43196

National Library of Canada | Bibliothèque nationale du Canada

CANADIAN THESES ON MICROFICHE | *THÈSES CANADIENNES SUR MICROFICHE*

NAME OF AUTHOR/*NOM DE L'AUTEUR* LINDA ANDERSON

TITLE OF THESIS/*TITRE DE LA THÈSE* SIMULTANEOUS INTERPRETATION : CONTEXTUAL AND TRANSLATION ASPECTS

UNIVERSITY/*UNIVERSITÉ* CONCORDIA

DEGREE FOR WHICH THESIS WAS PRESENTED/ *GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* MASTER'S OF ARTS

YEAR THIS DEGREE CONFERRED/*ANNÉE D'OBTENTION DE CE DEGRÉ* 1979

NAME OF SUPERVISOR/*NOM DU DIRECTEUR DE THÈSE* DR. TANNIS MAAG

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/*DATÉ* March 30, 1979 SIGNED/*SIGNÉ* Linda Anderson

PERMANENT ADDRESS/*RÉSIDENCE FIXÉ* 11 Salisbury Road
Pointe Claire P.Q.
H9S 3Z2

NL-91 (11-79)

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ MICROFILMÉE TELLE QUE NOUS L'AVONS REÇUE

SIMULTANEOUS INTERPRETATION: CONTEXTUAL AND TRANSLATION ASPECTS

Linda Anderson

A Thesis

in

The Department

of

Psychology

Presented in Partial Fulfilment of the Requirements
for the degree of Master of Arts at
Concordia University
Montreal, Quebec, Canada

March 1979

CONCORDIA UNIVERSITY

SIR GEORGE WILLIAMS FACULTY OF ARTS

Graduate Studies

This is to certify that the thesis prepared

By LINDA ANDERSON

Entitled SIMULTANEOUS INTERPRETATION: CONTEXTUAL AND TRANSLATION ASPECTS

Complies with the regulations of this University and meets the accepted standards with respect to originality and quality

For the degree of

MASTER OF ARTS IN PSYCHOLOGY

Signed by the final examining committee.

Chairman

Approved by the Dean of Faculty or Chairman of Department (or authorized representative)

18/5/ 1979

Name
Dean of Faculty or Department
Chairman

Name
Supervisor of Thesis

Name
Chairman, Committee of Graduate Studies

Abstract

Simultaneous interpretation: Contextual and Translation Aspects

Linda Anderson

The present study was designed to explore aspects of simultaneous interpretation which have received little experimental attention. Two experiments are reported which assessed the influence of linguistic and extra-linguistic context of a message on its interpretation by 12 professional interpreters. In Experiment 1, subjects interpreted formal texts read during a videotaped conference under three conditions: with the complete text available beforehand, a précis beforehand, and with no prior knowledge beyond the topic. In Experiment 2, interpretations were compared where the conference was visible on videotape and where it was not. The prior access to linguistic context/content and availability of visual information had no noticeable effect on interpretation quality. Reasons were advanced to account for this finding. A third experiment explored the central translation aspect of interpretation to ascertain the significance of the language change in the task. Output by the same 12 subjects was compared during interpretation requiring the input-output language switch (French to English) with interpretation where it was not (English to English), together with output during two control tasks involving verbatim repetition of French and English. Unilingual interpretation led to output of significantly greater intelligibility, but bilingual interpretation took no longer to perform and produced no significant losses in amount of information conveyed. The results were interpreted as suggesting that some "translation" procedure of matching up input and output language items occurs in normal simultaneous interpretation, but its extent is unclear.

ACKNOWLEDGEMENTS

The author's deep gratitude is extended to Dr. Tannis Maag for her unflagging interest and invaluable assistance during the extended process of realizing this most unwieldy of projects, particularly for her suggestions during the writing of this thesis.

I wish also to express my grateful appreciation to Mr. William Mundl of the Concordia University Psychology Laboratory for his efforts in resolving the technical problem of latency sampling, to Dr. Norman Segalowitz for his suggestions on the thesis, to Mr. Brian Harris of the University of Ottawa School of Translators and Interpreters for his interest and support and to Dr. David Gerver for his suggestions and insights into the experimental study of interpretation.

Special thanks go to Mr. Pierre Meurice for his willingness to undertake a thankless job and for his painstaking and uncommonly prompt evaluation of subjects' transcripts.

Table of Contents

Simultaneous interpretation is a phenomenon of both cognitive and linguistic interest: cognitive, because of the information processing involved in the task as well as the interpreter's apparent ease in juggling several concurrent operations; and linguistic, because of the type of information processing he does, the recoding of a message heard in one language into speech in another language. The present thesis is a report of three experiments conducted into both cognitive and linguistic aspects of interpretation. The first two look at contextual factors surrounding performance of the task, while the third focusses on its central recoding or "translating" stage. Before reporting on the three experiments, however, it seems useful, because the phenomenon is still poorly understood, first to clarify what simultaneous interpretation is and second, to look at the strategies previous investigators have used to study the phenomenon, as well as what they have found out about it.

Despite its misnomer, "simultaneous translation", simultaneous interpretation is not simply translation in a paced and oral form. In theory, it is distinguished from translation in that its focus is not on the words of a language, but on ideas which are expressed in speech. Its aim is not to establish linguistic equivalents between two languages, but to communicate the meaning of a speech. Being the reconstruction of oral messages, interpretation is in theory perhaps more akin to paraphrasing than to the code-switching of translation. In practice, though, interpretation is thought to be a mixture of the two (Seleskovitch, 1976).

In addition to being considered a "translator" of speech, the simultaneous interpreter may also be viewed as a complex information processing device confronted by a paced, auditory, tracking task, and required to monitor, store, retrieve and "translate" input in one language, while at the same time speaking in another language, and monitoring and occasionally correcting his own output (Gerver, 1971). Sequestered from the conference in his soundproof booth, he must analyze and comprehend a continuous stream of speech; shape, coax and contort the message into the mould of his output language, and then speak his version of it. He is at the same time analyzing the oncoming stream of new information and monitoring his own speech to ensure that it is properly paced, is intelligible to his listeners, and conveys the original meaning as he understood it, fully cognizant of the fact that there is no possibility, when he is in doubt, of asking for repetition or clarification of any point.

Despite general recognition that simultaneous interpretation can and does appear to work and has been in ever wider use since its inception over thirty years ago, relatively little is known about how or under what conditions it works, or about the parameters involved in the simultaneous interpretation of speech.

Most works on simultaneous interpretation (SI) are manuals giving general descriptions of the interpreter's task (Herbert, 1965; Seleskovitch, 1968; Van Hoof, 1962). The first lengthy discussion of it appeared in an unpublished master's thesis (Paneth, 1957), which did raise various issues of experimental interest, such as the time lag or ear-voice span separating interpreter output from source speaker input, the use that interpreters might make of pauses in the input speech

and the segmentation the interpreter himself performs on input, all topics taken up by later researchers. Kade and Cartellieri (1971), foreshadowing later research into attentional mechanisms in SI, discussed various methods which could help the interpreter cope with his parallel activities of listening to speech while producing it. Hanna (1966), discussing listening conditions and language redundancy, asserted that any interpreter could perform effectively without 40 per cent of what the source speaker says. The literature on simultaneous interpretation abounds in such statements of how the interpreter manages to do his job and the input conditions which help or hinder his task. These assertions may seem valid enough intuitively, but they are still largely untested experimentally. This is perhaps because it is still a relatively new phenomenon or because the presumed complexity of cognitive processes intervening between speech input in the source language and speech output in the target language can be daunting, dampening the ardour of anyone wishing to dissect the phenomenon with any degree of experimental rigour.

The earliest experimental research into SI often employed bilinguals as subjects rather than trained interpreters, and used simultaneous interpretation as a task incidental to the main concerns of the experimenter which were to investigate such cognitive phenomena as memory, attention or message processing (Goldman-Eisler, 1967; Lawson, 1967; Treisman, 1965). These were followed by general exploratory studies which focussed on various parameters of the simultaneous interpretation task, particularly temporal ones. Oléron and Nanpon (1965) attempted to study interpretation under real conference conditions, but then shifted to studying the performance of three professional interpre-

ters, working from and into different languages, under laboratory conditions. Barik (1971, 1972a, 1973, 1975) used six subjects, who varied in their interpretation ability, to focus on both temporal and linguistic aspects of interpreter performance. Gerver (1971), using greater numbers of professional interpreters as subjects, conducted the first large-scale attempt to manipulate various parameters of input experimentally.

In much of the research using simultaneous interpretation as an experimental task, the interpreter has been viewed primarily as an information processor, with the particular nature of his processing, the recoding of speech into another language, being secondary. The general strategy has been to manipulate some aspect of the input information and to observe the effects of the manipulation on the interpreter's output. Use of this strategy poses two methodological problems: one of measuring the effect of the input manipulations on the interpreted speech; and one of demonstrating that the effects observed are specific to the process of simultaneous interpretation itself and not to the activities of paced listening or speaking or performance of the two simultaneously.

One approach to the problem of measuring the speech output of the interpreter has been to use a latency measure of the time lag or ear-voice span (EVS) separating words in the interpreter's speech from corresponding words in the source speaker's speech. This interval is considered to reflect the time involved in the various stages of processing the incoming information (Barik, 1971, 1972a, 1973, 1975; Gerver, 1971; Oléron & Nanpon, 1965). The characteristic time lag interpreters show is, in information processing terms, considered to be due to the

accumulation of source language items in some sort of buffer storage while the central processor is working on previously received information. Longer lags presumably aid this processor as they afford larger chunks to operate on, but because of assumed limitations in capacity of the buffer store, a very long lag can lead to displacement of items from storage and result in omissions in output (Gerver, 1974a). Conversely, short lags can reduce cognitive load, but may lead to errors in output, since interpreters may be working on too small a semantic chunk for proper apprehension of a message (Barik, 1973).

The ear-voice span is an approximate measure only, since interpretation involves the transformation of ideas delivered orally into target language, and hence there is no guarantee that translation equivalents, of the nature of "chien" = "dog", will appear in the interpreted output. Further, given that two different languages are involved, matching the corresponding units of meaning in the source and target languages rests to some extent on the subjective judgement of the investigator. Despite these difficulties, a temporal measure of the interpreter output has been shown to be useful, since it varies with input manipulations such as information presentation rates and speech/pause ratios (Barik, 1973; Gerver, 1971; Oléron & Nanpon, 1965), speech length (Oléron & Nanpon, 1965; Barik, 1973, 1975) and input language (Oléron & Nanpon, 1965; Treisman, 1965). By itself, however, the time lag measure is not sufficient. There appears to be a relatively "characteristic" lag related to the task (Gerver, 1971; Treisman, 1965), as well as a maximum lag, which presumably reflects the limits of short-term memory. If task difficulty is such that subjects are already operating at maximum ear-voice span, further increases in task difficulty

will not affect EVS, but will instead result in poorer quality of interpretations (Barik, 1975; Gerver, 1971, 1974a).

The quality of the interpreted output must therefore be quantified in some fashion. Attempts to do this have used a variety of methods to come up with some numerical value for the amount of source input that comes through in interpreter output. The earliest investigators (Lawson, 1967; Oléron & Nanpon 1965; Treisman, 1965), although recognizing the importance of units of meaning in the interpreter's task, measured interpretation quality by counting the number or percentages of words correctly rendered in interpretations. The drawback of this word-count approach is that interpreters are not engaged in translating words, nor in translating a maximal number of the original words. Human and machine translation studies (Hockett, 1954; Yngve, 1955) have advocated even the sentence as the proper translation unit; and simultaneous interpretation researchers (Barik, 1975; Gerver, 1971; Goldman-Eisler, 1972b) have repeatedly found semantic/grammatical phrases to be the unit of analysis interpreters commonly work on. Some investigators compounded the problem of word-count analyses by comparing word differentials in interpreted versions with written translations of the same passages (Lawson, 1967; Oléron & Nanpon 1965). As has already been mentioned, the nature of the interpretation task precludes the type of lexical correspondences between the original and interpreted versions that one would expect in written translation.

Later investigators (Barik, 1971, 1975; Gerver, 1971, 1974a), refined the analysis of interpreter utterances by classifying output into general categories of omissions, additions and substitutions together with words correct, in the hope that these would provide clues to the

processes involved in the task. There does seem to be some agreement about the reliability of such categories in assessing translation quality (Brislin, 1976). However most researchers, despite the fact that any measure of the quality of interpretations must take account of whether they are actually understandable to listeners, excluded even intelligibility from their assessments (Barik, 1975; Oléron & Nanpon, 1965). In these, as in many attempts at content analysis in translation in general, the systems used have generally been subjective and tend to appeal to the experimenter's intuitions about accurate translation. Such subjectivity may be inevitable, but more objective techniques are possible (Brislin, 1976; Macnamara, 1967a). The best solution in simultaneous interpretation research so far seems to be the method Gerver (1971, 1974a) used: in addition to the categories of interpretation quality mentioned above, he used independent raters to evaluate interpreters' transcripts according to two nine-point scales developed by J.B. Carroll (1966) for comparing human and machine translations. These scales were designed to measure the intelligibility and informativeness of the transmitted message, and as such they reflect much more closely than do the above categories the actual aims of the interpreter at work.

The second methodological problem in studies of simultaneous interpretation, that of demonstrating that the effects of the input manipulations are specific to the interpretation process, is a difficulty only in studies which are concerned with establishing the nature of the process. Many investigators have examined the effects of various input parameters as they might occur in the practice of interpretation, without being concerned with whether these effects were on the interpreta-

tion process or not. However, if one wishes to draw conclusions about central processes peculiar to the interpreter's task, some sort of control is necessary. In fact, a small number of investigators have handled the problem by including a shadowing control group (Gerver, 1971, 1974a; Treisman, 1965). Shadowing, involving the word-for-word repetition of auditory input, controls for the more "peripheral" aspects of the interpretation task, that is, the processes involved in simultaneously speaking and listening to speech. Studies which have used a shadowing control have consistently observed that the ear-voice span for interpretation is longer than that for shadowing, suggesting that the level of processing required in interpretation is more complex (Gerver, 1974a). Treisman (1965) suggested that shadowing may involve only one central representation which is shared by input and output. A related finding, that the unit of analysis in interpretation is a major sentence constituent or phrase (Gerver, 1971; Goldman-Eisler, 1972b; Treisman, 1965), as opposed to a single word in shadowing (Gerver, 1971; Treisman, 1965), also points to differences between shadowing and interpretation that may be relevant to the processing involved in the latter. Thus, a shadowing control appears appropriate in any study in which inferences are being drawn about central processes of interpretation.

Turning now to a more detailed consideration of the findings on simultaneous interpretation, these can be subdivided into two broad categories, (a) characteristics of the input material and (b) characteristics of output and input/output relations. Most research by far has focussed on the first category.

Characteristics of the Source Material

Among the parameters of the input material which have been experimentally shown to affect simultaneous interpretation are semantic and syntactic redundancy of the input speech (Treisman, 1965), input presentation rate (Barik, 1972a, 1975; Gerver, 1971; Goldman-Eisler, 1967, 1972b; Oléron & Nanpon, 1965; Treisman, 1965), the length and nature of the input speech (Barik, 1975; Oléron & Nanpon, 1965), and the language in which it is spoken (Barik, 1975; Goldman-Eisler, 1972b; Treisman, 1965).

Redundancy. Writers on simultaneous interpretation have stressed the usefulness of source language redundancy in the interpreter's task, suggesting that it might facilitate decoding of input under high levels of signal distortion and interference (Hanna, 1966), or might help the interpreter make up time in his sentence comprehension (Kade & Cartellieri, 1971). Treisman (1965) experimentally manipulated source language redundancy and, using unskilled subjects, compared the effect of interword constraints, both syntactic and semantic, on the efficiency of performance in shadowing and interpreting. Information load, manipulated in passages with various orders of approximation to English and French, was found to have a greater effect on the number of words correctly transmitted in the interpreting task than in the shadowing. Semantic constraints were found to be more helpful in interpretation than syntactic ones, while the reverse was true for shadowing. Gerver (1976) referred to an experimental finding with Russian-English interpreters: 75 per cent of their responses to passages designed to instill a misleading set appeared to be according to subjects' expectations, rather than what they actually heard. Skilled interpreters have been found to

make errors that can be traced back to erroneous expectations about the development of sentences (Seleskovitch, 1968).

Presentation Rate. Interpreters themselves generally agree that there is an optimal and all-important rate for input speech, of roughly 100 to 120 words per minute (Seleskovitch, 1968; Thiéry, 1974). Speech rates both faster and slower than this optimum are considered to lead to decrements in performance (Seleskovitch, 1973). Experimental studies manipulating input presentation rate have all used rates faster than the presumed optimum. Across studies, rate has been manipulated in several different ways: as information rate in bits per second (Treisman, 1965); as broadly-defined speech/articulation rates and speech/pause ratios (Barik, 1973, 1975; Goldman-Eisler, 1967, 1972b; Oléron & Nanpon, 1965); and as controlled speech/pause ratios in compressed speech (Gerver, 1971).

Findings have not been consistent throughout, but subjects have generally been found to cope with faster presentation rates by increasing the length of the EVS somewhat and decreasing its variability (Barik, 1973; Gerver, 1971; Oléron & Nanpon, 1965). Most studies however report that between subjects, the average ear-voice span levels off at less than 5 seconds, normally 2 to 3 seconds or 4 to 5 words. There are marked individual differences among interpreters, but even those who work with long lags seem to increase it only up to a maximum of 10 seconds (Goldman-Eisler, 1967; Oléron & Nanpon, 1965). A relatively consistent time lag within subjects was also found in these studies and is considered to reflect the interpreter's having learned how fast he can process input without losing it through forgetting

(Barik, 1973; Gerver, 1971). Once the maximum limit on time lag was reached, further increases in input rate were found to be associated with number and length of omissions of input (Barik, 1975). Treisman (1965) found presentation rate to affect the interpreting group more than the shadowing controls in terms of words correctly rendered, but not in terms of ear-voice span.

The only extensive investigation of the effects of input rate on simultaneous interpretation, as performed by professional interpreters, was done by Gerver (1971) as part of his doctoral thesis. Subjects were assigned to either interpret or shadow French input at five controlled speech rates. He found first that the interpreting group corrected more and omitted more and longer passages than the shadowing group. Up to a presentation rate of 120 words per minute, the strategy that interpreting subjects appeared to use to keep pace with input was to speak more and pause less, thereby maintaining a fairly steady output rate. As input rates increased to higher levels, subjects in the interpreting task paused more and longer and spoke less than the shadowing controls, who continued to keep pace with input by pausing less and speaking more. Thus, the faster presentation rates had little effect on shadowing ear-voice spans, which ranged from 1.24 to 1.31 seconds, but did effectively increase the interpreting EVS, from 3.1 seconds at 95 wpm to 5.15 seconds at 164 wpm. With each increase in presentation rate starting at 95 words per minute, the amount correctly interpreted decreased and omitted material was longer (from 96.3% correctly rendered at 95 wpm down to 58.1% at 164 wpm), with proportionately fewer self-corrections being made of errors. From these findings, Gerver inferred that interpretation involves more processing than shadowing does and

concluded that faster rates affect not the perception but the processing of input. The optimal input rate for interpretation, according to Gerver's measures of errors of commission, omission and self-corrections, appeared to be between 95 and 120 words per minute. Past this rate, the intervals between items may have become shorter than the time needed to process them, and subjects thenceforth appeared to work in bursts, lengthening pausing time and speaking less.

Speech Length. Length of the input speech has been found to affect the accuracy of output and the response latencies of interpreters, with diminished accuracy and gradual lengthening of the ear-voice span up to a maximum of 10 seconds, the longer the source speech lasted (Oléron & Nanpon, 1965). Barik (1975) also found a tendency for his trained subjects to lag further behind with longer input speech. However, Oléron and Nanpon (1965) compared performance only between isolated words, sentences and paragraphs of various lengths, and confined themselves to counting the number of words added or omitted in the interpretations.

Source Language. Particular input languages have been shown to have an effect on the percentage of words correctly rendered as well as on ear-voice spans (Treisman, 1965), on the way interpreters segment input (Goldman-Eisler, 1972b), on the interpreter's speech rates (Barik, 1975), and on the number of errors and omissions in interpretations (Barik, 1975). However, since these studies used language groups that were small, heterogeneous and variable in their interpretation skills, such findings cannot resolve the question of whether interpretation from or into a particular language may be faster or somehow better.

Type of Input. Writers on simultaneous interpretation have continually stressed that notwithstanding the reading aloud of written texts that

occurs frequently in the conference situation, interpretation is properly intended for extemporaneous speech, in the same way as translation is for the written text (Seleskovitch, 1976). Barik (1973, 1975) manipulated type of input speech and found that the more prepared, less spontaneous presentations led to greater speech and articulation ratios of interpreter speech over original speech, with fewer additions and more, longer omissions occurring. Barik related these effects directly back to the generally greater speech rates and speech/pause ratios in the more prepared materials.

In summary, the features of the input material that have been experimentally investigated, such as speeded presentation rate, predictability and amount, type, and language of input have been demonstrated to have an effect on both the temporal characteristics and the contents of the interpreted versions. Where a shadowing control was used, input rate and redundancy have been shown to have greater effects on interpretation than shadowing (Gerver, 1971; Treisman, 1965), suggesting that such effects are on the complex language processing of simultaneous interpretation rather than on the peripheral activities of speech perception and production.

Formal Characteristics of Output and Input-Output Relations

Various investigators have examined the formal characteristics of interpreted speech and compared such features as speech/pause times and speech segments in the original and interpreted versions in an attempt to understand the attentional mechanisms of simultaneous interpretation. They have found, first of all, that simultaneous interpreters have their own particular way of segmenting the input speech: these segments are evidenced in the chunks of source language speech which the interpre-

ter's ear-voice span encompasses and are considered to be the units of meaning which the interpreter is processing (Goldman-Eisler, 1967, 1972b; Oléron & Nanpon, 1965; Treisman, 1965). Goldman-Eisler 1972b) found that interpreters had a "constructive" way of segmenting the input speech: in about 90 per cent of cases, interpreters tended to ignore input segmentation and impose their own, although identical segments in the two versions did occur more often for slow input speech than for fast. Although the type of chunking did vary somewhat with the input-language considered, English, French and German, 90 to 95 per cent of the time the minimum EVS segment consisted of a subject-predicate expression (NP + VP), with the VP appearing to be the critical determiner of meaning for the interpreter. The interpreter, in various studies (Goldman-Eisler, 1972b; Oléron & Nanpon, 1965; Treisman, 1965), seemed to wait to hear the syntactic and semantic structure of the phrase before beginning to interpret it. Treisman (1965) reported that her subjects attempted to interpret even random word strings as though they were phrases.

Writers on simultaneous interpretation have suggested that pauses in the input speech are very valuable to the interpreter, as they allow him to catch up in his processing of the source speech (Kade & Cartellieri, 1971). This raises the issue of how simultaneous "simultaneous" interpretation really is. A number of experimenters conducted synchronization analyses of speaking and pausing times in source and interpreted speech and found that interpreters do seem to be speaking during source language pauses more than would be expected, and to be expanding pause time beyond that of the source speech (Barik, 1973, 1975; Goldman-Eisler, 1967). The conclusion was drawn that the interpreter pauses

during the source speaking time in order to analyze the input speech, while crowding as much of his own output as possible into the source pausing time, in order that he may avoid having to listen and speak at the same time. In Goldman-Eisler's later (1972b) study, however, she found that simultaneity of operations was quite possible: the faster the input, the greater the simultaneity between input and output. This finding led her to speculate on how the various operations in interpretation were phased. Gerver (1976) had found, with professional interpreters as subjects, that over 85 per cent of input was correctly interpreted when an average of over 75 per cent of the total time was spent in simultaneous speaking and listening. As he pointed out, most of the unfilled pauses occurring in input speech which have been examined in simultaneous interpretation studies are too short, under one second long, for interpreters to use them actively for cramming output into. Few words could be crowded even into longer input pauses, as interpreters have commonly been found to speak at rates of between 1.6 and 2 words per second (Gerver, 1971).

Regardless of their usefulness for speaking in, pauses have been found to assist the interpreter in another way. The findings of Goldman-Eisler (1972a,b) and Barik (1973, 1975) suggested that the locus of pauses at grammatical boundaries in the source language texts, more frequent in less spontaneous speech, might be an important factor affecting interpreter performance, notably in terms of the number of omissions in his output. The suggestion was made that such pauses, by delineating units of meaning, might be aiding the interpreter to decode the meanings in the source speech (Barik, 1973). Gerver (1971) experimentally manipulated this issue by using input materials with and

without unfilled pauses, stress and intonation. He found that these variables did appear to assist the interpreter in decoding and/or recoding the source text, although whether these effects were specific to interpretation is not clear, since no shadowing control group was involved. The interpreters' ear-voice spans were not affected by the manipulation, but the number of errors increased and significantly less of the input was correctly rendered with proportionately fewer self-corrections being made. It appeared that to maintain a constant EVS to keep pace with the continuous input, more channel capacity was being diverted to the task of decoding and recoding the continuous input, and thus less was available for monitoring output.

The attention-sharing aspect of simultaneous interpretation is no longer really in doubt. It has repeatedly been shown, moreover, that concurrent tasks can be performed simultaneously, that attention may be shared and not inevitably switched between them. Performance on such verbal transcription tasks as typing, stenography, etc., can be "surprisingly" efficient, with attention being more efficiently shared, the more overtrained or automatic the skills involved (Peterson, 1969). Gerver (1974b) experimentally manipulated the time-sharing question, and tested the extent to which simultaneous listening and speaking itself impairs performance on cognitive tasks. With the tasks of interpreting, shadowing or just listening to recordings of French prose, and using English trainee interpreters as subjects, he found recall scores highest for subjects who simply listened, suggesting that the simultaneity of listening and speaking required by the other two tasks did interfere with comprehension and/or retention of information. What was interesting was that recall was significantly better after interpreting

than after shadowing: given concurrent cognitive tasks, it was still possible to comprehend and recall what was heard, and recall appeared to be enhanced by the more complex operations required in simultaneous interpretation.

Gerver (1971) proposed a model of the information processing occurring in simultaneous interpretation based on his own and previous research into mechanisms of attention allocation. He admits it can only be considered a first approximation, since systematic analysis has not been done on all aspects of the task. The model is a multi-channel one, with attention under normal conditions being shared between short-term memory operations, central recoding processes and output monitoring and correction. There are gaps in his model, as Gerver himself admits. Although comprehensive in many respects, the central portion of his flow chart that depicts the process of simultaneous interpretation simply features two boxes: one containing the instructions "decode and store source language", with an arrow pointing to the next operation which is labelled "encode and store target language". The flow chart provides for a "code book" which contains the source language lexicon and grammar, whereby source language input is decoded into its deep structure, which is then encoded, by means of the target language code book, into the surface structure of the target language. The model, however, says nothing about the way the structure of the source language message is mapped onto that of the target language, nor about how target language units are matched up with source language units.

The fact that Gerver's model of information processing in SI pays little attention to the special characteristics of the task, namely that of going from receiving a speech in one language to giving the

meaning of it in a different language, is perhaps not surprising, since the amount of relevant experimental data from adequately designed research is quite limited. As already mentioned, only two studies (Gerver, 1971; Treisman, 1965) employed a shadowing control group. In all other cases, the effects observed may have been central to the interpretation process or peripheral or a combination of both.

Furthermore, the methods used to quantify interpretation quality were in most studies somewhat wanting, since the adequacy of interpretation depends on how much of the meaning in the input speech is preserved in the interpretation, not how many words come through in it. The only investigator to go beyond the word-level approach was Gerver in his (1971) experiment on the effects of input noise. He was also the first researcher to take intelligibility of the output into account.

Not all studies used professional interpreters as subjects and where they were used (Barik, 1973; Gerver, 1971; Goldman-Eisler, 1967, 1972b; Oléron & Nanpon, 1965), the sample was often small and the degree of training varied. Thus findings may have been contaminated by the effects of insufficient training, since there are substantial enough differences in the skills of even experienced interpreters (Pinhas, 1974). Even when experienced interpreters were used, they were not always working into their usual target language (Barik, 1973, 1975; Oléron & Nanpon, 1965), with again the possibility of results being contaminated by the effects of inexperience in particular output languages. Few interpreters are officially "bilingual" in two output languages. The "reflex" interpretations (Seleskovitch, 1973) or transliterations, which have been reported in studies using inexperienced subjects (Barik, 1973, 1975), or subjects working in unaccustomed

directions (Barik, 1971, 1973, 1975; Oléron & Nanpon, 1965), can scarcely shed light on the way trained interpreters normally process their input.

Apart from these methodological considerations, if one examines simply the content of the small amount of experimental research that has been done, there are two obvious gaps in it. The first is the empty hole in Gerver's model, that is, the absence of any research on the central processes involved in the interpreter's recoding of input received in one language into another. This particular point will be picked up in the introduction to the third experiment of the present study. The second shortcoming is that simultaneous interpretation does not take place in vacuo, and yet most investigators have largely ignored the context in which interpretation is done. It is this context which the first two experiments of the present thesis propose to examine.

There are at least three aspects of the environment in which interpretation takes place which are universally assumed to affect the quality of interpretation, so much so that they are expressly incorporated into the interpreter's working conditions as laid down in his contract. The first requirement is that the booth be soundproof and free of environmental noise and that the sound system be adequate to the task; the second is that if texts are to be read in the course of the conference, an abstract or the complete presentation be made available to interpreters in advance of the conference; and the third, that the interpreter have a comfortable view of the speaker and audience.

The first aspect, the importance of optimal listening conditions, has been investigated experimentally and found to have a great effect on both the contents and delivery characteristics of the interpreter's

output. Gerver (1971, 1974a), using three noise levels and the tasks of interpreting and shadowing performed by professional interpreters, found that EVS remained fairly constant in both tasks under all noise conditions, but that noise increased pausing time during interpretation more than during shadowing. Noise had a greater effect on the interpreting group than the shadowing controls in terms of the number of omissions and errors, as well as corrections of errors, although not in terms of number of words correctly transmitted.

While poor listening conditions have been shown to lead to deterioration in interpretation quality, no one has yet investigated the two other contextual conditions that are clauses in the interpreter's contract.

Linguistic context at the sentence level has been found to affect performance during interpretation, particularly semantic context (Treisman, 1965), and interpreters have been found to make errors based on wrong expectations arrived at on the basis of linguistic and non-linguistic cues (Gerver, 1976; Seleskovitch, 1968). No research, however, has looked at the effects of linguistic context in interpretation beyond the sentence level, of the type that would be provided by prior reading of the text or summary of a speech for interpretation. Yet people do not regularly analyze sentences in isolation (Schank, 1972) and context effects go much beyond the sentence level (Bransford & Johnson, 1972). A message for interpretation could be considered to be not a sum of sentences, but an organic whole, in which each sentence is interpreted with respect to the broader linguistic and situational context in which it is uttered. Written translation has been shown to improve when the total context of the message for translation is

increased (Triandis, 1976). Even in the relatively simple speech processing task of shadowing, individuals have been found to make use of linguistic context well beyond the scope of a sentence (Lindig, 1976).

The presence of linguistic context surrounding a message would seem to be particularly important in simultaneous interpretation when the message is contained in a prepared text which is to be read aloud. Simultaneous interpretation as mentioned earlier is properly intended for spontaneous speech (Seleskovitch, 1976) and any interpreter will testify to the fact that interpreting written texts presented orally poses particular problems and is quite different from interpreting spontaneous speech. In addition to the tighter, more formal style of written presentations, speakers who read their texts generally speak more fluently and faster than when speaking "off the cuff" (Oléron & Nanpon, 1965; Seleskovitch, 1968). Speakers tend to hesitate more in spontaneous speech (Barik, 1975; Gerver, 1971; Goldman-Eisler, 1958, 1967), which could afford respite, processing or catch-up time for interpreters. For all these reasons, if speeches are to be read during conferences, an abstract that sets them in context, provides a general view of the main points and progression of arguments, or better still the complete text containing both form and content are thought to be invaluable aids to the interpreter.

The total content of communication includes, in addition, the extra-linguistic context surrounding the speech being delivered - the audience, the speaker's relation to it, his background, the purpose of of the conference, the "sociolinguistics" of the conference environment. This information is available in varying degrees to the interpreter who can process it as background to his analysis of the dis-

course itself. The importance of perceptual context in message processing in general was clearly demonstrated by Bransford and Johnson (1972). Non-verbal, visual information, such as the speaker's facial expressions and gestures and his audience's reactions to what he is saying, is considered to be an integral part of the overall communication system that supports and complements the verbal portion in complex ways (Argyle, 1972). It gives rhythm and effective colouration to speech, contours and punctuates sentences, provides a whole range of nuances that can supplement, attentuate, modify, correct or contradict the meanings of the verbal portion (Lurçat, 1972). Mehrabian (1971), studying the relative part played by verbal, vocal (prosodic and paralinguistic) and facial information in communicating feeling, had subjects interpret human emotions on the basis of these three types of cues. His results showed facial expressions to be the most powerful indicator of feelings.

A noted interpreter has stated flatly that interpretation will be unsatisfactory if the visual context is unavailable (Thiéry, 1974). An illustration of the degree to which meaning can depend on visual context is evidenced in an incident which occurred during a convention where the author was working. The chairman was heard to say firmly and quite inconsequentially, "I have not finished ... this afternoon.". The interpreter working was at a loss as to the meaning of these words: her view from the booth blocked her sight of the delegation of protesters who had at that moment entered the back of the hall and were attempting to interrupt the chairman to protest an increase in dues that was scheduled for public discussion that afternoon. The meaning of these words could only have been constructed if the interpreter had

some referential knowledge of the situation.

Experiments 1 and 2 of the present thesis were designed to evaluate the importance of these linguistic and extra-linguistic contextual factors on the performance of simultaneous interpreters. Experiment 1 was to ascertain whether prior information about either the context or total contents of a message would make any difference to the interpreter's performance of the message. The affects could be on the perception and/or encoding of the input speech, or on the recoding and/or production of output speech. Three conditions were used: a "text" condition, where subjects just before interpreting received a written copy of the speech, containing total form and linguistic/informational content; a "précis" condition, where subjects read beforehand a summary of the speech which set it in context, highlighted its main points and structure, but was devoid of any phrasing to be used[1]; and a "no information" control condition, where subjects had no prior knowledge of the speech beyond its topic.

Experiment 2 focussed on the effects of the visual context, over and above the verbal content, on interpreter performance. Two conditions were used: a "video on" interpreting condition, where visual information of the conference setting was available to subjects through a videotape recording; and a "video off" condition, where no such information was available.

Experiments 1 and 2

Method

Subjects

Subjects were 12 professional simultaneous interpreters who had

[1] The first two conditions used are typical of the "real" interpretation situation.

been working actively for a minimum of 5 years. Two were officially classified as English-French bilinguals, which means that they had two active "A" or output languages. The other ten were "English-booth" interpreters, with English as their "A" language and French as a passive or input language. The five-year criterion was a partial control for level of interpretation skills; in practice however, as in any profession, there are broad gradations in skills even among experienced interpreters (Pinhas, 1974).

Materials

The stimulus materials for interpretation in the two experiments were video-taped recordings made during an actual series of conferences conducted in English and French on the topic of university research, organization and funding. The conference was held in four separate sessions, each consisting of a formal presentation by a guest speaker, followed by a speaker-panelist discussion and an audience question-and-answer period. The passages used as input material varied widely in style and delivery rate. Excerpts were taken from the twelve-hour master tape and dubbed onto half-inch videotape for use in the present study. Written transcripts of all input materials were also made.

For Experiment 1, the input material was taken from the formal, prepared presentations made in French by three different speakers. Each separate passage was broken down into about 4 minutes of interpreting "warm-up" material, followed by about 6 minutes of material to be analyzed. All speakers read their presentations, but one (Speaker 3) did on occasion depart from his text to illustrate points. The three presentations were transcribed, with false starts, repetitions, etc., eliminated in order to simulate more closely the texts of prepared

speeches which interpreters may receive in real situations. These texts were for the "text" condition. One-page summaries of these texts were made for use in the "précis" condition.

Input material for Experiment 2 consisted of 6 passages (6 minutes long on the average) of spontaneous French speech on the same topic of university research, which were taken from the speech of 6 different panelists who spoke during the discussion periods following two of the formal presentations. In order to maintain the internal continuity of each discussion, it was decided to treat these 6 passages as two integral blocks of 3 speakers each, with passage order invariable within each block. In each block, the first 4 minutes of input speech were considered warm-up material.

The dependent variable being interpreter performance, it was decided to measure two separate aspects of performance, its intelligibility and its information content, by means of the two scales developed by Carroll (1966) and used by Gerver in his 1971 experiments. Carroll had found the scales to be highly correlated and to differentiate reliably between the human and machine translations used in his study. The nine-point Intelligibility Scale assesses the degree to which the translated version sounds like normal prose and would be understandable in the same way as if it had been originally spoken in that language. Since interpretations can be highly intelligible and the product of the interpreter's fertile imagination, a second scale, Carroll's nine-point Informativeness Scale, was used to measure the more conventional translation variable of "fidelity" to the original. This second scale turns the fidelity question around by asking how much new information the original message is perceived to convey after the translation has been

studied by raters. If it is judged to contain information that is missing or totally different from the message in the translation, the original is scored as highly informative. Conversely, if the original tells the rater nothing new or more than the translation did, it will be considered uninformative.

It will be noted that neither of these scales was developed for interpretation and that the assessment of performance they afford is global and takes no account of types of translation disruption or of style or "elegance" of interpretations. Since simultaneous interpretation is best viewed as "pragmatic", as opposed to "linguistic" translation (Brislin, 1976), its focus being on the effective transmission of a recoded message and the preservation of information in the process, the two scales were considered suitable for the present study.

Procedure

All stimulus tapes were played individually to subjects over Lloyds model Y707 earphones from a Sony AV 3600 videotape recorder, with each subject having a clear view of the Sony television screen. Subjects' output was recorded via a microphone on one track of the subject tapes by means of a two-channel Uher 8000 tape recorder, with input speech simultaneously recorded on the other track. In Experiment 1, subjects in the "text" condition were asked to read the text of the speech, but were not allowed to note down anything from the text. When they had finished, the text was retrieved, the VTR and television turned on and, after volume was set at a level comfortable to subjects, they were instructed to start interpreting. The "précis" condition followed the same procedure, with the text summary being given to subjects to read. In the "no information" control condition, subjects were only told of

the topic of the conference and, after an equivalent waiting period, were asked to start interpreting. Between each of the three experimental conditions, subjects were given a short rest period before beginning the next condition. In Experiment 2, for the "video on" condition, subjects saw the proceedings on the television screen; in the "video off" condition, the television monitor was turned off and only the sound track of the videotape remained on. Here again, the two conditions were separated by a short break.

Subjects were tested individually in one session lasting approximately two hours, with rest periods after each experiment. Half of the subjects were randomly assigned to do Experiment 1 first, and the other half, Experiment 2 first.

Transcripts of source and interpreter productions in the two experiments were given to two raters to score according to Carroll's Intelligibility and Informativeness Scales. It was decided, contrary to Gerver's (1971) practice, to transcribe interpreter tapes in full, with false starts, hesitations, repetitions, and gropings left in, since what listeners hear is the unexpurgated versions of the interpretations and it is on these versions that comprehension is based. An ideal, but prohibitively time-consuming way to have the interpretations rated, which would most closely simulate actual listener judgements, would have been to have judges listen to the source and interpreter tapes, thereby receiving the whole panoply of oral features which written transcripts cannot provide. In the present experiments, the two raters, one professional and one student interpreter, were instructed in the use of the scales, and then asked to bear in mind that passages had been conveyed orally and to read passages aloud as they scored them.

The scores which each rater assigned were pooled for each subject within experimental conditions. Subjects' scores in each of the two treatment conditions of Experiment 2 were collapsed across passages within blocks.

Design

Experiment 1 followed a one-way treatment x subject model, with each subject interpreting one passage in each of the three treatment conditions. The three passages were counterbalanced across conditions and the order in which treatment conditions were administered was counterbalanced across subjects. Experiment 2 was also a repeated measures design with the two blocks (of 3 passages each) counterbalanced across the two treatment conditions, but with the order of passages within blocks being invariable. Half of the subjects received the "video on" condition first, the other half the "video off" first.

Results

Experiment 1

Inter-judge reliability, assessed by means of the Spearman rank correction method, was judged to be satisfactory. The rho values for intelligibility scores on the three experimental passages ranged from .78 to .84, while those for informativeness ranged from .83 to .91.

Means and standard deviations of intelligibility and informativeness scores in each condition are shown in Table 1. High numbers on the nine-point Intelligibility Scale and low numbers on the nine-point Informative Scale denote superior performance. A rating of 0 on this latter scale is provided for cases where the translation conveys more information than the original. As can be inferred from this Table, analysis of the data showed no significant effect of prior information about the content of the speech on either the intelligibility measure

Table 1

Means and Standard Deviations of Intelligibility and Informativeness Scores as a Function of Degree of Prior Information about Speech Content

(Experiment 1)

Measure	Prior Information Condition					
	Text		Précis		No Information	
	M	S	M	S	M	S
Intelligibility	6.50	1.30	6.71	1.37	6.67	1.48
Informativeness	3.24	1.50	3.27	1.32	3.41	1.52

Note: Highest intelligibility score = 9.

Highest informativeness score = 0.

or the informativeness measure. (Both F ratios were less than 1). The source tables for the two analyses are given in Appendix 1.

Experiment 2

Spearman rank correlations again indicated high inter-judge reliability for ratings of both informativeness and intelligibility of the two blocks of passages, the reliability coefficients being .86 and .80 for intelligibility and .86 and .85 for informativeness.

Means and standard deviations for subjects' scores on the two scales in the "video on" and "video off" conditions are given in Table 2. While intelligibility and informativeness were both slightly better with the video on, as may be seen from the t values given in Table 2, in neither case was the difference statistically significant.

Discussion

The results of Experiments 1 and 2, focussing on the linguistic and extra-linguistic context of the interpretation situation, are somewhat surprising given the conventional wisdom regarding the influence of these variables on the quality of interpretation. Could the stipulations in the interpreter's contract with regard to having a good view of proceedings and receiving prepared speeches in advance of the conference be groundless? The present findings would seem to say so.

In Experiment 1, mean scores for each subject across prior information conditions ranged from 4.3 to 8.3 in terms of intelligibility of output, and from 1.97 to 5.86 in informativeness. In Experiment 2, scores ranged from 5.6 to 8.0 on intelligibility, and from 2.12 to 5.52 on informativeness. The maximum ranges possible in these Scales are from 1 to 9 for intelligibility and from 0 to 9 for informativeness. The ranges found afford some idea of a problem which was encountered

Table 2

Means and Standard Deviations of Intelligibility and Informativeness Scores as a Function of Viewing Condition

(Experiment 2)

Measure	Viewing Condition				t^a
	Video On		Video Off		
	M	S	M	S	
Intelligibility	7.09	.81	6.96	.69	.94 ns
Informativeness	3.20	1.08	3.35	1.14	.62 ns

Note: Highest intelligibility score = 9.

Highest informativeness score = 0.

[a]Critical value for $t_{11, .05}$ is 2.20

with the subjects used in the present study. There was considerable variability among them, despite the five-year experience criterion. Given this possibility of variability among interpreters, it would seem appropriate, in future research into simultaneous interpretation, to employ large samples of subjects or provide a stricter control of their level of skills, in order to avert the possibility of a Type II error. There was, however, no evidence in the data of the present two experiments suggesting that individual subjects were distorting the overall performance values under the different treatment conditions.

Appendices 2 and 3 contain subjects' mean intelligibility and informativeness scores as a function of the input passages which were used in Experiments 1 and 2. As can be seen from these Appendices, certain passages seem to have presented greater difficulty to subjects than others, in terms of intelligibility and informativeness scores, but the effects of the manipulations in both experiments seem to be randomly distributed across the passages. The intelligibility of output on the passages seems generally to follow the level of informativeness, but the passage where scores on the two Carroll scales diverge most is the third passage of block 1 in Appendix 3, where informativeness in both viewing conditions was best, but where intelligibility with the video on tended to be lower. It would be interesting, although beyond the scope of the present study, to see exactly where the greater difficulty in interpreting certain passages lay, and why for instance the information in the above passage should be conveyed well by subjects, but not very intelligibly.

A more detailed analysis of subject transcripts, beyond what the

Carroll scales allow, might reveal whether the effects of broad linguistic and extra-linguistic context are truly as negligeable on interpreter performance as they were found to be. It is quite possible that the two scales are not the most sensitive measures of interpreted output, but they do evaluate practical dimensions of performance which the interpreter himself must consider as he works on input speech, and as such, suggest that this performance was indifferent to such contextual effects.

The findings of Experiment 1 seem to indicate that prior access to the linguistic context or total information content of a prepared speech does not, as is universally claimed, inevitably make a difference to the interpretation of the speech. This finding is particularly interesting because the speeches were formal presentations which were read aloud, a condition which is supposed to make them particularly unsuitable for simultaneous interpretation. When interpreters beforehand had only a general idea of the topic of the speech, they were able to convey as much information and to convey it just as effectively as they could after having read the entire speech through.

These findings are not totally surprising, for interpreters are accustomed to plunging into source speeches "cold", in the sense of not being able to predict what is going to be said. Furthermore, subjects in the "no information" condition were not completely "cold", in that they knew the topic of the speech beforehand, and prior knowledge of even the title of messages has been found to be important in enhancing comprehension and recall (Bransford & Johnson, 1972). It is also possible that differences in performance due to prior knowledge about a speech might show up if a broader range or a more difficult set

of prepared texts were used. It may be that only when such formal presentations are particularly complex, technically or scientifically, or perhaps when input texts are read particularly fast, do interpreters need to have copies or précis of speeches ahead of time in order to interpret them properly. In the "ordinary run" of the task, of which the materials of Experiment 1 may be an example, preliminary "priming" with texts or abstracts may be superfluous. Future research could certainly be geared towards answering this question.

As far as visual context is concerned, it may be as Thiéry (1974) categorically stated, that closed circuit television cannot be considered an adequate substitute for a direct view of proceedings. Thus, in Experiment 2, it is quite possible that the situation of watching and interpreting a conference seen on a television screen, although it does on occasion occur in the real situation, might not be sufficiently naturalistic to show up differences in performance due to the presence or absence of the normal visual cues of a conference. Furthermore, it was not possible to control the scope and amount of visual information provided in the "video on" condition, but it was hoped that the spontaneous speech taken from six different speakers would provide a range of visual information which subjects could avail themselves of while interpreting. At least, the visual context was partially available to interpreters in the "video on" condition, whereas in the "video off" it was not, and the visual context provided by videotapes made for no better interpretation than when subjects could see nothing at all. It is up to future researchers into the interpretation environment to demonstrate whether, why and in what sense interpreters need to see the speaker and audience "live" in order to do a satisfactory job,

since in the present experiment they performed no less "satisfactorily" when input was solely auditory.

The subjects in Experiment 2 were markedly different in their apparent need to see proceedings on the screen. Some mentioned having experienced difficulty or discomfiture throughout when proceedings were not visible, whereas others occasionally were not looking at the screen when the visual information was there. Many subjects mentioned that they gradually got used to having the video off, finding after a while that the visual information was not needed. Indeed, while interpreting particularly fast or obscure passages, most closed their eyes to the television screen. This might conceivably have been an example of the potentially overloading effect that context, here visual, has been found to have (Schwartz, Singer & Macnamara, 1973).

It could be that, again contrary to popular belief, it is only when input conditions are particularly difficult, or when speakers draw abundantly from their store of non-verbal communication tools, or when there is substantial interplay between speakers and their situational context that interpreters in practice need the visual portion of the conference environment for effective interpretation. Abundant information can be garnered from the auditory medium alone, with listening to the radio being a case in point. Perhaps simultaneous interpretation can essentially be reduced to a task of listening comprehension. Theorists of interpretation have viewed the task of the interpreter as no different in essence from ordinary listening (Seleskovitch, 1968); all listeners "interpret" speech against the background of their knowledge and experience. The only difference between the two activities may be that the interpretation task requires the listener to go on to

reproduce what he has heard in his speech.

Experiment 3

We saw in the introduction to the present thesis that the special nature of the transformation the interpreter performs on input has been largely ignored by experimental investigators, so much so that Gerver's (1971) information processing model of simultaneous interpretation provides no information about what happens at the stages where the "translation" takes place. These stages, where the source language is monitored, encoded, stored and then recoded into target language, are nonetheless assumed to require the most processing capacity among the tasks confronting the interpreter (Gerver, 1971). The results of Experiment 2 raised the possibility that simultaneous interpretation might, at least at the input end, be essentially a task of listening comprehension, involving the monitoring, encoding and short-term storage of speech. If this is so, where then does the source-to-target language recoding, the "translation", fit in among the various tasks demanding the interpreter's attention? Does he even consciously translate input speech at all?

In the descriptive and experimental literature on simultaneous interpretation, the interpreter's recoding task has been viewed in two diametrically opposed ways. Some have considered the task essentially as one of establishing hook-ups between the two dictionaries the interpreter is presumed to maintain in memory: thus, his attention during interpretation is considered to be devoted primarily to the task of translating, i.e., finding words or phrases in language B that would match words or phrases received in language A (Oléron & Nanpon, 1965; Treisman, 1965). Opposed to this view is the one which holds that the

interpreter's task is to put thought or "meaning" into words, and the fact that the thoughts are to be phrased in another language is in some sense irrelevant (Goldman-Eisler, 1972b; Seleskovitch, 1976). In this view, the conscious matching up of input received in one language with output spoken in another is considered, if it occurs at all, to be more automatic and to demand less of the interpreter's attention than his primary, more complex task which is to analyze and comprehend the input message. In the present study, it is this dictionary match-up procedure that is meant by the term "translation", as used in the context of simultaneous interpretation.

Theorists discussing translation processes come out strongly against the relatively widespread notion that even translation is a direct recoding from one language to another (Leont'ev, 1973; Nida, 1969, 1976). Even in machine translation they feel that it is not words, but "bundles of componential features" which must be encoded, and kernel structures of language which are decoded and recoded into target language (Green, 1975; Nida, 1969, Schank, 1972; Yngve, 1955). Translation is facilitated by the fact that the kernel structures of different languages are "surprisingly" similar (Nida, 1969).

Theorists discussing simultaneous interpretation and interpreters themselves, as they consider the task to involve the reconstruction of ideas, not words, claim that interpretation does not really involve "translating" or the recoding of languages at all. What the interpreter does is merely give lexical expression to formless thought, "meaning", which is encoded to all intents and purposes non-verbally (Seleskovitch, 1976). The nature of the recoding transformations the interpreter performs on his input has never been investigated experi-

mentally, but there is a certain amount of indirect evidence from research on intepretation and from experiments into bilingualism which lends some credence to this latter view.

In simultaneous interpretation, even if input words were encoded in their verbal form, it is considered doubtful that interpreters could routinely recode input at a pace elicited by lexical units (Goldman-Eisler, 1972b; Seleskovitch, 1973). Even at input presentation rates which have been found to be optimal for interpretation, the word strings enter at a considerable clip, between 1.6 and 2 words per second (Gerver, 1971). Further, as interpreters have commonly been found to operate not on words, but on larger units, as evidenced by the way they segment input speech (Barik, 1973, 1975; Gerver, 1971; Goldman-Eisler, 1972b), such findings would suggest that the interpreter is working on broader conceptual units underlying surface words. Even if the interpreter could ignore concepts and encode/recode source langugage lexical units into target language units directly, not only would the intelligibility of output be questionable, but ready-made and permanent lexical matches in two languages are the exception rather than the rule (Seleskovitch, 1968).

The interpreter has, however, not always been found to be analyzing input into broader units of meaning. Some concepts, particularly mathematical, seem to be closely tied to their verbal form (Kolers, 1966b), and most interpreters admit to a mental language block when it comes to translating numbers. In a few isolated cases, direct, overlearned linkages between his mental dictionary items, such as "cinq" = "five", are useful and can be accessed without further analysis to convert lexical units directly from one language to another (Kade & Car-

tellieri, 1971; Leont'ev, 1973; Seleskovitch, 1973). Source language pronunciations or words have been occasionally found to intrude into professional interpreters' output (Barik, 1973; Oléron & Nanpon, 1965), which is clear evidence that surface-level words and not concepts are being encoded and directly accessed for output speech. As Gerver (1976) points out, however, the most interesting thing about these intrusions is how seldom they occur in the interpreter's total output.

Research on bilingualism has brought to light certain findings which are relevant to an understanding of the way the interpreter encodes and processes input speech. Early studies of the responses of bilinguals to word associations in their two languages had seemed to reveal separate, independent coding systems for the bilingual's languages, particularly for abstract words (Kolers, 1963, 1968; MacNamara, 1967b), which would suggest that the simultaneous interpreter would have to "translate" input speech, i.e., match input item A in one system with output item B in another. However, findings kept cropping up that semantically, phonetically and visually similar words, as well as concrete words in the two languages were closely connected, occasionally interfering in subjects' associations (Kolers, 1963, 1968; Taylor, 1971, 1976).

Studies of recall, both short and long term, by bilingual subjects clearly supported the notion that words in the two languages are encoded and stored primarily according to their semantic properties, with language being only an ancillary means of organizing information in memory (Kolers, 1966a,b; Lambert, Ignatow & Krauthamer, 1968; Nott & Lambert, 1968). In these experiments, bilinguals were often found to lose the language tag, i.e., retain the concept and forget the

language in which it was heard (Kolers, 1966a,b; Lambert et al, 1968; Nott & Lambert, 1968). In bilinguals' memories, there appeared to be one supralinguistic semantic representation for their two languages (Barnett, 1977; Glucksberg, 1975; Gumperz, 1972; Kolers, 1966a; MacLeod, 1976; Schank, 1972; Stromnes, 1974). In short, it appears that in the bilingual's brain, one system does the language processing (Barnett, 1977), as well as the "thinking" (Segalowitz, 1977), not two. Access from one language to the other therefore may be indirect, and mediated by the conceptual system (cf. Kolers, 1966a, 1968; Lambert et al, 1968; Nott & Lambert, 1968; Stromnes, 1974).

Whether or not there are two systems or one underlying language processing in the interpreter, he must, in order to accomplish his task, effect a switch from one language to another.

Again in the area of bilingualism research, Macnamara (1967a) postulated the existence of two switches in bilingual language functioning: one switch, more or less automatic, was at the beginning of the input or decoding process and the other, less automatic, was at the beginning of the output or speech production process. The series of language switching studies which ensued has suggested first that code switching at both ends of language processing can take some processing time, and processing capacity, with the effect however being more clearly demonstrated at the output end (Macnamara, 1967b; Macnamara & Kushnir, 1971; Macnamara, Krauthamer & Bolgar, 1968; Marsh & Maki, 1976; Kolers, 1966b). Kolers (1966b) for example concluded from his experiments that code switching was irrelevant to comprehension. Several investigators also found that language processing can take longer if input or output is in the weaker of the bilingual's two

languages (Marsh & Maki, 1976; Nott & Lambert, 1968; Schwartz et al, 1973).

Simultaneous interpretation has consistently been shown to take more time than shadowing, as measured in ear-voice span (Gerver, 1971; Oléron & Nanpon, 1965; Treisman, 1965). The extra time taken in interpretation, however, could be due to several activities, including the act of code switching itself as well as the recoding transformations being performed on input material, that is, the search in the interpreter's mental dictionary for the appropriate word to use in output speech. Gerver (1971) wished to separate out the language switch operation as it relates to the task of interpretation, but because of the "complexity" of the interpretation task, employed a simpler paradigm and studied the effect of language switching on interpreters' memory capacity for word lists presented and recalled in the same or a different language. The lists were constructed such that any difference in recall of "translated" lists, as compared with lists recalled in the language of presentation, would be due to the language switching and not to the process of "word search", or retrieving the word from memory. Using English-speaking trainee interpreters as subjects, he found no significant effect of the language switch on number of words recalled, although with longer lists, presentation of information in subjects' mother tongue tended to enhance recall. Code switching did not thus seem to lead to any diminished capacity to recall words, but it was not clear whether any different processing times might have been needed to store or recall lists with and without the language switch requirement.

The third experiment of the present thesis was another attempt to

ascertain the significance of the language switch in the intepreter's recoding processes, in order to find out indirectly whether he is engaged in translating input at all, i.e., recoding from one language to another, but it took a different approach from the one Gerver devised.

In many bilingual conferences, the situation frequently arises where two languages are spoken in rapid alternation. The interpreter is suddenly faced with a switch in the language of the discussion, requiring of course a converse change in the language of interpretation. It has happened countless times that the interpreter, apparently not noticing the language change, has continued working in his usual language, but now he is speaking the same language that the source speaker is using. It is sometimes surprisingly long before he is aware of his error. His output is by no means a straight playback of the speaker's words either; he is still "interpreting" the speech, relaying the message as he apprehends it in his own words. Gerver (1971) reported on interpreters who stated that they actually "heard" in the target language, which implicitly suggests that it is meaning that is encoded and lexical form dropped, that any "translation", if it occurs at all in interpretation, is somehow done without awareness upon receipt of input.

The paradigm for Experiment 3 was suggested by these instances of interpreters' inadvertently interpreting into the same language that the original speaker is speaking. It was decided to tackle the language switch issue by comparing performance on two simultaneous interpretation tasks which differed only in the requirement to switch from one language to a different one. One of the tasks was the interpreta-

tion of French input into English output, and the other was the interpretation of English input into English output. Although this latter unilingual condition is by no means the normal form of simultaneous interpretation, for the purposes of the present experiment it is referred to as "English interpretation", as distinguished from bilingual or "French interpretation". If simultaneous interpretation is, as interpreters themselves seem to believe, merely the expression of thoughts in words, then the fact of restating the thought in another language from the one in which it was received should, arguably, be of little account. Both of the present interpretation tasks involve essentially the same tasks -- monitoring, decoding, storage, recoding/word search, speech production and output monitoring for correctness -- but in the "unilingual" condition, the input and output languages are the same.

The experiment compares performance in the English and French interpretation conditions with performance in two shadowing tasks, shadowing of English and French speech. The shadowing tasks, as mentioned earlier, are a control for "peripheral" aspects of interpretation, to wit the simultaneous perception/production of English and French speech in the present instance. It also seemed reasonable to assume that the level of processing involved in shadowing would be minimal (cf.Cherry, 1953; Gerver, 1971, 1974a; Treisman, 1965). Although there is still some question about the amount of semantic processing shadowers do on input (Shaffer, 1973), the evidence seems to indicate that they do commonly extract relatively little meaning out of passages shadowed.

The comparison between the four experimental tasks could thus reasonably be broken down into the component stages or operations

illustrated in Table 3. As a reading of the Table reveals, the two interpretation conditions are distinguished from the shadowing tasks in the additional processing they require of input. Any differences in output, therefore, between the shadowing and interpreting tasks should be attributable to processes that are central to the simultaneous interpretation task. Similarly, since the only feature that differentiates French interpreting from all other conditions is the requirement to switch languages, any differences between French and English interpretation in particular could be considered to be due to language switching.

The experimental questions are first whether the act of crossing language borders, separate from the interpreter's other tasks, requires any processing time. Differences in response latencies among information processing tasks are considered to reflect differences in the number of stages involved in processing the information (cf.Carpenter, 1975; Trabasso, 1972). The fact that interpretation takes longer than shadowing has been attributed to the additional processing interpreters perform on input (e.g., Gerver, 1971). Similarly, the extra language switch operation in the bilingual interpretation condition should logically be expected to result in a further increase in response latencies over those in unilingual interpretation. One striking fact, however, arising out of all Gerver's (1971) research especially, is that there seems to be a "characteristic" latency attached to the task of interpretation itself, as well as a certain maximal limit, presumably reflecting short-term memory constraints, with the result therefore that the effects of manipulations of task difficulty have been less to alter the ear-voice span than to increase disruptions in interpreters'

Table 3

A Comparison between Processing Stages as a Function of Combinations of Input Language and Task

(Experiment 3)

	Input Language/Task Combination			
Processing Stages	English Shadowing	French Shadowing	English Interpretation	French Interpretation
	perceive Eng.speech	perceive Fr.speech	perceive Eng.speech	perceive Fr.speech
	encode Eng.	encode Fr.	encode Eng.	encode Fr.
			decode meaning	decode meaning
				code switch
			encode Eng.	encode Eng.
	produce Eng.speech	produce Fr.speech	produce Eng.speech	produce Eng.speech

output, or as Gerver (1971) termed it, to lower their error criteria.

The second question in Experiment 3 is therefore an indirect attempt to determine if the language switch occupies any processing capacity, by asking whether output in the two interpretation conditions will be different, in particular whether output in the French interpreting condition will convey different amounts of the original information and/or convey it more or less intelligibly than during English interpretation.

On any of these measures, an interaction between task and input language such that performance was significantly poorer/longer under French interpretation than under the other three conditions would indicate that the requirement to switch languages in the French interpretation task was commandeering some of the channel capacity available for processing the speech.

Method

Subjects

Subjects were the 12 conference interpreters previously tested in Experiments 1 and 2.

Materials

Input material was spontaneous speech excerpted from the same videotaped conference that provided materials for the first two experiments. Input consisted of questions from the floor, in English and in French, posed by different individuals to two of the keynote speakers, together with speakers' replies, in English and in French. Only the replies were to be considered for analysis, although longer portions of the tapes were played to allow subjects to hear the whole of the question-and answer sequence. Speech materials for use in the present experiment were unfortunately available from only two different speakers,

since in order that source speakers not be confounded with input language, it was necessary to take speech input from speakers who spoke both languages during the conference and spoke long enough to provide sufficient amounts of material. Both speakers were native "francophones" with a good command of English.

As unilingual interpretation is unnatural for a conference interpreter, an attempt was made to simulate the actual conference conditions where it has been observed to occur unwittingly. Hence, the English and French passages for administration were taken as far as possible from fast-flowing discussions where the two languages were spoken in alternation, avoiding as far as possible, any lengthy segments in one or other language. The mixed-language question and answer sequences for the two speakers were each roughly 35 minutes long. The order of passages for interpretation within each speaker sequence was invariable. The beginning of each speaker series was material in French and constituted the usual four minutes of warm-up time which was excluded from analysis. This was followed by a total speech sample in each language of at least 6 minutes, by each of the two speakers. In addition to the English and French passages to serve in the two interpretation conditions, two more passages, 3 minutes in length, were excerpted from the same two speakers' spontaneous material in each language. This material was to be used in the two shadowing conditions. All input materials were transcribed.

The Carroll Intelligibility and Informativeness Scales were again used to evaluate the quality of output in the four experimental conditions. However, the failure to obtain any significant treatment effects in Experiments 1 and 2 had raised the possibility that the two

Carroll scales might not be very sensitive measures of output in interpretation. For this reason, a second index devised by the University of Ottawa School of Translators and Interpreters for the detailed evaluation of written translations was used, primarily as a control, to measure the quality of output in the two interpretation conditions. In this index of "Informative Congruence" (October 25, 1974), the source language text is segmented into translation units according to a number of guidelines, and degrees of loss or addition of information in the translation units as compared with the original are calculated, yielding a percentage "I" of translation informative congruence.

Procedure

The input materials were played individually to subjects, in a separate session subsequent to the one involving Experiments 1 and 2, following the same procedure as in the two earlier experiments. Source input and subjects' output were also recorded in the same way as in the two earlier experiments, by means of the two channel tape recorder.

The experiment was carried out in two parts, separated by a rest interval, the parts corresponding to input material from the two speakers. Half of the subjects were randomly assigned to perform the Speaker 1 sequence first, the other half the Speaker 2 sequence first.

Subjects were first instructed to shadow ("repeat back everything you hear") a three-minute passage of French, followed by a three-minute passage of English, or vice versa. The order in which subjects received a particular language or passage to shadow was counterbalanced

across subjects. Following shadowing, training[1] was given in the English-English interpretation task, as it was assumed that the task would require practice before subjects could perform it consciously. To this end, a lengthy comment in English made at the same videotaped conference was played to subjects to practice on. They were instructed to relay the message in their own words, not to rack their brains if input words or phrasing was the way they themselves would express it. Subjects were told simply to get the source message across as clearly as they could, repeating, adding or clarifying anytime they felt it appropriate. All subjects said they understood and were able to perform the task comfortably before being allowed to proceed to the next phase of the experiment.

Subjects were then asked to interpret the speech they would be hearing and seeing. They were told that passages would be sometimes in French, sometimes in English, and that they were to carry on interpreting into English, whatever the source language. Following the interpretation sequence, subjects were again requested to shadow three minutes of English, followed by three minutes of French, or vice versa, spoken by the speaker whom they had just interpreted. The order of administration of passages and language of passage were counterbalanced

[1]Training in unilingual interpretation was interjected at this point in the experimental sequence to divert subjects from inadvertently carrying over the word-for-word repetition of shadowing into the unilingual interpretation.

here again[1]. At this point, subjects were given a rest period, after which the experiment was repeated using the sequence of input material from the other speaker.

To illustrate the rather involved time course of the experiment, a typical schema for a speaker sequence which a subject might perform is given below:

Eng.SH	Fr.SH	Eng.SI	Fr.SI (warm-up)	Mixed Fr/Eng.SI	Fr.SH	Eng.SH
(3 min)	(3 min)	(practice)	(4 min)	(12 min)	(3 min)	(3 min)

All subjects' speech output under the four treatment conditions was transcribed. The transcripts were scored for informativeness and intelligibility by two raters, professional interpreters, neither of whom had been subjects in the present experiments. Subjects' mean scores on the Intelligibility and Informativeness Scales were calculated for each passage and averaged across raters. These scores were then pooled for each subject within experimental conditions. In addition, a university translation student experienced in the use of the University of Ottawa Informative Congruence index rated subjects' output in the two interpretation conditions.

[1]The shadowing material was administered at the start and conclusion of the interpretation conditions in order to control for changes in this criterion measure over time. A further reason is that Gerver (1971) had found significant practice effects with subjects who performed interpreting and shadowing tasks successively.

To measure response latencies or time lags separating interpreter output from source input, the experimenter went through all typewritten transcripts of source language speech, marking off content (as opposed to function) words (cf. Cutler, 1977) at intervals of approximately every five or six words on the source transcripts themselves. Time lag was to be measured primarily between corresponding source and target language content words. A paper record of both tracks (original and interpreter) of each subject's tape was obtained by playing each tape on a Sony TC-252 D stereo tape recorder, while both tracks were being fed, via a voice integrating circuit using a 22 msec. time constant, into a Grass model 7 Polygraph. Source speech tracings were recorded on the top track of the paper record and interpreted speech appeared simultaneously on the bottom track. After setting the appropriate controls, the experimenter followed the source speech from the typewritten transcripts, while at the same time listening to the source speech track of one of the subject tapes. A marker pen was activated on the paper tracing whenever reference words were heard to occur in source speech. In this way, a number of reference marks were established along the pen tracing of the source speech, enabling the experimenter to match reference words in the source speech with the tracings on the source language channel of the paper record.

Subjects' speech was processed in the same way. Subjects' transcripts were marked off for reference content words which corresponded to the source content words. Where such target language words were absent, neighbouring content words were marked off to serve as target language reference words. The twelve subjects' tapes were then fed into the pen recording unit individually and words marked off and

located on the subject channels of the pen tracings in the same way as source words had been located. By comparing the tracings of the original speech on the top track with subject speech on the bottom track of the paper record, it was possible, through continuous cross-reference between channels, to compare words in the original speech with corresponding words in the twelve target language versions. As the speed of the pen recorder was set at 10 mm per second, it was possible to measure time lag, in seconds, as a function of position differences along the paper tapes between the source and target language reference words in the two channels. Despite the crudeness of the method (the reader is referred to Barik, 1972b, and Gerver, 1971, for a more sophisticated method of analyzing temporal parameters of speech), the method was considered sufficient for measurement of the time elapsing between the onset of source and target language words to an estimated accuracy of .2 seconds.

Ear-voice spans were estimated on the basis of 3 latency measures taken within one-minute periods of shadowing or interpretation. Three one-minute periods, located near the beginning, in the middle and near the end of each passage, were sampled for all passages shadowed or interpreted. The reason for sampling at three separate locations in each passage was to control for possible changes in ear-voice spans within a passage. The latency reported for a subject under a given condition is thus the average of 12 latency measures (4 passages per condition x 3 latency measures at a given passage location).

Design

The basic design of the experiment was a two-way treatment x subject model with two levels of task and of input language and with each subject participating in all conditions. With respect to latency

measures, the design was a 2 tasks x 3 input languages x 3 levels of latency location model.

Results

Spearman rank correlations indicated generally high inter-rater reliability for the informativeness scores of the English and French interpretation and shadowing material, .95 and .76, respectively for interpreted output, and .82 and .92, respectively for shadowed output. For intelligibility ratings, the corresponding values were .76 and .81 for shadowed output and .89 and .76 for interpretation output.

The Informative Congruence scores which were assigned to subjects' output in the two interpretation conditions were found to correlate highly with Intelligibility and Informativeness scores pooled across raters. This was particularly true of informativeness scores, which yielded rho values of .85 and .88 for English and French interpreted input, respectively. The values obtained with the Intelligibility Scale scores were .71 and .88 respectively, for English and French interpreted input. Use of the Carroll Scales hence seems to be vindicated. They appear to be measuring output quality in a manner commensurate with the detailed and time-consuming evaluations of information lost, added and substituted which the "I" Index affords. Since the I Index was used primarily as a control of the measures of output quality used, the "I" scores were not analyzed and results for the present experiment refer only to scores obtained on the Carroll Scales.

Latencies (EVS). Table 4 contains the means and standard deviations of latencies, in seconds, for the four combinations of input language and task, and for the three speech locations at which latencies were measured. The analysis of variance (see Appendix 5) revealed a highly significant

Table 4

Means and Standard Deviations of Ear-voice Span in Seconds as a Function of Temporal Location in Input, Task and Input Language

(Experiment 3)

Task	Input Language	Temporal Location in Input						
		Beginning		Middle		End		Mean
		M	S	M	S	M	S	
Shadowing	English	1.31	.25	1.39	.34	1.34	.31	1.35
	French	1.43	.51	1.56	.71	1.34	.57	1.44
Interpretation	English	2.93	.62	2.72	.77	2.74	.67	2.80
	French	2.98	.73	3.01	.86	2.94	1.04	2.98

effect of task, F (1,11) = 74.84, $p < .01$, on the ear-voice span of subjects. The mean latency in the two shadowing tasks was in fact 1.40 seconds, compared with a mean of 2.89 seconds for the two interpreting tasks. Neither the input language effect nor the Task x Input language interaction was significant, and there was no main effect or interactions involving location of latency measurement.

Intelligibility and Informativeness. Subjects' mean scores and standard deviations on the two nine-point scales obtained in the four combinations of input language and task are given in Table 5. (Note that high numbers in intelligibility and low numbers in informativeness denote better scores.) With regard to the intelligibility of output, the analysis of variance showed significant main effects of task, F (1,11)= 21.25, $p < .01$, and of input language, F (1,11) = 25.89, $p < .01$. As Table 5 shows, shadowed input was more intelligibly rendered than interpreted input and English input was rendered more intelligibly than French. The Task X Input language interaction for intelligibility was also significant, F (1,11) = 9.10, $p < .01$. Inspection of Table 5 indicates that the effect of input language was greater on the interpretation than on the shadowing intelligibility, and that French interpretation was poorer on this measure than all other conditions.

In respect of the informativeness of output, there were significant main effects once again of task, F (1,11) = 26.00, $p < .01$, and of input language, F (1,11) = 12.11, $p < .01$, but here the Task X Input language interaction was not significant, although informativeness scores were poorer in the French interpreting condition than in all others, as can be seen from a reading of Table 5. The Table also indicates that shadowed output was more informative than interpreted output and the

Table 5

Means and Standard Deviations of Intelligibility and Informativeness Scores as a Function of Task and Input Language

(Experiment 3)

Task	Input Language	Measure			
		Intelligibility		Informativeness	
		M	S	M	S
Shadowing	English	8.01	.33	1.33	.19
	French	7.31	1.12	2.03	1.29
Interpretation	English	7.28	1.11	2.01	.79
	French	5.86	1.32	3.09	1.37

processing of French input led to greater information loss in output than processing of English input. The analysis of variance source tables for the intelligibility and informativeness scores appear in Appendix 4.

It must be mentioned that in a small number of instances, some data, particularly latencies had to be discarded. Problems such as background microphone noise, a subject inadvertently unplugging the microphone, defective portions of subject tapes, etc., made some material unusable. Losses, fortunately, were minimal, less than an estimated 5 per cent of data.

Discussion

The only significant difference among experimental conditions in terms of ear-voice spans of subjects was due to the effect of task, which was substantial. The interpreting and shadowing latencies obtained are similar, moreover, to the values which Treisman (1965) obtained with her nonprofessional subjects and Gerver (1971) with his professional interpreters. The finding that latencies in the present experiment were differentiated only by task is not totally surprising, for two reasons.

First, other researchers (Gerver, 1971; Oléron & Nanpon, 1965; Treisman, 1965), as has already been mentioned, have consistently found different EVS ranges for shadowing and interpreting. The assumption of the present experiment that EVS length within limits reflects the number of processing stages of a task would seem to be borne out if we compare latencies in French and English shadowing with latencies in English interpretation, and ignore those of French interpretation. Despite the requirement in French shadowing for encoding into a language

which the majority of subjects were unaccustomed to producing in a paced task, response latencies during shadowing of French were considerably shorter than those in an interpretation task which did not even require a language switch (1.44 seconds vs 2.80 seconds).

The second reason for the latency findings being not totally surprising is that other investigators manipulating task difficulty within simultaneous interpretation have revealed relatively constant EVS under a wide variety of input conditions (Barik, 1973; Gerver, 1971; Goldman-Eisler, 1972b; Oléron & Nanpon, 1965; Treisman, 1965). Moreover, as these investigators had also found within subjects, so too did the subjects of the present experiment seem to have their own preferred EVS in both the bilingual and unilingual interpreting tasks, some exhibiting consistently shorter latencies (one presented a mean EVS of 1.78 seconds), while others seemed to prefer longer ones (the longest mean subject EVS for the two interpretation conditions was 4.20 seconds). What is more interesting here in regard to latencies in the two interpreting tasks is that apparently the "characteristic" between-subject EVS in interpretation holds up when subjects interpret both across and within language codes. The extra processing stage represented by the language switch requirement which differentiates the two interpreting conditions here was not reflected in any appreciably longer processing time: the language switching plus the recoding operation of French interpretation took only marginally longer to perform than English interpretation with the recoding operation alone, 2.98 seconds vs 2.80 seconds.

The similar ear-voice spans in the interpreting conditions could be considered as indicating that the processing load was the same in the two tasks, and therefore that the language switch stage in interpre-

tations imposes no additional burden on the information processor. However, since EVS appears to be constrained by immediate memory, it is also possible that the maximum capacity of this memory may have already been reached in both forms of interpretation. Hence, while interpreting English input into English, subjects could not perhaps work any faster than they could when interpreting French into English. It is also conceivable that the interpreter "preference" for a particular time lag might possibly be so strong as to outweigh the effects on EVS of many manipulations of the interpreting task. In any case, it would appear that any effect of the extra processing stage in the language switching of bilingual interpretation in the present experiment would have to be sought in some other aspect of performance, notably the quality of output.

The Task X Input language interaction for intelligibility scores indicates, as Table 5 shows, that subjects under the French interpretation condition did indeed convey the message significantly less intelligibly than during English interpretation. Output was also less informative, although the difference was not significant. If, as other investigators (Barik, 1973; Gerver, 1971) have assumed, interpreters do indeed learn how fast they can work without losing input due to forgetting, the interpreter's "trade-off" which Gerver (1971) referred to between short-term memory capacity and the processing demands of the interpretation task does appear to have been in evidence here to some extent. Subjects in the present experiment seemed to have been prepared to sacrifice the intelligibility, although not the information content, of their renditions while performing "French interpretation". The similar time lags in the two interpretation tasks in conjunction with

the extra operation of language switching in bilingual interpretation do seem to have meant somewhat diminished processing capacity being available in the latter task.

The language switching effect in interpretation seems to hold up within the individual passages interpreted in the two conditions as well, as a reading of Appendix 6 shows. English input passages were consistently better executed on both measures of performance quality than the French, with little overlapping of scores. There is also a tendency for informativeness scores on the different passages to follow intelligibility ratings, but these indications of passage difficulty seem to be unrelated to measures of ear-voice span. For example, the lowest score on both quality measures was found for Passage 4 of Speaker 2, which also resulted in the longest mean EVS (3.34 seconds). Interpretations of the same Speaker's Passage 5 yielded the second shortest mean EVS (2.76 seconds), together with low rankings on quality of output. These findings would appear to lend further support to the suggestion made earlier that measures of ear-voice span may provide only a rough indication of task difficulty.

In sum, despite the assertion of interpretation theorists that the only recoding done in the task is of meaning into speech, that statement seems not to tell the whole story in practice: the intelligibility of the interpreters' renditions suffered when the thoughts had to be restated in a different language, even with relatively general input material. Subjects did seem to be doing some "translating", but its extent is not clear since only one of the performance measures was affected significantly by the language switching. In point of fact, even an adamant proponent of the "non-recoding" theory of interpretation

found a code-switching routine to be occasionally operative among interpreters studied (Seleskovitch, 1973). Code switching could also reasonably be expected to present problems in practice when interpreting for example the arcane jargons of scientific or technical specialties.

Given that code switching in the present experiment had no significant effect on the amount of the original information which interpreters conveyed, one could speculate that, in light of the goal of interpretation which is first and foremost to relay information -- and without delay -- interpreters may, with heightened levels of task load, be more prepared in general to sacrifice the intelligibility rather than the information content of their renditions, in the hope perhaps that their listeners will perform the further task of properly reconstructing the bits of information given.

There are two points worth mentioning at this juncture in regard to previous findings comparing performance on shadowing and simultaneous interpretation. Tables 4 and 5 show that shadowing was significantly better executed than interpreting on both measures of output quality. This runs counter to Treisman's (1965) findings with her amateur bilingual subjects, who tended to perform shadowing and interpretation relatively similarly at lower information rates in terms of words correctly rendered in output. This had led her to speculate whether performance differences in shadowing and interpretation disappear with experienced interpreters, where, as she viewed it, the probability of having words in one language trigger their translation equivalents in the other may be so high as to approach certainty. The performance by the professionals of the present study seems to be

denying this possibility, at least in terms of the output measures used. It would appear unlikely that there is a sufficiently great number of such perfect and permanent translation equivalents for interpretation, even by experienced interpreters, to be equivalent to the more automatic task of verbatim reproduction of input words, albeit "reproduction" in another language.

Tables 4 and 5 also reveal that having to process French rather than English led to significantly poorer quality of output. This again runs counter to Treisman's (1965) results, where her bilinguals showed very similar performance when shadowing as well as when interpreting from French and English at various information rates. The subjects of the present experiment were highly variable in their ability to shadow French input. This seems to be further evidence in support of Gerver's (1971) statement that the interpreter's input languages appear to be more or less functionally equivalent, but that their training leads them to develop a particular set of encoding skills corresponding to their active or output language(s). The finding also points up the need in studies of interpretation to control the level of input and output language skills, even with purportedly "bilingual" interpreters.

A comparison of output quality in the French shadowing and English interpretation conditions reveals striking similarities in both intelligibility and informativeness scores which are rather surprising. Why should French shadowing, which ostensibly involves only verbatim repetition of input speech, be similar in latency to English shadowing but closer in quality to English interpretation, which involves more complex recoding transformations on input? One could speculate that

it might have been the task of encoding into French that diminished available processing capacity and had the effect of depressing output quality, for some reason down to the levels obtained in English interpretation.

Although shadowing ostensibly requires only "parrotting" of input speech, it is possible also that some semantic processing was being done even on the French material for shadowing. To some extent, therefore, the different processing stages in the shadowing and interpreting tasks may not be as clear-cut as they were assumed to be for the sake of the present experiment. Other investigators have found at least some level of semantic processing occurring during shadowing (Carey, 1971; Lindig, 1976; Mandler, 1973; Norman, 1969; Salter, 1973), and subjects have apparently been found to show excellent recall of passages presented at high speech rates, passages which they had shadowed completely unintelligibly (Lerner, 1975).

It is possible that the interpreter's training predisposes him towards semantic analysis of acoustic input even when not required to do so. There is evidence in subjects' shadowing transcripts, occurrences of phrase substitutions, additions, synonyms, etc., to suggest that some such processing was occasionally being done, more frequently in French shadowing. Subjects who had difficulty shadowing the French input may have occasionally resorted to paraphrasing to make up time. Also, English translations of French words, particularly at the beginning of French shadow passages, occasionally intruded into output. Latencies in shadowing also were found to drop as low as .4 seconds, which provides some evidence that expectancies could be operating at least occasionally. It would be interesting to see whether trained interpre-

ters tend to retain more information from shadowed passages than untrained individuals.

While the difference in interpretation results of the present experiment can reasonably be explained in terms of the language switch effect, there are other explanations which are possible. As the input languages were themselves different in the two interpreting tasks, performance differences could perhaps be due to intrinsic characteristics of the languages themselves, which have been mentioned by other investigators (Aaronson, 1974; Macnamara, 1967a; Treisman, 1965). Performance differences could also be due to subjects' differing abilities to decode in their first and second language, as bilingualism investigators have suggested (Nott & Lambert, 1968; Schwartz et al, 1973). Gerver's (1971) free recall experiment, it will be remembered, had revealed that despite his English trainee interpreter subjects' practice in receiving information in French for interpretation into English, they still recalled longer word lists better when they were presented in English. Given professional interpreters' virtually daily exposure to their second language, it would be unlikely however for such input language variables to account wholly for the different quality of output in the two interpretation conditions. A point to be remembered here too is that this difference in output occurred in spite of the oft-mentioned lexical similarity of English and French (Barnett, 1977; Taylor, 1976), which might have been expected to attenuate the language switching effect.

It could also be objected, in regard to the English-English interpretation condition, that the situation is too different and unnatural to be comparable to the normal French-English interpretation done in

the real environment. Indeed, a few subjects mentioned feeling ill-at-ease or tired while performing it. It could thus be argued that the processing demands and time required by "unilingual interpretation" are due not to normal processes of interpretation, but to subjects' relative lack of practice or difficulty in performing it.

One could also object that the processing operations of English interpretation, contrary to the French, might have been primarily devoted to monitoring output to ensure that it is not mere shadowing. It is true that the word search operation in unilingual interpretation could conceivably have been short-circuited, where a source language unit had the potential of being plucked directly for use in target language speech, and this could perhaps have presented a problem to subjects. However, output monitoring regularly occurs in "normal" bilingual interpretation: Investigators (Barik, 1971, 1975; Gerver, 1971, 1974a) have so routinely found interpreters to be monitoring and correcting their renditions that such "self-corrections" have been established as one of the categories of interpreted output. The syntactic and lexical similarity between interpreter output and source input during English interpretation varied between subjects, as was true of their output during the French interpretation, as is true in general of interpreters performing "normal" interpretation.

It is still possible that the attention load during the unilingual condition in the present experiment could have been artificially and excessively biased towards the testing of output for similarities with input. In light of the frequent occurrence of unconsciously unilingual interpretation in the real conference situation, together with the instructions that all subjects were given as to what was expected of

them during the task, it would seem unlikely however that the "testing for similarity" form of output monitoring during English interpretation would be so consuming as to make it completely incompatible with French interpretation.

There is one possible source of difficulty during the English interpretation task which was pointed up by two of the subjects of the present experiment. They mentioned having encountered difficulties with some English input passages, saying that with the two source speakers being both native speakers of French, their English was less spontaneous, harder to follow than their French speech. Another possible difficulty in the present English interpreting task is suggested by Lawson's (1967) study of selective attention for speech. It revealed greatest interference with simultaneous interpretation under conditions which could be considered as somewhat analogous to the situation of English interpretation where the input and output languages being monitored are the same.

If, however, inhibitory factors such as those mentioned above were in effect during English-English interpretation, one could reasonably have expected it to take longer to perform or to be less intelligible or informative than French-English interpretation, which was a highly polished and practiced skill of subjects. The reverse was found, particularly in regard to intelligibility.

Two words of caution are in order as to how far the present findings may reasonably be generalized. First, only two input languages were used and languages which were spoken by French-speaking bilinguals. Second, despite every attempt being made to simulate "real-life" interpretation conditions throughout the present study, the absence of such

non-verbal factors as the listeners (beyond the experimenter), the presence of colleagues, the anonymity of the booth, as well as the diminished situational context in videotapes of proceedings, made the situation necessarily artificial for subjects in the sense of not being experienced from an interpretation booth in actual conference conditions.

Investigations of simultaneous interpretation are still only at the ground-clearing stage. The present research is another exploratory attempt at gaining a rudimentary understanding of two aspects of simultaneous interpretation, one of practical, the other of theoretical interest, which had been almost completely ignored by researchers. The first two experiments tackled two of the most elementary questions about the influence of the interpretation environment and obtained results contrary to the countless findings that have emerged supporting the importance of both linguistic and extra-linguistic context in the comprehension of messages (Aaronson, 1974; Brain, 1961; Neisser, 1967). In the present case, neither increasing the redundancy of the message nor placing it in videotaped visual context was of any appreciable help to subjects interpreting the message. The latter finding could conceivably be interpreted as vindicating the strategy universally used by previous investigators of simultaneous interpretation who ignored the visual information of normal interpretation and presented input material only auditorily. Future research should, however, be geared towards finding out whether even the full, rich visual context under "live" conditions makes any difference to interpretations of speeches. If we could understand exactly when both linguistic and extra-linguistic context do facilitate simultaneous interpre-

tation, we might have new insights into the way we process information.

The third experiment was an attempt to gain a clearer understanding of the way the interpreter recodes his input, to ascertain whether the obligation to retransmit a message in a different language from the one in which it was received imposes any burden on his processing of the message. The results of the experiment revealed that the language change between input and output was not a negligeable factor in the interpretations done by the twelve subjects. It would thus appear that the interpreter does in fact "translate" and is engaged in recoding not simply abstract "thought" into speech, but speech in one language into speech in another. The extent of this translation recoding however, is unclear. If future researchers would turn towards the interpreter's "black box" in general, if they could determine exactly what does occur between the moments of encoding input units and of producing output speech and find out how the two are matched, it could afford fresh insights into semantic processing, and perhaps open up new avenues in the experimental study of semantics.

References

Aaronson, D. Stimulus factors and listening strategies in auditory memory: A theoretical analysis. Cognitive Psychologogy, 1974, 6, 108-132.

Argyle, M. Non-verbal communication in human social interaction. In R.A. Hinde (Ed.), Non-verbal communication. Cambridge University Press, 1972.

Barik, H.C. A description of various types of omissions, additions and errors of translation encountered in simultaneous interpretation, Meta, 1971, 16, 199-210.

Barik, H.C. Interpreters talk a lot, among other things. Babel, 1972, 3-10. (a).

Barik, H.C. Some innovations in a computer approach to the analysis of speech patterns. Language and Speech, 1972, 15, 196-207. (b).

Barik, H.C. Simultaneous interpretation: Temporal and quantitative data. Language and Speech, 1973, 16, 237-270.

Barik, H.C. Simultaneous interpretation: Qualitative and linguistic data. Language and Speech, 1975, 18, 272-297.

Barnett, G.A. Bilingual semantic organization: A multidimensional analysis. Journal of Cross-Cultural Psychology, 1977, 8, 315-330.

Brain, R. The neurology of language. Brain, 1961, 84, 145-166.

Bransford, J.D. & Johnson, M.K. Contextual prerequisites for understanding: Some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 717-726.

Brislin, R.W. (Ed.) Translation: Applications and Research. New York: Gardner Press, 1976.

Carey, P.W. Verbal retention after shadowing and after listening. Perception and Psychophysics, 1971, 91, 79-83.

Carpenter, P.A. & Just, M.A. Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review, 1975, 82, 45-73.

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

Cherry, E.C. Some experiments on the recognition of speech with one and with two ears. Journal of the Acoustical Society of America, 1953, 25, 975-979.

Cutler, A. & Foss, D.J. On the role of sentence stress in sentence processing. Language and Speech, 1977, 20, 1-10.

Gerver, D. Aspects of simultaneous interpretation and human information processing. Unpublished doctoral dissertation, Oxford University, 1971.

Gerver, D. The effects of noise on the performance of simultaneous interpreters: Accuracy of performance. Acta Psychologica, 1974, 38, 159-167. (a).

Gerver, D. Simultaneous listening and speaking and retention of prose. Quarterly Journal of Experimental Psychology, 1974, 26, 337-342. (b).

Gerver, D. Empirical studies of simultaneous interpretation: A review and a model. In R.W. Brislin (Ed.), Translation: Applications and Research. New York: Gardner Press, 1976.

Glucksberg, S. & Danks, J. (Eds.). Experimental psycholinguistics: An introduction. Hillsdale, N.J.: Lawrence Erlbaum, 1975.

Goldman-Eisler, F. Speech production and the predictability of words in context. Quarterly Journal of Experimental Psychology, 1958, 10, 96-106.

Goldman-Eisler, F. Sequential temporal patterns and cognitive processes in speech. Language and Speech, 1967, 10, 122-132.

Goldman-Eisler, F. Pauses, clauses and sentences. Language and Speech, 1972, 15, 103-113. (a).

Goldman-Eisler, F. Segmentation of input in simultaneous interpretation. Journal of Psycholinguistic Research, 1972, 1, 127-140. (b).

Green, T.R.G. Computer translation with paired grammars. Behavior Research Methods & Instrumentation, 1975, 7, 557-562.

Gumperz, J.J. The communicative competence of bilinguals: Some hypotheses and suggestions for research. Language in Society, 1972, 1, 143-154.

Hanna, B.T. Oral comprehension. Meta, 1966, 2, 46-52.

Herbert, J. Manuel de l'Interprète. Geneva: Georg., 1965

Hockett, C.F. Translation via immediate constituents. International Journal of American Linguistics, 1954, 22, 313-315.

Kade, O. & Cartellieri, C. Some methodological aspects of simultaneous interpreting. Babel, 1971, 17, 12-16.

Kolers, P. Interlingual word associations. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 291-300.

Kolers, P. Bilingualism and bicodalism. Language and Speech, 1965, 8, 122-126.

Kolers, P. Interlingual facilitation in short-term memory. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 314-419. (a)

Kolers, P. Reading and talking bilingually. American Journal of Psychology, 1966, 79, 357-376. (b).

Kolers, P. Bilingualism and information processing. Scientific American, 1968, 220, 3, 78-86.

Lambert, W.E., Ignatow, M. & Krauthamer, M. Bilingual organization in free recall. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 207-214.

Lawson, E.A. Attention and simultaneous translation. Language and Speech, 1967, 10, 29-35.

Leont'ev, A.A. Some problems in learning Russian as a foreign language. Soviet Psychology, 1973, 11, 1-117.

Lerner, S. A study of shadowing ability (Doctoral dissertation, The City University of New York, 1975). Dissertation Abstracts International, 1975, 35, 5120B.

Lindig, K.D. Contextual factors in shadowing connected discourse (Doctoral dissertation, The University of Chicago, 1976). Dissertation Abstracts International, 1977, 38, 934B-935B.

Lurçat, L. Du geste au langage. Bulletin de psychologie, 1972-73, 26, 501-505.

MacLeod, C.M. Bilingual episodic memory: Acquisition and forgetting. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 347-364.

Macnamara, J. The bilingual's linguistic performance: A psychological overview. Journal of Social Issues, 1967, 23, 58-77. (a).

Macnamara, J. The linguistic independence of bilinguals. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 729-736. (b).

Macnamara, J., Krauthamer, M. & Bolgar, M. Language switching in bilinguals as a function of stimulus and response uncertainty. Journal of Experimental Psychology, 1968, 78, 208-215.

Macnamara, J. & Kushnir, S. Linguistic independence of bilinguals: The input switch. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 480-487.

Mandler, G. & Worden, P.E. Semantic processing without permanent storage. Journal of Experimental Psychology, 1973, 100, 277-283.

Marsh, L.G. & Maki, R.H. Efficiency of arithmetic operations in bilinguals as a function of language. Memory and Cognition, 1976, 4, 459-464.

Mehrabian, A. Silent Messages. Belmont, Calif.: Wadsworth, 1971.

Neisser, U. Cognitive Psychology. New York: Appleton-Century-Crofts, 1967.

Nida, E.A. Science of translation. Language, 1969, 45, 483-498.

Nida, E.A. A framework for the analysis and evaluation of theories of translation. In R.W. Brislin (Ed.), Translation: Applications and Research. New York: Gardner Press, 1976.

Norman, D.A. Memory while shadowing. Quarterly Journal of Experimental Psychology, 1969, 21, 85-93.

Nott, C.R. & Lambert, W.E. Free recall of bilinguals. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 1065-1071.

Oléron, P. & Nanpon, H. Recherches sur la répétition orale de mots présentés auditivement. Année Psychologique, 1964, 64, 397-410.

Oléron, P. & Nanpon, H. Recherches sur la traduction simultanée. Journal de psychologie normale et pathologique, 1965, 62, 73-94.

Paneth, E. An investigation into conference interpreting. Unpublished master's thesis, London University, 1957.

Peterson, L.R. Concurrent verbal activity. Psychological Review, 1969, 76, 376-386.

Pinhas, R. Voici venue la fin d'une génération d'idées reçues en matière d'interprétation de conférence. AIIC Bulletin, 1974, 2, 8-9.

Salter, D. Shadowing at one and at two ears. Quarterly Journal of Experimental Psychology, 1973, 25, 549-556.

Schank, R.C. Conceptual dependency: A theory of natural language understanding. Cognitive Psychology, 1972, 3, 552-631.

Schwartz, J., Singer, J. & Macnamara, J. An analytic comparison of listening in two languages. The Irish Journal of Education, 1973, 1, 40-52.

Segalowitz, N. Psychological perspectives on bilingual education. In B. Spolsky & R.L. Cooper (Eds.), Frontiers of Bilingual Education. Rowley, Mass.: Newbury House, 1977.

Seleskovitch, D. L'Interprète dans les conférences internationales. Paris: Minard, 1968.

Seleskovitch, D. Langage et Mémoire: Etude de la prise de notes en interprétation consécutive. Paris: Minard, 1973.

Seleskovitch, D. Interpretation, a psychological approach to translation. In R.W. Brislin (Ed.), Translation: Applications and Research. New York: Gardner Press, 1976.

Shaffer, L.H. Latency mechanisms in transcription. In S. Kornblum (Ed.), Attention and Performance IV, New York: Academic Press, 1973.

Stromnes, F.J. Memory models and language comprehension. Scandinavian Journal of Psychology, 1974, 15, 26-32.

Taylor, I. How are words from two languages organized in bilinguals' memory? Canadian Journal of Psychology, 1971, 25, 228-240.

Taylor, I. Similarity between French and English words - a factor to be considered in bilingual language behavior? Journal of Psycholinguistic Research, 1976, 5, 85-94.

Thiéry, C. Can simultaneous interpreting work? AIIC Bulletin, 1974, 2, 3-5.

Trabasso, T. Mental operations in language comprehension. In J.B. Carroll & R.O. Freedle (Eds.), Language Comprehension and the Acquisition of Knowledge. Washington, D.C.: V.H. Winston, 1972.

Treisman, A. The effects of redundancy and familiarity on translating and repeating back a foreign and a native language. British Journal of Psychology, 1965, 56, 369-379.

Triandis, H. Approaches toward minimizing translation. In R.W. Brislin (Ed.), Translation: Applications and Research. New York: Gardner Press, 1976.

Van Hoof, H. Théorie et pratique de la traduction. Munich: Max Hueber, 1962.

Yngve, V.H. Sentence-for-sentence translation. Mechanical Translation, 1955, 2, 29-37.

Appendix 1

Analysis of Variance Source Tables for Experiment 1

A. Intelligibility Scores

Source	SS	df	MS	F
Treatment (T)	.2942	2	.1471	.225
Subjects (S)	49.0233	11		
T X S	14.3725	22	.6533	
Total	63.6900	35		

B. Informativeness Scores

Source	SS	df	MS	F
Treatment (T)	.2098	2	.1049	.076
Subjects (S)	39.2060	11		
T X S	30.2360	22	1.3744	
Total	69.6518			

Appendix 2

Mean Intelligibility and Informativeness Scores as a Function of Input Passage and Degree of Prior Information about Speech Content

(Experiment 1)

Measure	Passage	Prior Information Condition		
		Text	Précis	No Information
Intelligibility	1	7.00	7.38	7.12
	2	6.62	6.62	7.12
	3	5.88	6.12	5.75
Informativeness	1	2.81	2.23	2.92
	2	2.55	3.20	2.66
	3	4.36	4.38	4.67

Appendix 3

Mean Intelligibility and Informativeness Scores as a Function of Input Passage and Viewing Condition

(Experiment 2)

Measure	Input Block	Passage Position in Block[a]	Viewing Condition	
			Video On	Video Off
Intelligibility	1	1st	6.00	6.67
		2nd	7.42	7.58
		3rd	7.08	7.33
	2	1st	7.75	7.04
		2nd	7.75	7.42
		3rd	6.17	5.42
Informativeness	1	1st	3.60	3.71
		2nd	2.71	2.61
		3rd	2.56	2.33
	2	1st	3.25	3.70
		2nd	2.92	3.23
		3rd	4.38	4.63

[a] Passage order was invariable within blocks.

Appendix 4

Analysis of Variance Source Tables for Experiment 3

Intelligibility and Informativeness

A. Intelligibility Scores

Source	SS	df	MS	F
Input language (I)	13.462	1	13.462	25.888+
Task (T)	14.214	1	14.214	21.247+
I X T	1.519	1	1.519	9.096+
Subject (S)	32.860	11		
I X S	5.718	11	.5200	
T X S	7.359	11	.6690	
I X T X S	1.838	11	.1671	

B. Informativeness Scores

Source	SS	df	MS	F
Input language (I)	9.470	1	9.470	12.110+
Task (T)	9.048	1	9.048	26.000+
I X T	.444	1	.444	2.018
Subject (S)	31.427	11		
I X S	8.606	11	.782	
T X S	3.823	11	.348	
I X T X S	2.422	11	.220	

+$p < .01$

Appendix 5

Analysis of Variance Source Table for Experiment 3

Latencies

Source	SS	df	MS	F
Input language (I)	.675	1	.675	3.750
Task (T)	80.073	1	80.073	74.835+
Latency location (L)	.20	2	.10	2.326
Subject (S)	33.715	11		
I X T	.064	1	.064	.244
I X L	.158	2	.079	1.039
T X L	.238	2	.119	1.008
I X S	1.980	11	.180	
T X S	11.770	11	1.070	
L X S	.935	22	.043	
I X T X L	.110	2	.055	.932
I X T X S	2.881	11	.262	
I X L X S	1.664	22	.076	
T X L X S	2.606	22	.118	
I X T X L X S	1.296	22	.059	

$^{+}p < .01$

Appendix 6

Mean Intelligibility and Informativeness Scores and EVS as a Function of Input Language and Passage Interpreted

Measure	Speaker 1			Speaker 2				
	English passages		French passages[a]	English passages		French passages		
	1[b]	2	3	1	2	3	4	5
Intelligibility	7.33	7.96	6.75	6.71	7.12	6.42	4.79	5.50
Informativeness	2.00	1.61	3.05	2.35	2.06	2.82	3.64	2.83
EVS (seconds)	2.73	2.81	2.84	2.82	2.78	3.03	3.34	2.76

[a] Although the total amount of speech for French and English interpretation provided by the two Speakers was about equal, the number of discrete passages within the French input differed. Speaker 2 thus provided 3 shorter French passages which together were comparable in length to the one long French passage of Speaker 1.

[b] The actual order in which subjects interpreted passages spoken by Speaker 1 was: 3, 1, 2; for passages spoken by Speaker 2 the order was: 3, 4, 1, 5 2.