DOCUMENT RESUME

ED 084 906		FL 004 392

AUTHOR	Gerver, D.
TITLE	Simultaneous and Consecutive Interpretation and Human Information Processing.
INSTITUTION	Durham Univ. (England).
SPONS AGENCY	Social Science Research Council, London (England).
PUB DATE	Dec 71
NOTE	122p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Acoustical Environment; Auditory Perception; Experiments; Human Services; *Information Processing; *Interpreters; Listening Comprehension; Psychological Patterns; Reaction Time; Statistical Analysis; Tables (Data); Time Perspective; *Translation; Verbal Stimuli

ABSTRACT

The aim of this study is to investigate the effects of noise on the performance of simultaneous conference interpreters, and to carry out a detailed examination of verbal and temporal aspects of their output in relation to source language input. A further aim is to compare the relative effectiveness of simultaneous and consecutive interpretation in conveying information to the listener under both normal and adverse listening conditions. A bibliography of references is included. (Author/DD)

Social Science Research Council

Grant No. HR566/1

Simultaneous and consecutive interpretation and human information processing

Dr. D. Gerver
Department of Psychology
University of Durham

December, 1971

Table of Contents

	Page
Abstract	i
Acknowledgements	viii
Noise and simultaneous interpretation	1
The effects of noise on the temporal characteristics of the performance of simultaneous interpreters	26
Level of processing and comprehension of prose	50
A comparison between simultaneous and consecutive interpretation	58
Discussion and critique	71
Appendix I	79
Appendix II	84
Bibliography	97

ABSTRACT

i) University: University of Durham

ii) Project title: Simultaneous and consecutive interpretation and human information processing

iii) Period covered by this report: September, 1968 - July, 1971

iv) Director of project: Dr. D. Gerver

v) Research staff: Miss Carol Sherrard (Research assistant)

vi) Research aims:

The principal aim of the research was to investigate the effects of noise on the performance of simultaneous conference interpreters, and to carry out a detailed examination of verbal and temporal aspects of their output in relation to source language input. It was also hoped that the results of such an analysis would contribute towards an understanding of the complex information processing involved in simultaneous interpretation.

A further aim of the research was to compare the relative effectiveness of simultaneous and consecutive interpretation in conveying information to the listener in both normal and adverse listening conditions.

(vii) Research methods:

a) Noise and simultaneous interpretation.

12 simultaneous interpreters were asked to shadow or interpret simultaneously into English pre-recorded French passages presented at 3 signal-white noise ratios. Subjects' output was analysed for errors, omissions, corrections, ear-voice span, and various temporal characteristics involving speech and pause times. Both pen recordings and on-line computer analyses were employed for the latter measurements. Two independent judges also rated subjects' interpretations on scales of intelligibility and informativeness and the Maudsley Personality Inventory was administered to all subjects.

b) Simultaneous interpretation and recall of material interpreted.

9 trainee simultaneous interpreters were asked to listen to, shadow, or simultaneously interpret into English pre-recorded passages of French prose. After each passage subjects were asked to answer 10 questions pertaining to the text.

c) Simultaneous and consecutive interpretation.

Groups of undergraduates with little or no knowledge of French were placed in a conference-like situation and asked to listen to passages of French prose together with either simultaneous or consecutive interpretation which they received through headsets. After each passage subjects were asked to reply to 10 questions relating to the passage. The comparison was made in both quiet and noisy listening conditions.

(viii) Research summary:

a) The effect of noise on the performance of simultaneous interpreters was studied in an experiment in which 12 professional conference interpreters shadowed or simultaneously interpreted into English 6 pre-recorded passages of French prose at 3 signal-white noise ratios: No noise, +5 db, -2db. The principal findings were that, though both shadowers and interpreters omitted more as noise increased, interpreters made more errors than shadowers at each signal-noise ratio. Furthermore, the ratio of interpreters' self-corrections to errors remained almost the same at the 3 signal-noise ratios in spite of the increase in errors. It was also found that, though interpreters had greater ear-voice spans than shadowers, the ear-voice spans for both tasks remained fairly constant under the 3 experimental conditions. It was suggested, therefore, that in order to maintain a constant ear-voice span under noisy listening conditions, simultaneous interpreters were prepared to sacrifice accuracy by lowering their response criteria.

Analysis of the temporal characteristics of subjects' performance demonstrated a retarding effect of noise on simultaneous interpretation as compared with shadowing. This was shown in decreased output rates and in increased pause times. It was also found that the proportion of input pause time utilised by shadowers and interpreters, as well as the proportions of total input-output time spent in simultaneous listening and speaking, did not differ significantly either within or between experimental conditions and tasks.

A Yerkes-Dodson type effect was found in the relationship between interpreters' performance (words correct) at each signal-noise ratio and scores on the M.P.I. Neuroticism scale; no correlation being found in the No noise condition, a positive (but not quite significant) correlation at +5 db, and a significant negative correlation at -2db. The range of subjects' Neurooticism test scores was within the limits for the normal group in the test standardization samples. In summary, the higher the neuroticism score, the better the performance under slight stress, but the greater the decrement in performance under greater stress. No relationship was found between Neuroticism scores and the temporal measures of output mentioned above.

b) Simultaneous interpreters' retention of material translated was examined in an experiment in which 9 trainee simultaneous interpreters were asked to listen to, shadow and simultaneously interpret into English 3 pre-recorded French prose passages. After each passage Ss were asked 10 questions relating to the passage. Significantly higher comprehension scores were obtained in the listening condition than in simultaneous interpretation, which in turn yielded significantly higher scores than shadowing. These results are discussed in terms of the level of processing required for each task, and its subsequent effect on channel capacity.

c) The relative effectiveness of simultaneous and consecutive interpretations in conveying information to the conference listener was examined in two experiments. In the first experiment 15 undergraduate subjects with minimal knowledge of French listened under simulated conference conditions to 3 French prose passages together with simultaneous interpretations of the passages, and 15 similar subjects listened to the same 3 passages followed by a consecutive interpretation of each passage. After each passage subjects were asked 10 questions relating to the passage. No differences were found in retention scores for the simultaneous and consecutive conditions.

In order to examine the effect of adverse listening conditions upon an audience's ability to retain information presented in either simultaneous or consecutive interpretation, the first experiment was repeated with 4 groups of subjects and two types of adverse listening conditions. In one condition both source language text and the interpretation thereof were heard against a background of white noise in the lecture theatre. In the second condition subjects heard source language texts and interpretations against a background of noise in their headsets. Once again no significant difference was found between conditions, though higher average comprehension scores were obtained for consecutively than for simultaneously interpreted passages.

Though it was not possible to compare the results of the two experiments by means of any statistical test, a comparison of the two sets of results appears to show that, as might be expected, noise in the headphones caused a slightly greater decrease in comprehension scores for simultaneously than for consecutively interpreted passages.

d) Finally, the report includes a complete description of the computer system and programme developed for the on-line analysis of temporal characteristics of twin-track tape-recordings of speech.

(ix) Papers and publications:

a) Conference papers

"How many things can we attend to at once? The problem of simultaneous interpretation." Presented to the British Association at Durham, September, 1970.

"Noise, attention and simultaneous interpretation." Presented to the British Psychological Society at Exeter, April, 1971.

"Noise and the simultaneity of simultaneous interpretation." Presented to the Experimental Psychology Society, London, January, 1972.

b) Publications - submitted and in preparation

Level of processing and comprehension of prose. (Submitted for publication).

Noise, attention and simultaneous interpretation. (In preparation).

Noise and the simultaneity of simultaneous interpretation. (In preparation).

ASPA - Automatic speech-pause analyzer. A suite of computer programmes for monitoring patterns of speaking and pausing from twin channel tape recorder output.

Submitted for publication.

ACKNOWLEDGEMENTS

The author wishes to thank the Social Science Research Council for their generous support of this project.

Thanks are also due to Mr. Newmark of The Modern Languages Department of Central London Poly-technic (Holborn College), and to Mr. T. Pilley of the Linguists Club, London, for practical assistance in carrying out the studies on simultaneous interpretation. Mrs. Patricia Longley of Central London Poly-technic, through her continued interest, encouragement, and practical help has been of particular assistance in all of the author's research with simultaneous interpretation.

Finally, thanks are due to the subjects themselves who gave of their valuable time to take part in the experiments, and to Miss Carrol Sherrard who was responsible for obtaining and analysing a large part of the data.

I

Noise and simultaneous interpretation

Introduction

There are two types of translation services provided at international conferences or at small gatherings where the participants may not speak one another's languages. At larger conferences one is more likely to find simultaneous translation where the interpreter translates out loud as he listens to the source language message. Since, as Miller (1963) has pointed out, it is an almost universal characteristic of human verbal behaviour that a person will rarely both speak and listen at the same time, the very fact that simultaneous interpreters appear to be able to do just this for up to 15-20 minutes at a time justifies a psychological study of the performance of simultaneous interpreters.

The other form of conference interpretation is consecutive interpretation in which the interpreter first listens to the source language speaker, then delivers a translation when the source speaker has finished. Each type of interpretation can be expected to impose different loads on the cognitive capacities of both the interpreter and his audience. The studies to be discussed in this report represent attempts to study the processes involved in simultaneous interpretation, particularly when carried out under poor listening conditions, and to assess the relative effectiveness of these two forms of interpretation in conveying information to the listener.

Noise and simultaneous interpretation

By any standards the performance of skilled simultaneous interpreters is a remarkable feat. They appear to be able to monitor, store, retrieve, and translate source language input whilst at the same time uttering their translations in a target language. What is more, since they correct themselves as they go along, often without interrupting the flow of speech, they must also be monitoring some form of feedback of their translation, and comparing it with what they can recall of the original message.

Continuous skilled tasks involving attention switching and/or data transformation are particularly likely to be susceptible to interference from environmental noise, as has been shown by Broadbent (1958), Jerison and Wing (1957), Woodhead (1958) and others. It is not surprising, therefore, to find that simultaneous interpreters are particularly sensitive to environmental noise, and will often refuse to work in conditions which do not appear to be particularly stressful to the observer. There would appear to be two possibly interlinked explanations for any detrimental effects that noise may have on interpreters' performance:

. . .

The first is suggested by some experiments by Rabbitt (1966). In the first of these Rabbitt presented Ss with tape recorded word lists either with or without pulse modulated white noise. Ss repeated the words aloud as they heard them, and at the end of each list they were tested for recognition of these words as part of a second list. Rabbitt found that noise levels which did not affect intelligibility reduced both the efficiency of recognition, and Ss' confidence in their recognition judgments. Rabbitt (1968) went on to demonstrate that, not only recognition, but also recall was poorer when test material was presented at noise levels which did not affect intelligibility. Finally, and perhaps more pertinent still to the problem of noise and simultaneous interpretation, Rabbitt (1968) carried out an experiment involving the "understanding" of connected discourse rather than recall of word lists. He found that the introduction of noise in the second half of auditorily presented prose passages interfered with recall of material from the first half of the passage heard without noise. Rabbitt interpreted his results as "demonstrations that increased difficulty of recognition of speech through noise may interfere with other activities (conveniently called rehearsal) which may be necessary to efficiently retain data in memory".

The implication of these studies for the effects of noise on simultaneous interpretation would be, then, that it is not so much perception of the input message which may be affected by moderate noise levels, but rather the processes involved in recall and transformation (translation) of the source language message. The second explanation of the effects of noise on simultaneous interpreters' performance is in terms of individual differences in arousal level. The relationship between arousal level and performance is usually assumed to take the form of an inverted U, insofar as performance improves up to an optimal point (depending on the task) as arousal level increases, but if arousal level increases even further performance begins to fall off. This use of the term "arousal level" is often defined in terms of a state of general arousal, and anxiety in particular. Broadbent (1957) suggested that highly anxious individuals would have a low tolerance for noise, and Stennet (1957), for instance, using an auditory tracking task, palmar conductance, EMG, and monetary and shock incentive conditions found an inverted U relationship between performance level and level of arousal. The test used to investigate this factor in the present study was the Eysenck Personality Inventory. The E.P.I. is an easy to administer scale, and the Neuroticism scale is reported by Jensen (1965) to correlate .76 with the Taylor Manifest Anxiety Scale - itself devised as a measure of level of emotional response, and therefore of drive. Jensen (1962) also reports that Ss with high N scores stood up less well in a laborious serial learning task.

Finally ear-voice spans in shadowing and simultaneous interpretation were examined. Treisman (1965), using bilinguals rather than trained interpreters, found a greater ear-voice span for interpreting than shadowing. This, together with Pollack and Rubenstein's (1963) finding that it takes longer to respond to words in noise, suggested that any increase in the time taken to respond to words in noise might be reflected in corresponding increases in ear-voice spans, particularly in interpretation where S has additional operations to perform on the input.

Method

Materials

6 French passages of approximately 500 words each were recorded on tape by a male native French speaker at a rate of about 120 words per minute. The passages were taken from recent issues of the UNESCO Courier, and were examples of the type of subject matter frequently encountered by interpreters at conferences.

Subjects

The subjects were 12 professional conference interpreters, with many years experience of working simultaneously from French to English.

Procedure

A 3x3 Greco-Latin square design was repeated four times: all subjects shadowed three passages, and interpreted three passages. Each passage was either shadowed or interpreted at one of three signal to white noise ratios: no noise, +5db, -2db. Six subjects translated before shadowing, and 6 subjects shadowed before translating. The passages shadowed by the first group were translated by the second group, and vice-versa. Separate stimulus tapes were prepared for each subject according to the Greco-Latin square design.

The stimulus tapes were played to Ss individually over headphones from a Uher Universal Teaching tape recorder, and their responses were recorded on the bottom tracks of their stimulus tapes via a boom microphone attached to the headset. After preliminary trials with interpreters to determine the approximate level at which they preferred to listen to the passage, the approximate mean sound pressure level was read from a Dawe Sound Level Meter set against the headphones in a soundproof room. This level was approx. 80 db re 0.0002 dyne cm^2. The output from the white noise generator, the band width of which was tailored to the speech spectrum, was then fed via a step attenuator (set at 10 db) to the headphones, and the gain increased until the sound level meter read 80 db. This established 0 db signal-noise ratio. The tape recorder signal was then fed to a sound mixing unit, while the signal from the noise generator was fed to the mixing unit via the step attenuator. In this way the appropriate signal-noise ratio could be adjusted before the mixed signal was finally fed to Ss' headphones.

The signal-noise ratios +5db and -2db were selected after pre-testing with a number of interpreters, who were asked whether they might encounter various noise levels in the actual interpretation situation, and if so whether they would find them tolerable. It should be noted, of course, that there is rarely as much noise in the conference amplication system as in this study; the most likely source of annoyance during conferences being the random scrapings and bangings due to audience movement during speeches, as well as coughs, sneezes, etc. At +5db the interpreters said that they could manage to work, but would not like to do so, while at -2db they reported that they would find it very difficult to work, and would soon refuse to carry on. All of the Ss actually employed in the experiment itself reported that they could not work at -2db, and would refuse to begin if they actually encountered such poor listening conditions at a conference.

. . .

Treatment of Results

1. Ear-voice span

For both shadowers and interpreters ear-voice span in words was calculated at every 5th word of the input in terms of the number of words not yet correctly repeated or interpreted by the subject. Words omitted entirely in shadowing or translation were counted as part of the ear-voice span until the subject had passed beyond the point at which they could have been shadowed or meaningfully interpreted (i.e. in the correct context). Provided that some reasonable connection could be inferred between the interpreter's output and the original message, an error in translation was counted as a correct translation and not as part of ear-voice span. Paraphrase was taken into account, and there were specific rules relating to the number of words which could be meaningfully translated into English. For instance, "ne...pas" was counted as one word from the "ne", and articles in the original were not counted when they would not normally have been translated.

2. Number of words correct

In assessing the correctness of interpretations a reasonable paraphrase was accepted as correct, since a word-for-word translation was not expected, and indeed would not have been a good translation from the interpreter's point of view.

3. Evaluation of translations

Two independent judges evaluated each translation on two scales used by J.B. Carroll (1965) in order to compare human with machine translations. Both scales were 9 point scales, and the first "Intelligibility" scale was employed to assess the degree to which the interpreter's translation of a passage sounded like normal well-thought-out prose, and would be understandable in the same way as if it had been originally spoken in the target language. Though the translated version of a passage may sound (or read) well, it may bear little relationship to the original in terms of the meaning conveyed. Carroll approached the question of the fidelity of a translation by converting it into the complementary question of whether the original could be found to contain no information that would supplement or controvert information already conveyed by the translation. The scale of "Informativeness", then, pertains to how informative the original passage is perceived to be <u>after</u> the translation has been seen and studied. If the translation already conveys a great deal of information, the original may be said to be low in informativeness relative to the translation being evaluated. If, however, the translation conveys a limited amount of information, it may be that the original conveys a great deal more, in which case the original is high in informativeness relative to the translation being evaluated. Using 36 raters, Carroll found that the scales reliably differentiated between translations employed in his study.

The two raters employed in the present study were both graduate students in the Department of Education at Durham University, with some experience in marking translations from French up to A level standard. The translations were transcribed, and in order to prepare the translations for evaluation all hesitations and false starts were omitted, and the resulting texts were given to the two independent judges. The judges were not informed of the experiment, and were simply asked to evaluate the texts by assigning them to points on the separate 9 point scales of Intelligibility and Informativeness developed by Carroll (see Appendix I).

Results

i) Ear-voice span

The results of the ear-voice span measurement at each signal-noise ratio are shown in Table 1:

	No noise	+5db	-2db	$\bar{X}$
SHADOWING	1.55	1.87	2.58	2.00
INTERPRETING	5.29	5.60	6.09	5.66
$\bar{X}$	3.42	3.73	4.33	

Table 1: Mean Ear-voice spans (words).

The analysis of variance showed that there was a significantly greater ear-voice span for interpreters than for shadowers ($F = 121.72$; $df = 1,10$; $p < .01$). There were no significant main or interaction effects for signal-noise ratio.

ii) Words correctly shadowed or interpreted

The number of words correctly shadowed or interpreted by each subject was transformed into a proportion of the number of words in the passage concerned, and then into arc-sin form for the analysis of variance. These results are shown in Table 2.

	No noise	+5db	-2db	$\bar{X}$
SHADOWING	0.97	0.91	0.83	0.90
INTERPRETING	0.86	0.77	0.73	0.78
$\bar{X}$	0.91	0.84	0.78	

Table 2: Mean proportions of words correct.

Significantly more words were correctly shadowed than interpreted ($F = 36.66$; $df = 1,10$; $p < .001$). Signal-noise ratio had a significant effect on proportions of words correctly shadowed or interpreted ($F = 14.28$; $df = 2,20$; $p < .001$). There was no significant interaction between task and signal-noise ratio.

iii) <u>Words omitted, errors of commission, corrections</u>

Subjects' protocols were further examined for omissions, errors of commission, and corrections.

a) Omissions

Mean numbers of words omitted in shadowing and interpreting are shown in Table 3 below. Significantly more words were omitted in interpretation than in shadowing ($F = 21.63$; $df = 1,10$; $p < .001$), and there was a significant effect of signal-noise ratio ($F = 8.50$; $df = 2,20$; $p < .01$). There was no significant interaction effect.

	No noise	+5db	-2db	$\bar{X}$
SHADOWING	3.75	19.67	36.67	20.03
INTERPRETING	27.42	30.08	51.83	36.44
$\bar{X}$	15.58	24.88	44.25	

Table 3: Mean number of words omitted.

b) Errors

Table 4 below shows the mean number of errors committed (i.e. words wrongly translated or shadowed). Significantly more errors were made in interpreting than in shadowing ($F = 11.03$; $df = 1,10$; $p < .01$). Signal-noise ratio had a significant effect on error score ($F = 25.65$; $df = 2,20$; $p < .001$), and the significant interaction between signal-noise ratio and task verifies the impression gained from inspection of the table that the lowest S/N had a greater effect on interpreters than shadowers.

	No noise	+5db	-2db	$\bar{X}$
SHADOWING	5.75	16.58	25.08	15.81
INTERPRETING	14.33	25.92	46.33	28.66
$\bar{X}$	10.04	21.25	35.71	

Table 4: Mean number of errors.

c) Corrections

The mean number of corrections, or revisions, of words in interpreting are shown in Table 5. The number of corrections made in shadowing was so small as to make any comparison with interpreting unnecessary. Even though the mean number of corrections did not fall as S/N decreased, when the means are viewed as proportions of the errors made (see Table 5 below) it can be seen that corrections fell proportionately as errors increased at lower S/Ns.

	No noise	+5db	-2db
INTERPRETING	6.5	7.35	7.35

Table 5: Mean number of corrections.

iv) E.P.I. neuroticism and extraversion scores

a) Neuroticism

The mean neuroticism score for the group of simultaneous interpreters employed in this study ($\bar{X}$ = 10.5; S.D. = 6.20) places these Ss within the range of Eysenck's standardization sample from the normal population. Rank correlations (Spearman's rho) were calculated between neuroticism scores and the measures of performance: ear-voice span, words correct, etc. Except for the relationships between neuroticism and words correctly shadowed or interpreted at different signal-noise ratios shown below, all correlations proved to be insignificant (i.e. close to zero). As can be seen from Table 6 below, there is a slight negative correlation between neuroticism and shadowing performance at each S/N, whereas the relationship between neuroticism and interpreters' performance is more complex. At +5db, rhos for both shadowing and interpreting miss significance at the .05 level ($r_s .05 = 0.506$), while rho between neuroticism and interpreters' performance at -2db just misses significance at the .01 level ($r_s .01 = 0.712$).

	No noise	+5db	-2db
SHADOWING	-0.1861	-0.4380	-0.2702
INTERPRETING	-0.0262	+0.4091	-0.6940

Table 6: Coefficients of rank correlation (rho) between neuroticism scores and words correctly shadowed or interpreted.

b) Extraversion

Mean scores for subjects ($\bar{X} = 12$; S.D. = 4.59) are within the range of Eysenck's normal standardization sample. Rank correlations were calculated between extraversion scores and performance measures, and all proved to be close to zero.

v) Evaluations of intelligibility and informativeness

a) Intelligibility

The mean scores assigned by each judge at each signal-noise ratio are shown in Table 7 below. A three factor repeated measurements analysis of variance (Order x S/N x Judges) (Winer, p. 319) was employed, and the results of the analysis showed no significant difference between orders of presentation of the translation task ($F = 4.321$; $df = 1,10$), a significant effect of signal-noise ratio ($F = 4.396$; $df = 2,20$; $p < .05$), and a significant difference between judges ($F = 32.752$; $df = 1,10$; $p < .001$). None of the interaction terms were significant. Since there is no effect of Order, only the means for Judges and signal-noise ratios are shown in the table below.

	No noise SN_1	+5db SN_2	-2db SN_3	
Judge 1	6.8	6.2	5.4	6.1
Judge 2	5.3	4.7	4.2	4.7
	6.1	5.4	4.8	

Table 7: Mean intelligibility scale scores.

Scheffé's test showed that the difference between No noise and +5db was significant at the .05 level ($F = 4.68$, $df = 1,20$), that between No noise and -2db was significant beyond the .01 level ($F = 17.58$), while the difference between +5db and -2db was not quite significant at the .05 level ($F = 4.12$; $F_{.05} = 4.35$).

b) Informativeness

A similar repeated measurements analysis of variance was carried out on Informativeness scale scores. There was a significant difference between orders of presentation of the translation task ($F = 7.085$; $df = 1,10$; $p < .05$), showing that subjects who translated first scored higher on the Informativeness scale, than those who translated after shadowing. There was a significant effect of signal-noise ratio ($F = 4.688$; $df = 2,20$; $p < .05$), but no significant effect for judges ($F = 4.495$; $df = 1,10$). None of the interaction terms were significant. Table 8 below shows the mean scores for order and signal-noise ratios.

	SN_1	SN_2	SN_3	
Order 1	5.1	6.5	6.5	6
Order 2	3.3	3.8	5.1	4.1
	4.2	5.2	5.8	

Table 8: Mean informativeness scale scores.

Scheffe's test showed that the difference between No noise and +5db was significant beyond the .05 level ($F = 6.769$; $df = 1,20$), that between No noise and -2db was significant beyond the .001 level ($F = 18.48$), while there was no significant difference between +5db and -2db ($F = 2.88$).

DISCUSSION

These results confirm Treisman's finding of greater ear-voice spans for simultaneous interpreting than for shadowing (which is not very surprising in view of the extra work interpreters must do). It must, after all, take more time to transform than simply to repeat the input.

The lack of a significant effect of noise on ear-voice span may be due to the fact that, as Miller, Heise and Lichten (1951) showed, context facilitates perception of speech in noise, and this may have counteracted any increased in time that might have been needed to respond to individual words in noise. On the other hand, it might also be hypothesized that, as time for word recognition increases Ss might attempt to maintain a constant ear-voice span, sacrificing accuracy in their attempt to do so. Indeed the results shown in Table 2 tend to bear this out, all the more so in the case of interpretation where one simply cannot afford to lag too far behind.

As can be seen, significantly more of each passage was correctly shadowed than interpreted, and noise had a significant adverse effect on performance of both tasks. The lack of significant interaction between task and noise level, however, appears to indicate that noise did not have a greater effect on interpretation than on shadowing, contrary to what would have been predicted from Rabbitt's experiments.

On the other hand, closer examination of subjects' protocols for errors and omissions showed that noise did have differential effects on shadowing and interpretation, the deterioration in the quality of the translations as noise increased being also reflected in judges' ratings.

Subjects both committed more errors, and omitted more, in interpretation than in shadowing, but though there were both more errors and omissions in noise for both experimental tasks, noise had a significantly greater effect on errors in interpretation than in shadowing. Since there was generally better performance in shadowing of the same passages, it seems unlikely that interpreters' errors could be due entirely to misperception of the source message in noise. Furthermore subjects appeared less able to monitor their errors at lower S/N's, and it seems reasonable to conclude therefore that difficulty in perceiving source language input has, in fact, resulted in less channel capacity being available for translation and monitoring of output by the interpreter.

As a matter of interest the following are examples of the type of correction Ss can make:

Source	Interpreter
La Carte	the charter - the map
.....est imprime.....	is imprinted - is printed
.....sur tous les continents.....	especially on the continents - on all the continents
Aussi bornee que cette activite.....	Just as limited as this - however limited this activity
.....sur de tres grandes distances.....	over great distances - very great distances

In these and other examples of self-correction the interpreters appear to be carrying out a monitoring procedure similar to Miller, Galanter and Pribram's (1960) TOTE or "test-operate-test-exit". S generates a target language response, which passes a first test and is uttered, the utterance is then tested again. If the second test is passed, the interpreter proceeds to the next item, if not he "operates" again by generating further response, and so on. Repeated operations are possible because the original message remains available until the translation passes whatever tests S decides to carry out.

The tests referred to above would directly reflect a subject's response criterion, and may operate at phonetic, syntactic, as well as semantic levels of analysis. It is suggested, then, that apart from any effects on channel capacity per se, noise may also lead S to lower his normal criterion and accept inadequate responses.

The importance of individual differences in response to stressful working conditions is demonstrated in the relationship between E.P.I. Neuroticism scale scores and the proportions of the input texts correctly shadowed or interpreted. With a simple task such as shadowing there is only smallish (and insignificant) correlation between Neuroticism scale scores and performance at different noise levels. In the more complex task of interpreting, however, an inverted U shape relationship is found with no correlation between N and performance under good listening conditions, an improvement in performance with increase in N under slight noise stress, but a significant fall-off in performance for subjects with higher N scores when noise stress is increased.

A tendency to high arousal under stress (insofar as it is assessed by the E.P.I. N scale) appears, therefore, to be an advantage for simultaneous interpreters when working under moderately stressful listening conditions, but a liability under more stressful listening conditions.

Since the -2db condition employed in this study would never be encountered in practice by professional conference interpreters (they would refuse to work in such adverse conditions), whereas the +5db condition might well be encountered from time to time, these results could be born in mind in the selection and training of simultaneous interpreters.

II

The effects of noise on the temporal characteristics of the performance of simultaneous interpreters

Introduction

Two hypotheses were advanced to account for the possible effects of noise on the performance of simultaneous interpreters in the introduction to the first part of this study:

1) The first hypothesis was based on two experiments by Rabbitt (1966, 1968) on the effects of noise on recall, and suggested that levels of noise which would not necessarily impair perception of the input would interfere with the processes involved in retrieval and transformation of the source language message. One would, therefore, expect a greater decrement in the performance of simultaneous interpreters than shadowers as input noise increased.

2) The second hypothesis suggested that the difficulty of perceiving speech in noise would be reflected in increased response latencies, which would in turn be reflected in increased ear-voice spans.

Broadbent's recent discussion (Broadbent, 1971) of reaction times and response criteria suggests an elaboration of this last hypothesis: "...reaction is delayed until there is sufficient evidence to exceed some critical value and cause one of the possible actions to occur...The efficiency of the response will then depend on the precise setting of the criteria. If criteria for all reactions are set rather cautiously, performance will be slow but accurate. If the criteria are set to allow response on relatively little evidence, performance will be fast but inaccurate."

Increased reaction time to speech in noise could, therefore, have two rather different effects on interpreting, depending on a subject's response criterion. If S maintains his normal criterion of correct response, one would expect little deterioration in the quality of what was translated as noise increased, together with a slower but steady output rate. This, however, could only be attained at the cost of an increase in ear-voice span, together with an increase in omissions, but not of errors. If, on the other hand, S attempts to overcome increased reaction time by lowering his response criterion in order to maintain a fairly constant ear-voice span, one would expect almost unchanged ear-voice span and output rate, but an increase in errors (but not necessarily omissions). The two strategies for the simultaneous interpreter to adopt, therefore, would be either to sacrifice simultaneity for accuracy, or to sacrifice accuracy for simultaneity.

In the present context simultaneity refers to the fact that the simultaneous interpreter attempts to translate as he listens. In fact the translation itself is not simultaneous, and will always lag slightly behind the input. Listening and speaking can, however, be simultaneous in so far as a large proportion of the total source language input time may be taken up with simultaneous output by the interpreter, and an examination of the actual proportion of total input-output time spent in simultaneous listening and speaking is proposed in the present study. It has been suggested, for instance, by Goldman-Eisler (1968) that:

> "...the intermittent silence between chunks of speech in the speaker's utterance is a very valuable commodity for the simultaneous interpreter; for the more of his own output he can crowd into his source's pauses, the more time he has to listen without interference from his own output."

A similar suggestion has been made by Barik (1969). Though intuitively this appears to be a reasonable strategy, it is doubtful whether simultaneous interpreters can make such use of input pauses, since the average length of speech pauses would in itself preclude such a strategy. Goldman-Eisler (1968), for instance, found that the majority of pause lengths were 1 second or less in a variety of speaking situations. Barik (1969) examined pause durations in various types of source language input (different languages, texts, spontaneous speech, rehearsed text etc.), and found average pause durations of from 0.92 to 1.98 seconds. The difference between the Barik's and Goldman-Eisler's average pause times can be accounted for by the difference in the criterion for pauses: Barik's being .60 seconds, while Goldman-Eisler's was 0.25 seconds. The mean input pause time in the passages employed in the study to be discussed was .75 seconds. Obviously such figures will vary according to speaker, language, topic, and a variety of other factors, but equally obviously there is not much a speaker (i.e. simultaneous interpreter) can do to avoid filling such short gaps if he is already speaking.

In any case, it is doubtful whether simultaneity of input and output in itself presents the simultaneous interpreter with much of a problem: it is rather the additional information processing load imposed by the translation task that is likely to lead the simultaneous interpreter to attempt to vary his pattern of speaking and pausing - particularly when noise adds to the cognitive load.

Both this suggestion, and Goldman-Eisler's regarding the use interpreters might make of input pauses can be tested by comparing the distributions of speech and pause times in shadowing and in simultaneous interpretation, as well as the relative proportions of input pause times utilised in the two tasks. A related question concerns the simultaneity of input and output in simultaneous interpretations. If simultaneous interpreters do attempt to optimize their use of input pauses, or to adopt some other strategy to cope with the load of continuous throughput, such as pausing frequently themselves, then one might expect a reduction in the time spent in simultaneous listening and speaking both in comparison with shadowers, and as the difficulty of their task increases.

Method

The present analyses were carried out on subjects' tapes obtained for the first part of the study on the effects of noise on simultaneous interpreters' performance discussed in the previous section of this report.

In order to analyze the temporal characteristics of the experimental tapes both pen recordings and computer analyses were made of the tapes:

1) Pen recordings: In order to analyze speech times, pause times etc. both tracks of each subject's tape were transcribed on a paper record of a pen tracing of each channel, as in Figure 1. The pen recordings were obtained by replaying subjects' tapes on a stereo tape recorder, and feeding the output from each channel to two modified speech trigger units (derived from circuits provided by Ramsay and Law, 1966). The frequency response of each unit was tailored to a bandwidth of between 100Hz and 4500 Hz in order to filter out unwanted low and high frequency noise. The output from the trigger units was then fed to a modified Marconi E.E.G. pen recorder.

In effect, each trigger unit operates relays which are ON when there is speech on a channel, and OFF when there are silent periods (unfilled pauses). Since delay in onset, and delay in offset of the unit varies with signal level (Ramsay and Law, 1966), it was necessary to monitor the offset delay in order to minimize the error due to increases in loudness. A sudden rise in level before a pause could cause an over-estimation of speech time and underestimation of pause time, similarly too low a level could lead to underestimation of speech time and overestimation of pause time. The output from each channel of the trigger circuit was, therefore, monitored on a multi-channel display oscilloscope whilst the tape recorder output was monitored through separate loudspeakers. The auditory signals could then be matched with visual traces of the operation of the trigger units to produce optimal onset sensitivity, and minimal offset delay by adjusting the sensitivity of the trigger circuit. By feeding a signal of known duration into the circuit and systematically varying the level, it was estimated that errors of offset delay (i.e., the relay switching off too early or on too late) would have been within the range +5 to 10%, depending on the level, and the rate of decay of the last signal before the pause. The circuit is illustrated in block diagram form in Figure 1.

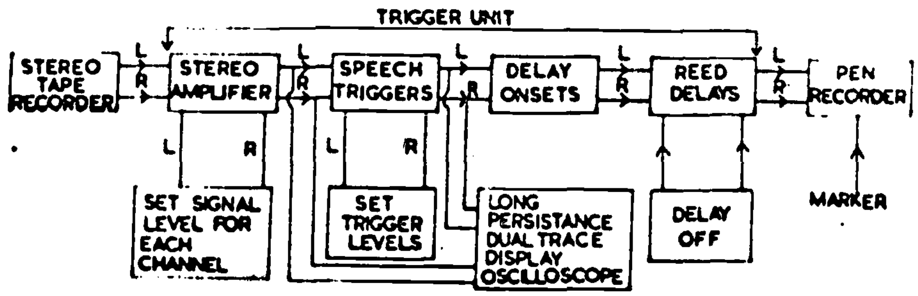

Figure 1: Block diagram of trigger unit input and output.

After setting the appropriate level controls, the experimenter followed the text from a typewritten copy, and activated a marker pen on the EEG recorder at approximately every 5th word of the input text. As can be seen from Figure 2, this practice provides a number of reference points against which to match the tape recording with the pen recording of each channel. The speed of the pen recorder was set at 3 cm. per second. Both tape recorded tracks were then transcribed by hand onto the pen recordings, and the measurements of ear-voice span, words per utterance, speech time, pause time etc. were then also made by hand. By continuous cross-reference between channels it was often possible to obtain quite accurate location of words within utterances.

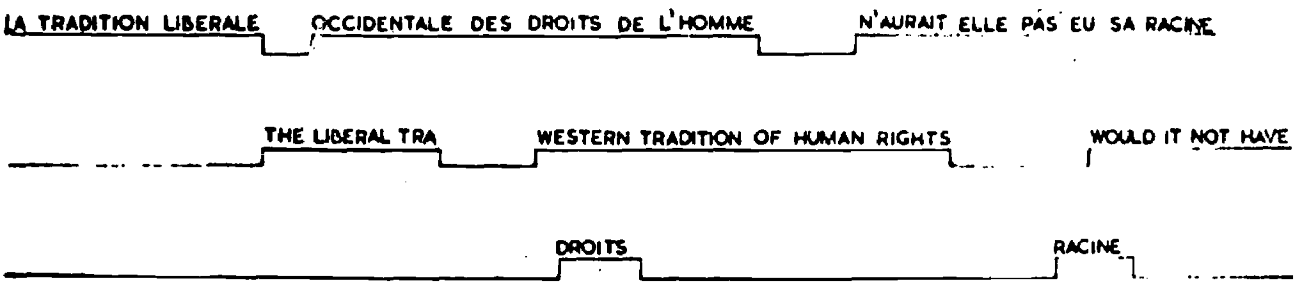

Figure 2: Pen tracings of source tape (top track) interpreter's version (middle track) and markers (lower track).

For the purposes of the present study it was decided to adopt Goldman-Eisler's (1968) criterion of 250 msec. to define the minimal interval for an unfilled pause. The most important reason for this was the variable nature of the recordings being assessed. As Ramsay and Law (1966) point out, the method employed to obtain pen recordings in the present study is sensitive to variations in record level. A relatively soft signal will cause the trigger circuits to go on late, and off early, relative to the actual onset and offset of the signal, while a louder than usual sound (i.e. once a particular trigger level has been set) will lead to accurate onset but late offset. As Ramsay and Law (1966) demonstrated, under-or-over-estimation of speech and pause times will result from variations in signal level. Since it was not possible to control the level of recordings obtained under language laboratory conditions, it was decided to adopt Goldman-Eisler's criterion of 250 msec. for unfilled pauses. This would allow for the 10% maximum error in measurement estimated for the tapes and equipment employed in the present study, and still be within the 200-500 msec. range for listener discrimination of pauses described by Boomer and Dittman (1962). As Martin (1970) confirmed, 250 msec. is within the range of silent intervals upon which both machine and listener tend to agree.

Other investigators of "unfilled pauses", such as Maclay and Osgood (1969), Martin (1967), Martin and Strange (1968), and Tannenbaum, Williams and Hillier (1965) have employed listeners' judgments to identify pauses. In these studies judgments of pauses were based on such criteria as "an abnormal hesitation in speech" (Maclay and Osgood, 1959), "hesitations between words judged as abnormal by the speaker yielding the utterance" (Martin, 1967), or "silences of unusual length". No strict criterion defining pauses was employed in these studies, and no time values were provided by these authors. Physical records of pauses have been employed by such authors as Boomer (1965), Boomer and Dittman (1962), Goldman-Eisler (1968), Martin (1970) and Suci (1967). Suci (1967) provided no definition of "pause" in terms of the minimal interval of silence required before a pause could be judged present. Boomer and Dittman (1962) demonstrated that unfilled pauses can be well discriminated by listeners at durations above 200 msec., and almost perfectly above 500 msec. In fact, Boomer (1965) used 200 msec. as his criterion. Goldman-Eisler (1968) employed a criterion of 250 msec. in a number of experiments discussed in her book, arguing that pauses up to 200-250 msec. might occur as part of ritardando effects or articulatory shifts between plosives. Martin (1970) employed speech spectography and listeners' judgments to identify pause location, and found listener-machine agreement over a range from 50 to 4970 msec. Over 21% of the listener-machine pauses were less than 200 msec., whereas unheard silent intervals (i.e. recorded by the machine, but not noted by the listeners) ranged from 50 to 110 msec.

Computer analysis

Basically this analysis involved the sampling (at a rate of 100 times per second) of each track of the subjects' tapes in order to determine whether or not speech was present. Any silence on either track of a quarter of a second or more was recorded as a pause, whereas the time value of any silence of less than a quarter of a second was added to the times of the speech segments on either side of the interval.

The final analysis provided data regarding the total times and the distribution of times the source speaker-simultaneous interpreter "system" was in each of 4 dyadic states: both channels silent, source on-interpreter off, source off-interpreter on, both on. Since the description of the equipment and the programming involved in this analysis is rather lengthy, a full description is provided in Appendix II of this report.

Treatment of results

Where appropriate, the data from the above measurements were analyzed by analysis of variance according to the experimental design described in the first study. Where the data was presented in terms of percentages or proportions they were transformed to their arc-sin values in order to preserve homogeneity of variance.

RESULTS AND DISCUSSION

Output rate

As can be seen from Table 9, output rate was faster for shadowing than for interpreting ($F = 9.95$; $df = 1,10$; $p < .05$). There was also a significant difference between output rates at the different signal-noise ratios ($F = 5.15$; $df = 2,20$; $p < .05$). The interaction between task and signal-noise ratio (S/N) was almost significant at the .05 level ($F = 4.10$; $df = 2, 20$; $F_{.05} = 4.35$) suggesting a greater fall in output rate for interpreting than shadowing.

	No noise	+5db	-2db	$\bar{X}$
Shadowing	118.33(1.94)	116.80(1.95)	111.31(1.86)	115.48(1.92)
Interpreting	120.64(2.10)	106.76(1.78)	96.04(1.60)	107.81(1.79)
$\bar{X}$	119.49)1.99)	111.78(1.86)	103.68(1.73)	

Table 9: Mean output rates w.p.m. (secs.).

These results could be due to the fact that interpreters were omitting and pausing more when noise was present (see below), thus spreading less speech over the same period of time. In order to shed some light on whether or not these results reflect omissions and pauses, or a genuine decrease in output rate, the mean output rate per utterance was estimated by dividing the mean numbers of words per utterance by the mean utterance times, and is shown in Table 10 below.

	No noise	+5db	-2db	$\bar{X}$
Shadowing	139.20(2.32)	130.20(2.17)	136.80(2.28)	135.60(2.26)
Interpreting	127.20(2.12)	107.40(1.79)	107.40(1.79)	115.20(1.92)
$\bar{X}$	132 (2.20)	116.40(1.94)	1.21(2.21)	

Table 10: Estimated mean output rate within utterances w.p.m. (secs.).

Comparison of the two tables suggests the contribution of an increase in omissions and pausing to the differences in overall output rate shown in Table 10 above, particularly in the case of shadowing. It would appear, then, that output rate in itself is a more sensitive measure than ear-voice span of the retarding effect of noise on simultaneous interpreters' performance, and that such an effect has been demonstrated in the present experiment.

Since it is not possible to partial out the contribution of omissions to pause time these variables are confounded in the data presented in Table 11 below, which also demonstrate a significant interaction between task and noise level ($F = 4.33$; $df = 2,20$; $p < .05$). The effect of noise level was also significant ($F = 5.20$; $df = 2,20$; $p < .05$), but there was no significant difference between tasks.

	No noise	+5db	-2db	$\bar{X}$
Shadowing	1.05	1.01	1.13	1.06
Interpreting	1.00	1.20	1.33	1.18
$\bar{X}$	1.00	1.11	1.23	

Table 11: Mean pause times (secs.).

These results would appear to confirm the retarding effect of noise on interpreting. The effect cannot be solely attributed to difficulty in recognition, since it is significantly less when shadowing. It is interesting to note that, when working in the optimum condition, there is virtually no difference between the pause times for shadowing and interpreting. Whatever function pauses may serve in normal speech, they do not appear to reflect the relative complexity of simultaneous interpretation compared with shadowing. Goldman-Eisler (1968), on the other hand, maintains that hesitation is "the behavioural concomitant of cognitive activity". If this is so then it appears reasonable to expect the complexity of such activity to be reflected in either the length of hesitation, or in the total amount of pause time relative to total speech time, but as can be seen both from Table 11 above, and Table 12 below, this was not the case.

	No noise	+5db	-2db	$\bar{X}$
Shadowing	0.22	0.24	0.28	0.25
Interpreting	0.18	0.23	0.32	0.24
$\bar{X}$	0.20	0.24	0.30	

Table 12: Mean ratios of total pause time to total speech time (P/S).

This table shows that approximately the same ratio of pause time to speech time is maintained in either shadowing or interpreting. The only significant difference in the above table is between noise levels ($F = 4.833$; $df = 2,20$; $p < .05$): total pause time increasing as noise increases. The effect of noise on both tasks was to decrease speaking relative to pausing. It is interesting to compare the ratios of pause time to speech time for the two experimental tasks with input P/S in order to see to what extent each task involves a redistribution of the input pause and speech times, as Goldman-Eisler (1968) has suggested is frequently the case in simultaneous interpretation. It can be seen from Table 13 below that, when interpreting, subjects tended to have similar P/S to the input passages, except at -2db where they paused more and spoke less than the input. In shadowing, on the other hand, subjects appear to consistently exceed the input P/S.

	No noise	+5db	-2db
Shadowing	1.10	1.08	1.31
Interpreting	.95	.96	1.15

Table 13: Ratio of output P/S to input P/S.

In shadowing, therefore, it seems that subjects can afford to pause more under all conditions, and still correctly reproduce more than simultaneous interpreters at each noise level.

With reference to the use simultaneous interpreters made of input pauses, it can be seen from Table 14 that there were but small differences between either tasks or noise levels. Only a slightly higher percentage of input pause time being used in interpretation at -2db. In the No noise condition, the proportions for both tasks are quite similar, and suggest that, given fairly short input pauses (the mean input pause time for the three passages employed in this study was .75 sec.), there is not much the interpreter or shadower can do to vary the "use" he makes of input pause times. One factor which is not taken into account by the present measures, however, is the possibility that simultaneous interpreters might increase their output rate within input pauses. This hypothesis seems rather unlikely, since,(except for longer input pauses) not much could be crammed into the average pause time of less than 1 second, part of which would be taken up by S's decision that a pause had occurred.

	No noise	+5db	-2db	$\bar{X}$
Shadowing	.84	.86	.84	.85
Interpreting	.87	.88	.82	.86
$\bar{X}$	.87	.87	.83	

Table 14: Percentage of input pause time interpreter/shadower speaking.

If subjects could not cram more output into input pauses, did they manage to reduce the total amount of time spent in simultaneous listening and speaking, or the average period of time spent in simultaneous input and output? Did noise, for instance, lead subjects to try to break up their output more, so avoiding relatively long stretches of simultaneous work?

As can be seen in Table 15 shadowers paused significantly more often ($F = 5.03$; $df = 1,10$; $p < .05$) than interpreters. Though there was an increase in the number of pauses in both tasks as noise increased, the effect was not significant, neither was there any interaction between task, noise and number of pauses.

	No noise	+5db	-2db	$\bar{X}$
Shadowing	42.42	46.17	47.33	45.31
Interpreting	34.67	35.08	40.33	36.70
$\bar{X}$	38.55	40.63	43.83	

Table 15: Mean number of pauses.

Though subjects paused more often when shadowing than when interpreting, similar proportions of total input-output time were devoted to simultaneous listening and speaking in each task, and noise had no significant effect on the proportion of total input-output time spent in simultaneous listening-speaking ($F = 2.98$; $df = 2,20$). These results are shown in Table 16. The results for mean simultaneous listening-speaking times (Table 17) appear to show a slight interaction between noise and task in that simultaneous listening and speaking times become shorter for interpreters than for shadowers as noise increased. None of the F tests, however, even approached unity.

	No noise	+5db	-2db	$\bar{X}$
Shadowing	.67	.66	.65	.66
Interpreting	.67	.65	.62	.65
$\bar{X}$	.67	.655	.635	

Table 16: Proportion of total input-output time spent in simultaneous listening-speaking.

	No noise	+5db	-2db	$\bar{X}$
Shadowing	1.52	1.53	1.51	1.52
Interpreting	1.72	1.58	1.50	1.60
$\bar{X}$	1.62	1.555	1.505	

Table 17: Mean simultaneous listening-speaking times (secs.).

The results of the present analysis must be considered in the light of data presented in the first part of this study, in which it was found that there were significant effects of both task and noise level on the number of words correctly shadowed or interpreted. More words were correctly shadowed than interpreted, and the number of words correctly processed decreased significantly as noise increased. Further analyses revealed that, though significantly more errors and omissions were made in both tasks in noise, there was a significantly greater effect on the number of errors made when interpreting. This result, together with the related finding that the proportion of interpreters' self-corrections relative to the number of errors committed decreased as noise increased, suggested that one of the effects of noise on interpreters' performance was to cause them to lower their response criterion. It had also been suggested that one of the effects of noise on simultaneous interpreters would be an increase in ear-voice span due to the greater response latency to words in noise (Pollack and Rubenstein, 1963). Though ear-voice spans were greater for interpretation than for shadowing there was very little effect of noise on ear-voice spans in either task. It would appear, then, that subjects in the present study adopted the strategy of sacrificing accuracy for simultaneity when working under noise stress. In this context simultaneity refers to the number of words the interpreter is prepared to lag behind the original. Within the limitations of the present experimental situation Ss appeared to regard the payoff from attempting to lag no more than between 5 and 6 words behind the original to be preferable to the risk involved in omissions and errors.

Analysis of the temporal characteristics of subjects' performance demonstrated the retarding effect of noise on the performance of interpreters, as opposed to shadowers. This was shown in decreased output rates and increased pause times. Since interpreters did not omit more than shadowers as noise increased, and the difference in shadowers' output rates as noise increased disappeared when within utterance rates were estimated, this result indicates, that noise slowed down performance on the more complex task. In other words, when they were speaking simultaneous interpreters spoke more slowly, and the pause ("processing") time between utterances also increased. It is interesting to compare these results with those from an experiment by S. Fisher (Unpublished, quoted in Broadbent, 1971) who, with a continuous serial reaction task, found that bursts of noise at 80 db produced a slowing of the particular reaction in progress when the burst arrived. Here there was a delay of half a second or so. If we regard simultaneous interpretation as a continuous auditory tracking task involving a sequence of decisions regarding the source language message, then a rapid succession of noise bursts might be expected to have a similar effect to that found by Fisher. Although purely in the auditory modality, and involving a continuous noise source, the situation is perhaps not so different from Fisher's since the noise would have a different effect on the words perceived according to their predictability as well as the variable level of intensity of the recording of much of the source language text. In other words, the subjective effect for the listener may well have been of a series of bursts of noise scattered among more readily processed "chunks" of input.

Other aspects of the results have been discussed above, and the general implications will be discussed in the final section of this report.

III

Level of processing and comprehension of prose

Introduction

Within the context of his discussion of a single channel theory of human information processing, Welford (1968) suggests that simultaneous interpreters manage to carry out their task because they learn to ignore the feedback from what they are saying: "Simultaneous translators seem to acquire the ability to do this (i.e. speak and listen simultaneously) after long practice...ignoring the feedback from their own voices. In consequence their speaking voices are often strange, and they themselves report that they have little idea of what they are saying or confidence that it is correct."

The analysis of interpreters' output in the previous study however, showed that simultaneous interpreters must be attending to some form of feedback since there were corrections of their own output at various levels; phonological, syntactic and semantic. One can but conclude that Welford's informants were either being modest, or were simply bad interpreters (or perhaps simply not aware that they were able to correct themselves, and of the implications of this ability for a psychological analysis of the task). The current evidence suggests, therefore, that some interpreters at least are able to monitor their own output to the extent of correcting themselves as they carry out their task. This in itself implies that they must match their output against some model of the segment of input being translated. Whether or not they understand the information they are transmitting can be tested in two ways: (i) by evaluating the translations, and (ii) by asking the interpreter to take a comprehension test on the material just translated.

It is the second method which will be employed here. If simultaneity of listening and speaking in itself hinders understanding, then one would expect shadowers to have poorer performance than listeners on comprehension tests after shadowing or listening to prose passages. In fact, Carey (1971) did not find this to be the case in an experiment in which subjects either listened to, or shadowed, prose passages recorded at 1, 2, or 3 words per second. Subjects were then given tests of word recognition, syntactic recognition, and semantic retention. At the slowest rate shadowers' word recognition and semantic retention scores were somewhat higher than those for listeners, but these differences disappeared at faster rates (i.e. at approximately the normal speech rate of 2 - 3 words per second).

If simultaneity of listening and speaking does not preclude recognition and retention, then it could be argued that, since simultaneous interpretation involves a more complex level of analysis of the input message than shadowing, simultaneous interpreters would retain more than shadowers transmitting the same messages. In order to examine both Welford's suggestion and the above hypothesis it had originally been planned to compare the comprehension-retention of 3 groups of subjects after each group had either listened to, shadowed or simultaneously interpreted passages of French prose. Unfortunately the small number of suitably skilled subjects available necessitated a repeated measurements design in which subjects knew they were going to be tested after having carried out the experimental task. Since the simultaneous interpreter does not normally expect to be quizzed after a session in the booth the present experiment does not actually test what might be termed incidental learning in simultaneous interpreters, but it is hoped that it does examine the effect of complexity of input processing on retention and comprehension.

METHOD

Materials

Three French texts of approximately 300 words each were recorded on tape by a native French speaker at a normal speech rate. The texts, dealing with road safety in winter, marine resources, and the channel tunnel, were taken from "General Studies French" (Light and Howitt, 1966), and were chosen because they lent themselves to the construction of questions on their factual content, and because they made no special demands on subjects' knowledge of technical vocabularies. Ten questions, requiring comprehension and recall, were prepared for each passage, and were typed on individual 8" x 5" index cards. It should be stressed that the questions could not be answered from general knowledge, and were specifically related to the content of each passage.

Subjects

The subjects were 9 members of a class completing their training as simultaneous interpreters, working primarily from French to English. All subjects were well practised in shadowing as well as simultaneous interpretation.

Procedure

A repeated measurements latin square design was employed in which each passage was either listened to, shadowed, or simultaneously interpreted from French to English according to the order set out in the design.

The stimulus tapes were played to subjects individually over headphones from a UHER Universal Teaching tape recorder at a comfortable level. Subjects' shadowing or interpreting responses were recorded on the bottom tracks of their stimulus tapes via a boom microphone attached to their headsets, and were monitored by the Experimenter in order to ensure that the task was being carried out.

As soon as each subject had listened to, shadowed, or interpreted each passage, the ten question cards were placed face down in front of him in random order. Subjects were then instructed to write the answer to each question on the question card, and to place the card face down on a separate pile when the question was completed. Subjects were self-paced on this task, and were allowed a further two minutes to review their replies on completing the ten questions.

RESULTS

Since the different questions involved the recall of one to four basic points from the text, it was decided to score answers in terms of a percentage of the total possible number of points covered for each question. These percentages are shown in Table 18 below:

LISTENING	INTERPRETING	SHADOWING
58	51	43

Table 18: Mean percent correct responses.

Subjects' scores were transformed to their arc-sin values for the analysis of variance, in order to ensure homogeneity of variance. The differences between scores on the three tasks were significant beyond the .001 level ($F = 11.3143$; $df = 2,12$). There were no significant differences between either passages, or the order in which the task was performed. Orthogonal comparisons between treatment means (Edwards, 1968) showed that the differences between means for shadowing and interpreting, and between interpreting and ~~shadowing~~ listening were significant beyond the .025 level, while the difference between means for listening and shadowing was significant beyond the .001 level.

Discussion

Though, as already stated, this experiment is not a direct test of Welford's suggestion that simultaneous interpreters "have no idea of what they are saying", within the limitations imposed by the experimental design these results lend further support to the evidence from interpreters' self-corrections for the hypothesis that simultaneous interpreters do continuously monitor their own performance. These results do not, however, confirm Carey's finding of no difference between shadowing and listening retention scores at normal speaking rates, and it could be argued that in the present situation simultaneous listening and speaking did impair subjects' comprehension and recall when shadowing scores are compared with listening scores. When shadowing scores are compared with interpreting scores, however, it becomes apparent that it is not simultaneous listening and speaking per se that impairs performance on the recall task, but rather the nature of the processing subjects are required to carry out on the input message. Since shadowing does not require the complex analysis and transformation of input which is necessary in interpretation it seems fair to say that only a comparatively simple form of processing is carried out in order to transfer input from an auditory to a vocal mode. Simultaneous interpretation, on the other hand, requires analysis of both input and output at a number of levels. It may well be that the need to monitor both input and output can be viewed as a more intensive form of rehearsal than shadowing, and also makes more demands on channel capacity. When the listener is able to devote his full channel capacity to input processing, without having to share attention between multiple tasks (as in shadowing or interpreting), one would expect better comprehension and recall than in attention sharing situations. It would appear, therefore, that comprehension and retention of complex verbal input are a function of the channel capacity used by the subject in processing the message.

IV

A comparison between simultaneous and consecutive interpretation

Introduction

Simultaneous interpretation can be viewed as a paced tracking task with transformation of information by the interpreter. The task of the interpreter's audience (who only have to cope with input, and not with simultaneous output as well) is somewhat similar if less complex. Whereas the simultaneous interpreter may not infrequently encounter noisy listening conditions, the conference listener will always be attempting to track the interpretation against a highly variable level of background noise. This may come from inside or outside the conference hall, but will be mainly due to interference from the voice of the floor speaker whose speech is being interpreted. Under these circumstances it would appear that, given an equivalent translation (in terms of the information conveyed), consecutive interpretation, which is carried out when the floor speaker has finished, should be more effective in conveying information to the conference listener than simultaneous interpretation. Furthermore, insofar as it avoids interference from the additional noise of the floor speaker's voice, consecutive interpretation might be expected to stand up better to noisy listening conditions than would simultaneous interpretation as far as the listener's ability to absorb information is concerned.

In order to study these questions two experiments were carried out comparing listeners' responses to tests on simultaneously and consecutively interpreted material heard under both quiet and noisy listening conditions.

EXPERIMENT NO. 1

Ease of Comprehension of Simultaneous and Consecutive Interpretations

Method

Three French passages of approximately 350 words each, dealing with the topics of the Channel Tunnel, undersea mining, and the hazards of winter driving were recorded in both simultaneous and consecutive interpretation in English by a professional conference interpreter. Since the emphasis in this experiment was to be on the listener's, rather than the interpreter's, performance, the interpreter was given the opportunity to produce his best interpretation by allowing him to practise the material before the recordings of simultaneous and consecutive interpretation were made.

Ten questions were devised for each passage, and were typed on separate 8" x 10" index cards. These were placed face down in front of subjects in a different random order for each subject. Separate answer sheets were provided for each question.

Subjects

The subjects were 30 university undergraduates all of whom had little or no knowledge of French, and had English as their mother tongue. 3 groups of 5 subjects listened to simultaneous interpretations, and 3 to consecutive interpretations.

Procedure

The experiment was conducted in a large university lecture theatre. The voice of the French speaker was relayed to listeners from the front centre of the hall from an electrostatic speaker at lectern height at an average sound pressure level of 75 db re. .0002 dyne cm^2 which was judged to be approximately the level of delivery of a person speaking in order to be heard by a large audience. The simultaneous interpretation was relayed to subjects through headphones at a level of between 78-80 db re. .0002 dyne cm^2 from the main control tape recorder and amplifier. Subjects were tested in groups of 5, and were seated along one row of seats in the centre of the hall.

The following instructions were given to subjects, depending on the type of interpretation to be presented.

i) Simultaneous interpretation.

"You are going to hear 3 examples of the type of interpretation you would hear if you attended an international conference. You will listen to simultaneous English interpretations of speeches in French.

You will hear the English translation through your headphones, while the French will come at the same time from the loudspeaker in the front of the hall. After each short passage you are to answer 10 questions which are typed on the cards which are placed face down in front of you. Answer the questions on the sheets provided. You will be allowed 45 seconds in which to answer each question, after which a buzzer will sound for you to go on to the next card. If you cannot answer a question in the time allowed please carry on to the next one when the buzzer sounds. You will be allowed extra time at the end of the 10 questions to review your answers."

ii) Consecutive interpretation.

Except for the following alteration the instructions were the same as for the simultaneous interpretation condition:

"You are going to hear 3 examples of the type of interpretation you would hear if you attended an international conference. You will listen to consecutive translations in English of speeches in French. You will first of all hear the French passage from the loudspeaker in the front of the hall, after which you are to put on your headphones and listen to the English translation of the passage. After each short passage..."

Experimental Design and Analysis of Results

A repeated measurements Latin Square design (Winer, 1962; p. 549) was employed. Each of 6 groups of 5 subjects received a different order of presentation of passages: 3 in consecutive translation, and 3 in simultaneous translation. Since the answers to each question covered from 1 to 4 points from each passage, subjects' responses were scored in terms of a percentage of the total number of possible points covered for each question. These scores were then transformed to their arc-sin values in order to ensure homogeneity of variance.

Results

The results, shown in Table 19 below, showed no significant difference in the effect of type of interpretation on listeners' scores ($F = 0.52$). No significant difference was found for passages ($F = 0.67$) or for the interaction between passages and type of interpretation ($F = 0.23$). Though there was a significant practice effect demonstrated by the significant F for Trials ($F = 7.27$; $df = 2,48$; $p < .01$), there was no significant interaction between Trials and type of interpretation ($F = 0.54$).

	Trial 1	Trial 2	Trial 3	$\bar{X}$
SIMULTANEOUS	0.50	0.60	0.57	0.56
CONSECUTIVE	0.51	0.61	0.63	0.58
$\bar{X}$	0.51	0.61	0.60	

Table 19: Proportions of correct responses.

EXPERIMENT NO. 2

The Effects of Noise on Comprehension and Recall of Simultaneous and Consecutive Interpretation.

Method

The same recorded materials and test questions as were used in Experiment No. 1 above.

Subjects

Four groups of ten subjects each. All subjects were undergraduates with little or no knowledge of French, and with English as their mother tongue.

Procedure

Two noise sources were employed in this experiment. In one, white noise from a Dawe white noise generator was relayed to the lecture theatre through the P.A. system at a level of approximately 70 db, or 4-5 db less than the level of the voice from the speaker. In the second condition, noise was relayed to subjects' headphones; maintaining a similar signal-noise ratio of +4/5 db for the interpretation in the headphone. Strictly speaking these signal-noise ratios are not the same, since the close fitting headset would serve to attenuate the noise from the hall. No artificial ear was available, however, to assist with the calibration of exact sound pressure levels at the headphone itself.

Two groups of ten subjects heard interpretations against a background of hall noise; one group listening to simultaneous, the other to consecutive, interpretations. Two groups heard interpretations against a background of noise in the headphones; one group listening to simultaneous, the other to consecutive, interpretations.

Apart from the above changes, the experimental procedure was the same as for the previous experiment.

Experimental Design and Analysis of Results

Due to: (i) the limited number of subjects available with little or no knowledge of French; (ii) the fairly complex design problems involved in counterbalancing all levels of all factors in an experiment of this type; (iii) the fact that the previous experiment had not revealed any differences between passages, or interaction between passage and type of interpretation, it was decided to employ a repeated measurements factorial design (Winer, 1962; p. 341).

Once again, subjects' responses were scored as percentages of the total possible score for a passage, and these scores were transformed to their arcsin values.

Results

Noise source		PASSAGE 1	2	3	$\bar{X}$
Headphones	Sim.	.47	.46	.53	.49
	Con.	.59	.53	.56	.56
L. Theatre	Sim.	.45	.51	.51	.49
	Con.	.49	.55	.51	.52
$\bar{X}$		.50	.52	.53	

Table 20: Proportions of correct responses for simultaneous and consecutive interpretations under different noise conditions.

Though once again the scores for listeners to consecutive interpretation were slightly higher than those for listeners to simultaneous interpretation, the difference was not significant ($F = 3.7841$; $df = 1,36$; $F_{.05} = 4.11$). None of the other main or interaction effects were significant either.

Discussion

Though subjects' comprehension/retention scores were marginally higher after listening to consecutive than to simultaneous interpretations the differences were not significant either in good or noisy listening conditions. Within the limitations of the design of these two experiments, therefore, the interference from the source language speakers' voice did not prevent subjects from monitoring simultaneous interpretations as successfully as they did consecutive interpretations. Given equivalent translations in terms of the amount of source language information conveyed by each form of translation, there appears to be little difference in the effectiveness of consecutive and simultaneous interpretations in conveying information to the listener.

Certain points regarding consecutive interpretation are worth noting at this stage. Van Hoof (1962) points out that, in recent times, consecutive interpretation has gradually given way to simultaneous interpretation at international conferences, but that it will usually be the preferred type of interpretation at small meetings, such as comittees and round-tables, where two languages suffice.

There are 2 forms of consecutive interpretation:

(i) Continuous, where the interpreter waits until the source speaker has finished his entire speech before delivering his version (ii) Discontinuous, in which the interpreter delivers his version at breaks in the source speaker's output.

Though consecutive interpretation is often thought to be superior to simultaneous interpretation in terms of accuracy and style, its use will inevitably considerably lengthen conference proceedings. It should be remembered, however, that the process may in itself be less stressful for the interpreter than simultaneous interpretation insofar as he does not have to listen and speak simultaneously relying constantly on short-term memory. The basic skill developed by the consecutive interpreter is that of taking rapid notes in various "shorthand" forms (shorthand from the point of view of mnemonic codes rather than a system of shorthand as such), and reconstructing the original message on the basis of his notes. Though this may often enable the interpreter to deliver a more accurate and stylistically acceptable translation than would be the case with simultaneous interpretation delivered under time pressure, the present results show that there is no reason to prefer one form of translation over the other solely on the basis of the listeners ability to gain information from either form of translation.

Discussion and Critique

Discussion

To sum up the main conclusions so far:

i) Noise appears to slow down the performance of simultaneous interpreters, and to cause a deterioration in the quality of translations which is reflected in increases in errors and omissions. It was hypothesised that, in order to maintain a constant ear-voice span under noisy listening conditions, interpreters were prepared to sacrifice accuracy by lowering their response criteria.

An inverted U-shaped relationship was found between Eysenck Personality Inventory Neuroticism Scale scores and interpreting performance under the different listening conditions.

ii) Though the experiment was not an ideal test of incidental learning during simultaneous interpretation, it was found that simultaneous interpreters recalled more than shadowers after "processing" the same passages. The import of this result is that the high level of information processing demanded by simultaneous interpretation does not preclude some of the interpreters' information handling capacity being available for storage (temporarily at least) of some aspects of the message being transmitted.

iii) It was found that simultaneous and consecutive interpretation were almost equally effective in conveying information to listeners, whether under perfect or noisy listening conditions.

Practical implications

Since the results of the second experiment are of mainly theoretical interest, only the first and third experiments will be discussed below - in reverse order.

i) Since either form of interpretation was found to be equally effective under both good and poor listening conditions, there is no a-priori reason for choosing the one over the other in the actual conference situation. Other considerations, such as the expense involved, size and type of meeting, number of languages required, facilities available etc. must determine the choice of type of interpretation. In the final analysis however, it is the effectiveness of the individual interpreter, whatever method he uses, which is most important.

ii) Bearing in mind the small sample size, the relationship between E.P.I. Neuroticism scale scores and interpreting performance under stress suggests that a certain level of arousal (insofar as arousal is indicated by N scale scores) is advantageous to the interpreter under moderately noisy (stressful) listening conditions.

Though no relationship was found between N and performance under perfect listening conditions, the above should be born in mind in the selection of interpreters, since the actual conference situation will almost invariably involve a number of sources of stress (e.g. speaker variables such as rate of speaking, accent, loudness etc., as well as the pressure of working for an actual audience).

Furthermore it would be useful when selecting interpreters, as well as trainees, if:

i) Group norms were available for "successful" professional interpreters on measures of personality, cognitive ability, and performance in their various languages.

ii) The present study were extended to cover a large group of subjects, and a broader range of speaking and stress situations.

Theoretical considerations

Apart from the work of Barik (1969) and Goldman-Eisler (1968) there has been almost no attempt at analyzing the complex information processing behaviour involved in simultaneous interpretation.[1] Gerver (1969) examined the effect of source language presentation rate on simultaneous interpretation. Quantitative as well as qualitative aspects

[1]Two earlier studies (Oleron and Nanpon, 1965; Treisman, 1965) dealt mainly with the time lag between source language presentation rate and interpretation.

of performance were discussed in terms of strategies for coping with information overload. An analysis of omissions, errors, and corrections was carried out and led to an analysis of behaviour under speed stress very similar to that given for performance under noise stress in section 1 of this report. Barik (1971) also presented a descriptive account of various types of omissions, additions and errors of translation.

As can be seen from the discussion of omissions, errors, and self-corrections in the first section of this report, a close analysis of the type of process involved in producing such "discontinuities" in output might provide a number of fruitful insights into the simultaneous interpreters' behaviour. Furthermore, analysis of the temporal characteristics of simultaneous interpreters' performance under various input conditions may shed some light on the question as to what extent sequential and/or parallel processes are involved.

It is interesting to note, for instance, that two of the subjects in the first study were simultaneously listening and speaking for over 80% of the total input-output time, and two more for over 75% of that time. That they could do this whilst correctly interpreting over 85% of the input, and at the same time monitor, and correct their own output must surely have implications for the study of memory and attention in general. None of the current theories of attention and skilled behaviour (Welford, 1968; Broadbent, 1971) appear to be able to account in themselves for such behaviour.

The writer is at present preparing a model of the simultaneous interpreters' performance which incorporates both sequential and parallel processing, and is based on the present research as well as recent research and theory in memory and attention.

Critique

From the point of view of the conference interpreter, the conditions under which these experiments were carried out might be regarded as so artificial as to render meaningless any conclusions regarding "what really goes on" in the interpreters' booth, or how the conference listener copes with simultaneous or consecutive interpretation. The following considerations should be born in mind, however, when evaluating the work:

a) The languages studied are limited to those in which the writer is himself fluent, and the results may t'be true of translations to, or from, other languages.

b) Though differences in general style of speech may be expected according to whether a source language passage is delivered spontaneously, or read from a text, the use of read texts for interpretation is not an entirely artificial situation for the interpreter. Many speeches and papers delivered at international conferences are, in fact, read from prepared texts. Apart from the questions of the availability and quality of recordings, the need to control speaker variables, as well as the general considerations of the experimental design, precluded the use of actual conference material. The conclusions regarding the effects of noise on interpreters' performance, however, may well not be typical of different source language speaking situations (or of different noise sources).

c) The difficulty of obtaining suitably qualified subjects in sufficient numbers (for any particular combination of languages) imposes a severe restriction on the design and execution of experiments in this area. Even where advanced students of interpretation were employed as subjects, it was not possible to obtain enough of them to study the effect of "level of processing" in independent groups. As noted in the discussion of that study, the effect of a repeated measurements design was to prime subjects in advance for the test following each passage. Since simultaneous interpreters must rarely, if ever, feel that they will be questioned regarding the passages they translate this situation hardly provides a true test of incidental learning during simultaneous interpretation.

d) Finally, due to limitations imposed by the nature of the present research it has not been possible to investigate the effect of source language linguistic variables (such as grammatical structure) on the interpreters performance. An investigation of the effect of syntactic and semantic factors is, however, currently in progress.

APPENDIX I

Scales of intelligibility and informativeness employed in Noise experiment

SCALE OF INTELLIGIBILITY

9. Perfectly clear and intelligible. Sounds like ordinary speech; has no stylistic infelicities.
8. Perfectly or almost clear and intelligible but contains minor grammatical or stylistic infelicities and/or mildly unusual word usage that could, nevertheless, be easily "corrected".
7. Generally clear and intelligible, but style and word choice and/or syntactic arrangement are somewhat poorer than in category 8.
6. The general idea is almost immediately intelligible, but full comprehension is distinctly interfered with by poor style, poor word choice, alternative expressions, untranslated words, and incorrect grammatical arrangements.
5. The general idea is intelligible only after considerable effort but given the effort one can be fairly confident that one understands. Poor word choice, grotesque syntactic arrangement, untranslated words, and similar phenomena are present but constitute mainly "noise" through which the main idea is still perceptible.
4. Masquerades as an intelligible passage, but actually is more unintelligible than intelligible.

Nevertheless the idea can still be vaguely apprehended. Word choice, syntactic arrangement, and/or alternative expressions are generally bizarre, and there may be critical words untranslated.

3. Generally unintelligible; it tends to sound like nonsense, but one can at least hypothesize the ideas intended in the passage.
2. Hopelessly unintelligible, yet does not appear total nonsense.
1. Both totally unintelligible and nonsensical.

SCALE OF INFORMATIVENESS

9. Extremely informative. Makes "all the difference in the world" in comprehending the meaning intended. (A rating of 9 should always be assigned when the original completely changes or reverses the meaning conveyed by the translation.)
8. Very informative. Contributes a great deal to the clarification of the meaning intended. By correcting structure, words, and phrases, it would make a great change in the listener's impression of the meaning intended, although not so much as to reverse or change the meaning completely.
7. Between 6 and 8.
6. Clearly informative. Would add considerable information about the sentence structure and individual words, putting the listener "on the right track" as to the meaning intended.
5. Between 4 and 6.
4. In contrast to 3, adds a certain amount of information about the sentence structure and syntactical relationships. It may also correct minor misapprehensions about the general meaning of the sentence or the meaning of individual words.

3. By correcting one or two possibly critical meanings, chiefly on the word level, it gives a slightly different "twist" to the meaning conveyed by the translation.
2. No really new meaning added by the original, either at the word or grammatical levels, but the listener would be somewhat more confident that he apprehends the intended message.
1. Not informative at all, no new meaning added, nor would the listener's confidence in his understanding be increased or enhanced.
0. The original contains, if anything, less information than the translation. The interpreter appears to have added certain meanings, in order to make the passage more understandable.

APPENDIX II

ASPA - Automatic speech-pause analyzer

A suite of computer programmes for monitoring patterns of speaking and pausing from twin channel tape recorder output.

Introduction

The aim of the programmes to be described below was to monitor, in real time, twin channel tape recorings of speech (source language speaker, and simultaneous interpreter), in order to produce as output a digital record of the presence or absence of speech on each channel, together with the length of time the channels remained in any particular state:

Channel 1 on - Channel 2 on

Channel 1 on - Channel 2 off

Channel 1 off - Channel 2 on

Channel 1 off - Channel 2 off

The following points had to borne in mind when developing the programmes:

i) In order to provide a more consistent wave-form for sampling, the envelope of the speech wave-form rather than the speech wave-form itself was to be sampled.

ii) In order to reduce the amount of data handled by the programmes data should be initially processed in the sampling programme.

iii) Since the signal being sampled (i.e. the envelope of the speech waveform) would be of a low frequency, and in order to reduce the amount of data handled, a relatively slow sampling rate (100 times per second) was decided upon.

iv) For speed and convenience the data were to be stored in disk files.

v) The criterion for a pause was to be ¼ sec. Any break in speech on either channel of less than ¼ sec. was to be discounted and the interval regarded as continuous with the speech signal on either side of it on the relevant channel.

vi) For simplicity, and in order to facilitate debugging, the programmes should be developed in modular form.

The flow chart overleaf illustrates the sequence of 8 programmes developed. These are described more fully below.

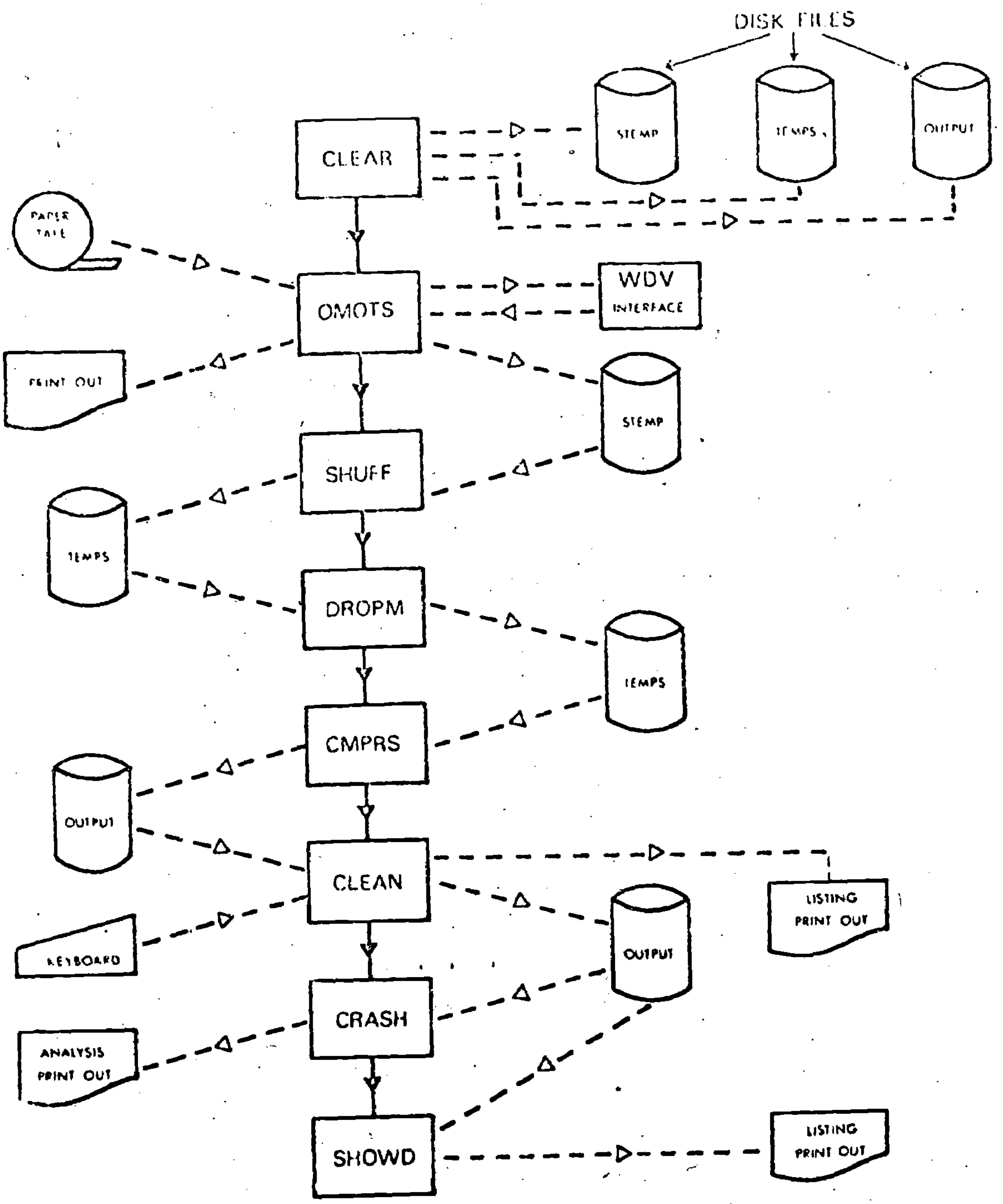

Figure 1: Sequence of programmes and operations in ASPA.

Equipment

i) Computer: IBM 1130, with 8,000 16 bit words of core storage and disk drive holding 512,000 words. 1134 paper tape reader, and 1131 console typewriter output.

ii) Interface: WDV, with analogue input and output, variable rate interrupt generator, and clock.

iii) Oscilloscope: 4 track slow scan.

iv) Stereo tape recorder: Phillips Pro-12.

v) Two envelope follower circuits each consisting of a simple rectifier with capacitor smoothing and a decay resistor.

ASPA - The programmes

Since the actual details of interface sampling, disk storage routines etc. will depend upon the user's own particular computer configuration the programmes are presented below in flow chart form with textual descriptions. The actual programmes, written in Fortran IV can be obtained, together with operating instructions, from the author.

i) CLEAR

CLEAR works entirely with disk output. The sole purpose of CLEAR is to clear all disk data files used to zero. This is necessary at the beginning of each run in order to avoid pick-up of data from previous runs. See Figure ii below.

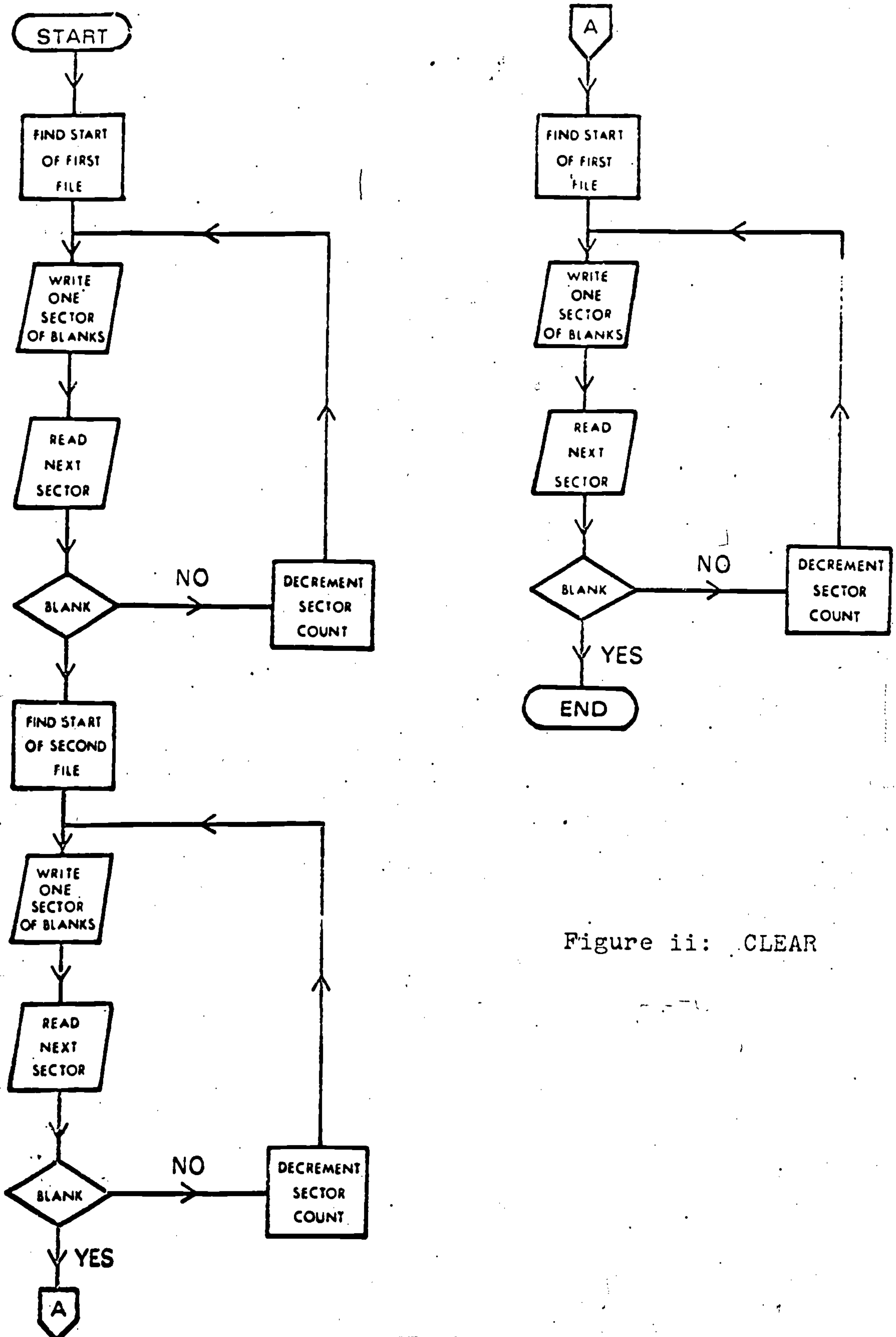

Figure ii: CLEAR

ii) OMOTS

OMOTS consists essentially of three programmes which sample the digitized envelope of the speech waveform at the interface, and store the results of the sampling on disk. The configuration of equipment necessary for operating this programme is shown in Figure iii and is explained in the text following.

The composite programme first reads a piece of paper tape which defines the maximum number of disk file sectors which can be used, the trigger levels and bias for sampling the envelope of the speech signal. The programme then prepares itself to accept interrupts and waits. This allows the operator to prepare the equipment, i.e. advance the experimental tapes to the point required and start the replay. When the Programme Start key on the computer keyboard is pressed, the generation by TIMON of level 3 interrupts at the interface is initiated. These are timed to arrive at 10 msec. intervals. On receipt of an interrupt the programme samples two of the interface analogue inputs, and decides whether they exceed a certain critical level, which is independently variable on both channels by means of the trigger levels, and bias defined on the paper tape. The programme then decides whether one or both channels have changed from one state to another since the previous sample.

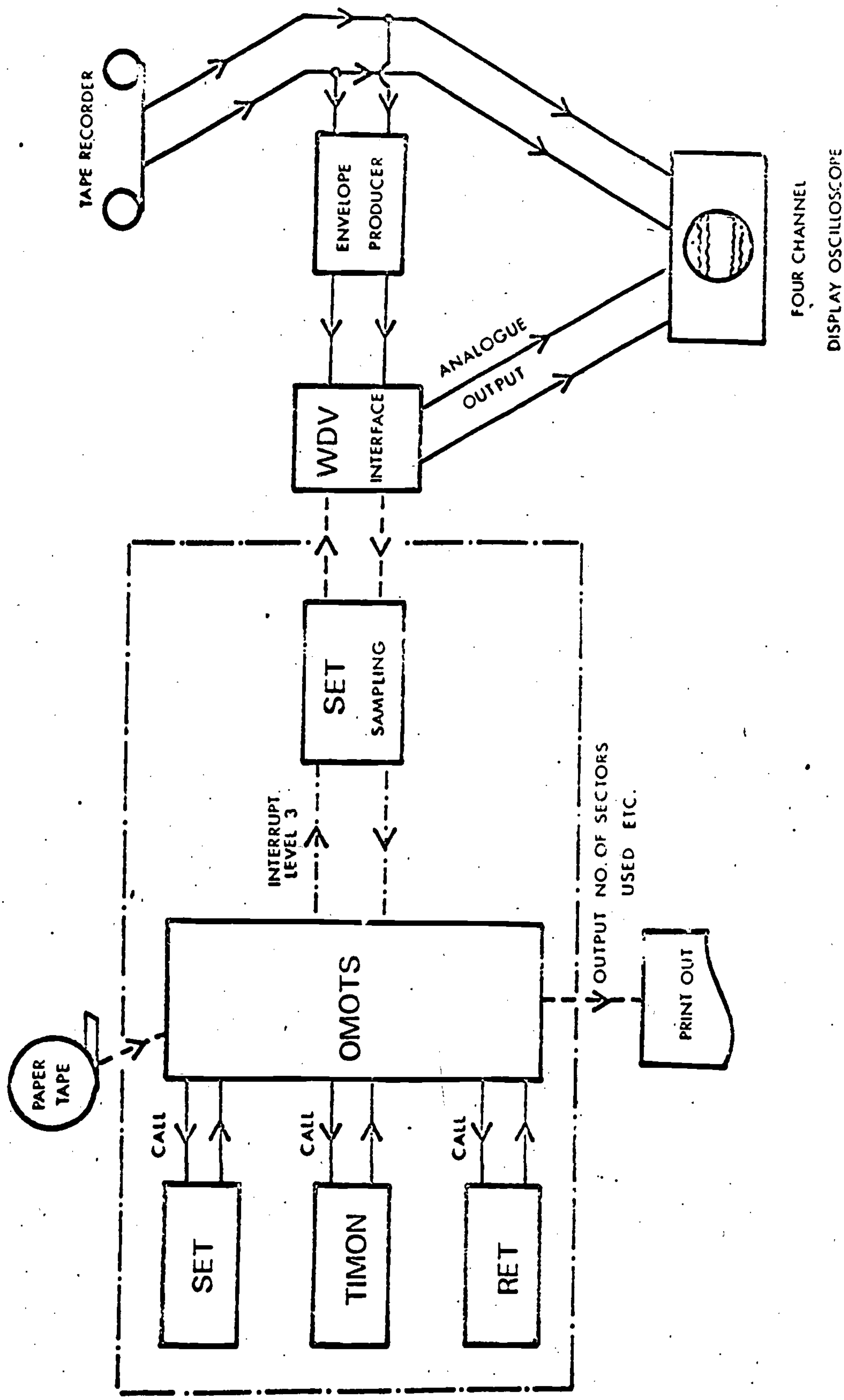

Figure iii: Input configuration.

It constructs a data set of the form TTTS and stores it in a buffer in the programme; where TTT = time in 1/100 sec. since the last change, and S = present state of the channels:

0 = both off

1 = Channel 1 on, 2 off

2 = Channel 1 off, 2 on

3 = both on.

This programme has two buffers so that when the first is full it can interchange them, storing data in the second whilst writing the contents of the first onto disk. This process of double buffering enables the programme to continue sampling unimpaired even though it is simultaneously sampling and transferring data to disk, and despite the fact that data is being produced from an irregular sampling of the channels. On receiving a signal from the keyboard (i.e. on termination of the relevant portion of tape being analyzed) the programme commands the interface to stop generating interrupts, writes the contents of the last buffer onto the disk and ends, leaving the computer ready for the next programme. The programme also has a facility for clamping both channels so that neither can change state more frequently than a pre-determined rate (in this case 5/100 sec.): this was incorporated to reduce the amount of data produced. OMOTS and the sequence of sub-programmes are shown in Figures iv and v following.

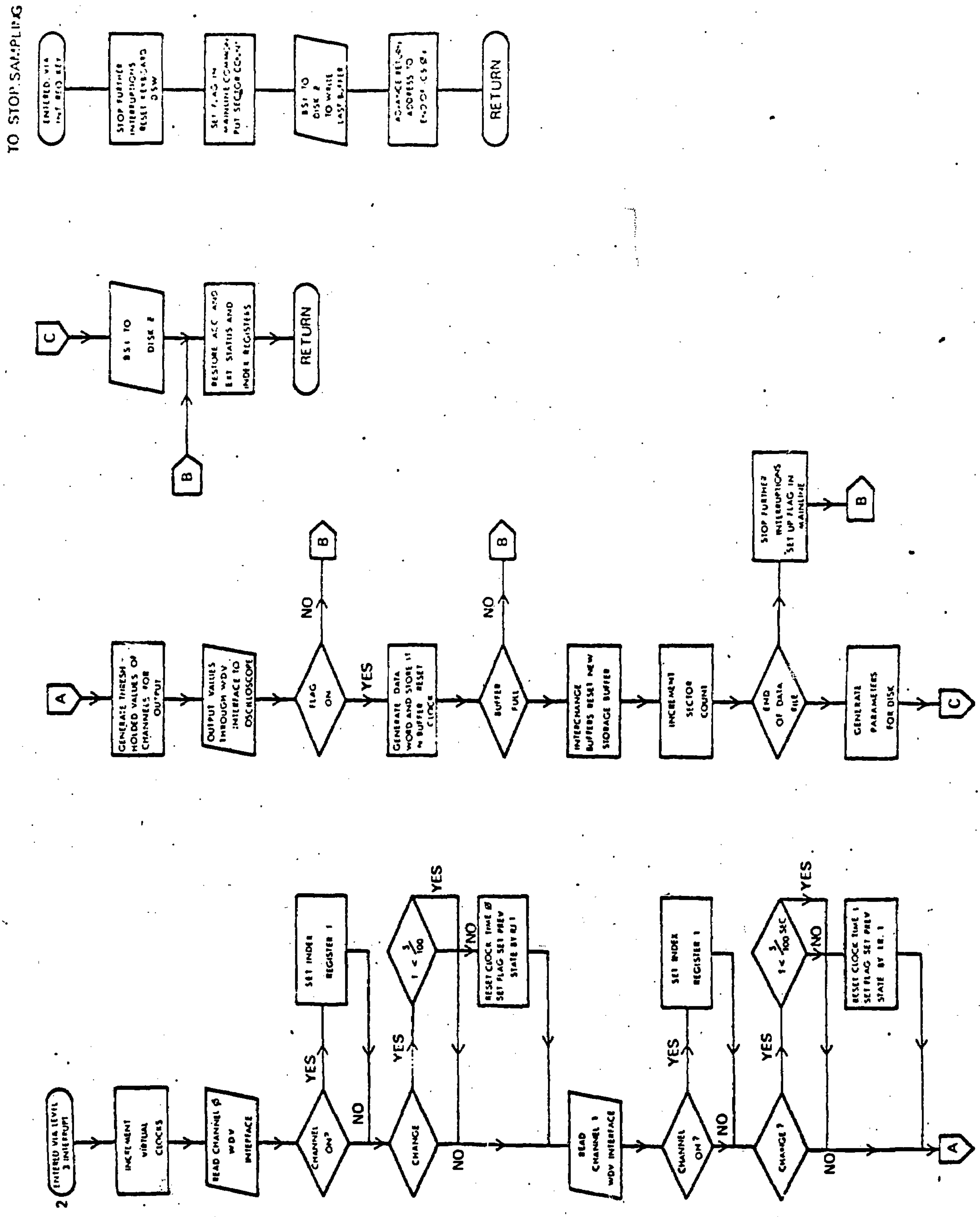

Figure v: OMOTS (SET, TIMON, RET).

The following inputs and outputs are employed with OMOTS:

i) Paper tape input describing various parameters: length of disk file to be used, trigger level, bias for differential trigger levels if required.

ii) Analogue inputs on the interface for monitoring tape recorder output. It should be stressed that the tapes to be analyzed should be as free as possible of any background noise. Where necessary a band-pass filter should be used to eliminate unwanted high or low frequency interference.

iii) Disk Output, for recording the data.

iv) Analogue outputs from the interface. These (as shown in Figure iii) are displayed on two channels of the 4 channel oscilloscope and provide the operator with a picture of what the programme "thinks it is hearing". These can be compared with the auditory signals and their displays on the other two oscilloscope channels, and the level of the tape recorder output can be adjusted until the operator is satisfied that the computer is satisfactorily tracking the recorded signals.

Before analyzing each tape OMOTS is run in order to set the appropriate levels. This is done using a test tape of known signal-silence times, and the actual experimental tape in question.

On setting up the equipment for use the test tape should be prepared and run through ASPA comparing the final output with the known test tape events and event times.

iii) SHUFF

This programme (see Figure vi) reads from disk and writes back onto disk.

It inverts the data produced by OMOTS, converting an assembler array into a FORTRAN compatible array. It also converts the data words into a new form TTTS, by building up a new word, using the time from the present data word, and the state from the previous data word. One data word is lost in this process. This programme is necessary since the data produced by OMOTS represent the time the system was in a state <u>before</u> it changed to its current state, whereas the final analysis requires the time the system is actually in a particular state.

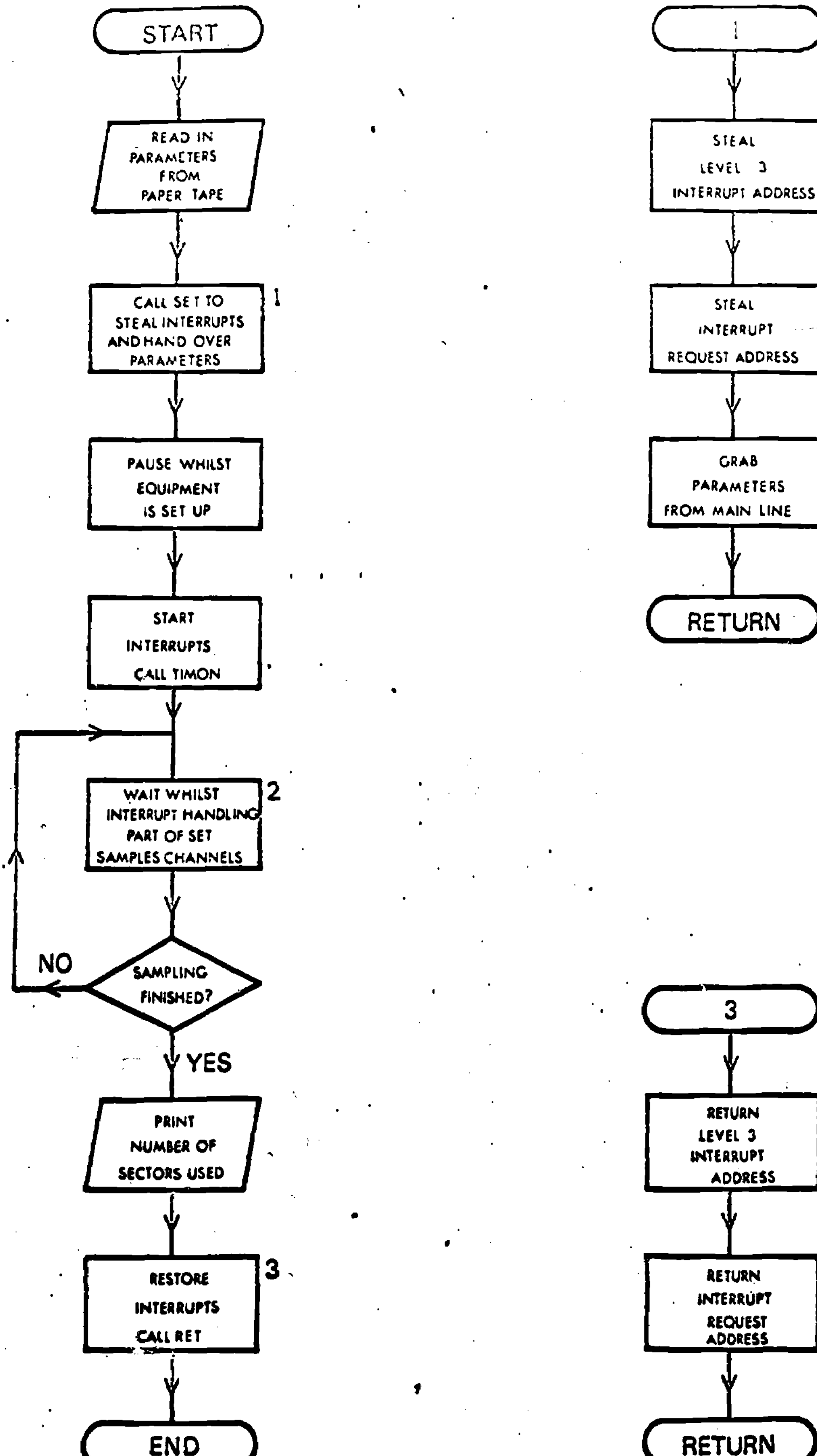

Figure iv: OMOTS.

vi) CLEAN

CLEAN is used to eliminate noise at the start of the record. This could be due to switching on of the tape recorder, noise on the tape etc. The programme (Figure ix) prints out the start (e.g. the first 10 events and times) of the data produced by the previous programmes, and waits for a number to be typed in. The number is selected by the operator on inspecting the data printout. For instance, in the present experiment all data sets should have started with a short interval (e.g. .8 seconds) with Channel 1 on, Channel 2 off. If the data set showed 0102 0053 0041 0033 0801, the operator would type in the last figure and all preceding data words would be eliminated for the final analyses.

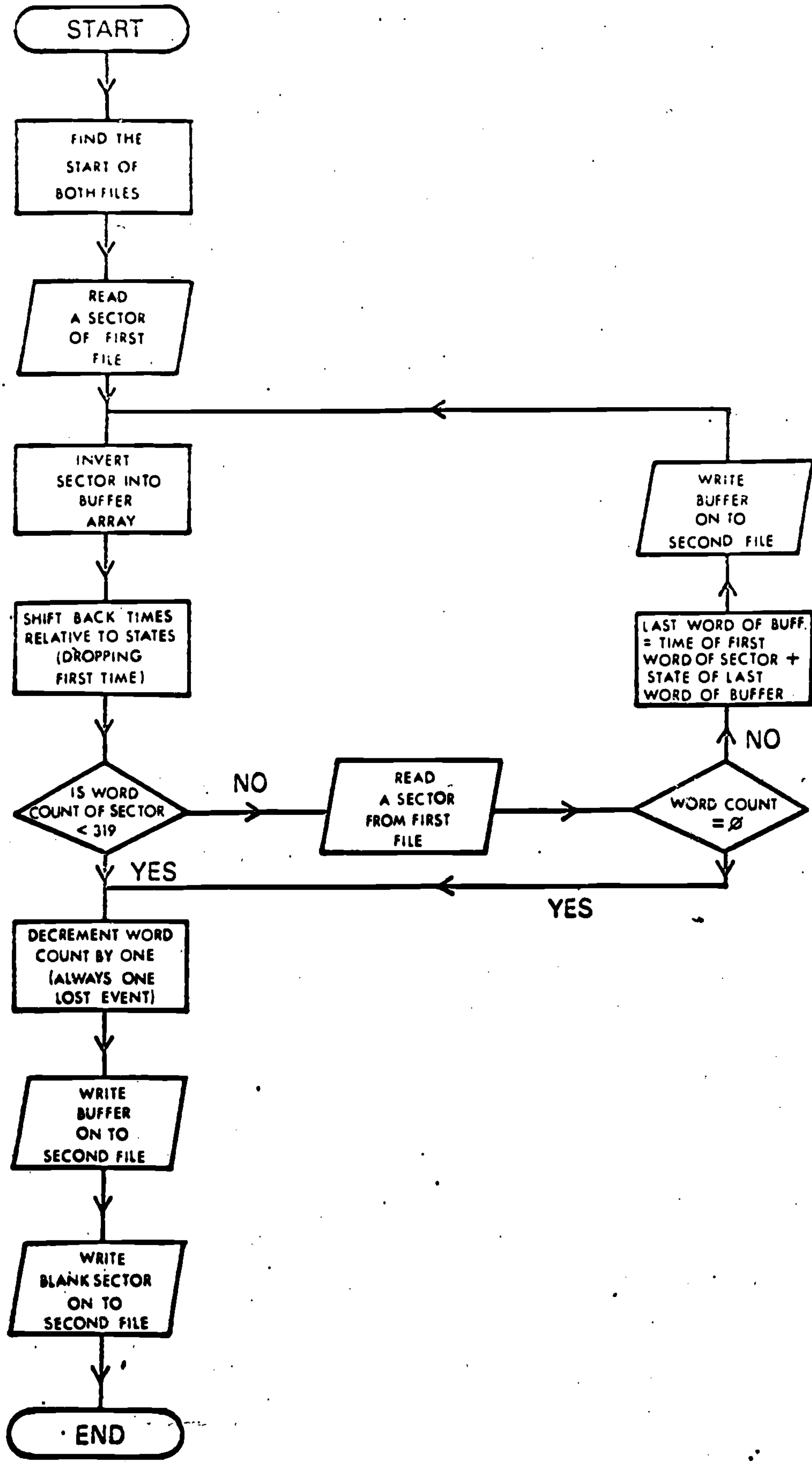

Figure vi: SHUFF.

v) CMPRS

This programme (Figure viii) also works from disk to disk, taking the data from DROPM and compressing any sequences of data words which have the same state into one data word which has as its time the sum of sequence times.

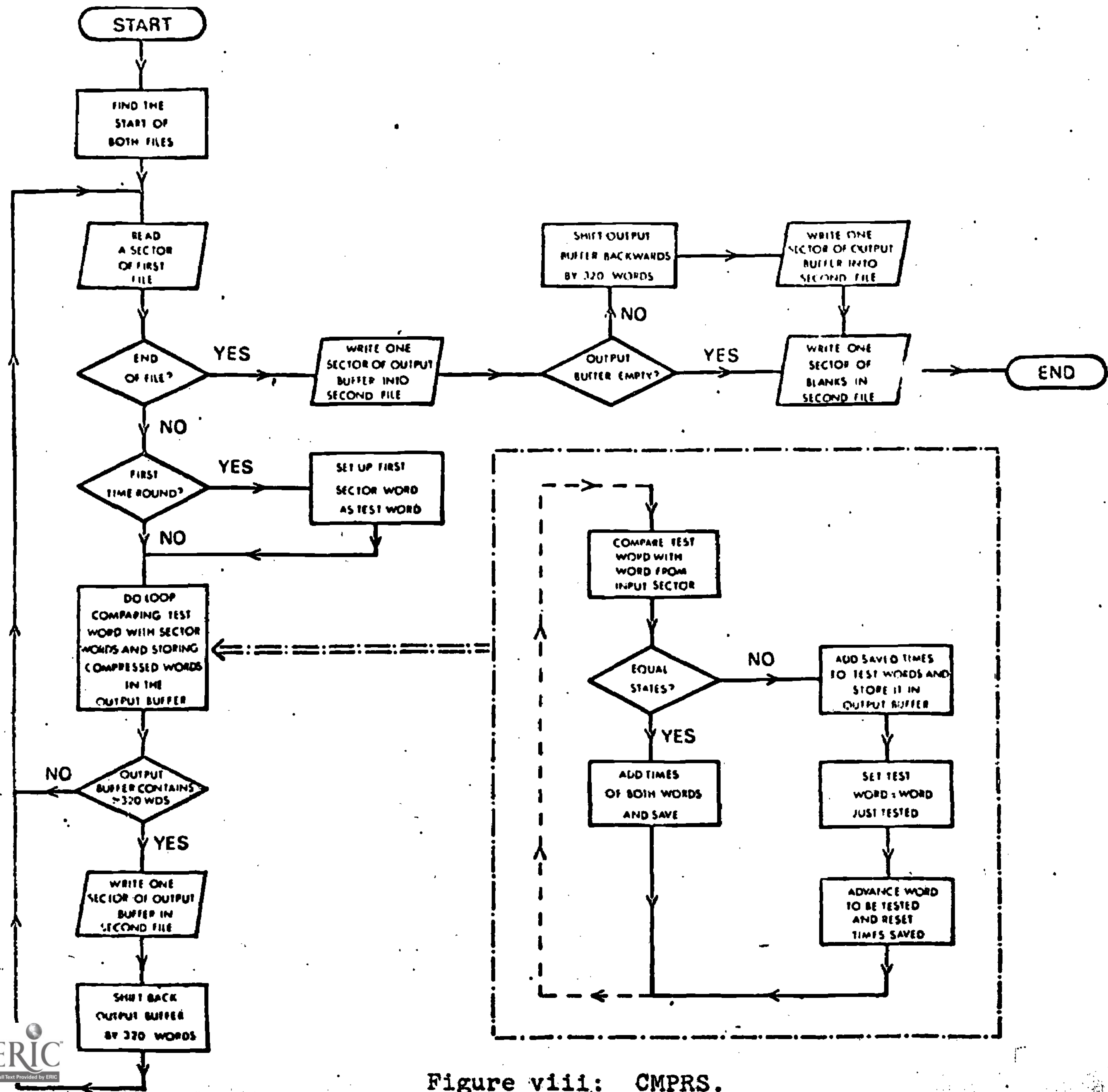

Figure viii: CMPRS.

iv) DROPM

This programme (Figure vii) reads from disk and writes back onto disk, preparing the data from SHUFF so that pauses of less than ¼ sec. can be eliminated, by effectively switching back on any channel which has switched off for less than ¼ sec. This interval can, of course, be varied to suit the user.

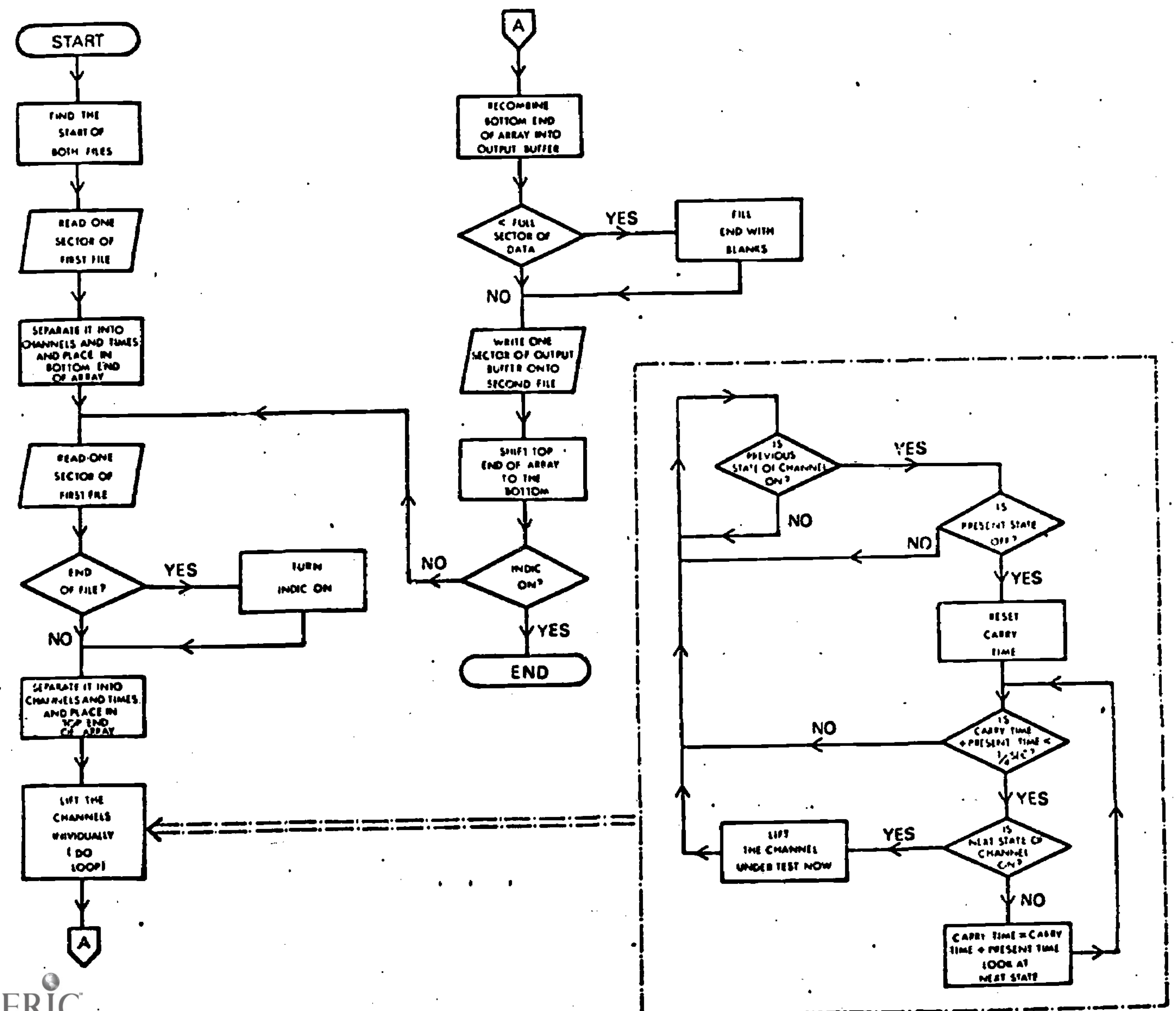

Figure vii: DROPM.

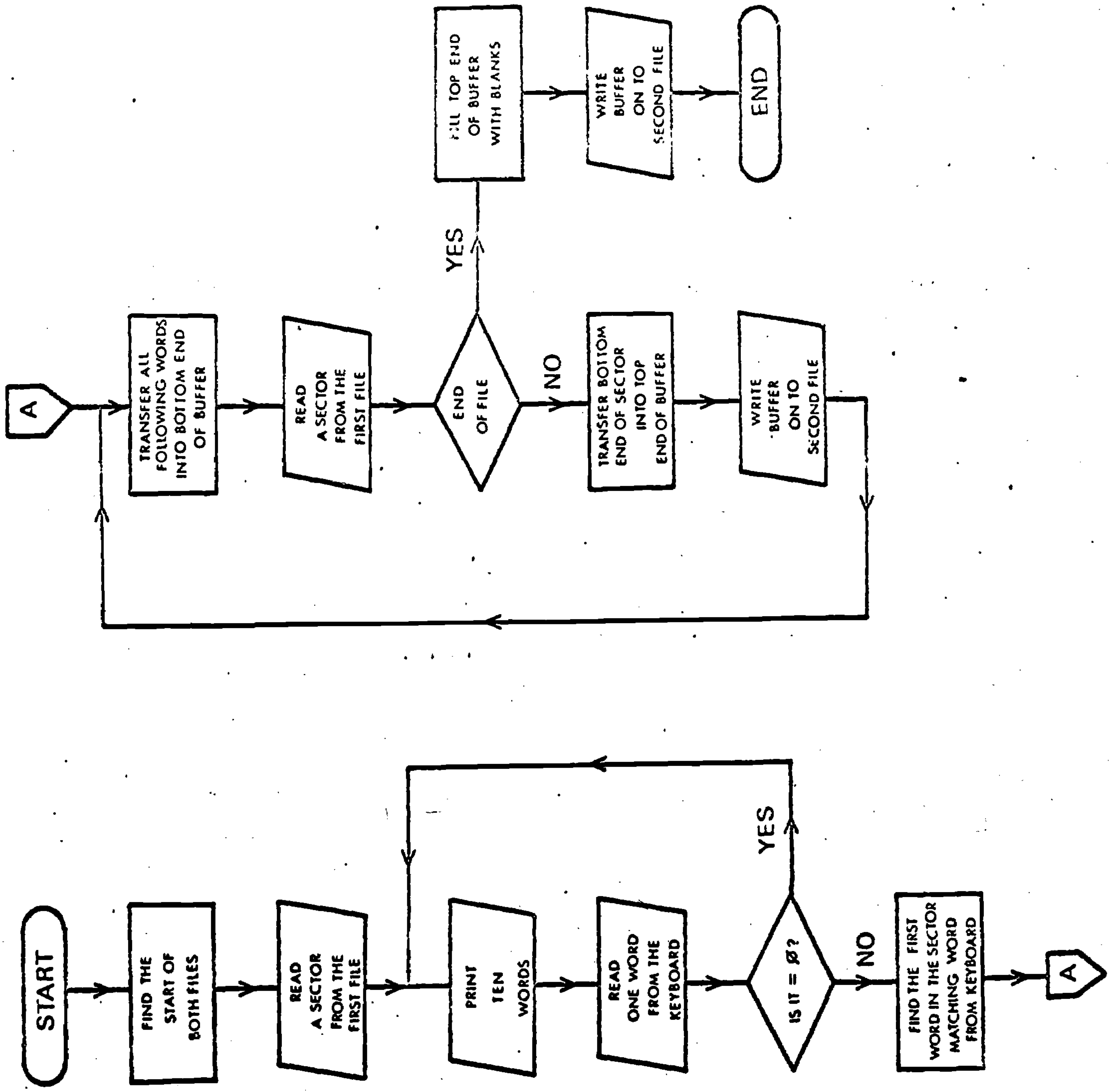

Figure ix: CLEAN.

vii) CRASH

CRASH reads from disk, analyzes the data and prints the results. Since the type of analysis required (e.g. time x state frequency distributions, means and variances of times spent in each state, etc.) depends on the user no description of this programme is provided here.

viii) SHOWD

Reads from disk files and prints the stored data. Again, this depends on the user's individual requirements and no further description is provided.

BIBLIOGRAPHY

Barik, H.C. (1969) Unpublished Ph.D. Thesis, University of North Carolina

Barik, H.C. (1971) A discussion of various types of omissions, additions and errors of translation encountered in simultaneous interpretation. Meta. Journal de traducteurs. 16, 199-210

Boomer, D.S. (1965) Hesitation and grammatical encoding. Language and Speech, 8, 148-158

Boomer, D.S., Dittman, A.T. (1962) Hesitation pauses and juncture pauses in spontaneous speech. Language and Speech, 5, 215-220

Broadbent, D.E. (1957) Immediate memory and simultaneous stimuli. Quart. J. exp. Psychol. 9, 1-11

Broadbent, D.E. (1958) Perception and Communication. London: Pergamon Press

Broadbent, D.E. (1971) Decision and Stress. London: Academic Press

Carey, P.W. (1971) Verbal retention after shadowing and after listening. Perception and Psychophysics, 9(1-13) 71-83

Carroll, J.B. (1966) An experiment in evaluating the quality of translations. Mechanical Translation, 9, 55-66

Edwards, A.L. (1968) Experimental Design in Psychological Research. New York: Holt, Rinehart, and Winston

Gerver, D. (1969) The effects of source language presentation rate on the performance of simultaneous interpreters. In Proceedings of the 2nd Louisville Conference on Rate and/or Frequency Controlled Speech. University of Louisville, Louisville, 1969.

Goldman-Eisler, F. (1968) Psycholinguistics: Experiments in Spontaneous Speech. London: Academic Press

Jensen, A. (1965) Review of M.P.I. in D.K. Burros (Ed.) The Sixth Mental Measurements Yearbook. Highland Park: Gryphon Press. pp 288-291

Jerison, H.J., & Wing, S. (1957) Effects of noise and fatigue on a complex vigilance task. WADA tech. rep., 57-14

Light, E., & Howitt, J.B. (1966) General Studies French. Oxford: Pergamon Press

Maclay, H., & Osgood, C.E. (1959) Hesitation phenomena in spontaneous English speech. Word, 15, 19-44

Martin, J.G. (1967) Hesitations in the speaker's production and listener's reproduction of utterances. J. verb. Learn. verb.Behav., 6, 903-909

Martin, J.G. (1970) On judging pauses in spontaneous speech. J. verb. Learn. verb. Behav., 9, 75-78

Martin, J.G., & Strange, W. (1968) The perception of hesitation in spontaneous speech. Perception and Psychophysics, 3, 427-438

Miller, G.A. (1963) Review of J.H. Greenberg (Ed.), Universals of Language. Contemp. Psychol., 8, 417-418

Miller, G.A., Galanter, E., & Pribram, K.H. (1960) Plans and the Structure of Behaviour. New York: Holt, Rinehart, and Winston.

Miller, G.A., Heise, G.A., & Lichten, W. (1951) The intelligibility of speech as a function of the context of the test materials. J. exp. Psychol., 41, 329-335

Oleron, P., & Nanpon, H. (1965) Recherches sur la traduction simultanee. Journal de Psychologie Normal et Pathologique, 62, 73-94

Pollack, I., & Rubenstein, R. (1963) Response times to known message-sets in noise. Language and Speech, 6, 57-62

Rabbitt, P.M.A. (1966) Recognition: Memory for words correctly heard in noise. Psychon. Sci., 6, 383-384

Rabbitt, P.M.A. (1968) Channel-capacity, intelligibility and immediate memory. Quart. J. exp. Psychol., 7, 176-195

Ramsay, R.W., & Law, L.N. (1966) The measurement of duration of speech. Language and Speech, 9, 96-102

Stennet, R.G. (1957) The relationship of performance level to level of arousal. J. exp. Psychol. 54, 54-61

Suci, G.J. (1967) The validity of pause as an index of units in language. J. verb. Learn. verb. Behav., 6, 26-32

Tannenbaum, P.H., Williams, F., & Hillier, C.S. (1965) Word predictability in the environment of hesitations. J. verb. Learn. verb. Behav., 4, 134-140

Treisman, A.M. (1965) The effects of redundancy and familiarity on translating and repeating back a foreign and annative language. Brit. J. Psychol., 56, 369-379

Van Hoof, H. (1962) Theorie et pratique de la traduction. Munich: Max Hueber

Welford, A.T. (1968) Fundamentals of Skill. London: Methuen

Winer, B.J. (1962) Statistical Principles in Experimental Design. New York: McGraw-Hill

Woodhead, M.M. (1958) Effects of bursts of loud noise on a continuous visual task. Br. J. industr. Med., 15, 120-125

Addendum

Jensen, A.R. (1962) Extraversion, neuroticism a serial learning. Acta Psychologica 20, (2) 69-77.