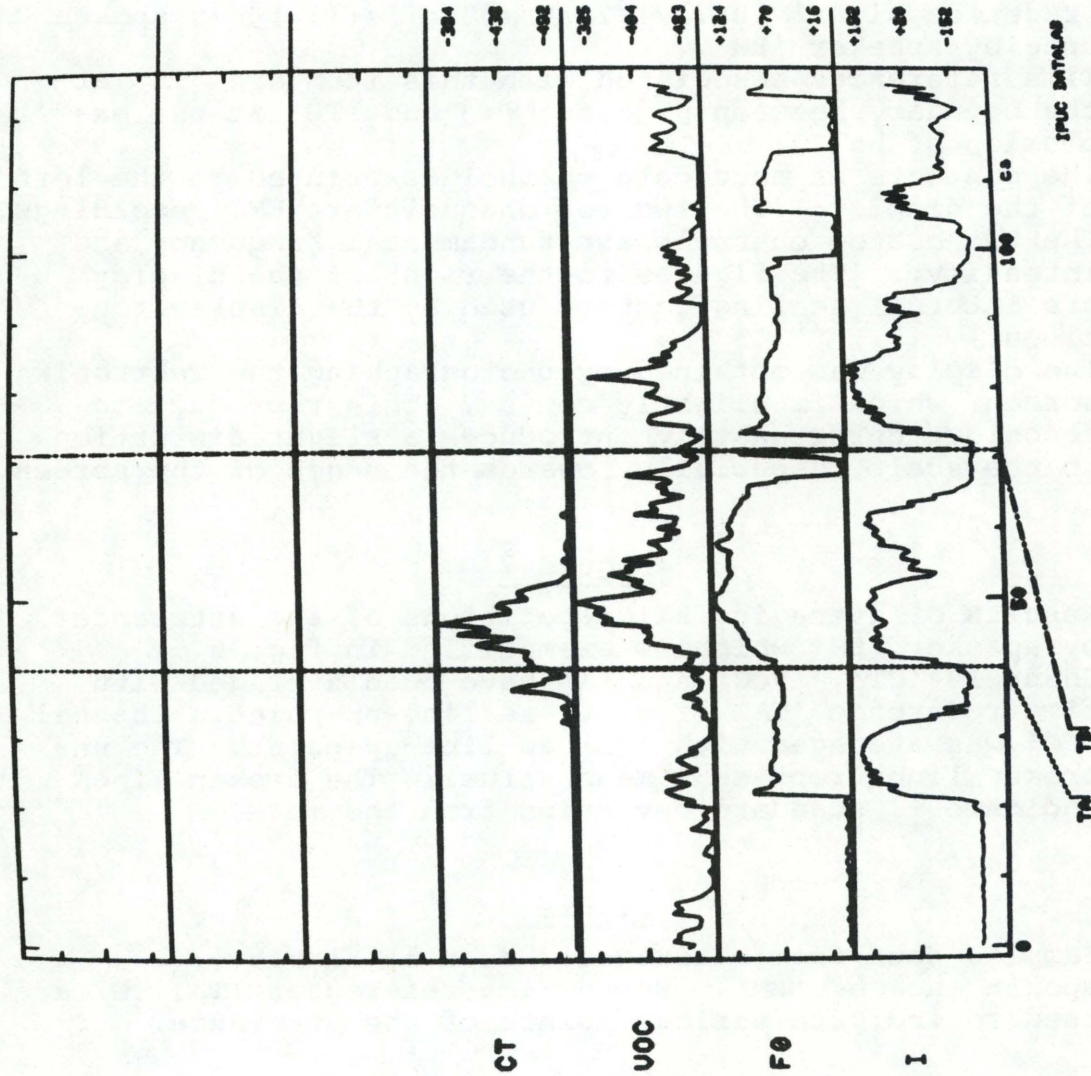
Channels 'CT', 'VOC' and 'I' have been averaged with time reference 'TA' (fig. 6) as line-up-point. Channel 'FØ' was averaged with 'TB' as line-up-point. The unbroken lines represent mean values. The broken lines indicate ± 1 standard deviation from the mean.

Figure 8

Sampled data from utterance: [de:bibi'bi:bibi(...)] spoken once by 'BM'. Seven time references (TA...G) are used to indicate various points of the utterance.

Figure 9

Average of all repetitions of the utterance exemplified in fig. 8. Time reference 'TF' was used as line-up-point for all channels.



DATADISPLAY
FROM FILE:

FX: WIE2NZ
FP: BM
RC: 01

CONTENTS:
WIE2NBR02DET

Figure 6

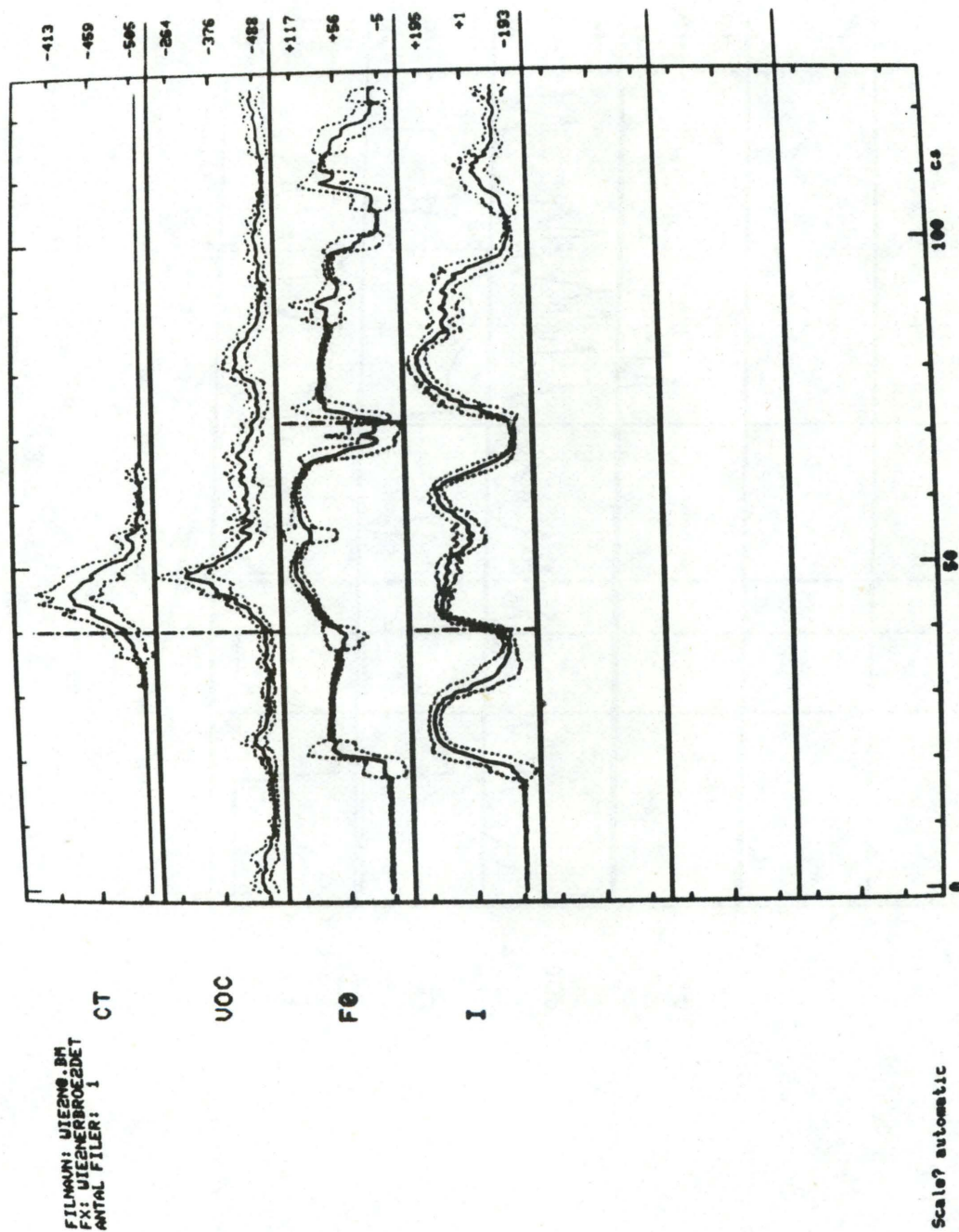


Figure 7

DATADISPLAY
FROM FILE:

FX: BI7ZZZ
FP: BM
RC: 03

CONTENTS:
BIBI'BI:BI17

Scale? automatic

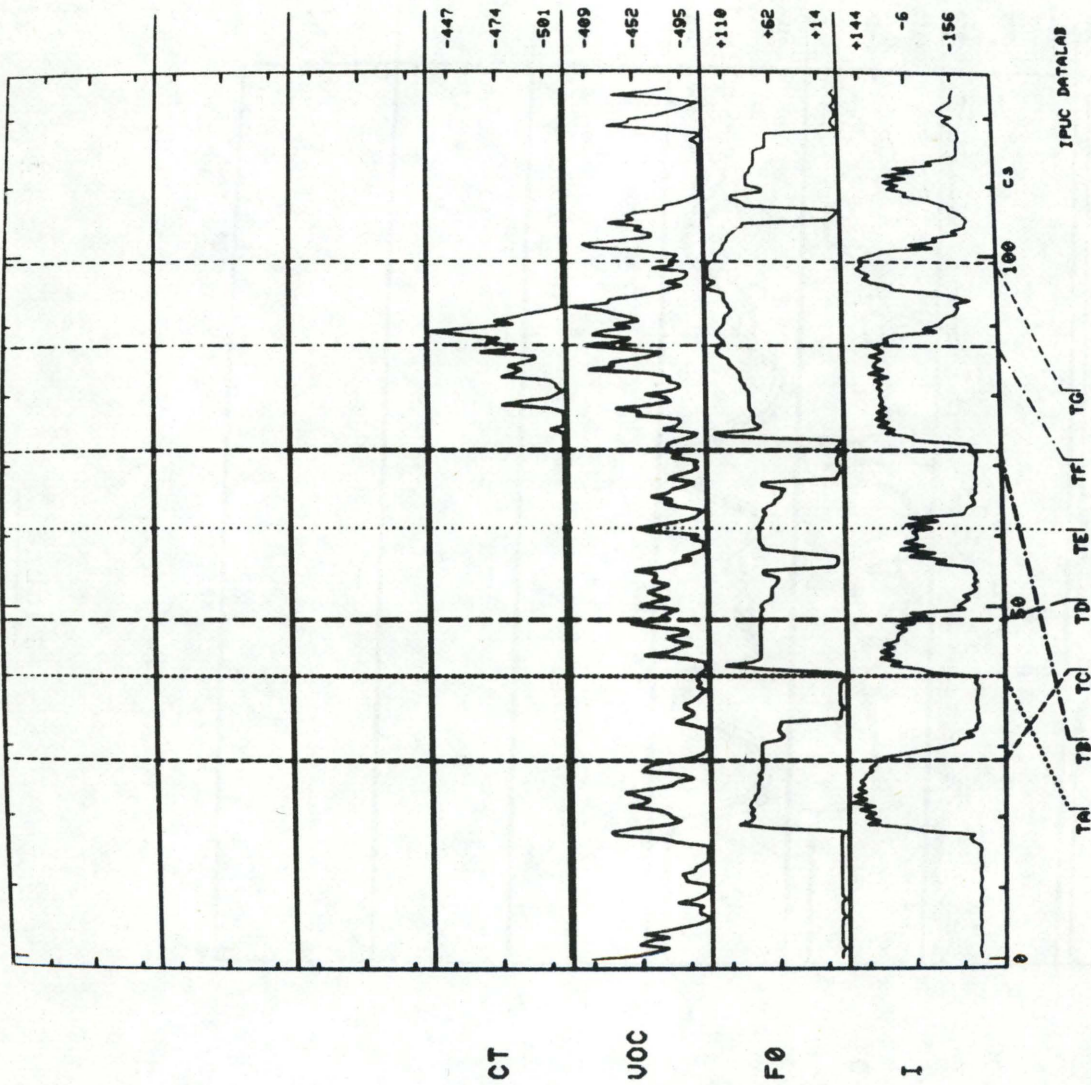


Figure 8

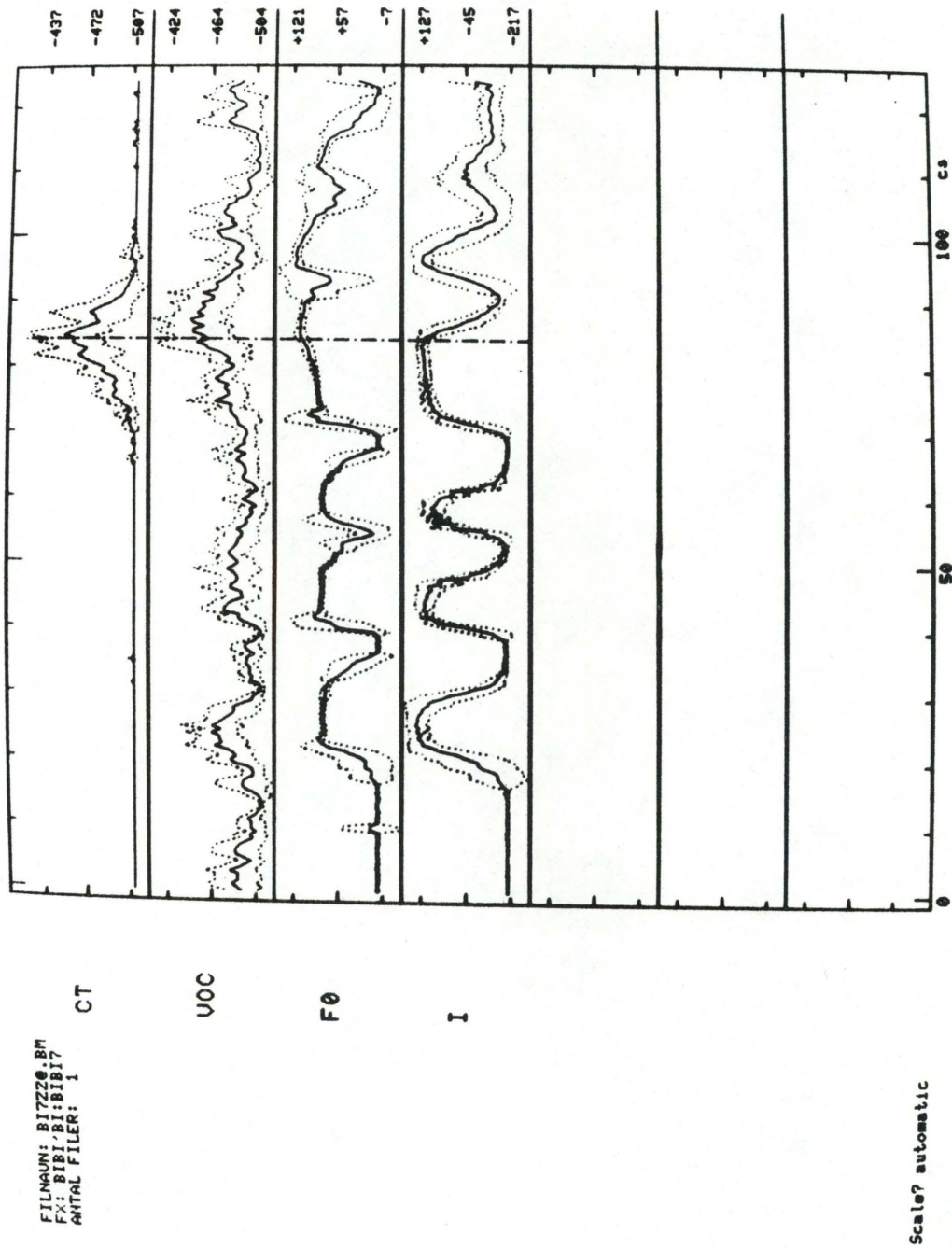


Figure 9