

A Psychological Approach to Simultaneous Interpretation

David Gerver

Volume 20, Number 2, juin 1975

URI: <https://id.erudit.org/iderudit/002885ar>

DOI: <https://doi.org/10.7202/002885ar>

[See table of contents](#)

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (print)

1492-1421 (digital)

[Explore this journal](#)

Cite this article

Gerver, D. (1975). A Psychological Approach to Simultaneous Interpretation. *Meta*, 20(2), 119–128. <https://doi.org/10.7202/002885ar>

A Psychological Approach to Simultaneous Interpretation *

In this paper I will review some recent research on simultaneous interpretation, and discuss one way of describing certain aspects of the on-the-job performance of simultaneous interpreters. My own interest in the subject stems from an interest in cognitive behaviour in general, and in particular in the light that the study of such complex behaviour as simultaneous interpretation can shed on more general aspects of human attention, memory, and linguistic behaviour. Recent psychological research on their skills has not been particularly favourably received by some interpreters. Their reserve stems, I feel, from a misunderstanding of what psychologists are attempting to do. When faced with complex behaviour, the psychologist's first task must be to describe before he can explain. The research to date has been basically concerned with such description, and admittedly the information so far available is patchy and over-simplified. My aim in presenting this paper has been to foster an interest both in the field and in increasing our knowledge of such complex behaviour. The model discussed is but one of many, and would need a great deal of elaboration before constituting an adequate account of a simultaneous interpreter's performance.

As you know from hearing interpreters at work at international conferences, simultaneous interpreters possess the rare capacity of speaking and listening simultaneously for reasonable lengths of time. While they listen to one language, they carry out complex transformations on the segment of the message they have just heard, while uttering their translation of the previous portion of the message. To lapse into the psychologists' jargon, the interpreter's task is a form of complex human information processing involving the perception, storage, retrieval, transformation, and transmission of verbal information. In a sense it is also a paced tracking task.

In spite of the fact that the simultaneous interpreter's skills would appear fertile ground for psychologists interested in, say, the study of memory, attention, bilingualism, or the analysis and development of skilled behaviour, surprisingly little research has been carried out in this area by psychologists — and none at all on the selection and training of interpreters. A thorough literature search, and

* A slightly altered version of a paper presented at the 18th International Congress of Applied Psychology in Montreal in July 1974.

requests for information to all of the major schools of conference interpretation throughout the world, revealed only a dozen or so authors whose work was available for review, and of these only six were psychologists, only four of whom were directly concerned with the study of simultaneous interpretation per se.

In the Soviet Union, Professor Leont'ev has recently deplored the lack of work in this field, and called for research on the mechanisms involved in simultaneous interpretation (Leont'ev, 1973). Some work has been carried out in Russia by Professor Chernov and his colleagues of the Thorez Institute for Foreign Language Study in Moscow, and I will be referring to Professor Chernov again later in this paper. Apart from Oléron and Nanpon in France, Henri Barik in North America, and Goldman-Eisler, Treisman, Lawson, and myself in the United Kingdom, almost no authors of papers or books on human skills, cognitive, psychology, or psycholinguistics even mention simultaneous interpretation. The two exceptions are Welford, and Neisser. Welford (1968) discusses simultaneous interpreters' ability to listen and speak simultaneously within the context of a discussion of single channel theories of human information processing. Welford suggests that simultaneous interpreters acquire the ability to listen and speak at the same time because, after much practice, they learn to ignore the sound of their own voices. He also cites presumably anecdotal evidence that interpreters often have little idea of what they are saying, or confidence that it is correct. As will be seen, this is not quite the case, since inspection of translations by interpreters shows that they must be attending to some form of feedback since they will occasionally correct words or phrases they have just uttered.

The second exception is Neisser (1967), who refers to simultaneous interpretation within the context of a discussion of the "motor" theory of speech perception. Neisser argues, in the following excerpt, that simultaneous interpretation provides evidence against the "motor theory": "In a sense simultaneous translation is a form of "shadowing". However, it is not words, or articulatory movements that are shadowed. The translator, who is obviously attending to, and understanding, the incoming stream of speech cannot possibly be imitating the speaker's vocalisations. His own vocal tract is occupied with an entirely different output."

However regrettable the neglect of the subject, the reasons are perhaps not hard to find. Firstly, the phenomenon is a comparatively recent and rare one, for though conference interpretation has been with us for some time, simultaneous interpretation only began to come into its own after its use at the Nuremburg trials, and with the growth in international organisations, and conferences such as the one we are attending here in Montreal. Secondly, the task is extremely complex, and the experimental psychologist used to neat, easily executed and analysed experiments might consider the pay-off from such research not worth the price paid in lengthy data analysis. Apart from the problem of defining both independent and dependent variables, the relatively small number of professional simultaneous interpreters with a particular language combination available in any one place at any one time does not make for simple experimental design. It would be wrong to claim that research in simultaneous interpretation is at anything but a preliminary, descriptive stage; even so it is not too early to attempt a step-by-step description

or model of the process, bearing in mind that the studies I will mention have been concerned with interpretation to or from English and French, English and German, or English and Russian. There have been, so far, no replications of studies in other language combinations. This last point is obviously quite important, for as was pointed out in a recent review of Soviet Psycholinguistics (Prucha, 1972), many of the findings of Anglo-Saxon psycholinguists working in English have not been replicated in Russian, and the same may no doubt be true for other languages.

As Barik has pointed out (Barik, 1973), a number of questions spring immediately to mind in relation to the simultaneous interpreters' task. How does the interpreter carry out all the activities involved in simultaneous interpretation? (One might add, what are the activities involved?). What strategies must be employed in order to distribute his attention most effectively between input, processing, and output? What aspects of the source language message affect his performance, and how does he segment almost continuous input? What difficulties arise in the course of the task? To what extent can the study of the interpreter's speech patterns, and the content of his translations provide cues to the processes involved? Finally, one might ask what personality and situational variables affect performance?

The first research on simultaneous interpretation was by Oléron and Nanpon in a paper published in 1964. They calculated ear-voice spans (the number of words, or the time, the interpreter's output lags behind the input of what he has just uttered) from recordings of a number of interpreters obtained in the laboratory, and at conferences, translating different lengths and type of passage from and to a variety of languages. They found that delays could range from 2 to 10 seconds (others have found ear-voice spans ranging from 2 to 8 words). Oléron and Nanpon suggested that, depending on listening conditions, input rate and type of text, the extent of the delay is determined by the relative difficulty in organizing the input — the interpreter must grasp a certain amount of material before he can translate. In a recent paper Goldman-Eisler suggests that the Predicate Phrase is the minimal unit required before translation can commence (Goldman-Eisler, 1972). On the other hand, because of limitations in short-term memory capacity, interpreters cannot afford to lag too far behind.

I only have time to mention two of my own experiments: the first on the effects of source language speech rate (Gerver, 1969), the second on the effects of noise on simultaneous interpreters' performance (Gerver, 1974). Shadowing¹ and interpretation were compared in order to ensure that any differences in interpreters' performance occurring in different experimental conditions could be attributed to their effect on the processes involved in simultaneous interpretation rather than difficulties in perception alone.

In the first experiment six interpreters simultaneously interpreted into English, and six shadowed in French, a recording of a French passage which increased gradually in presentation rate (the tape was processed by mechanical means so that, though the speed increased, the pitch remained the same). It was found that

1. Shadowing is a task often employed in psychological experiments on attention, and involves the immediate vocalisation of auditorily presented speech stimuli.

as input rate increased, and presumably became faster than processing rate, interpreters lagged further behind than shadowers, but they also made significantly more errors, and more omissions, than shadowers at faster rates. In fact the optimal input rate appeared to be that often quoted by simultaneous interpreters — between 100 and 120 words per minute. While shadowers were able to increase their output rate in terms of words correctly transmitted to a maximum of about 130 words per minute, simultaneous interpreters appeared to have an optimum rate of 100 words per minute.

In the second experiment (Gerver, 1974), 12 simultaneous interpreters either interpreted or shadowed six pre-recorded passages of French prose from UNESCO reports in three listening conditions: no noise, quite noisy, and very noisy (the noise was from a white noise source). Again interpreters lagged further behind than shadowers under noisy listening conditions, but this also involved more omissions, and significantly more errors. Analyses of temporal aspects of interpreters performance led to the conclusion that the main effect of noise was to retard the processing of input.

There is, unfortunately, no time to discuss the more detailed analysis of errors of commission and omission carried out on the data from the two experiments I have just mentioned, or that by Barik, who suggested that the analysis of such errors might provide cues to the processes involved in simultaneous interpretation (Barik, 1971). I will, however, mention the self-corrections that interpreters will often make, without pausing in their speech. Some examples are shown below.

TABLE 1

<i>Source</i>	<i>Interpreter</i>
(i) ... est imprimé is imprinted — is printed ...
(ii) Sur tous les continents ...	Especially on the continents — on all the continents
(iii) Aussi bornée que cette activité	Just as limited as this — however limited this activity
(iv) ... for psychologists pour des pschoo-psychologues ...
(v) ... in physical education dans l'édifis—l'éducation physique ...
(vi) ... which will be published this month	... qui doit sortir d'ici—au cours de ce mois-ci.

The first three are from experiments, the second three from recordings made at conferences. In these and other examples of self-corrections interpreters appear to be carrying out a procedure similar to Miller, Galanter and Pribram, 1960). The interpreter generates a target language response which passes a first test, and may be uttered. It might then be tested again. If the second test is passed, the interpreter may proceed to the next item; if not, he can operate again by generating a further response to the same stimulus. Soon I will come to the way this type of checking procedure can be accounted for in a model of the process of simultaneous interpretation.

This evidence suggests, then, that contrary to Welford's remarks, simultaneous interpreters may be aware of what they are saying, and of whether it is correct. Whether they do correct themselves must depend on their expertise in the area being translated, on their response criteria, and whether they judge that the time is available.

Furthermore, they may also recall what they have said. In a recent study (Gerver, 1974) I found that when interpreters were asked to listen to, shadow, or simultaneously interpret passages of French prose, after which they were asked 10 factual questions about each passage, though there was significantly better recall after merely listening than in the other two conditions, there was also significantly better recall after interpretation than after shadowing.

Finally, before discussing a model of the process, we come to the question of the strategy employed by interpreters in order to cope with the demands of the task of listening and speaking simultaneously — a rare occurrence in everyday life! Both Goldman-Eisler and Barik have suggested that, in order to avoid the strain of simultaneous listening and speaking, interpreters will try to make use of the brief pauses in the source language message (Goldman-Eisler, 1968; Barik, 1973). Barik concluded that simultaneous interpreters do make greater use of source language pauses than would be expected if one assumed that the interpreter's intervals of speaking and listening were independent of the intervals of speaking and pausing in the source language speakers speech. Though intuitively this is an attractive hypothesis, an analysis of the distribution of pause times in a number of speeches made at conferences revealed that there are very few pauses the interpreters could have taken advantage of in the way suggested. Employing a pause criterion of 250 msec. the following distribution was obtained from a computerised analysis of 3 to 5 minute segments of recordings of 10 English speakers made at conferences: out of a total of 804 unfilled pauses.

TABLE 2

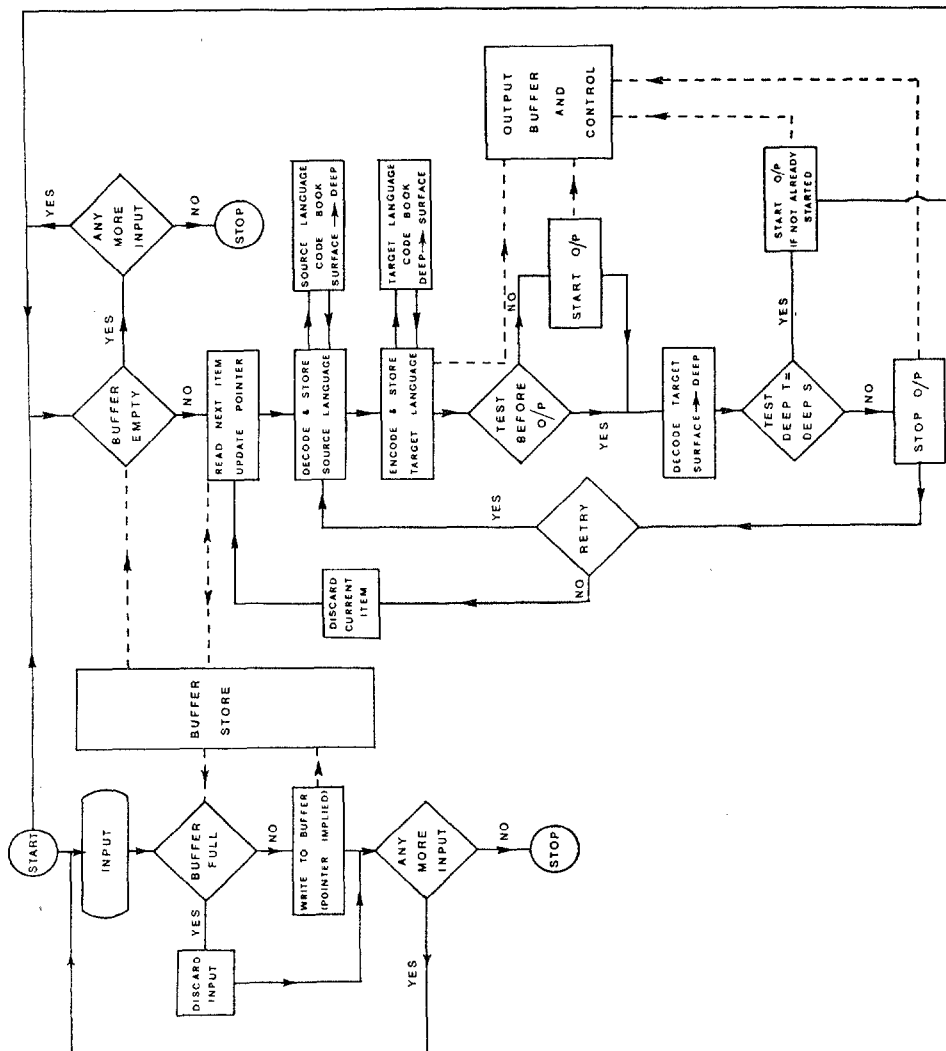
Distribution of unfilled pause times (in seconds)

1/4" — 1/2"	48%	
1/2" — 3/4"	23%	
3/4" — 1"	12%	83% 1" or less
1" — 2"	13%	
2" +	4%	17% over 1"

If the interpreter was already speaking at a rate between 96 and 110 words per minute (rates calculated from a number of interpreters) there is obviously not much he can fit into most pauses, but then neither can he avoid filling them if he is already speaking. This is not to deny that he may occasionally make use of the few longer pauses in the way suggested. In any case, it is doubtful whether simultaneous listening and speaking per se is that much of a problem for experienced interpreters. Difficulties are more likely to occur through the paced nature of the task, a speaker's particular accent or style of delivery, whether or not the interpreter can see the speaker, whether the subject matter or vocabulary are obscure, whether the interpreter has seen the script in advance, how he feels on that particular day, and so on.

In the experiment on the effects of noise, in the no noise listening condition over 85% of the source language text was correctly interpreted by some of the subjects even though they were simultaneously listening and speaking for over 75% of the total input-output time.

A computerised analysis (employing a pause criterion of 250 msec.) of 14 conference recordings from 5 to 20 minutes in length and of six interpreters working from French to English showed that on the average they spoke and listened simultaneously for 65% of the total time. Which brings me to my final discussion of how it might all be done. Perhaps this is best illustrated by the following flow chart — please forgive the complexity — but it is after all fairly complex behaviour !



The flow chart shows such permanent structural features as a short-term buffer storage, long-term memory for knowledge of languages, etc., and an output buffer. The diamonds represent control processes which can be selected at the option of the interpreter — and which may also determine the distribution of attention to the different components of the task.

Input Procedures

Source language input is received in a short-term buffer storage which is monitored, and is under the control of what might be called “input routines”, which are sensitive to the state of the environment (e.g. whether source language input is available), to the state of the buffer store — for instance whether further input can be accepted, or whether recent input is fading, and also implement whatever strategy the interpreter employs to segment the input.

“Operational” — or “Working” Memory

The term “operational memory” was suggested by Posner (1967) for memory involved in the active re-installment of permanently stored information — which seems a particularly appropriate term for the type of process involved when an interpreter accesses information regarding both source and target languages, and carries out the operations involved between perception of the source language message and vocalisation of the translation.

Decoding and Encoding

Kade and Cartellieri (1971), and Chernov (1973) propose a statistical or Markovian model of the translation process at this stage: an interpreter makes probabilistic inferences about the development of a sentence on the basis of what he has just heard, his own utterances are then based on these predictions, becoming more determined as the source language sentence proceeds. Now there can be little doubt that interpreters will make use of redundancy in order to predict sentence endings, but there are a number of problems with such a superficial model, apart from the familiar objections to Markovian models of language users cited by Miller and Chomsky (1963). In short there are far too many possible grammatical sequences of even 20 words for a person to have heard them in a lifetime, and since he could not predict without experience of similar combinations there would be many sentences he could never interpret. Finally there would be many structures such as embedded sentences which could not be constructed or predicted by a Markovian model of language.

Furthermore, how would such a model cope with the following type of sentences ?

J'ai lu la critique de Chomsky

- (1) I read Chomsky's critique
(criticism)
- (2) I read the critique (criticism)
of Chomsky

or

J'ai trouvé ce fruit délicieux

(1) I found that delicious fruit

(2) I found that that fruit tasted
delicious

Though context might provide cues, such sentences could only be disambiguated (as could the familiar "Flying planes can be dangerous") by reference to the deep structure, or underlying meaning of the sentence. I use the terms "surface" and "deep" not to imply any particular generative model of language, but to distinguish the sounds, words, and sentences heard by the interpreter from his understanding of their underlying meaning.

"Output Procedures"

In the flow chart it is suggested that the source language message — as understood by the interpreter is stored for comparison with the target language transform. While interpreting, the interpreter has the choice of beginning output immediately, or of checking whether his version of the segment of the source message is a (to him) satisfactory version of that message before starting his utterance. The evidence for such testing before output comes from conversations with interpreters who have stated that on occasion they may hold up output in order to check it (presumably against their recall of the source message, and/or whether it fits the context). Even if the interpreter does begin output without testing, testing may still occur, the matching of output with some template of the original can proceed. This may occur at either surface or deep levels, it may involve decoding the translation and matching the derived meaning with that of the original — as understood by the interpreter. If the match is satisfactory, the interpreter can proceed, if not output can be halted. The interpreter can then decide whether or not to try again, or on occasion he might even decide that he's getting too far behind or that the error is not critical, this would mean leaving the output as it was, and proceeding to the next portion available in store.

Gerver (Gerver, 1971) suggested that one possible model for linguistic, and other, aspects of simultaneous interpretation might be similar to Halle and Stevens analysis-by-synthesis model of speech recognition (Halle and Stevens, 1964). Halle and Stevens suggested that speech recognition involves the internal synthesis of speech according to certain rules, and the matching of internally generated speech patterns against the pattern under analysis. The generative rules for speech perception would be similar to those used for speech production. As Neisser (1967) has pointed out, such rules need not be restricted to the analysis of speech sounds alone, but can encompass larger units as well: "Auditory synthesis can apparently produce units of various sizes. The listener can ask himself "What sounds were uttered?" or "What was meant?" In each case he must have a set of rules: phonetic, syntactic, semantic or what you will... Hearing an utterance the listener constructs one of his own in an attempt to match it. Such matching may go on at various levels... that is in terms of segment sizes. If a single noise-masked word is presented, the listeners preliminary speech analysis may pick out distinctive

features or syllables which suggest a tentative answer ; various related words are then synthesized until one fits. If the stimulus is an entire sentence, a few words tentatively identified by the preliminary system may guide the synthesis of whole constituents as units, or even of the whole sentence”.

After synthesis of the source language message, then, the interpreter might generate his translation in an analogous way, except that the situation is more complex than with one language alone. The interpreter will not only have to know the rules at the different levels of analysis for each of the languages involved, but he will also have to know how, and at what levels, they correspond. The process of interpretation, then, would involve the continuous synthesis and analysis of possible translations of the source message. Continuous generation, monitoring, and testing of the translation against the source language as understood by the interpreter. According to this view, monitoring and possible revision are an integral part of the process, rather than an additional activity *after* translation.

I wish to argue that, though the main focus of the interpreter's activity and attention will be on the actual translation of a message, information may be acquired simultaneously in a buffer storage while a running comparison is carried out between former input and output. The model is consistent with recent discussions of theory and research in memory and attention by Norman (1968) and by Kahneman (1973). Norman, for instance has suggested that a number of storage mechanisms may be involved in human information processing — just as in other complex data processing devices temporary storage mechanisms are needed in order to maintain the results of intermediate steps in analysis. Kahneman reviews recent research on the problem of attention and concludes that, though attention is limited, the limit varies from moment to moment, and the amount exerted at any one time depends primarily on the demands of concurrent activities. Attention can, then, be shared among a number of tasks, contrary to the single channel hypothesis discussed by Welford (1968) — allocation of attention being a matter of degree, and under the control of a person's temporary intentions as well as permanent dispositions. At high levels of task load, however, attention becomes almost unitary. Thus in simultaneous interpretation an interpreter can, under normal circumstances (e.g. good listening conditions, moderate speech rate, easily grasped source language material) easily divide his attention between the various functions outlined in the flow chart, but when, for instance, coping with difficult listening conditions, and/or difficult subject matter, attention will be focussed on decoding and encoding, and monitoring of input and/or output might suffer. This was in fact found in the noise experiment (Gerver, 1974) in which the ratio of corrections to errors declined as noise increased. In other words, interpreters' response criteria declined, or rather they were unable to maintain them, and it is suggested that this was due to the increased difficulty in having to cope with degraded input.

Well, you might well ask, what does all this add up to? Certainly not a simultaneous interpreter, and an interpreter would no doubt tell me so in no uncertain terms. I have only attempted to describe some of the information handling properties of a mechanism that can cope with a number of tasks at once, and I've suggested that, though some aspects of the task may be carried out sequentially,

the task as a whole must involve some degree of attention sharing. I've also tried to hint at one or two of the properties of the linguistic analysis involved. There has not been time to mention, more than in passing, motivational and situational aspects of simultaneous interpretation. It is I hope a beginning, and certainly I would not agree with those interpreters who say you have to be one to know one, for just as I think the interpreter's skills can tell psychologists much about such topics of psychological endeavour as memory, attention, bilingualism, and the interaction of personality with task performance, so I hope that psychologists can eventually contribute to the interpreter's understanding of his own skills.

DAVID GERVER

REFERENCES

- BARIK, H. (1971) A description of various types of omissions, additions, and errors encountered in simultaneous interpretation. *Meta*, 16, 199-210.
- BARIK, H. (1973) Simultaneous interpretation: temporal and quantitative data. *Language and Speech*, 16, 237-270.
- CHERNOV, G. V. (1973) Towards a psycholinguistic model of simultaneous interpretation (in Russian). *Linguistische Arbeitsberichte*, (Leipzig), 225-260.
- GERVER, D. (1969) The effects of source language presentation rate on the performance of simultaneous conference interpreters. In E. Foulke (Ed.) Proceedings of the 2nd Louisville Conference on Rate and/or Frequency Controlled Speech. University of Louisville, pp. 162-184.
- GERVER, D. (1971) Simultaneous interpretation and human information processing. Unpublished D. Phil. Thesis. Oxford University.
- GERVER, D. (1974a) The effects of noise on the performance of simultaneous interpreters: Accuracy of performance. *Acta Psychologica*, 38, 159-167.
- GERVER, D. (1974b) Simultaneous listening and speaking and retention of prose. *Quart. J. exp. Psychol.*, 26, 337-342.
- GOLDMAN-EISLER, F. (1968) *Psycholinguistics: Experiments in spontaneous speech*. London, Academic Press.
- GOLDMAN-EISLER, F. (1972) Segmentation of input in simultaneous interpretation. *J. Psycholinguistic Research*, 1, 127-140.
- HALLE, M. and K. N. STEVENS (1964) Speech recognition: A model and a programme for research. In J. A. Fodor and J. J. Katz (Eds.), *The structure of language*. Englewood Cliffs, Prentice-Hall.
- KADE, O. and C. CARTELLIERI (1971) Some methodological aspects of simultaneous interpretation. *Babel*, 17, 12-16.
- KAHNEMAN, D. (1973) *Attention and effort*. Englewood Cliffs, Prentice Hall.
- LEONT'EV, A. A. (1973) Some problems in learning Russian as a foreign language. *Soviet Psychology*, 11, 1-117.
- MILLER, G. A., E. GALANTER and K. H. PRIBRAM (1960) *Plans and the structure of behaviour*. New York, Holt, Rinehart and Winston.
- MILLER, G. A. and N. CHOMSKY (1963) Finitary models of language users. In R. D. Luce, R. R. Bush and E. Galanter (Eds.), *Handbook of Mathematical Psychology*, Vol. II, New York, Wiley.
- NEISSER, U. (1967) *Cognitive Psychology*. New York, Appleton-Century-Crofts.
- NORMAN, D. A. (1968) Toward a theory of memory and attention. *Psychol. Rev.*, 75, 522-536.
- OLÉRON, P. and H. NANPON (1964) Recherches sur la traduction simultanée. *J. de psychologie normale et pathologique*, 62, 73-94.
- PRUCHA, J. (1972) Soviet Psycholinguistics, *Janua Linguarum Series Minor*, 143, The Hague, Mouton.
- WELFORD, A. I. (1958) *The Fundamentals of Skill*. London, Methuen.