Intonation Analysis System

- Users Guide -

Stand: 31.10.1981

Verf.: Peter Wittenburg, Gunter Senft

Inhalt:

1. Intention

The intonation analysis system serves to calculate resp. display the suprasegmental parameters of a given piece of natural speech.

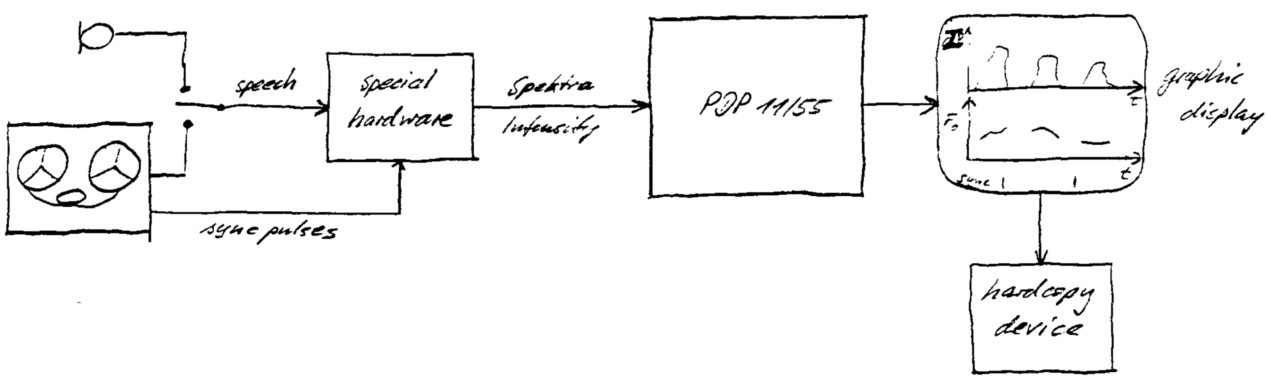
These are: - basic frequency (f_o) ⇒ pitch
- amplitude ⇒ intensity
- durations ⇒ time pattern

It also can serve to test pitch detection algorithms based on short term Fourier representations and to display series of these Fourier representations (f-domain representations) to detect other parameters.

The set-up was chosen so that nearly real time performance is achieved, i.e. that the user will get an output without any large delay. So the system works rather efficiently both in "interactive" sessions with short speech material and with large amounts of speech to be analyzed.

2. Global layout

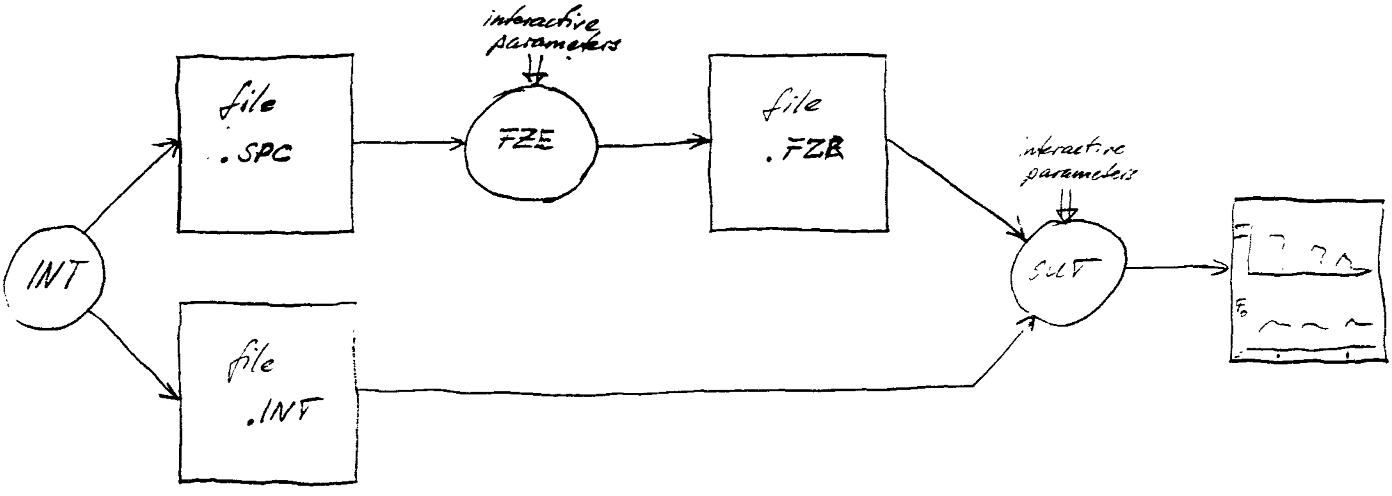
a) Hardware



The speech can be entered via Micro or TR. It passes some fast, special purpose electronics, so that a conversion from the original speech into series of spectral representations and a coding of the intensity. Also synchronization purposes pulses can be entered into the information, which can be seen later on on the graphic output.

On the computer, three programs serve for data acquisition, analysis and output. The output is first generated on a graphic display and then a copy can be made on paper.

b) Software



The software is set up by starting the indirect command file @ DL:INT, which resides on the disc DL:Intonation.
The acquisition is done by starting the indirect command file@INTST. This program creates a file .SPC which contains series of spectra and a file, .INT which contains the sampled intensity values. The derivation of the basic frequency is performed by the program FZE. FZE checks first if the data is acquired correctly and then calculates per spectrum the F_o-value which is stored sequentially in a file .FZE. Therefore some analysis parameters have to be defined. The program OUT performs the output of both F_o and intensity (if given also sync marks are drawn). Some parameters defining the form of the output have to be entered.

Figure 1 shows the correct cabling of the system. The Intonation Analysis System (IAS) can only work properly, if the cabling is done correctly. This is, of course, a rather trivial statement, but experience has shown that troubles start when people try to do the cabling. Therefore make sure that:

1st the envelope generator (2) is connected with
- the attenuator (3)

and with the computer via the
- AD 11K cable

2nd the attenuator (3) is connected with the preamplifier (4)
(and, of course, with the envelope generator (2))

3rd the preamplifier (4) is connected with the
tape-recorder
and with the spectrum analyzer
(and, of course, with the attenuator (3)).

4th the system is connected with the computer:
therefore check at the computer whether the two cables (label 125) are connected with the DR 11 B via the front side connectors.
If you have to do that connection at the computer, handle the cables very carefully!

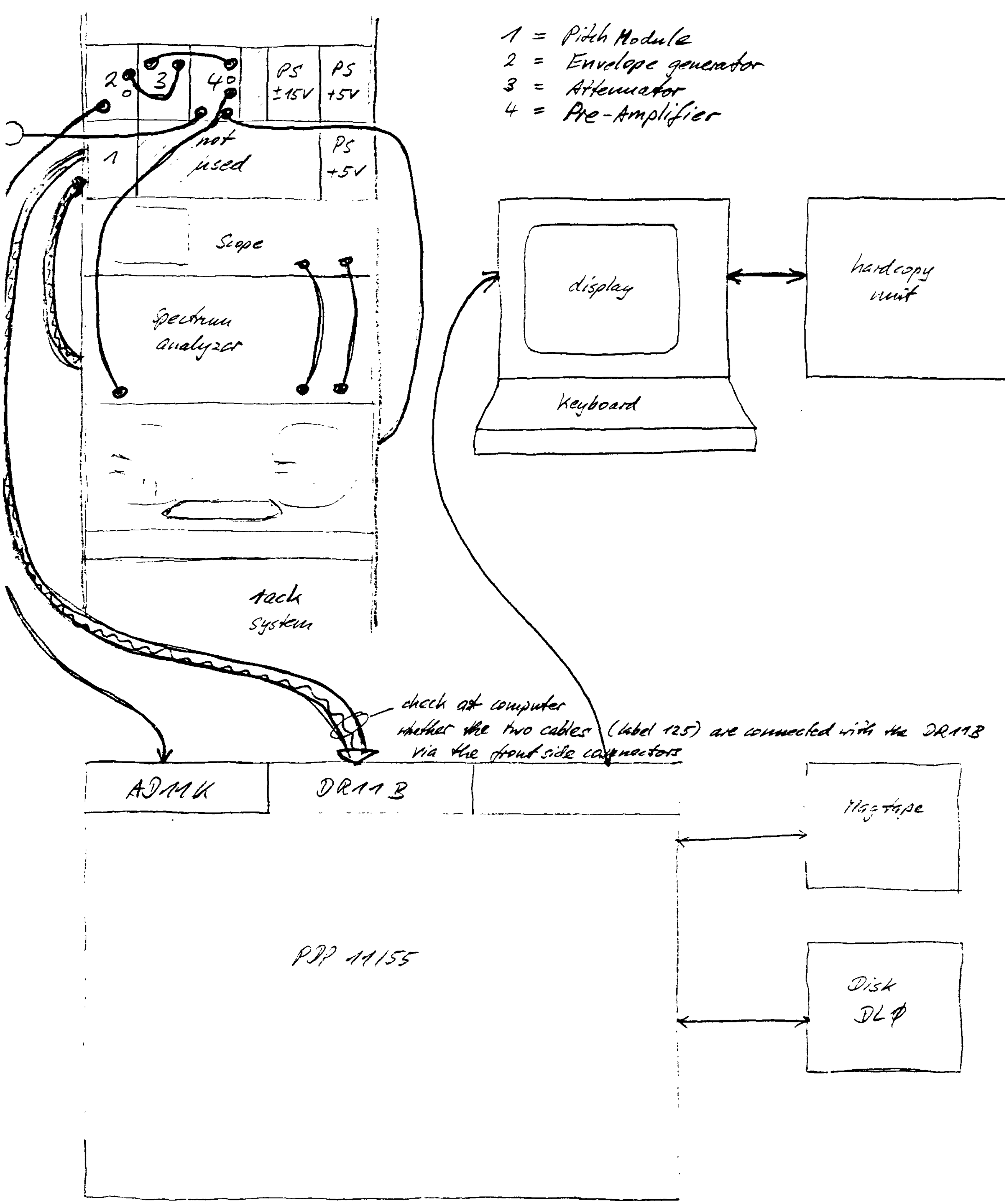
5th the computer has to be loaded with the INTONATION DISC
If you have to load the computer with the intonation disc

- ask someone from the technical group, how and where you can load the disc.
- press the load button
- wait till load button lights up
- change the discs, that is put in your intonation disc and make sure that it fits correctly in the machine.
- press the load button
- wait till the "ready" button lights up.

Figure 1

3. hardware/cables

check carefully



The whole setup of the hardware should normally be chosen as indicated on the photo (see photo 1).

If you have done the cabling, switch on all the units needed for the intonation analysis; these are:

- the tape recorder
- the spectrum analyzer
- the scope
- the terminal

  and
- the hard copy unit

Please make sure after having finished your analysis that all these systems are switched off before you leave the room.
If all the units remain switched on for a very long time (e.g. over night) they may cause severe damage to the units.

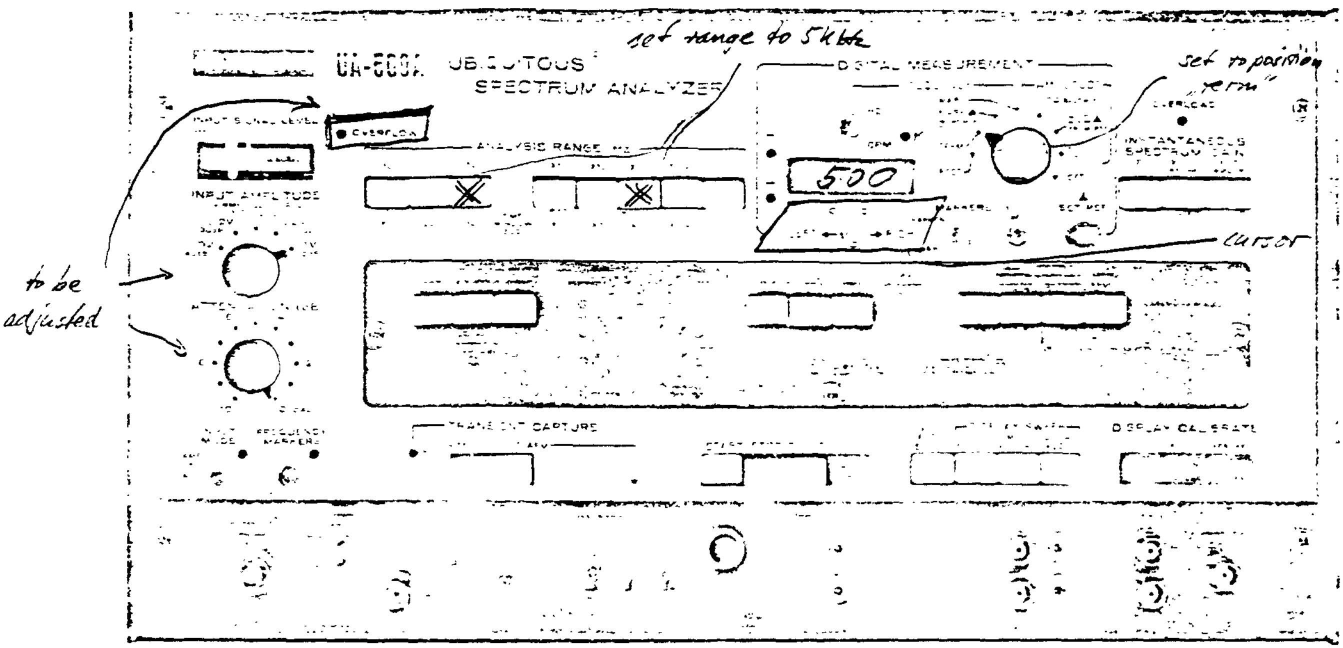
The most important point is to adjust all auditive levels so that the full dynamic range of all components is optimally used, i.e. test for a sentence that the amplifiers resp. attenuators are adjusted so that the overflow/overload lamps on spectrum analyzer, amplifier, and envelope generator just do not come up.

Adjusting the amplifier needs turning up or down the volume control knob.
Adjusting the attenuator needs setting up or down the number indicated at the top of this unit.

A correct adjustment of all auditive levels is the basis of a good intonation analysis. To do it properly needs some experience with the IAS. So do some tests, "play" with the system, and do not give up too soon.

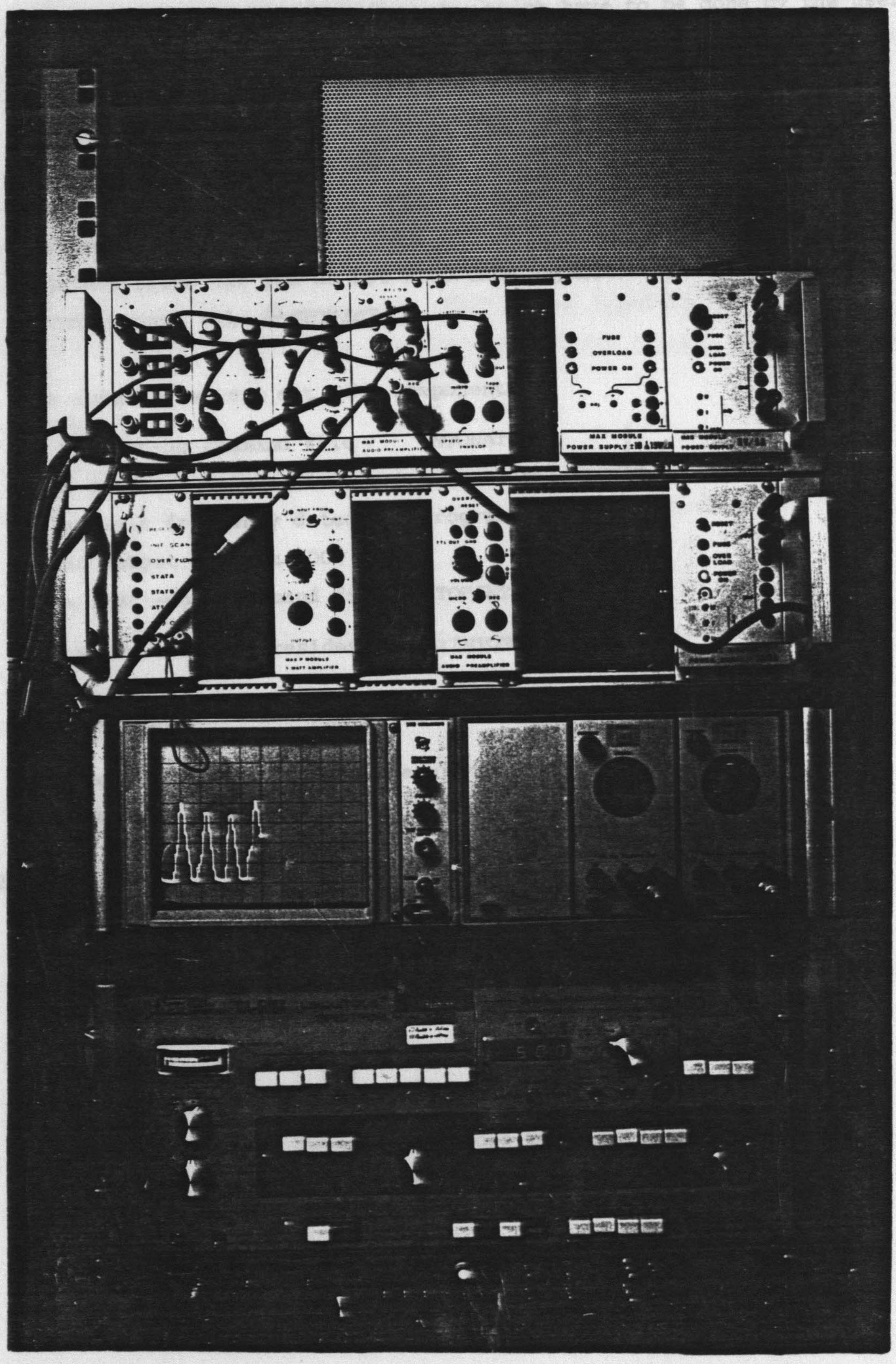
On the emplifier (4) you can chose whether you want to use the microphone or the tape recorder as "speech source", i.e. check the position of the switch before you start.

On the spectrum analyzer you should check whether the settings are exactly as indicated on the photo, except the input attenuation which has to be adjusted by the user.
When switching the spectrum analyzer on you have to position the cursor until you have got a number between 500 and 1000 with the last digit being a zero (examples 500, 680, 970, ...; not allowed: 79<u>1</u>).
This number divided by 10 is used in the software as the parameter N!!
The normal value should be 500, so N = 50.



The oscilloscope gives you a kind of first hint about your input by its visual display.

However, do not take this display too serious. You can adjust the base line of the display by turning the two variable knobs on the right of the unit.


FUSE
OVERLOAD
POWER ON
MAX MODULE
POWER SUPPLY
AUDIO PREAMPLIFIER
SPEECH ENVELOP
INIT SCAN
OVER FLOW
STATA
STATB
MAX P MODULE
1 WATT AMPLIFIER
MAX MODULE
AUDIO PREAMPLIFIER

4. Software/parameters

Software operation includes 7 steps, which have to be done by the user:

- login and mount disc
- run command file @ INTONAT
- run acquisition program @ INTST
- run analysis program FZE
- run output program OUT
- delete files (if you do not need them anymore)
- run command file @STOPINT

The software exists out of 5 components:

1. a start-up command file INTONAT.CMD
2. a data acquisition indirect command file @INTST
3. a basic frequency detection program FZE
4. an output program OUT
5. a shut -down command file STOPINT.CMD

The whole software resides on DL:INTONAT

1. INTONAT.CMD

Drivers are loaded and tasks are installed.

Program INT is started.

2. @INTST

INT samples data from spectrum analyzer and envelope generator. Spectral data is stored in a file .SPC while the intensity data (sampling frequency 100 Hz) is stored in a file .INT.

INT at first asks for a file specification: {DL:INT } < name >

use DL: if you do not have to much sentences to analyze at one step
use MT: if you have several sentences to be analyzed at one step
Name can be chosen by the user (not more than 7 letters).

then INT asks for three parameters: NSEC, N, RANGE (all integers)

NSEC you have to input an estimate of the lenght of your stimuli (in sec)

N you have to divide the number indicated on the spectrum analyzer by ten.

RANGE band setting on spectrum analyzer (normally 5 letters)

INT then opens a file, if opening was successful INT asks whether everything is OK, where you can answer Y or N (instead of Y also return) INT then prepares everything and prompts with "RETURN TO START SAMPLING". It starts immediately sampling after having typed in "return". If you work with the tape-recorder, type first return, then start the tape.

3. FZE

FZE checks the correctness of the data stream (synchronity check), decodes the bit pattern and calculates for every spectrum the basic frequency F.

It at first asks for a file specification {DL:/MT} < name >
So you just enter the same as you have entered in the program INT

FZE then opens the files: {DL:/MT:} < name > .SPC } old
DL: < name > .INT
DL: < name > .FER new

From file .INT it reads from the first block the parameters N, RANGE as default values.

FZE then asks whether you want to input parameters; if you type N, it uses the default values N = 50, RANGE = 5, PAUSE = 1000, FTHR = 8, NTHR = 20, YTHR = 1.7, XTHR = 5.0.

These settings turned out to work pretty well for a wide range of material.

Meanings of some parameters:

.FTHR serves as an ideal high pass filter, since FTHR = 8 means that all frequency components smaller than 80 Hz are ignored. For female child speech you can make FTHR even higher.

.NTHR serves as a noise rejection level (in voltage), so depending on the sound level you can vary this parameter.

.YTHR and XTHR are parameters used by the pitch detector 1.7 and 4.0-6.0 turn out to deliver good results.

.PAUSE serves to identify pauses, i.e. when the number of spectra with all values < NTHR exceeds PAUSE than a pause is indicated and calculation starts again when a value > NTHR.

4. <u>OUT</u>

OUT is the program which performs the output of intensity, F_o and sync pulses (if present). With help of parameters you can vary the format of the output.

OUT at first asks for a file name only: < name > since the device always is DL:. It opens two files: DL: < name > .INT }old

DL: < name > .FZR

It then asks if you want to have the intensity curve in dB (D) or metric scale (V). If you don't want to use the standard default values, you have to enter new parameter-values:

- POINTS/STEP is a factor which just stretches the output format on the display.

.DELTAT determines the time resolution of your output, i.e. if DELTAT = 10 then you get for each 10 ms a value of both I and F_o (if necessary a simple interpolation is done).

.FTHR and PAUSETHR again serve to identify pauses to reduce the amount of output

.FLOW and FHIGH determine the F_o sealing.

.SCALE-STEP determine the sealing on the time axis in msec.

5. Procedure/Examples

- Turn on power of all units at least 5 minutes before you start working!! i.e. switch at the side of the rack system, switch of hardcopy unit, switch of spectrum analyzer and scope.
- press erase on display
- check all settings, as indicated in this documentation
- check all cabling, as indicated in this documentation (if the lights are still on on the Pitch Module, the cable of this computer is not connected properly)
- check the microphone (TR switch)
- adjust the audio levels so that no overflow occurs
- check, if you want to use a Mag Tape, if it is installed
- then login, mount the disc and start the software (INTONAT)

operation

- when finished stop the software (STOPINT)   (BYE)
- switch power off on all units

type

```
HEL 7,1/PETER

          RSX-11M BL26   MULTI-USER SYSTEM

GOOD AFTERNOON
15-NOV-81 15:35 LOGGED ON TERMINAL TT4:

Welcome to RSX-11M V3.2 timesharing

*
******   You are logged onto the PDP 11/55, our mainframe!!!   ******
>MOU DL:/OVR/VI
MOUNT ** VOLUME INFORMATION **
CLASS       = FILES 11
DEVICE      = DL0:
LABEL       = INTONATION
UIC         = [1,1]
VOL PRO     = [RWED,RWED,RWED,RWED]
FILE PRO    = [RWED,RWED,RWED,R]
CHAR        = []
ACP NAME    = F11ACP
>
```

type

check if label "Intonation" is printed

type

```
@INTONAT
>UNL SP:
UNL -- DEVICE SP0: NOT IN SYSTEM
>UNL XB:
>UNL MT:
>LOA LA:
LOA -- DRIVER ALREADY RESIDENT
>LOA XB:
>LOA MT:
>REM INT
>REM FZE
>REM OUT
>INS DL:INT/TASK=...INT/PRI=200./CKP=NO
>FIX INT
>INS DL:FZERA/TASK=...FZE
>INS DL:INTOUT/TASK=...OUT
>@ <EOF>
>
```

```
type      @INTST
          >REM INT
          >INS INT/TASK=...INT/PRI=200.
          >INT
type      SPECIFY DEVICE AND FILENAME (NO EXTENSION) : DL:SATZ1   (or DL: Peter 1   etc )
                                                                       disc  name to be
                                                                             chosen by the user
          INTFIL : DL:SATZ1.INT }
          F0FIL : DL:SATZ1.SPC  } two files are created
type      SPECIFY NSEC , N AND RANGE : 4,50,5
           DMA PARAMETERS :
          NBUF :      9 NSEC :     4 BUFSIZ :  4096
           LPA PARAMETERS :
          NBUF :      1 NSEC :    10 BUFSIZ :  1024
           ALL PARAMETERS OK ?
type*      Return
           RETURN TO START SAMPLING
type      Return

          4                 50                  5 (kHz)
          nr. of seconds    (500/10)            setting of range
                            nr. of points
```

*Return means: type " Return ↵ " key on the terminal

```
DMADON
DMBFDN
DMADON
DMBFDN
DMADON
DMBFDN
DMADON
DMBFDN                          input in progress
DMADON
DMBFDN
DMADON
DMBFDN
DMADON
DMBFDN
DMADON
DMBFDN
DMADON
DMBFDN
DMADON
ALL DMA BUFFERS FILLED          end of input ; stop speaking
DMBFDN
LPADONE
LP-DK-DONE
ALL REQUESTS WELL DONE          input was successful
...INT  --  STOP  END OF INTONATION SAMPLING

>@ <EOF>
>
```

```
type    FZE
type    SPECIFY DEVICE AND FILENAME (NO EXTENSION) : DL:SATZ1
        DL:SATZ1.INT
        DL:SATZ1.FZR   new file is created
        DL:SATZ1.SPC
type    PARAMETER INPUT WISHED ? Y
type    N,RANGE,PAUSE : 50,5,1000
type    FTHR,NTHR,YTHR,XTHR : 8,20,1.7,5.0
        START AT BLOCKNUMBER :       17
         NEW BUFFER READ
         NEW BUFFER READ
         NEW BUFFER READ
         NEW BUFFER READ
         NEW BUFFER READ
         NEW BUFFER READ
         NEW BUFFER READ
         NEW BUFFER READ
         EOF REACHED
         FZER TERMINATED *********************
        >
```

make sure that you use the same name as in ⑨ INTST

nr. of points in spectrum, range setting, nr. of f_0 points after which a pause is identified

8 : (= 80 Hz) lowest expected frequency

20 : metric noise rejection level ; use a higher level, if your tape is noisy

1.7
5.0 } algorithm dependent variables

```
type   OUT
type   INPUT FILENAME : SATZ1
       INTENSITY FILE : DL0:SATZ1.INT
       FZERO FILE : DL0:SATZ1.FZR
type   INT.FILE IN VOLT- OR DB-SCALE(V/D) : D
type   PARAMETER CHANGE WISHED(Y/N) : Y
type   INPUT POINTS/STEP,DELTAT : 4,10
type   INPUT FTHR AND PAUSETHR : 1,1000
type   INPUT FLOW,FHIGH,SCALE-STEP : 50,250,100
type   AUTOMATIC COPY(Y/N) : N
type   MEAN(Y/N) : Y
```

name

D = Logarithmischer Output
V = metrischer Output

if N: default values are taken

4: stretches (use then 5,6, or any higher integer) or compresses (use 3 or any lower integer) the output on paper

10: 10 ms per drawn point f_0

if you have a number of sentences, pauses are identified and not drawn on paper

50-250 = upper or lower boundary for frequency scaling

high voices (female voices) most often need a higher boundary than 250

100 = 100 ms scaling on time axis

if Y automatic hardcopy is made

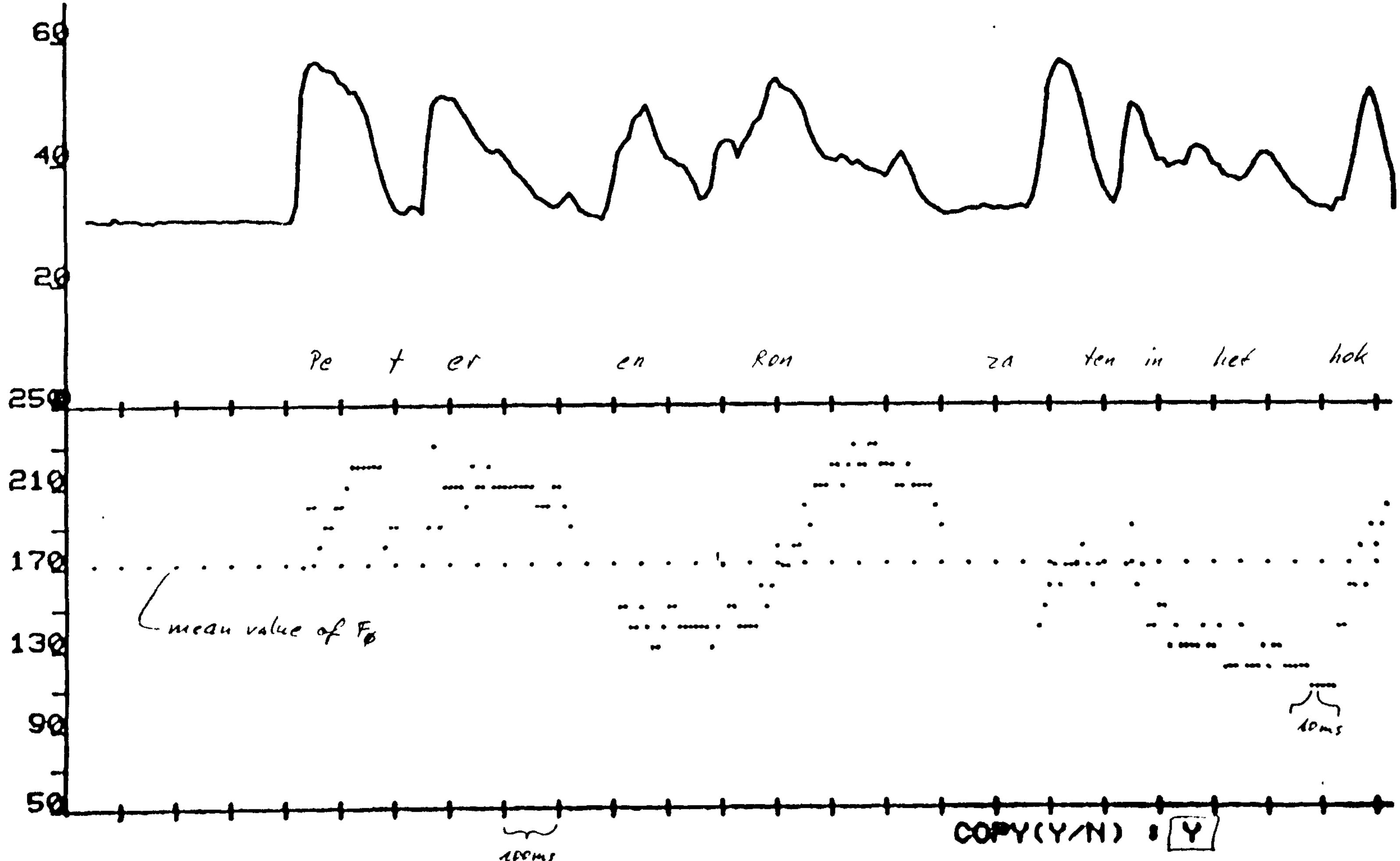
mean value is calculated and drawn as dotted line

The terminal then shows the analysis in such a display :

```
FILE : PETER1          DELTAT :      10 ms
DELTAF :    6.23 ms    DELTAI :   10.00 ms
SAMPL.FREQ. :    100    SCALE :     100 ≙ 100ms
PAGE :     1
```

page indication

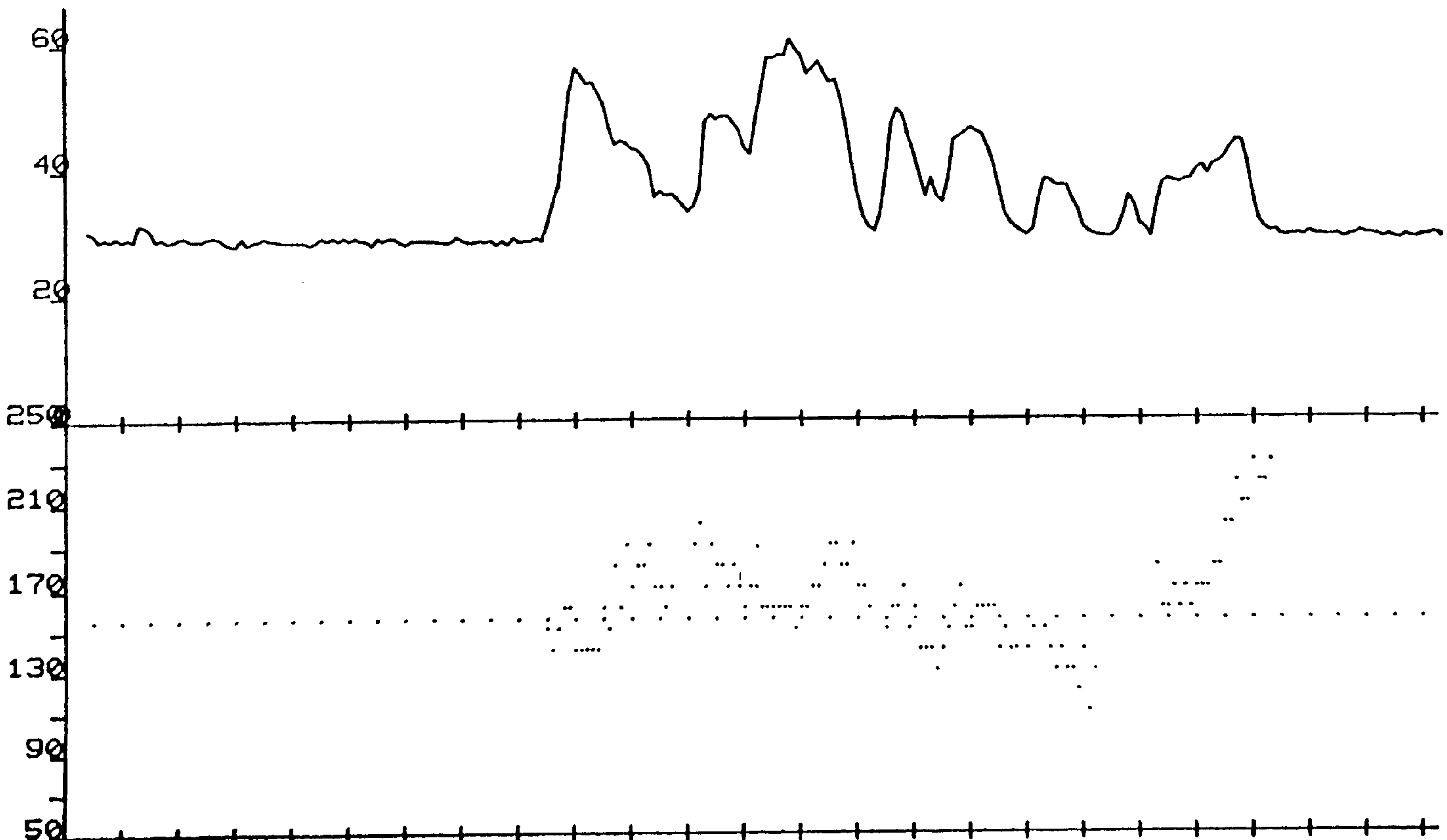
envelope resp. intensity is sampled with 100Hz



type y if you want a hardcopy

TYPE

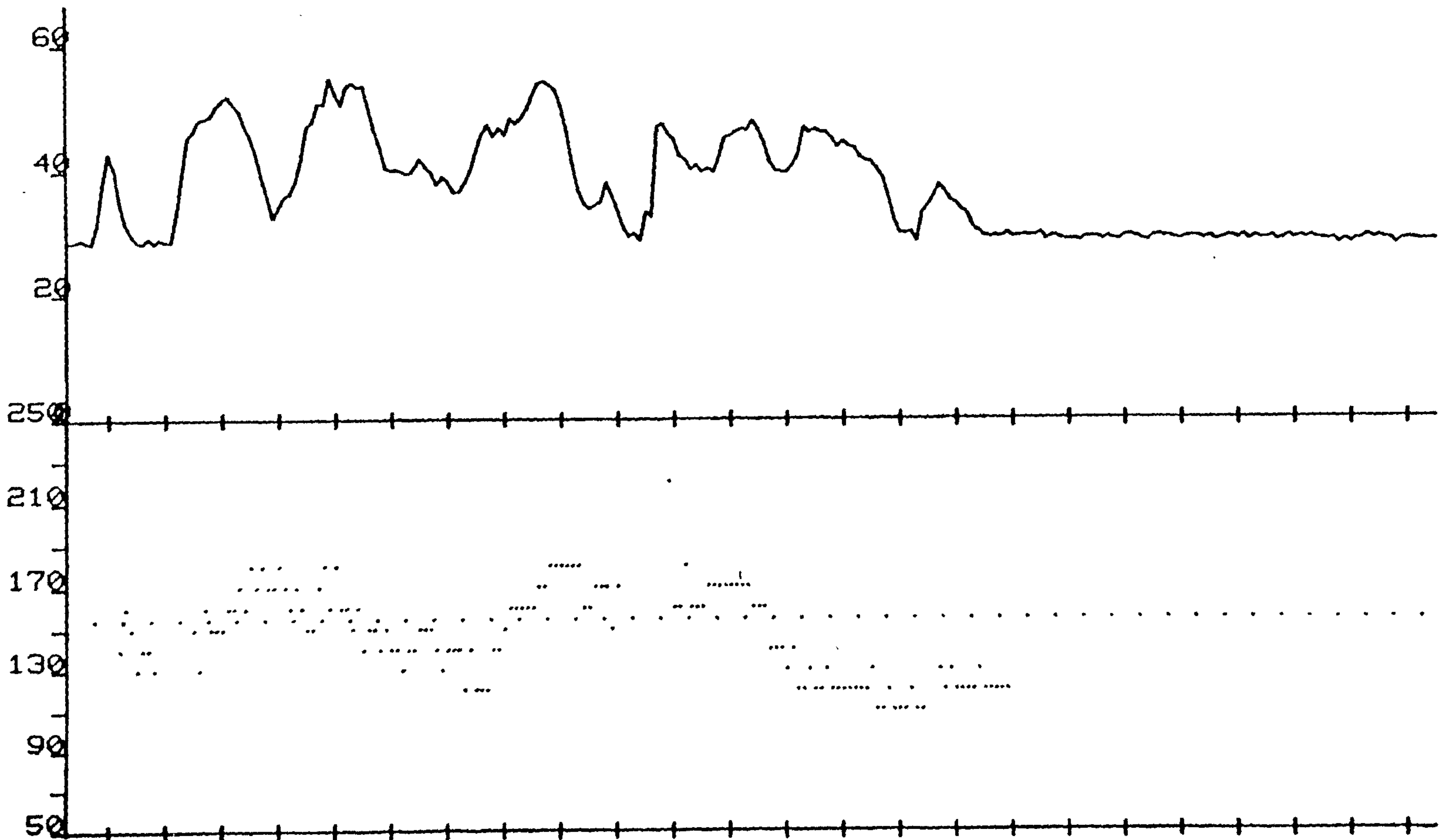
FILE : GUNTER1 DELTAT : 10
DELTAF : 6.23 DELTAI : 10.00
SAMPL.FREQ. : 100 SCALE : 100
PAGE : 1



COPY(Y/N) : Y

manche Leute reden tief

FILE : GUNTER1 DELTAT : 10
DELTAF : 6.23 DELTAI : 10.00
SAMPL.FREQ. : 100 SCALE : 100
PAGE : 2



COPY(Y/N) : Y

.. hohe Stimme haben,

type

```
END OF FILE,NEW FILE WISHED(Y/N) : N
...OUT  --  STOP END OF OUTINT
```

if the data are successfully sampled in INT you can always go back either to FZE or to OUT and change your parameters there to improve your output

```
>
> if you are content with your output,
> delete the created file by typing:
>
>PIP SATZ1.*;*/DE
>
>    blank !
>
>@INTST
```

type PIP SATZ1.*;*/DE — Return (before typing "return" key, check your deletion command !)

type @INTST — if you want to go on with your analysis and proceed as indicated above – (id est : DL : Satz 2 .... etc)

```
END OF FILE,NEW FILE WISHED(Y/N) : N
...OUT  --  STOP END OF OUTINT

>PIP SATZ10.*;*/DE
>
>
>
>BYE
>
HAVE A GOOD AFTERNOON
15-NOV-81 16:13 TT4:  LOGGED OFF
>
DMO -- TT4:    DISMOUNTED FROM DL0:    *** FINAL DISMOUNT ***
```

type — if you have finished your work

SWITCH OFF ALL UNITS !

19 b

If you have any problems please contact Peter (Tel: Institute: ext. 122
Private: 226066)

We are also looking forward to any suggestions for improving our system. Check, for instance, whether you think it useful to get with your printed output some information like:

- $\overline{X}$ of F_o
- Ranges of F_o
- s d
- or on filters and how they were used to suppress noise.

Have fun and success with your intonation analyses.

Gunter and Peter