

A MICROCOMPUTER PAUSEMETER FOR LINGUISTS

Jönsson G, Manderesson B and Sigurd B.

INTRODUCTION

The existence of pauses at certain points of utterances has long been noted. Some references from the latest decades are the following: Lounsbury (1954), Goldman-Eisler (1958), Boomer (1965), Gårding (1967), Beattie (1979). Current interest in speech production has focused on pauses as they seem to be important clues to speech planning activities, see e g Butterworth (1980), Kempen and Hoenkamp (1982). Pause behaviour in simultaneous interpretation has been studied by e g Goldman-Eisler (1972), Barik (1973) and Vamling (1982). The relation between pauses and anxiety is treated by Cook (1969). Differences in pause behaviour between men and women are treated by Einarsson (1978).

Measuring pauses and speech segments by hand on spectograms, oscillograms or their digital equivalents gives detailed information about the speech signal. But it is time-consuming and requires good training in phonetics. Many linguists would welcome an automatic pausemeter as a simple tool for determining the pause characteristics of utterances. An automatic pausemeter may process the speech signal and display or print diagrams and tables showing the durations of pauses and speech segments.

Automatic pause measuring equipment developed in recent years is mentioned by e g Barik (1973), Beattie (1979). It has not, however, been discussed much among linguists and phoneticians

to whom such instruments should be of great interest. Automatic pausemeters do not belong to the standard equipment of linguistic or phonetic laboratories, although their potential for speech research is very promising.

The purpose of this paper is to describe a simple automatic pausemeter which can be built at any department where a microcomputer and some standard electronic components are available. The pausemeter to be presented exists in two versions: One (1) is a display (real time) version which shows the duration of the pauses or speech segments as you produce them.

The other (2) is a more exact version that stores measured time values for subsequent use for various computations, displays and print-outs. This second version prints tables of the duration of all pauses and speech segments, computes their average durations and standard deviations and plots various diagrams.

The hardware has been developed by Jönsson, the programs (in Basic) have been written by Sigurd and Mandersson has assisted in the conception and design of the system. Sigurd has written the body of this paper. Jönsson and Mandersson have written the description of the pause-speech decision unit and contributed to the discussion of the problems of distinguishing pauses in speech.

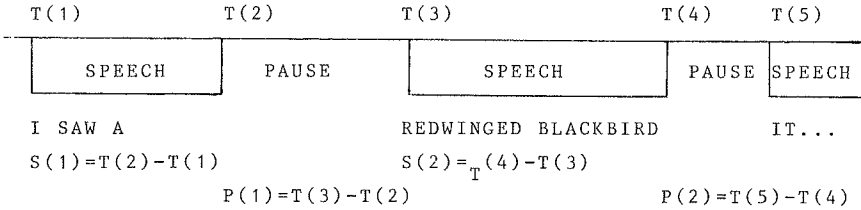
PAUSES AND SPEECH SEGMENTS

To the discourse analyst the pauses of focal interest are those between paragraphs, clauses, phrases and words. We will call these discourse pauses. Such pauses of interest are marked by (p) in the following example:

I saw a (p) redwinged blackbird (p) It was (p) perching in
a (p) redbud tree.

The beginning of the utterance can be depicted as in fig 1.

UTTERANCE



T(i): Time reading number i
 S(j): Speech segment number j
 P(k): Pause number k

Fig 1 Idealized view of the pauses and speech segments in speech.

The basis of the pausemeters are the time (clock) readings at moments when there is a change from pause into speech or vice versa (T(i)). The duration of a speech or pause segment can be derived by subtracting a preceding time reading from the current reading. Generally the duration of a speech segment (S(j)) can be computed by deducting uneven readings from even ones and pauses (P(k)) by deducting even readings from uneven ones. Given the duration of all speech and pause segments further calculations can be undertaken.

The decision whether there is a pause or a speech segment is a binary decision and it is not easy to take. A human listener may mark pauses by ear and even measure them roughly by a stop

watch, but as is seen on any spectrogram or in experimenting with automatic pausemeters it is not easy to design an automatic device which achieves results in accordance with human preception.

DETERMINING THE PAUSES OF SPEECH

A pause may be defined as a time segment where there is no energy present. Disregarding the so called filled pauses, which of course cannot be defined in this way, discourse pauses should be discoverable by looking for segments of no energy. Several problems, however, turn up in practice.

One problem is the fact that there are segments where there is no energy even within individual speech sounds, e g in stops such as p, t, k in such words as apa, ata, aka, or even longer in appa, atta, acka. The duration of the silent phase of the stops in the first words can be as much as 175 msec, the second 200 msec, when pronounced by a Swede. These pauses could of course be of interest and the pausemeter to be presented can easily measure the duration of pauses in such stops if set to do so. But such pauses should not be counted among the discourse pauses.

Another problem is the low energy of such sounds as f, th, h. A pause detection device designed to measure discourse pauses must be able to skip the pauses inherent to certain speech sounds such as stops, but still react to weak speech sounds. If the device finds pauses at stops and weak speech sounds within words it will register too many pauses. If it does not note the existence of weak sounds in the beginning or at the end of certain words the durations of utterances and pauses will not be correct.

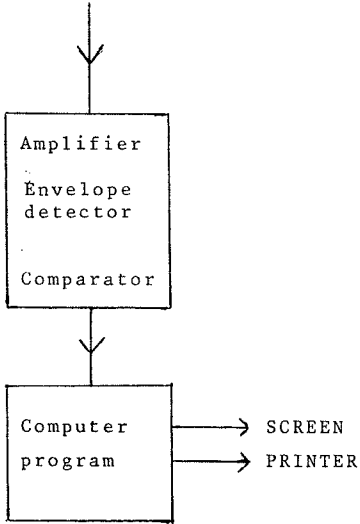
Rabiner & Schafer (1978) have discussed the difficulty of determining the beginning and end of words and state that they are difficult to locate if there are

- 1 Weak fricatives (f, th, h) at the beginning or end
- 2 Weak plosive bursts (p, t, k) at the beginning or end
- 3 Nasals at the end
- 4 Voiced fricatives that become devoiced at the end of words
- 5 Trailing off of vowel sounds at the end of an utterance

Rabiner & Schafer describe an algorithm that is intended to discover the beginning and end of utterances (speech segments) on the basis of two time-domain measurements - energy and zero-crossing rate. We have to rely on simpler methods in the pausemeter to be described. The general construction of the pausemeter is shown in fig 2.

TAPE
MICROPHONE

Function



Decides whether
pause or speech
segment is at hand

Measures changes
between pause and
speech.
Computes (and
displays) the
durations of
pauses and speech
segments.
Draws diagrams
and prints tables.

Fig 2 Block diagram showing the overall concept and function of the pausemeter

THE SPEECH-PAUSE DECISION UNIT.

The speech-pause decision unit distinguishes between speech and pause by comparing the signal envelope with a predefined threshold. The signal envelope is determined by an envelope detector and it is assumed that the envelope suitably approximates the short-time signal energy. The principle of the unit is shown in fig 3a. The speech signal is first amplified. Then the signal envelope is determined by the envelope detector where the time constant is variable. The output from the envelope detector is compared with a threshold in the comparator. The output from the comparator is either 0 or 1 and this signal is tested by the microcomputer every 20 ms.

Fig 3b shows the circuit diagram and component values. Even non-engineers should be able to build it. The diode at the data output limits the output voltage to positive values. A LED circuit is also used to indicate the result of the comparison and facilitate the use of the device.

The input signal to be processed is taken from a tape recorder either from a microphone or a tape. The tape recorder level meter is used for speech level adjustments. The time constant can be varied. A time constant of about 33 ms is found to be suitable in most cases. This value is also used in a similar device mentioned in the literature. However, it is not only the time constant but also the threshold value that affects the speech/pause decision. This level is not variable. Instead the input amplification can be varied and is adjusted so that the LED circuit indicates that the unit is working properly.

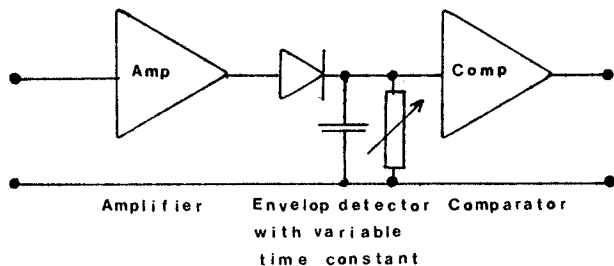


Fig 3a

PAUSEMETER CIRCUIT DIAGRAM

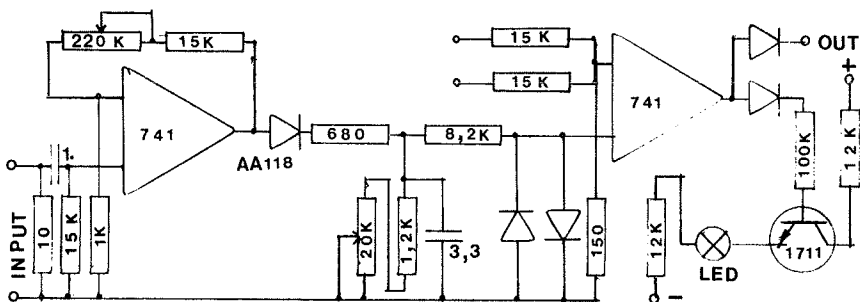


Fig 3b

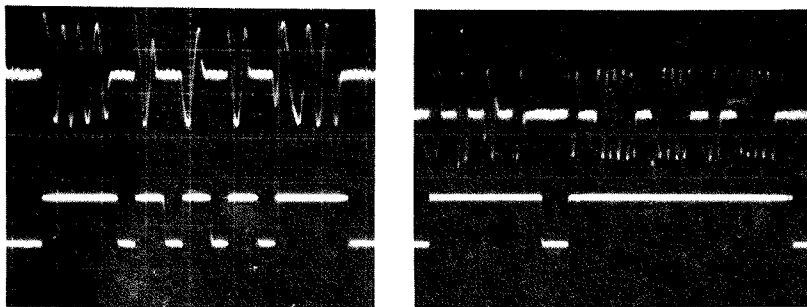


Fig 3c

Fig 3a Main features of speech-pause decision unit.

3b Circuit diagram of speech-pause decision unit.

3c Lower curves show the decisions of the speech-pause decision unit for two cases with small and large time constant respectively. The signal shown here is a Morse signal which has been used in calibration and testing.

THE REAL-TIME PAUSEMETER

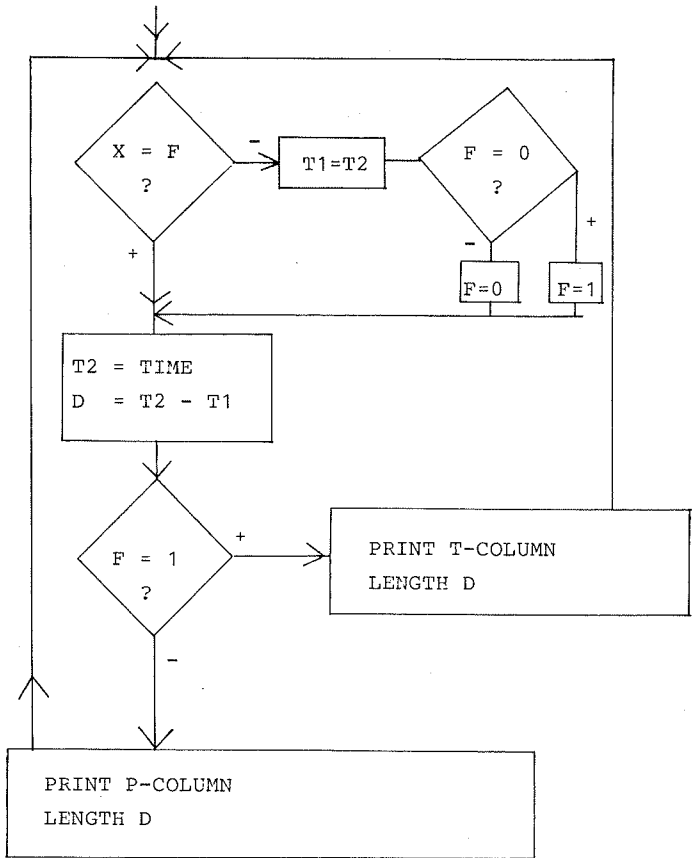
The overall design of the pausemeter is shown in fig 4. The real-time version is characterized by the following features (cf the print-out of the program at Appendix I).

The program starts by asking for the number corresponding to speech (LÅT TAL VARA PÅ OCH TRYCK PÅ J, "LET SPEECH BE ON AND PRESS J"). This number is then used as response to the question SKRIV IN DENNA SIFFRA VID ? ("WRITE THIS NUMBER AT ?")

The number may change with the equipment (e g printer) connected to the computer; that is why it is necessary to ask this question.

The computer is now ready for measuring as long as J is held down (TRYCK J FÖR MÄTNING, "PRESS J FOR MEASUREMENT"). The real-time pausemeter displays the duration of the successive speech and pause segments by building a vertical diagram on the screen where the columns grow to the right. The columns grow as the speaker speaks or makes a pause. When a pause or speech segment is finished its duration is given by the last figure and the computer starts measuring and displaying the growth of the next (pause or speech) segment. Speech segment columns are filled with T's while pause columns are filled with P's. The screen may be filled with 22 columns at a time and therefore a short utterance can often be shown in full on the screen. One may take a photo of the screen (no print-out facilities are included in this real-time version) to get a record of the pause structure of the utterance recorded.

The precision of the real-time version is limited and the results should be used with care. The real-time version is, however, very good for displaying typical pause-speech relations. One may e g talk spontaneously and record the pauses and speech segments immediately for presentation in class. Typically speech segments last 2-3 seconds while pauses last 0.5 second and the corresponding graphic pattern is easily seen. The real-time version of the pausemeter may



X: INPUT (0/1)
 T: SPEECH (1)
 P: PAUSE (0)
 F: TEST VARIABLE

Fig 4 Flow chart of real-time pausemeter

be used to show how long one is able to speak without pausing. It may be connected to a radio to show differences in pause behaviour between speakers, languages etc. It is important to remember that this version does not store any values and the speech must be recorded if it is to be examined further.

The function of the real-time pausemeter is very simple. The input number appears in INP(58%). If this number equals the number in the test variable F% the system assumes that speech is going on, the time is read and printed at the screen at the end of the column of T's (line 55). If the input signal is not the number for speech the system reads the clock and starts building the pause column (line 60).

The program includes a scale factor (0.05, lines 55,60) which can be changed if the columns of the diagram are required to grow at a different speed. The time is read in PEEK(65008%) and is assumed to be long enough for this version of the meter.

THE PRINTING PAUSEMETER

The flow chart at fig 5 shows the main features of the pausemeter. A print-out of the program is found in Appendix II. The computer starts by asking how long pauses (or speech segments) should be skipped (HUR LÅNGA PAUSER OCH TALSEGMENT SKALL ÖVERHOPPAS?) If 1 is the response segments shorter than 20 msec are skipped; if 2, segments shorter than 40 msec etc (line 15). The system then asks for the numbers for speech (tal) and for pause which vary according to the equipment attached and therefore need to be specified by the operator (lines 17-19).

The system then asks whether it is to start measuring (SKALL JAG BÖRJA MÄTA J/N). If J is pressed it starts and goes on until CTRL-A is not pressed. (It is thus not necessary to hold J down for the whole measuring period, as is the case with the real-time version.)

In line 110 the counter for time values (N) is set to 0 and the test variable F is set to the pause number previously entered by the operator, to make the meter start by measuring the change from the pause value (= the beginning of a speech segment). In the flow chart fig 5 the test variable has been given the values 0 and 1.

Line 120 sets the clock by inserting the values 255 in both cells from which the count down starts. The measuring loop is 130-170. Line 130 checks whether there is a change in the input signal, as this is what the system looks for. If so the clock is read by lines 143-147. A certain correction is done in lines 143, 144, due to the design of the computer clock. Line 148 checks whether enough time has passed since the last measurement was taken. If this is the case the time reading is stored by line 150, if not the system returns to line 130. Line 155 prints a T (tal, speech, talk) or a P (pause) preceded by a number on the screen for the information of the operator. Line 160 inverts the value of the test variable. Line 170 checks whether more than 1000 values have been stored, which is the maximum of the present version. If not, the process continues. If more than 1000 time values have been stored or the process has been stopped by pressing CTRL-A line 180 prints the stored values on the screen. Line 190 asks whether a print-out is required (ÖNSKAS UTSKRIFT). If a print-out is required this is handled by 200-220.

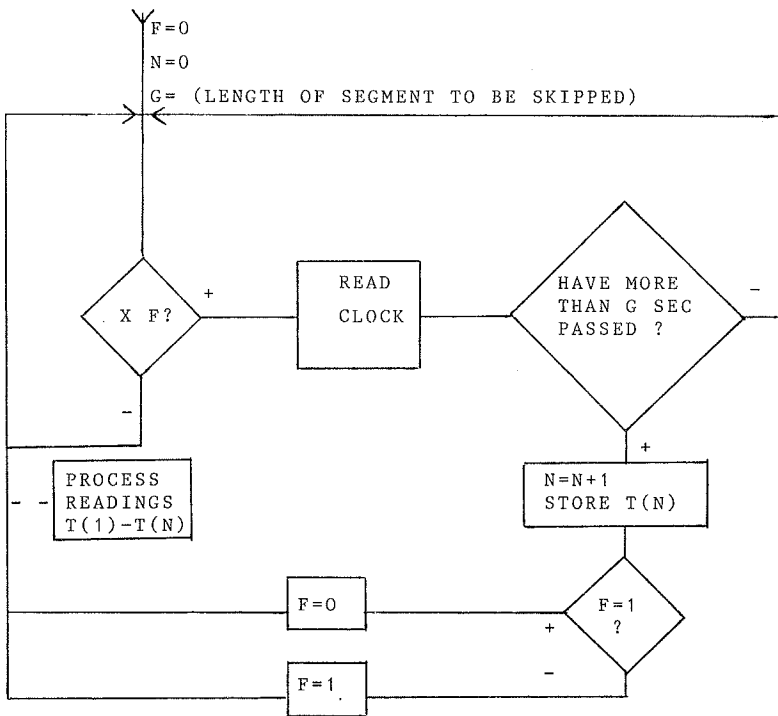


Fig 5 Flow chart of pausemeter. X is an input variable varying between 1 (speech) and 0 (pause) and given by the comparator (pause detector). F is a help variable varying between 1 and 0 for comparison. G is a variable set by the operator and used to avoid measuring too short pauses and speech segments. The system notes changes in the input signal and reads the time whenever a change is noted. The time readings (T(1) to T(N)) are stored to be processed later.

Line 225 asks whether a reduction of the values is required. If 0 is pressed no reduction is made, if 1 or higher number is pressed a reduction of 20, 40, 60 etc msec is made. This reduction affects the end of all subsequent calculations of speech segments as can be seen from line 227 where every other value is reduced by the value given. A print-out is also offered. Note that the original time values cannot be retrieved once they have been changed.

The next parts of the program give various tables and diagrams, calculated on the basis of the time values obtained. Lines 229-280 calculate the duration of the successive time segments, the total duration of speech segments, the average duration and the standard deviation. If required a print-out is made. Lines 340-470 offer the same information for pauses. Lines 480-570 offer a table of the duration of successive pause and speech segments. Lines 600-660 offer the same information in the shape of a histogram (STOLPDIAGRAM).

Lines 700-845 offer a diagram showing the difference between the duration of a speech segment and the preceding pause. The values are plotted as a star that moves around 0. It moves to the right (is positive) if the speech segment is longer than the preceding pause and to the left if the pause is longer than the following speech segment. Normally people talk more than they pause and the stars are consequently generally found to the right of the zero line. This diagram is offered as an example of the kind of diagram a program can be designed to present. The diagram was devised to study the relation between the duration of pauses and speech segments. For example, Beattie (1979) has hypothesized that alternating cycles of fluent and hesitant speech would be reflected in pause durations being shorter or longer than speech durations. This would appear in our plot as alternating concentrations to the right and left of zero respectively.

The print-out for version (2) is illustrated at Appendix III.

CONCLUSION

Automatic pausemeters will facilitate the investigation of overall pause behaviour in speech, but for detailed information about the speech signal spectrograms are still required. The features of pausemeter programs can easily be modified to meet special requirements (custom made). The versions presented have been found helpful in certain kinds of investigations. The real-time version is a convenient pedagogical tool for presenting typical or individual pause-speech patterns in class or for explorative purposes. The printing version has been used by Kärina Vamling for detailed studies of pause behaviour during simultaneous interpreting. Such studies could hardly be undertaken without the aid of an automatic pausemeter. The meters presented will be used in experiments in psycholinguistics and phonetics and it is assumed that new and better versions will be developed.

APPENDIX I

Real-time pausemeter

```
1 ; "REALTIME PAUSMETER SEPT 1982"  
5 ; "LET SPEECH BE ON AND PRESS J"  
10 GET VX ; ; INP(58)"=TAL"  
12 ; "WRITE THIS NUMBER AFTER ?" ; INPUT TX  
15 ; "LET SILENCE BE ON AND PRESS J" ; GET VX ; ; INP(58)"=PAUSE" ; INPUT PX  
20 PDKF 65008%,255% ; POKF 65009%,255% ; FX=PX ; T1X=0X ; T2X=0X ; NX=0X  
25 ; CHR(12) ; ; "PRESS J FOR MEASUREMENT"  
30 GET XX  
35 IF XX("J" GOTD 20  
40 IF INP(58)=FX THEN 45  
41 T1X=T2X ; IF FX=TX THEN FX=PX ELSE FX=TX  
42 IF NX(22X THEN NX=NX+1X ELSE ; CHR(12) ; NX=0X  
45 AX=PEEK(65009%) ; AX=PEEK(65008%)-1X ; IF AX=-1X THEN AX=255X  
50 T2X=256X*(255X-AX)+255X-AX ; DX=T2X-T1X  
55 IF FX=TX THEN ; CUR(NX,0X);STRINGX(DX*.1,84).02*DX ; GOTO 30  
60 ; CUR(NX,0X);STRINGX(DX*.1,80).02*DX ; GOTO 30
```



```

PAUS"KX": "P1(KX)" SEK" : NEXT UX
380 FOR ZX=1X TO KX : H=H+P1(ZX) : NEXT ZX : ; "TOTAL PAUSLANGD:"H" SEK" : B=H/K
X : ; "MEDELLANGD:"B" SEK"
390 FOR ZX=1X TO KX : C=C+(P1(ZX)-B)ØZX : NEXT ZX
400 W=SQR(C)/KX : ; "STANDARDAVVIKELSE:"W" SEK"
410 : "ØNSKAS UTSKRIFT?J/N" : INPUT ØX
415 IF ØX()="J" THEN 480
420 OPEN "PR:" ASFILE 1
430 FOR ZX=1X TO KX : ; *1,"PAUS"ZX": "P1(ZX)" SEK" : NEXT ZX
440 : *1,"TOTAL PAUSLANGD:"H" SEK"
450 : *1,"MEDELLANGD:"B" SEK"
460 : *1,"STANDARDAVVIKELSE:"W" SEK"
470 CLOSE 1
480 : "ØNSKAS PAUS-TALTABELL?J/N" : INPUT ØX
490 IF ØX()="J" THEN GOTO 600
500 FOR ZX=1X TO KX+1X
505 : "TAL "ZX": "T1(ZX)" SEK"
510 IF ZX()KX+1X THEN : "PAUS"ZX": "P1(ZX)" SEK"
515 NEXT ZX
545 : "ØNSKAS UTSKRIFT?J/N" : INPUT ØX
547 IF ØX()="J" THEN GOTO 600
548 OPEN "PR:" ASFILE 1
550 FOR ZX=1 TO KX+1
560 : *1,"TAL "ZX": "T1(ZX)" SEK"
565 IF ZX()KX+1X THEN : *1,"PAUS"ZX": "P1(ZX)" SEK"
568 NEXT ZX
570 CLOSE 1
600 : "ØNSKAS STOLPDIAGRAM?J/N" : INPUT ØX
610 IF ØX()="J" THEN 700
612 : SPACEX(4)"1"SPACEX(4)"2"SPACEX(4)"3"SPACEX(4)"4"SPACEX(4)"5"SPACEX(4)"6"SP
ACEX(4)"7"SPACEX(4)"8 SEK"
615 FOR ZX=1X TO KX+1X
620 : STRINGX(T1(ZX)*5X,84)ZX
625 IF ZX()KX+1X THEN : STRINGX(P1(ZX)*5X,80)ZX
630 NEXT ZX
635 : "ØNSKAS UTSKRIFT?J/N" : INPUT ØX
640 IF ØX()="J" THEN 700
643 OPEN "PR:" ASFILE 1
645 : *1,SPACEX(4)"1"SPACEX(4)"2"SPACEX(4)"3"SPACEX(4)"4"SPACEX(4)"5"SPACEX(4)"6
"SPACEX(4)"7"SPACEX(4)"8 SEK"
647 FOR ZX=1X TO KX+1X
650 : *1,STRINGX(T1(ZX)*5X,84)ZX
655 IF ZX()KX+1X THEN : *1,STRINGX(P1(ZX)*5X,80)ZX
660 NEXT ZX : CLOSE 1
700 : "ØNSKAS TAL-PAUSDIFERENS DIAGRAM?J/N" : INPUT ØX
710 IF ØX()="J" THEN 800
715 : SPACEX(11)"5"SPACEX(6)"-0+"SPACEX(6)"5"
720 FOR ZX=2X TO KX+1X
730 A1X=2X*T1(ZX)-2X*P1(ZX-1X)+19X
740 : SPACEX(A1X)"*T"ZX"-P"ZX-1X
750 NEXT ZX
800 : "ØNSKAS UTSKRIFT?" : INPUT ØX
810 IF ØX()="J" THEN 870
815 OPEN "PR:" ASFILE 1
820 : *1,SPACEX(11)"5"SPACEX(6)"-0+"SPACEX(6)"5"
825 FOR ZX=2 TO KX+1X
827 A1X=2X*T1(ZX)-2X*P1(ZX-1X)+19X
830 : *1,SPACEX(A1X)"*T"ZX"-P"ZX-1X
840 NEXT ZX
845 CLOSE 1
870 STOP

```

APPENDIX III

the measurement of short sequence with 5 pauses

T 1: 24
 T 2: 112
 T 3: 137
 T 4: 220
 T 5: 223
 T 6: 280
 T 7: 289
 T 8: 376
 T 9: 430
 T 10: 513
 T 11: 541
 T 12: 621
 T 13: 625

TALSEGMENT 1: 1.76 SEK
 TALSEGMENT 2: 1.66 SEK
 TALSEGMENT 3: 1.14 SEK
 TALSEGMENT 4: 1.74 SEK
 TALSEGMENT 5: 1.66 SEK
 TALSEGMENT 6: 1.6 SEK
 TOTAL TALLANGD: 9.56 SEK
 MEDELLANGD: 1.36571 SEK
 STANDARDAVVIKELSEN: .223139 SEK

PAUS 1: .5 SEK
 PAUS 2: .06 SEK
 PAUS 3: .10 SEK
 PAUS 4: 1.08 SEK
 PAUS 5: .56 SEK
 TOTAL PAUSLANGD: 2.38 SEK
 MEDELLANGD: .476 SEK
 STANDARDAVVIKELSE: .159138 SEK

TAL 1: 1.76 SEK
 PAUS 1: .5 SEK
 TAL 2: 1.66 SEK
 PAUS 2: .06 SEK
 TAL 3: 1.14 SEK
 PAUS 3: .18 SEK
 TAL 4: 1.74 SEK
 PAUS 4: 1.08 SEK
 TAL 5: 1.66 SEK
 PAUS 5: .56 SEK
 TAL 6: 1.6 SEK

1 2 3 4 5 6 7 8 SEK

TTTTTTT 1
 PP 1
 TTTTTTT 2
 2
 TTTT 3
 3
 TTTTTTT 4
 PPPPP 4
 TTTTTTT 5
 PP 5
 TTTTTTT 6

5 -0+ 5
 *T 2-P 1
 *T 3-P 2
 *T 4-P 3
 *T 5-P 4
 *T 6-P 5

REFERENCES

- Barik, H (1973), "Simultaneous Interpretation: Temporal and Quantitative Data". *Language and Speech* 16.
- Beattie, G W (1979), Planning units in spontaneous speech: some evidence from hesitation in speech and speaker gaze direction, in conversation. *Linguistics* 17-1/2, 61-78.
- Boomer, D S (1965), Hesitation and grammatical encoding. *Language and Speech* 8.
- Butterworth, B (ed) (1980), *Sentence production*. New York: Academic Press.
- Cook, M (1969), Anxiety, speech disturbance and speech rate. *British Journal, social and clinical psychology* 8.
- Einarsson, J (1978), *Talad och skriven svenska*. Lund: Studentlitteratur.
- Goldman-Eisler, F (1958), The predictability of words in context and the length of pauses in speech. *Language and speech* 1.
- Goldman-Eisler, F (1972), Segmentation of input in simultaneous translations. *Journal of Psycholinguistic Research* 1:2, pp 127-140.
- Gårding, E (1967), Prosodiska drag i spontant och uppläst tal. In: G Holm (ed), *Svenskt talspråk*. Uppsala (1967).
- Kempen, G & Hoenkamp, E (1982), Incremental sentence generation. *Proc Ninth Intern conf comp linguistics*. Prague.
- Lounsbury, F G (1954), Transitional probability, linguistic structure and systems of habit-family hierarchies. *Psycholinguistics*. Baltimore.
- Rabiner, L R & Schafer, R W (1978), *Digital processing of speech signals*. Englewood Cliffs: Prentice-Hall.
- Rochester, S R (1973), The significance of pauses in spontaneous speech. *J Psychling research* 2.
- Vamling, Karina (1982), Experiment med simultantolkning. In: Kulick, D, Henningsson, L-Å, Vamling, K. *Om tolkning. Praktisk lingvistik* 7. Lund. Department of linguistics.