ON THE REPRESENTATION OF VERBAL ITEMS IN SHORT-TERM MEMORY

A. J. W. M. Thomassen

ON THE REPRESENTATION OF VERBAL ITEMS IN SHORT-TERM MEMORY

PROMOTOR

PROF. DR. J. KREMERS

ON THE REPRESENTATION OF VERBAL ITEMS IN SHORT-TERM MEMORY

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN
DOCTOR IN DE SOCIALE WETENSCHAPPEN
AAN DE KATHOLIEKE UNIVERSITEIT TE NIJMEGEN,
OP GEZAG VAN DE RECTOR MAGNIFICUS DR. G. BRENNINKMEIJER,
HOOGLERAAR IN DE FACULTEIT DER SOCIALE WETENSCHAPPEN,
VOLGENS BESLUIT VAN DE SENAAT IN HET OPENBAAR TE VERDEDIGEN
OP DINSDAG 23 JUNI 1970 DES NAMIDDAGS TE 2 UUR

door

ARNOLD JOANNES WILLEM MARIE THOMASSEN

geboren te Eindhoven

Druk: Drukkerij Schippers
Nijmegen
1970

Acknowledgements

It is pleasure to acknowledge the help of the many who contributed in some way to this investigation and to its report in the present shape. Special thanks are due to those whose assistance and whose cooperation was obtained in the experiments. I also wish to thank the members of the Vakgroep Funktieleer for their support and for their patience during the past year. I am indebted to Miss W.M. Hoogstraten and Miss M.J. Dénis who did the typing. I should like to express my most sincere gratitude to my wife Jennifer Anne for her work on the English text and for her care of many typographic details. Special acknowledgement is also due to Dr. D.E. Broadbent, Director of the MRC Applied Psychology Unit, Cambridge, England. His hospitality during my visits in 1966 and in 1967 provided a stimulating opportunity for experimentation and for discussion. These visits to the APU were made possible by a fellowship from the Netherlands Organization for the Advancement of Pure Research (ZWO).

Contents

INTRODUCTION

Introduction

There are only very few people who cannot repeat a five-digit number spoken to them, or who cannot react adequately to a simple question such as "Have you had any breakfast yet?" The problem how these skills are performed is clarified to some extent by questions of how the verbal information contained in the two messages is analysed by the perceiver, how this information is retained in the correct order or related to existing knowledge about one's past, and how the required answer is assembled and produced. It would be a very ambitious enterprise to try and find the solution to these three questions all at once and so far no one has made the attempt. Conversely, it would be fruitless to set out solving any single one of the problems related to the above achievements without the conviction that, whatever aspect of this behaviour is being studied, it is always a form of processing of verbal items (e.g. words) that one is concerned with. When more is known about the representation of verbal items in any of these stages, advances may also be expected with respect to more complicated problems such as how different types of errors can be explained, what is the maximum amount of verbal information a person can take in at once, and what is the best way of presenting it to him.

This study is concerned with verbal material exclusively. Dealing with verbal items supposes the activity of speech mechanisms, which must be assumed to differ from non-speech mechanisms on neurophysiological grounds. This assumption is further supported by the fact that the sparse short-term memory studies performed with non-verbal items seem to indicate clearly that, in a large number of aspects, they are processed differently from verbal items (see Adams 1967; Sanders 1968).

Let us assume for the moment that there are three different levels of representation of verbal items, the lowest occurring at a very early stage as a prerequisite for adequate perception and thus preceding perception, the second occurring at a stage during which the items are perceived and actively held under attention, and the highest level at a more permanent stage when the items together with their meaning are put in a store from where they may be conjured up at any time if required. During the past decade a variety of attempts have been made to describe stores of the latter kind and laws governing item retieval from them (Mandler 1968; Kempen 1970) as well as item production in grammatical sentences. This will not be attempted here. In contrast, the present study is directed towards understanding how verbal items are represented at the two levels described as lower than the latter one. In other words, how are verbal items represented before their perception has taken place and what is their representation during the stage of active attention during and following perception and preceding the overt answer required?

Of course, these questions also have a history of almost ten years; they obviously belong to the domain of the study of short-term memory, which is also the area of interest in the present study. Short-term memory experiments deal with single presentations of usually verbal items and their recall or recognition by the subject immediately or soon after the last item presented. If the verbal items are meaningless or arranged in a meaningless order, these experiments provide an opportunity for studying the two lowest levels of representation mentioned above, without too great risks of performance being primarily based on the third or highest level where there is also a meaning component attached to their representation. Technically, short-term memory, and certainly the trigram STM, has come to denote a set of varying but specific assumptions about possible mechanisms. Therefore, although in the first chapter reference is made to several experiments which are in fact STM experiments, the designation 'short-term memory' will be avoided or used in a non-technical sense until by the second half of the first chapter the implied mechanisms are more precisely specified. It should be made clear that the term representation of a verbal item denotes a set of internal replications, each constituting its functional counterpart at a certain level of processing, and thus providing the cues for further processing. It is assumed that each replication involves a neurological analogue of the item, either as a specific firing pattern or as an increased level of activity in the specific set of neurons corresponding to that item.

The present study too has a history of its own. It was started as an investigation into the causes of inferior short-term retention in deaf children. Gradually its focus has shifted from short-term memory as such to the 'acoustic' factors found in verbal short-term memory experiments and to the superior capacity of the auditory system for assimilating discontinuous sequential information. These acoustic aspects were, again, shifted to a second plane as evidence seemed to decide for articulatory representation in short-term memory and in other modes of dealing with verbal items.

It is possible to retrace the sketched history in the individual chapters and in their order. Chapter 1 discusses the function of immediate storage, provides arguments for assuming different stages of assimilation of information and deals with some different types of approach in the field of immediate memory. The chapter is closed with a qualitative model to serve as a framework for most of the experiments in the following chapters. Chapter 2 supplies a number of empirical data and experiments of theoretical importance with respect to the 'acoustic' mode of representation of verbal items in short-term memory. The data include findings on modality effects in the recall of non-verbal signals and on the immediate recall of verbal items by the deaf. The chapter is concluded with some evidence in favour of articulatory representation. This line is continued in Chapter 3 where speech perception and speech production are discussed. The kernel of the chapter is formed by an elaborated model derived from the model of the first chapter. Experimental tests of details of the model deal with verbal short-term memory and other verbal tasks. The discussion points at a number of unsolved problems and indicates possibilities for further approach.

CHAPTER ONE

Chapter 1

The study of verbal memory over short intervals

1.1 Notes on the function of immediate memory

The present section is concerned with a number of situations in which the active retention of verbal information over a relatively short interval is required for adequate performance. In order to illustrate the general importance of such retention, the examples are partly chosen from outside the laboratory. Furthermore, they illustrate two different aspects in the function of immediate storage. One is related to maintaining the availability of information after it has been perceived and the other, which is of great importance in the perception of speech, enables the perceiver to delay his perceptual decisions so that these can be made with greater efficiency. A specific characteristic of the function of immediate storage is perhaps best described as one of a rough copying register from which information may be transferred into a more permanent store with different retrieval properties. As such it prevents overload of these permanent storage and retrieval mechanisms.

The relevance of an immediate memory system for human behaviour may be demonstrated by a few of the many sequential tasks performed in everyday life. An example which is often quoted in this context is dialling a telephone number. The number has to be looked up in the directory and subsequently retained for some time during the operation of the dial. For a five-digit number the operation will take approximately eight seconds. This 'retention interval' is filled for the most part by turning the dial, by waiting for it to return and by selecting the next digit. During this interval, and especially due to the latter activity forgetting will occur. This in turn may be counteracted by subvocal rehearsal, but any interruption by an outside event requiring our attention during this process will still endanger the retention of further digits to be selected. It may necessitate the whole sequence to be started over again, beginning by a repeated search for the same number in the telephone directory. Under less disturbing conditions, however, the capacity of the immediate memory system will be sufficient for this kind of task.

Retaining information over a period of seconds is equally essential, however, in tasks far more fundamental than the one just described. Especially verbal communication requires a kind of storage which enables the subject (S) to hear what is 'being said' either by himself or by someone speaking to him. What is needed is an extension of the actual present formed by the split-second which is too short even for one syllable to be uttered. One syllable in isolation would often not be recognizable. The intelligibility of a spoken message is, indeed, greatly enhanced by its sequential redundancy and in order to profit from this redundancy provided by the very recent past S must have the information of recent verbal events available to him. This will enable him to recognize phonemes, morphemes and sentences.

The advantage of a kind of buffer storage in speech perception has been emphasized e.g. by Miller (1962). He suggested that less time for perception is required if perceptual decisions are postponed:"In order to comprehend messages spoken at 150 words/min, we would presumbly have to make about a dozen phonemic decisions every second and perhaps 100 phonetic decisions ... A single delayed decision would require far less time than would a series of immediate decisions" (pp. 81 - 82). Especially if there are considerable sequential constraints in the message (such as in normal speech) S will do better if he stores several elements of information in order to accumulate evidence until sufficient information is available to make a single perceptual response. It is still an unsolved problem what the smallest elements are that are stored until a decision is made leading to the next stage of processing. Perceptual delay at the macro-level of words and phrases has been demonstrated by Moray and Taylor (1958). They asked their Ss to shadow one of two dichotic messages and found that in many instances they responded in a 'discontinuous' fashion, i.e. by groups of words separated by silent periods. The authors reported an increase of this tendency with greater redundancy of the stimulus material and a tendency toward 'continuous' word-for-word responding with uncertain messages.

Another method of studying amount of delay before perceptual processing of speech units is suggested in an experiment by Ladefoged and Broadbent (1960). They presented a verbal sequence with somewhere in the sequence a click or an s-sound. The Ss generally indicated the location of this signal several words before the place in the sequence where it actually occurred. These authors interpret the effect in terms of selective attention to the signal, leading to 'prior entry' of this stimulus; but another interpretation, which is in agreement with the above findings by Moray and Taylor (1958), is that the subjects located the signal not at the word that was presented along with it, but at the word that was being identified while the click or s-sound occurred. The additional finding that the click was indicated even earlier if it was presented late in a sentence, and that the difference between actual and perceived occurrence decreased with non-redundant digit messages, further support this interpretation in terms of preperceptual storage.

In fact, even when the discussion is restricted to speech perception, one may recognize both types of storage discussed above in this single process. First, there is storage leading to greater amounts of information which, after necessary delay, can result in a relatively easy identification of words or phrases. Second, there is storage of words and phrases already identified which is a prerequisite for the understanding of the complete message. It is very likely that these two kinds of storage are, in fact, merely two stages within a complete storage hierarchy ranging from the smallest phonetic elements with minimal duration to lengthy and complex sentences.

Apart from being a prerequisite for the assimilation of sequential information, especially in the case of verbal communication, the immediate memory system may also be considered as a mechanism which prevents undue overloading of permanent memory. It is characteristic of many tasks such as the ones mentioned that either the appropriate response to every element is given immediately after cessation of the message (e.g. dialling a telephone number) or that there is only one response in more general terms to a message of some length (e.g. answering a question). In neither case is there any need for permanent storage of the detailed information originally present in the message once it has been dealt with. Long-term storage of such details seems to be restricted to messages that are either repeated more often or that arouse exceptional interest or emotion. If all verbal information were stored permanently in its original elementary form, chances of interference from earlier similar verbal experiences would be great. The original format, moreover, would have to be paralleled by another kind of storage which allows direct access to the semantic content of these messages. Absence of such a parallel system would require a complete running through of the original verbal sequences in search for their meaning whenever this would be wanted. From these considerations the adaptive significance of a separate immediate system acting as a buffer mechanism, which itself leaves little or no permanent traces, can be understood. In fact, even if such permanent traces could be formed under the conditions that prevail during speech communication, one might wonder whether the memory space as defined by the number of neurons and their interconnections in the human brain would be sufficient in view of the enormous amount of verbal stimulation dealt with in the course of a lifetime.

The succeeding section (1.2) will discuss arguments for such a distinction between temporary and permanent types of memory. It is not customary to consider storage preceding perception as a form of memory. Correspondingly, with respect to preperceptual storage there will be little doubt as to its being distinct from permanent storage. Any deficiencies in dealing with preperceptual (visual) information,such as are known e.g. in the literature of certain brain damaged patients are usually discussed within the framework of (visual) perceptual deficiencies, unrelated to forms of permanent memory. Similarly, in the following pages the distinction between temporary and permanent modes of storage is restricted to the distinction between the two postperceptual types of storage, one temporary and one relatively permanent.

1.2 Evidence on different neural storage mechanisms

After having pointed out the functional properties of an immediate storage mechanism and having indicated the importance of distinguishing a temporary buffer type store from a more permanent memory mechanism we shall now look at sources of evidence from which it appears that the latter differentiation probably reflects a distinction between different neural mechanisms. The reason for providing this evidence is not so much to show that immediate memory as studied in experiments with verbal material is a neurological entity as to demonstrate the principle of qualitatively different stages in processing and storage in the central nervous system (CNS). As examples are chosen the memory defects shown by older people, by patients suffering from the Korsakoff syndrome, and by people showing the effects of surgical removal of certain portions of brain tissue. Subsequently there will be reference to neurophysiological findings related to the different stages to another. These four paragraphs do not have the status of systematic reviews; they merely point at the fact that there are instances in which permanent and temporary memory mechanisms may be affected differentially and at the possibility of controlling permanent storage by experimental manipulation at the level of the CNS. Furthermore, the discussion does not enter into the question whether retrieval rather than storage is the explanation of the memory defects.The section will be concluded by notes on the relation between neural storage mechanisms and immediate memory for verbal material.

1.2.1 Memory dysfunction in ageing

The increasing difficulty experienced by ageing people of permanently retaining new impressions - rather than of remembering old events - may be considered as a source of evidence for the distinction under investigation. The literature on this matter, however, is far from conclusive. Decreased retention of new information over longer intervals is a common result, but some authors relate it to an inferior registration or inadequate coding and short-term storage (e.g. Welford 1958), while others, who find an almost equal immediate memory performance in young and old people (e.g. Heron and Craik 1964) are inclined to localize the memory defect in the transition to permanent memory or in the latter mode of storage exclusively The problem here is twofold. Firstly, a considerable number of experiments on ageing have used complex memory tasks, which do not fail to show the effects of increasing age. Secondly, older people sometimes show certain memory defects which - although often accompanying senescence - are not specific for old age. Traditional, simple memory span tasks are much less sensitive to age differences than more complex immediate memory situations. In those cases where defects are reported they may, more often than not, be explained by the complex nature of the task used. Some examples are the following. There are greater difficulties for older people in tasks involving dichotic listening (Inglis and Caird 1963; MacKay and Inglis 1963) or bisensory presentation requiring rapid alternation between eye and ear (Broadbent and Gregory 1965). Similarly, when the material is a complex visual sequence (Wallace 1956), when interference material is present (Kay 1959; Broadbent and Heron 1962), when many items are presented, and in tasks requiring organization at perception (Heron and Craik 1964; Craik and Masani 1967) the aged are at a disadvantage.

In all these cases there is reason to interpret the inferior results of older people not in terms of a decreased short-term capacity to hold information per se, but rather in terms of an inability to switch attention in order to organize the incoming information at perception. This interpretation is in agreement with Welford's (1958) suggestion about a reduced channel capacity in aged people, which need, however, not rule out normal immediate recall of relatively simple material presented at a moderate rate. Such normal performance has indeed been reported (e.g. Gilbert 1941; Doppelt en Wallace 1955; Broadbent and Gregory 1965; Heron and Craik 1964;

Craik and Masani 1967) and this evidence may be taken to suggest that with increasing age there is not necessarily a loss of the capacity for adequate perception, encoding, rehearsal, and recall. Any defects become apparent only when, somehow, complexity introduces extra interference. If it is assumed that this is because of a lack of spare capacity for organization or recoding, the same limitation may explain an inadequate permanent storage by older subjects.

The latter interpretation thus locates the retention defect at the transition from temporary to permanent storage and points at problems arising from a limited capacity which is adequate for immediate memory performance as such but insufficient for additional recoding operations during perception and rehearsal to ensure more permanent storage. Our discussion, however, is not quite satisfactory. Ageing is a multidimensional process involving a number of factors. Moreover, performance in older people especially is likely to reflect the personal level of motivation and attention and the individual history of the subject. Pure ageing does not seem to exist and therefore memory disorders merely due to ageing are hard to describe. Equally difficult in the interpretation of results obtained from old people is the isolation of pure memory effects.

1.2.2 The Korsakoff syndrome

In the clinical literature on memory defects there is frequent reference to the Korsakoff syndrome. The syndrome, which may appear in patients as a result of brain damage due to chronic alcolholism or to brain injury (thalamus, hippocampal region, mammilary bodies) has as its main characteristic a severe amnesia. Intelligence defects are usually absent in the patient and he retains the capacity to retrieve and apply knowledge and abilities acquired before the onset of his illness. New information, however, cannot be held for more than a few seconds or minutes: memory span tasks may be performed adequately but retention over longer intervals is severely impaired, if not impossible (Williams and Zangwill 1952; Barbizet 1963). This memory defect is sometimes related to that resulting from old age (Kral 1958) and to that incurred by the type of brain operation which is mentioned below (Milner 1959), but it would seem difficult to determine to what extent each of these reflect the same amnestic syndrome. Clinical tests and observations thus tend to support the memory distinction also in the Korsakoff syndrome, but the other symptoms ('temporal and spatial disorientation' and 'confabulation'), although certainly better defined than the symptoms of ageing, are equally likely to obscure the amnestic effect. Here may lie one of the reasons that the Korsakoff syndrome has received comparatively little attention outside the fields of psychiatry and clinical psychology.

1.2.3 Bilateral hippocampal lesions

Strong arguments come from the study of memory in patients suffering from lesions in the hippocampal region (Milner and Penfield 1955; Penfield and Milner 1958). Relief from severe focal epilepsy may in some cases be obtained by the surgical removal of portions of the temporal lobes. Typically, these ablations involve the deeper structures, situated at the mesial aspects of the temporal lobe (hippocampus, hippocampal gyrus, uncus, amygdala). If the lesions are bilateral, however, or if the other lobe was damaged before the operation, the consequence may be a clear-cut amnestic defect. After such an operation the patients, especially those with larger excisions, may show an almost complete loss of the ability to store new information permanently. The only new learning that seems possible is perhaps the acquisition of new motor skills. Earlier knowledge and skills and other functions, including intelligence, attention and immediate memory span as measured with formal tests remain, however, unchanged. Small amounts of new material, if capable of verbalization, may be retained for several minutes by deliberate rehearsal, which accounts for the normal span performance; but any distraction from such concentrated rehearsal activity immediately leads to forgetting. Material which does not allow verbal coding is lost in about 30 sec even without interruption by outside events. Moreover, the postoperative disruption of long-term storage of new information is apparent regardless whether the manner of testing is recall, recognition, or relearning (Milner 1959, 1966, 1969). The small number of patients displaying the effect in its pure form, as well as the clinical nature of the observations, might cast doubt on the reported findings. However, there are some convincing data to support the generality of the results. The symptoms can be reproduced temporarily in patients with unilateral lesions if an intracarotid injection of sodium amytal is applied to inactivate the other, intact hemisphere and not if the damaged hemisphere itself is inactivated (Milner 1966). Succesful attempts have furthermore been made to demonstrate that the postoperatve amnesia in one of the patients described earlier by Milner (1959), while practically ruling out all permanent acquisition left his immediate memory performance (viz. the rate of short-term decay) undisturbed (Wickelgren 1968).

Together with the evidence on ageing and the Korsakoff syndrome, the above neuropsychological findings on bilateral hippocampal lesions provide a basis for the distinction we wish to make. On the one hand there is no interference with immediate recall or recognition, or with prolonged rehearsal; nor is there a disturbance of the availability of old habits established before the operation. On the other hand new material presented repeatedly or rehearsed succesfully for some time does not leave a permanent trace. The latter effect justifies the assumption of two distinct stages in the normal history of information entering human memory, separated by a transition process. The transition from one stage to the next, which normally occurs if rehearsal is possible, is under the control of hippocampal activity. Once consolidation has taken place, the activity of the intact hippocampus is no longer required. Such dual storage interpretations of memory have been proposed by Milner (e.g. 1959). They have been reemphasized by Weiskrantz (1964) and several others, most recently e.g. by Baddeley and Warrington (1969).

1.2.4 Some neurophysiological data

Electroconvulsive shock (ECS) has been applied frequently in animal experimentation as a method of interfering with activity in the CNS. If rats are trained to avoid a charged grid and if an ECS is given immediately after each training trial, there is little evidence of any learning. If, however, the interval between these trials and ECS is extended, the interference with learning gradually declines over a period of about 1 hour (Duncan 1949; Thompson and Dean 1955). In order to control for the aversive effect of the ECS the most reliable results are obtained in single-trial avoidance situations (Madson and McGaugh 1961; Heriot and Coleman 1962). The effect may be explained by a disruption of 'reverberating' activity at various stages of a consolidation process which lasts about 30 min to 1 hour. After this period, interference by ECS is no longer possible, unless the process of consolidation is deliberately slowed down by reducing brain activity (Gerard 1953; Leukel 1957). Conversely, a reduction of the ' fixation time' was found to result from enhanced neural activity (McGaugh 1961).

The interference effects are most probably related to the restricted retrograde amnesia covering approximately the last 30 min before concussion (Russell and Nathan 1946) and before electroshock treatment in psychotic patients (Williams 1950). The effects provide supporting evidence for Hebb's original assumption of 'activity traces' which gradually bring about permanent storage, consolidated in 'structural traces' (Hebb 1949, pp. 60-66). The above findings furthermore suggest that the early memory processes involve electrophysiological events, while the permanent storage resulting from these events is probably of a non-electrical kind.

The problems concerning the exact mode and place of storage in memory are of great current interest: they are by no means solved. The mode of storage is being pursued along structural and biochemical lines; arguments for both intra and interneuronal changes have been put forward (See Grossman 1967; Gurowitz 1969). With respect to the place of storage, the conception of a side-by-side location of specific 'engrams' has been doubted since Lashley (e.g. 1929) failed to demonstrate a direct relation between destruction of specific portions of brain tissue and the retention of any type of learned behavior in rats and monkeys. As an illustration of such a complex neurophysiological representation of memory may serve the tentative model that has recently been suggested by Pribram (1969). The model describes storage in memory terms of holographic interference patterns, resulting from interacting fields of firing neurons, distributed throughout the brain, including the frontal and posterior associative cortex and the brain stem.

The data presented in the last four paragraphs have a widely varying source, background, and significance. They show coherence, however, to the extent that they all point out that permanent storage in memory does not occur at once, but rather in qualitatively different stages. Immediately upon the entrance of new information, there is a dynamic activity, partly taking place in an autonomous fashion; this activity is enhanced by repeated presentation or (which may be equivalent) deliberate rehearsal. As long as a subject can devote his attention to this information, the material can be retained and used for recall or recognition, provided there was no interference and not too much presented at once. This first process seems to involve electrophysiological events, which do not require a completely intact cerebrum. Simultaneous with these processes that make immediate retrieval possible, a consolidation process is started, which ultimately results in permanent storage. The consolidation process which may involve the function of the temporal lobe, is based on an electrophysiological activity with a normal duration of 30 min to about 1 hour. The long-term storage it brings about is structural after probably being first mediated by chemical processes, which in turn were started by the reverberating neural activity.

Although this has not been stated explicitly, it should be noted, perhaps, that the normal consolidation phase probably extends for some time after the rehearsal activity (as a deliberate process) has ceased. It is unlikely that a few minutes of rehearsal will directly result in truly permanent storage from the very moment that rehearsal stops. Much more probable is the assumption that rehearsal induces the kind of activity which enables the hippocampal structures to bring about - through a neural process of much longer duration - a final, permanent trace. This assumption is in better agreement with findings on the critical periods involved in retrograde amnesia in humans and in animal experimentation. More attention to the concept of rehearsal will be given in chapter 2 (2.1.4).

The nature of the trace is not the only way in which relatively permanent storage is different from temporary storage. The former is also much more closely integrated into existing networks that stretch out over large parts of each hemisphere. Integration into existing networks is probably promoted by activating certain networks during presentation and rehearsal, which is what happens when the subject recodes the information. In older people there is not much spare capacity left for this purpose, so that the non-integrated information is soon lost. In Korsakoff patients and in the postoperative amnesia described (perhaps also in ageing people when their temporal lobes are subject to degeneration) there is a reduction or absence of the consolidating activity so that the chemical and, ultimately, structural changes are not produced. The older networks, however, are retained, and retrieval of information from these networks remains possible.

1.2.5 Neural and psychological storage mechanisms

A number of questions remain regarding the relation between the different storage mechanisms discussed in the two preceding sections. In the former section two 'immediate' and relatively raw forms of storage were mentioned, both concerned with tasks such as the perception of speech. They can be described as a preperceptual storage mechanism for adequate perception and a postperceptual storage mechanism for retention and further processing of the verbal message. Both these mechanisms were distinguished from a third storage mechanism to which processed information is transferred and stored according to its semantic content. In the latter section also such a distinction was made between temporary and permanent storage mechanisms. It appeared possible that the temporary store exerts its function relatively undisturbed even if the transition to the permanent store is seriously handicapped by senescence, lesions, or disturbed brain activity. The question is whether in both sections the same mechanisms are discussed or that only an analogy exists between them. An important point in this discussion is formed by the duration of the interval required to attain a permanent form of storage. According to Rosenzweig (1969) some physiologists use the term 'short-term memory' to designate a consolidation period of hours and even days, while experimental psychologists reserve the term for processes of less than a minute. It is possible that at the behavioural level, and especially in verbal behaviour the observable transition from temporary to permanent is achieved relatively fast, or rather that the aftereffect of verbal stimulation vanishes completely in a matter of seconds if during this short interval no associations have been made with larger semantic networks. It is in correspondence with the discussion thus far if the verbal items in immediate storage are assumed to have a very short life indeed. Storage time may be increased by recoding at perception and rehearsal. These operations would thus result in a more stable immediate trace but they would also increase the probability of the start of an integration into semantic networks. These networks themselves may as a result of rehearsal also be activated, and performance such as recall might soon be based on these semantic structures as well as on immediate memory. This, however, does not mean that integration has already been attained and that truly permanent storage already exists. It is quite probable that many further minutes and even hours or days of consolidating activity are needed for this, and that it still remains possible to prevent permanent storage during this period. Characteristic for the first few seconds is, at any rate, that such integration has not yet, or only barely, begun.

A further question is concerned with the relation between the two 'immediate' forms of storage. They have both been labelled 'relatively raw' storage mechanisms because of their lack of semantic integration, but obviously postperceptual storage is much less 'raw' than preperceptual storage, because during perception, considered as a recognition process, some classification is applied. Preperceptual storage in auditory speech perception may then be regarded as a stage during which the evidence for a certain classification accumulates and postperceptual storage is concerned with holding a set of classified units under attention until a further perceptual (or classificatory) decision can be made. Indeed, it is quite arbitrary at what level one may decide that a verbal unit has been 'perceived'. Thus there is some arbitrariness in naming certain stages preperceptual and other stages postperceptual. It is, however, reasonable to consider as the smallest verbal units capable of perception those units for which (learned) verbal responses are available in a permanent verbal category store, and to expect that these units are activated during rehearsal. Not always will the smallest units be used for coding and rehearsal; especially if the (auditory) material presented is redundant, such as in speech, or if the manner of presentation suggests larger units (e.g. two-digit numbers rather than single digits) there will be an extra preperceptual delay reducing the number of verbal units to be rehearsed. Thus, at least in auditory tasks there is difficulty in deciding what items are still in a preperceptual stage and what items are already in a postperceptual stage of processing. In speech perception and in auditory tasks used to study immediate memory they will both be involved in a temporal overlap.

1.3 Various approaches in the study of immediate memory

The last section was closed by a discussion of the relation between immediate preperceptive and postperceptive storage mechanisms and it was concluded that in speech perception and immediate memory tasks both mechanisms are involved during the first seconds after presentation. The present section deals with some of the experimental and theoretical approaches of immediate memory for verbal material. No comprehensive account of the large amount of theoretical and experimental work in the field of immediate or short-term memory will be given. Such reviews are available in the literature both in article and in book form (e.g. Melton 1963; Peterson 1963; Posner 1963; Brown 1964; Sanders 1964; Süllwold 1964; Aaronson 1967; Adams 1967; Murdock 1967; Neisser 1967; Atkinson and Shiffrin 1968; Norman 1969). Here only a number of topics will be dealt with. They include in a historically determined order the simple 'immediate memory span' (IMS) tests dating from the beginning of the century and the results of factor-analytic studies performed since the thirties. Both these approaches demonstrate on the one hand the relatively great amount of interest for immediate memory and on the other they provide a justification for studying as a separate 'ability' the retention and reproduction of verbal material over short intervals. Modern ways of approach have to some extent developed outside this tradition. Some were inspired by information and communication theory in the fifties. The communication models, however, similarly stress the discontinuity between temporary and permanent storage principally by restricting themselves to the former mechanism exclusively. The mathematical models dating from the sixties are in part based on the information and communication models but they often do not include the distinction between short-term and long-term storage and if they do the distinction is not as much a discontinuity as it is in the communication models.

1.3.1 Studies of immediate memory span

Immediate memory span (IMS) is the maximum number of items that in 50 per cent of the trials can be recalled by S in their original order immediately after a single presentation. For verbal material IMS ranges between 5 and 9. A review of the early literature on memory span is provided by Blankenship (1938). The vast majority of the 146 references listed in the bibliography of this review deal with memory span as a means of assessing a S's individual capacity and most contributions to the literature covered in this review are made by psychologists engaged in 'mental testing'. Important questions such as whether or not memory span is a specific ability, what are the factors that affect IMS and what is the position of memory span in intelligence and clinical testing, are all touched upon in this literature. However, practically all these questions remain unanswered, in part as a result of large variations in the manner of collecting the data and in part because of a lack of theortical issues on the subject. Among the few things on which most of the early investigators agree is a definition of IMS and on the evidence that with various types of material relatively high reliability coefficients for memory span performance may be obtained, while, moreover, IMS may be measured for all modalities. Several authors, e.g. Martin and Fernberger (1929) demonstrated a large practice effect on IMS. In their case this was the result of practice spread out over several months, but these investigators also indicated that any recall of more than five units should be attributed to subjective grouping on the part of S. An unambiguous effect of rate of presentation on IMS could not be determined, but most authors agree that there is an optimal rate, which is set e.g. by Terman (1916) at slightly faster than one item per second.

Points of interest that are raised in the literature before 1938 are the processes involved in memory span and the distinction between afterimage, memory span, and permanent memory. 'Attention' and 'associability'- by which is meant the ability of S to group the series of elements together - are repeatedly mentioned as requirements for performance in a memory span task. A further factor is 'language ability' needed for the reproduction of the series. Lastly, 'imagery' is another process claimed to be involved; it was recognized by Richet as early as 1886 that this process differs from the mere use of an afterimage by S. The distinction between permanent memory and memory span was pointed out by Binet in 1894.The differences referred to may be summarized as differences in permanence and amount of material in the two memory systems. Also the fact that under certain conditions memory span may be impaired and memory not, or vice-versa, is quoted as evidence for this distinction.

From a great number of publications in which correlation among intelligence and memory tests are discussed, the conclusion may be drawn that immediate recall of verbal material constitutes a separate factor. Several studies report or suggest the specificity of such an 'immediate memory span' faculty on the basis of mere inspection of raw performance data or correlations, (e.g. Smith 1903; Radossawljewitch 1907; Compton and Young 1933; Hall 1936; Garrett 1938; Balinski 1941; Thomas and Young 1942; Johnson 1955). It is of interest to note that the earliest of these articles, which are quoted by Süllwold (1964), were published right at the beginning of the testing era when, of course, no factoranalytic or similar techniques were available. In later studies a formal factoranalytic approach has often been followed (e.g. Woodrow 1939; Brener 1940; Thurstone 1941; Kelley 1954; Katzenberger 1965). The results of all these studies support the existence of a separate immediate memory factor with which verbal tasks such as the classic 'digit span' are highly correlated and which, itself, shows only a moderate correlation with tests measuring learning ability and intelligence.

1.3.2 Information theory and verbal short-term memory

Immediate memory with its relatively well defined capacity has attracted research by many information theorists. Typical for a communication system, as described in terms of information theory, is a limited capacity channel. At first sight, the limitations of the immediate memory span would fit such a description. However, from the very early applications of information theory to immediate memory it has been clear that the other requirement of an information channel, namely that its capacity is limited to a maximum amount of information (expressed as a maximum number of bits per sec) is not met by the facts. If a subject can retain 8 decimal digits in immediate memory, he transmits 26.6 bits of information; the corresponding 26 or 27 binary digits would, however, clearly be beyond the span capacity of this subject. It is implicit here, that the digits are recalled in their correct order. If it is assumed that serial order is a different kind of information which has to be retained separately, better agreement with information models may be attained (Crossman 1960, 1961), but the price thus paid is the neglect of the specific characteristic of immediate memory span that the digits are retained in the correct order. The same problem is encountered when digit and letter spans are compared. Roughly equal amounts of information are carried by 8 decimal digits and by 5 to 6 letters from an alphabet of 26.

The differences that do exist between digit and letter span are,however, smaller (Hayes 1952) and these can be explained better by reference to (acoustic) properties of the symbol vocabulary than by the size of the vocabulary, which defines the information content of the symbols (Conrad and Hull 1964; See Chapter 2: 2.1.2).

The fact that some letters are more frequent in the language than others leads to the prediction that there will be differences among letters in their probability of immediate recall. High frequency letters carry less information and are therefore more easily retained. This was confirmed in a study by DiMascio (1959) who found a positive correlation between letter frequency and memorizability and in an experiment reported by Underwood and Schulz (1960) showing that frequent letters are learned easier as response items in a paired-associate (PA) task. In another study, however, using visual letter-by-letter presentation of six-consonant sequences for immediate ordered recall no effect of letter frequency was found (Conrad, Freeman, and Hull 1965).

If sequential dependency is taken into account a large effect of redundancy is to be expected. Such effects have indeed been observed. Miller and Selfridge (1950) presented lists of words with different contextual restraints. They found that after a single presentation longer lists (up to 50 words) were recalled better with higher order approximations; criterion performance on the longest lists was reached with connected text only. Short lists (10 words), however, required only second-order approximation for recall performance at the level of connected text. To the extent that extrapolation to even shorter lists is allowed, these results indicate that recall of word lists of approximately span length only minimally reflects the redundancy effect of contextual constraints. In this study the constraints were, of course, provided by the rules of language. Other rules may also be learned, as has been demonstrated by Aborn and Rubenstein (1952) who trained their subjects during three days to learn certain constraints in the composition of their material (nonsense syllables). This resulted in subsequent superior recall of lists made up according to these rules, and an approximately constant rate of information transmission was found across redundant and non-redundant lists. However, above a certain limit the amount of information recalled from redundant sequences was considerably less than that from random sequences (Rubenstein and Aborn 1954).

In fact there is no single study on immediate memory for verbal material which demonstrates the role of information content on the amount recalled without also suggesting that this role is far from decisive and subject to a number of conditions. For example the high correlations found by Miller, Bruner, and Postman (1954) and by Baddeley (1964a) between the correct report of simultaneously presented whole eight-letter sequences and their sequential dependency are not observed when a sequential letter-by-letter presentation is used. This suggests that the effect of redundancy is dependent on conditions which facilitate the detection of the 'lawful' structure of the material such as in the visual, simultaneous presentation of sequences as a whole. It may be that some extra memory capacity is required to detect redundancy in successively presented material, so that the net gain due to redundancy after successive presentation is much smaller than expected on the basis of sequential dependency only. However, there is an effect of sequential dependency, also after successive presentation. Under free recall conditions and also after longer retention intervals a large effect of sequential redundancy was observed by Miller (1958). Significant correlations in the order of $\underline{r}_s = .30$ between predictability of six and seven-consonant sequences (in terms of digram frequency rank order scores) and probability of correct ordered recall have been reported by Baddeley, Conrad and Hull (1965) and by Conrad, Freeman, and Hull (1965).

Miller (1965a) suggested the principle of recoding into chunks as an explanation for the fact that, after some practice (or with redundant material) the memory span seems to increase. The maximum number of chunks, then, would still be limited to about seven by the memory span capacity, but the number of original stimulus items (and therefore the amount of information) recoded into each chunk depends on the degree to which they allow such organization. These ideas are closely related to the observations on practice and grouping by Martin and Fernberger (1929) and other authors as mentioned in the preceding section of this chapter (1.3.1).

1.3.3 Communication models of verbal short-term memory

Historically the most important communication model of immediate memory is the 'mechanical model' proposed by Broadbent (1957). It assumes a filter which selects information from only one among several input channels and passes it on into a limited capacity mechanism for further processing. A channel is chosen on the basis of its physical characteristics; switching between channels is possible, but it takes time. Short-term storage of information in an input channel may be required if the processing system cannot handle all the information presented on that channel or if the information is presented in a channel which is not being selected. There is rapid decay (in the order of a few seconds) of the trace in the input channel, but once a signal is attended to, rehearsal may put its fading trace back into the system, which is represented as a 'recurrent circuit'. The model is based to a large extent on the results of multi-channel experiments which indicated temporary storage of non-attended auditory information (Cherry 1953; Cherry and Taylor 1954; Broadbent 1956) and on the research on immediate memory by Brown (1954, 1955), which included experiments with controlled opportunity for rehearsal.

The main features of the model are preselective storage, later called the S system (Broadbent 1958), the selective operation by the filter, the limited capacity processing mechanism, or P system, and the rehearsal cycle. Broadbent's (1957, 1958) early models are the only formal communication models of immediate memory in the sense that they apply not only the terminology of communication theory but also state the properties of e.g. the P system in informational terms; it is the rate of information, not the amount of stimuli, which determines the spare capacity in the memory system. In later versions by others, and also by Broadbent (1963), the formal informational approach has been left. The latter model also makes a clearer distinction between the storage of unselected material and material that has been accepted and received at least one response. Broadbent's suggestion is, however, that both the preselective and the postselective store are subject to decay with time, be it at different rates.

On the preselective side some evidence for such decay comes from the dichotic listening tasks already mentioned: there is a rapidly deteriorating short-term storage of unattended auditory signals. That there is also a visual analogon of this kind of storage has been shown in experiments using tachistoscopic exposures of complex visual arrays (Sperling 1960; Averbach and Sperling 1961; Averbach en Coriell 1961). It was shown in these experiments that a large amount of visual information is available in a very short-lived store (a fraction of a second) from which the subject may select only a limited amount for recall before the information has decayed completely. The rate of decay could also be determined in these experiments. On the postselective side Broadbent's (1963) arguments for autonomous decay with time can be summarized as demonstrating that perceived but unrehearsed material is forgotten as a function of time rather than as a function of the nature of the activity by which rehearsal is prevented. A strong point is made with respect to the distinction between, on the one hand, this kind of postselective short-term memory, which is mainly subject to decay occurring when rehearsal is impossible or delayed and, on the other hand, long-term memory, resulting from repeated presentation or prolonged rehearsal, where interference with storage by other stored material is a function of the similarity between the two different materials and not a function of time. To this issue we shall return in the next chapter (2.1.2).

Several contributions to the building of communication-type models of short-term memory were made by Sperling (1963, 1967). In the latter paper three 'successive approximations to a model for short-term memory' are suggested. The third of these, which incorporates most experimental data on the recall of tachistoscopically presented letter arrays, is characterized by a visual information store (VIS), a scan-rehearsal component, and an auditory information store (AIS). A scanning mechanism reads information at high rate from VIS. The scanned visual image is fed into a recognition buffer where it is converted into a programme of motor instructions which are stored in this buffer. The rehearsal mechanism may execute the programme; this process is, however considerably slower than setting up the programme. Rehearsal output is entered and temporarily retained in AIS. If recall is delayed, a second rehearsal may be executed: now AIS is scanned and the image which is converted into a programme of motor instructions by the recognition buffer now is an auditory image. An important assumption is ... "that the sound-image of a letter can enter AIS directly from subvocal rehearsal without the necessity of actually being converted into sound" (Sperling 1967, p. 287). The visual scanning aspect of the model is related to findings on the number of items recalled from a 50 msec array in function of the interval between the array and the appearance of an erasing stimulus to stop the scanning process: a scanning rate of 10 msec per item has been proposed (Sperling 1963). The speech and auditory aspects of the model are based on the generally less formal observations that several subjects were audibly or visibly engaged in rehearsal during retention and recall and that most subjects reported subvocal rehearsal. Overt rehearsal was enhanced both by loud noise and by playing a recording of the subject's own voice reading letters. A significant - though small - decrement in recall performance was, however, only observed under the latter condition. Some subjects described the task as one of selective listening to events 'inside' while neglecting the events coming from 'outside' (Sperling 1962). Finally, the interesting but unsystematic observation was made that errors in the written report of the tachistoscopically presented arrays of letters sometimes showed acoustic similarity to the items that were in the array (Sperling 1963). To this phenomenon we shall pay ample attention in Chapter 2 (2.1.1).

1.3.4 Mathematical models of verbal short-term memory

The last section discussed STM models in a qualitative manner. It focused on mechanisms, processes, and their logical relations rather than on quantitative predictions based on mathematical equations. Our purpose at present is to indicate a number of areas where a quantitative approach has been applied succesfully Mathematical models of STM mainly date from the latter half of the past decade. The models may, of course, be categorized according to the mathematics used in their construction or in the procedure of estimating parameters. Another subdivision is according to the aspects of memory which are central in the models. Thus, there are models centred round acquisition, retention, or round retrieval. There are also models focusing on coding or on rehearsal. Moreover, some models deal with PA learning, others with serial learning or with free recall.

It is of interest to note that most mathematical models, although historically related to the information and communication approach discussed, do not explicitly distinguish between temporary and permanent storage. Parsimony is frequently served in the models by the neglect or the deliberate exclusion of specific immediate memory factors in the experimental data. The circumstance that some models do not differentiate between verbal and non-verbal information and that in many experiments recognition tests are used or interitem associations tested is probably related to this fact.

In the present section a selection among many alternatives has been made. First, one model (Wickelgren and Norman 1966) is chosen to represent studies regarding the strength of the trace as a single, unidimensional entity. Typical of this model, though not of strength models as such, is also the use of memory operating characteristics (MOCs) to determine trace strength Second, a complementary approach (Bower 1967) is discussed, which conceives of the memory trace as composed of a number of elements, each with a different decay history. Third, a memory system (Atkinson and Shiffrin 1968) is mentioned in which the STM-LTM distinction is worked out to considerable detail and which pays special attention to strategies pursued by the subject. Our selection is further justified by the problems on coding and rehearsal that we shall meet in Chapter 2, and by observations on the internal representation of verbal STM material in Chapter 3.

Strength models in general do not attempt a qualitative description of the trace that represents an item in memory, rather they describe the quantitative history of the decaying memory trace after an increase at presentation. Thus, important assumptions in the strength models proposed by Wickelgren and Norman (e.g. 1966) are the following: Presentation of an item results in a certain increase α of the initial strength α_0 of the internal representation of that item. If rehearsal between presentation and test of an item is prevented, there is exponential decay to a proportion ϕ of this strength as a function of the number of interpolated items or of the duration of the interval. Furthermore, it is assumed that the criterion rules and normal distribution assumptions of signal-detection theory apply to the recognition processes involved when 'old' and 'new' items are presented as test items. Memory operating characteristics (MOCs) are thus obtained by plotting for a number of confidence levels the probability of false recognition of new items. Trace strength values $\underline{d}$ are estimated from the MOCs and the parameters are then estimated from $\underline{d}$ values by testing goodness of fit.(The application of decision models to memory tasks is discussed in Norman and Wickelgren 1965 and in Green and Swets 1966). Strength models have been applied to short-term recognition memory for three-digit numbers (Wickelgren and Norman 1966) and for the pitch of pure tones (Wickelgren 1966a; 1969a). In the former study various alternative strength models were developed and tested. A simple decay model, as described, and an additional assumption concerning an extra high value for the first item in the sequence (STM Acquisition-Primacy) were supported. As indicated above, the recognition situation allows determination of the trace strength of an item itself. It is also possible, however, to test the association strength between any two successively presented (or paired) items $\underline{i}$ and $\underline{j}$. In this case, the subject may be required to recognize, choose, or recall the correct item $\underline{j}$ after being probed by item $\underline{i}$. Satisfactory results have been obtained with these various methods for short-term memory of lists of four digit pairs (Norman and Wickelgren 1969).

A model in which the nature of the memory trace and the cause of forgetting are specified in greater detail has been presented by Bower. In his multicomponent model Bower (1967) attempts a description of the encoded stimulus as stored at presentation in a number of elements which are in part lost during the retention interval. The conceptual analysis, thus, is at a more elementary level than that of the unidimensional strength model by Wickelgren and Norman discussed above. The muliticomponent model assumes that first a stimulus analyzing mechanism, through a kind of pattern recognition process, extracts a number of attributes, on which primary encoding of the features of the corresponding meaningful stimulus as recognized is performed.

Thus, an item is represented in memory by an ordered list of attributes, which in the model is represented mathematically as a multicomponent vector. It is assumed that during forgetting each of the components may be lost in all-or-none fashion, either independently or hierarchically with the least important components being most liable to forgetting. Errors in recall or recognition are then due to wrong reconstructions made by a decision mechanism on the basis of insufficient information left from the remaining components. Similarly, many errors are non-random and portray the presented (correct) item with respect to the components retained. The model, though in an incomplete stage of development, has been applied in the same paper by Bower (1967) to a great variety of memory experiments including recall, recognition, and multiple-choice conditions. In all these cases the theoretical predictions, obtained in rather complicated mathematics are fitted with great accuracy by the experimental data.

A special trait of the memory system proposed by Atkinson and Shiffrin (1968) is the important function ascribed to control processes strategically adopted by the subject. Structural features such as sensory register (SR), short-term store (STS), and long-term store (LTS) form one aspect of memory; flexible processes selected by the subject (coding, rehearsal, and search strategies) constitute an equally important second aspect. For a number of situations, especially continuous PA learning tasks, the following assumptions about the memory system are made. Information from SR is selected and with a certain probability α it is entered into a rehearsal buffer (RB) within the STS mechanism. RB has a fixed number of positions $\underline{r}$. Items tested while in RB are retrieved with perfect accuracy. Items in STS but outside RB can only be recalled at the next trial: retrieval is then perfect. While an item resides in RB it builds up a long-term trace. This transfer to LTS occurs at a constant rate of θ per trial. When more items are adopted for rehearsal than there are positons in RB, an older item is knocked out of RB. The long-term trace built up in function of its stay in RB is then subject to exponential decay at a rate of τ per trial as new items are presented. In the continuous PA and other experiments reported in the same paper (Atkinson and Shiffrin, 1968) a number of variables such as stimulus and response ensemble, number of reinforcements, overt vs. covert practice are all manipulated with obvious success.

1.4. A framework for research in verbal short-term memory

The preceding section discussed a variety of approaches in the study of verbal short-term memory. The approach to be followed in the present section will be of the type of the communication models. The points referred to in the context of the span studies, information theoretical studies, and the mathematical models will, however, be similarly related to our approach. The succeeding paragraphs will discuss a definition and a model of verbal short-term memory. The model constitutes a qualitative description of mechanisms and processes assumed to be operative in verbal short-term memory. At present it does not seem feasible to present a formal model, or to make quantitative predictions at the level of analysis pursued in this study. The existing quantitative models (some of which were briefly discussed above) that have been applied with relative success to the area of short-term memory do not touch specifically upon the mechanisms and processes involved in the internal representation of verbal items. Their parameters include trace strength, decay, probability of rehearsal and in one case there is reference to different and relatively independent components making up the trace, but the models do not describe what components constitute the trace in short-term memory, what are the rehearsal processes and how a trace is strengthened or weakened. The difficulties of quantification will become apparent as we discuss the details of the model in a further paragraph (1.4.3).

1.4.1 General characteristics of verbal short-term memory

In the different contexts of the preceding section a number of characteristics of verbal short-term memory have been mentioned. The most important are the distinction between preperceptual and postperceptual storage, coding and rehearsal as based on speech responses, and forgetting due to decay and overwriting by new information. The other aspects of immediate memory performance may all be related to these characteristics The requirement of attention, mentioned already in several of the articles reviewed by Blankenship (1938) may be translated in terms of filtering and scanning as described by Broadbent (1957) and Sperling (1963). Attention, resulting from context or instruction, is involved in the selection of information available in the form of sensory registration in an input channel and in the selection of the input channel itself, discarding the information available in other channels. At recognition an identification response is made which, as a verbal response to a verbal item, must be seen as the primary form of coding in immediate memory. The role of speech was stressed already by the early authors (see Blankenship 1938) and by many later investigators of whom e.g. Sperling (1963) and Atkinson and Shiffrin (1968) were mentioned above. Speech units do not only provide a ready opportunity for coding individual items, but speech responses may also readily chained together so that the assimilation even of meaningless verbal sequences is greatly facilitated by their transformation into a sequence of speech units. This applies a fortiori to rehearsal of coded items. In the process of rehearsal the factor associability (see Blankenship 1938) and grouping (Martin and Fernberger 1929) play a role of importance: an originally homogeneous sequence of items may acquire a structure which facilitates recall. Especially as a result of practice, rehearsal strategies may become more effective in this respect (Martin and Fernberger 1929). Silent rehearsal may be faster than overt rehearsal or recall (Sperling 1963) but its speed remains limited, certainly if concurrently new items must be coded by the same (speech) mechanism. The latter restriction does not apply to the very first item in a sequence, for which indeed an extra strong trace was observed by Wickelgren and Norman (1966). The sensitivity of immediate memory performance to the rate of presentation (e.g. Terman 1916) and the limited capacity of the short-term processing mechanism (Broadbent 1957) are most probably related to the maximum rate of coding and rehearsal. In any case, IMS is determined (though not defined) by the number of items which may be rehearsed in a single repetition (Atkinson and Shiffrin 1968). If the primary coding response is in terms of speech units, the cues on which recall depends are speech cues, also in visual tasks. The 'imagery' required (Richet 1868) may therefore primarily be the capacity to make use of these speech cues, and errors may primarily reflect their incomplete availability or usage when recall is required. This is in agreement with the acoustic confusions reported by Sperling (1963) and with the more systematic observations by Conrad (1962, 1964) to be discussed in the next chapter (2.1.2). If the speech responses provide their cues not in an all-or-none manner, but according to different, more or less independent attributes, the attributes being formed by the distinctive features of speech, then this would be in agreement with on the one hand the multicomponent approach by Bower (1967) - although he proposes components of a different type - and on the other hand with the above finding by Sperling (1963) that certain errors in recall from a visual multiletter array typically are non-random.

Two further points should be made. One is that the fact that immediate memory constitutes a relatively independent system (e.g. Katzenberger 1965) may be related to the early stage of processing involved. This point has been made earlier in the context of the relation between neural and psychological storage mechanisms (1.2.5). It was suggested that the integration into semantic and associative networks is not a factor of importance in adequate immediate memory performance. The relevance of such integration for long-term performance is, however, obvious. The other observation is concerned with the information theoretical approach of verbal short-term memory. If the interpretation provided by information theory proved not quite satisfactory in the case of immediate memory, it is because a limited number of items can be retained regardless of their information content, and because with longer sequences of items only very rough predictions based on higher order approximations can be made. One may ask now whether more memory capacity is occupied by less probable items or whether certain coding and rehearsal responses occur with greater facility if they follow certain other responses. These interpretations are not essentially different, but the latter places more emphasis on the response sequence while the former merely looks at the statistical properties of the stimulus items. Because of the large role attributed in the present approach to coding and rehearsal, which both require the ready availability of responses, it seems plausible to situate the locus of the effect of sequential dependency, and of the other factors affecting immediate recall performance as described above, at the response side rather than at the stimulus side. Since active verbal responses follow the path of articulation it may be primarily articulatory skills which are brought to bear in the experimental situation requiring immediate recall of meaningless verbal material. A related suggestion has been made by Baddeley (1964b) concerning the results of a PA learning experiment. He found that S-R pairs are learned better if the last letter of the S item and the first letter of the R item form a digram which is frequent in the language. Baddeley's suggestion is that this kind of verbal learning can be regarded as the "acquisition of a motor skill in which existing language habits are transferred to a new sequence of verbal responses."

1.4.2 A definition of verbal short-term memory (STM)

The commonest type of definition of STM[1] is an operational description of the experimental manipulations which intend to measure it. Several such definitions have been given, e.g. by Wickelgren (1965a) " short-term memory is being studied when the list is presented once at a rate of less than 2 sec per item and retention is assessed after less than 30 sec" (p.53). This certainly is a definition on which there is common agreement in the literature, but the exact reasons for the single presentation, the specified rate, and the 30 sec interval have not been stated. Single presentation as opposed to repeated presentation (which is a common procedure in the learning of longer lists) is the most consistently obeyed practice by all experimenters in the field of STM. Most of them will agree, however, that rehearsal constitutes an extra presentation but rehearsal is not excluded by the definition. A fast presentation rate is efficient as a means of preventing mnemonic strategies to be used and long-term associations to be formed, but the optimal rate to prevent this is of course dependent on the type of material presented. Two sec per item may thus be much too slow for single-word sequences but hardly sufficient for an adequate perception of multi-letter items. The 30 sec interval is probably based on an estimate of the maximum duration of a memory trace for a certain amount of meaningless verbal material under controlled rehearsal conditions. Apart from the fact that also other estimates have been made (from several sec to several min), the definition does not specify the fact that material should be meaningless. An unrehearsed seven word sentence will normally be associated to a meaningful context even if it is recalled after only 15 sec.

Our comments on Wickelgren's definition have made clear that verbal STM is not dependent primarily on a certain period or a certain presentation mode but rather a specific set of mechanisms and processes that have certain temporal characteristics. Verbal STM might thus also be described as a primary processing stage following presentation and preceding permanent and associative storage of verbal items. Studies of STM are, correspondingly, concerned with an early stage in the processing of verbal information, immediately upon its presentation and preceding its integration into already existing networks. This processing stage is characterized by preperceptual storage aiding perception, by coding at perception, and by subsequent rehearsal. These characteristics may also serve to define verbal

---

[1] From now onwards the notion 'short-term memory' will be used in a more specific sense, denoting certain mechanisms and experimental manipulations, as discussed in the present section. In general, the abbreviation STM will be employed. Correspondingly, the trigram LTM (long-term memory) by exclusion of STM roughly denotes more or less permanent modes of storage and experiments with repeated presentations.

STM in a minimal manner: a processing mechanism in which verbal items are registered in a preliminary store from where they may be read into a categorization mechanism applying a verbal code which is stored for some longer time and which may be rehearsed until a response is required.

1.4.3 A preliminary model of verbal short-term memory

We shall now discuss in some detail the characteristics of verbal STM, its components and their functional relationships. A flow diagram is presented in Figure 1.1. The properties of this preliminary model are not controversial. Most of them were mentioned in the previous sections of the present chapter. In Chapter 3 a more detailed version of the present model will be proposed in a wider framework and aspects of it will be tested in a series of experiments. Of the present model the aspects preperceptual storage, recognition and coding, and rehearsal are the main characteristics.

Preperceptual storage. Both the visual and the auditory system are assumed to possess a sensory register (SR) into which raw stimulus material is copied and retained over a very short interval for further processing. These registers are called VSR and ASR, referring to visual and auditory sensory registration respectively. Both VSR and ASR are subject to rapid decay and to masking or replacement by succeeding stimulation. Information not read out before a critical amount of decay has occurred cannot be processed any further. One essential distinction between VSR and ASR is that the former has spatial characteristics which may, in principle, be scanned in any order, while ASR represents the input in a fixed temporal sequence, so that read-out can only occur in the order of arrival of the stimuli. Another more important difference is that the rate of decay in ASR is considerably slower (extending over period of seconds) than that in VSR (a fraction of a second). Under the influence of the experimental instruction or of context variables any of the SRs receives attention and within the SR selected there is a search for specific aspects of the SR content. These aspects are also determined by instruction or context. In VSR the items to be selected may be specified e.g. by their location in the array, or by their shape or colour; in ASR there may be a specific search e.g. for words, rather than digits, or for the items spoken by a male rather than by a female voice. Selection of an item in ASR does not wipe out its sensory registration. Thus, an item may be both in ASR and in a further stage of processing. Similarly, VSR information of a visual item is not wiped out by its selection, but VSR decay is too fast for its sensory registration to be available at later stages.

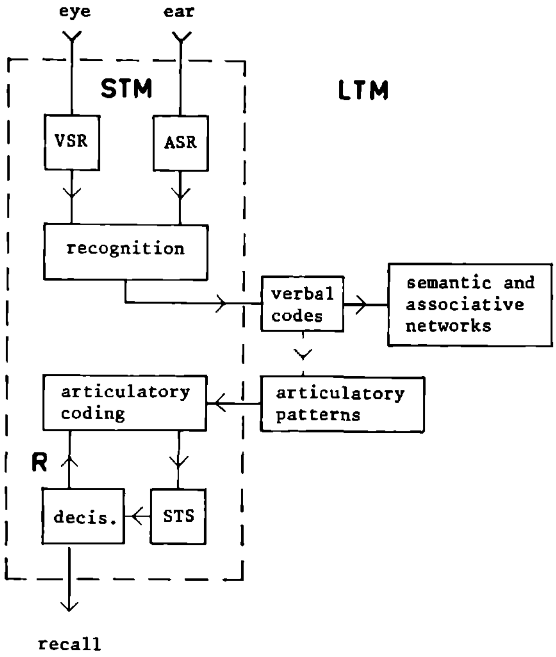


Figure 1.1. Preliminary model of verbal short-term memory. VSR = visual sensory registration; ASR = auditory sensory registration; STS = short-term store; R = rehearsal cycle.

Recognition and coding. At the next stage recognition is assumed with respect to the items read out of VSR and ASR and the active application of verbal codes to the recognized verbal items. It is assumed that this involves the articulatory mechanism, though not necessarily its overt activity. The articulatory patterns corresponding to the verbal codes are available in a separate store for ready application. Because of the close relationship between audition and articulation in speech it is assumed that the articulation patterns are more closely related to auditory stimulation than to visual forms of stimulation. The application of verbal codes does not necessarily entail at this stage the application of semantic and other associations attached to these verbal codes. This requires two different paths leading from the verbal code store: one to its semantic components stored permanently in an 'abstract' manner, the other to articulatory patterns which are also permanently stored as programmes. The actual application, either vocally or subvocally, of articulatory patterns is achieved by an articulatory coding mechanism. A code applied according to these specifications is assumed to be stored for only a short time, although longer than the SR storage time. The store for coded verbal items will be called short-term store (STS). It is assumed that coded items in STS may be retrieved and that, since the code is articulatory, and items are recognized in STS by reference to the articulatory patterns, there is no differentiation in STS between modalities of presentation.

Rehearsal. If the articulation mechanism is not engaged in the application of codes to new items entering from the verbal code store, a decision mechanism may enter (groups of) such items for repeated coding and STS storage. This process of rehearsal (R) may occur several times. Correct rehearsal is only possible after successful STS retrieval. This is less probable if there are many items, if items are articulatorily similar, or if there is a prolonged interval between coding or last rehearsal of an item and its attempted STS retrieval. Rehearsal of certain items may delay or obliterate the coding and STS storage of later items. Rehearsal will normally be in the order of coding. To the extent that this is the 'correct' order, rehearsal increases the probability of recall of the items in their correct order.

Recall of coded items. Upon a decision to that end, an item can be recalled if its original or rehearsed code can still be recognized successfully in STS. At fast recall rates it is possible that STS is cleared before the information in ASR has decayed beyond recognition. The ASR information may then be coded and recalled without rehearsal or it may support the recall of items that are both in ASR and in STS. The decay rate for VSR information is assumed to be too fast for such delayed recognition, coding, and recall.

CHAPTER TWO

Chapter 2

Acoustic factors in short-term memory

The present chapter is primarily concerned with the coding response at presentation and during rehearsal in STM tasks. This response, performed either as an overt or, more frequently, as a covert vocalization by the articulatory coding mechanism from the model of the previous chapter determines the amount retained and the qualitative type of errors that will occur when the information is retained incompletely. Articulatory coding transforms the nominal verbal stimulus, e.g. a visually presented trigram, into the functional stimulus which may be described in terms of speech parameters, i.e. in terms of audition and articulation. The generality of these factors and their effect on qualitative and quantitative aspects of recall will be discussed in the first section which includes four experiments. The next section is concerned with interactions between verbal and non-verbal material on the one hand and presentation mode on the other. The superior qualities of the ear for successively presented information and the opportunity for recoding also of non-verbal items are discussed in connection with experimental data. Lack of adequate articulatory coding, of speech experience, and of audition as a primarily sequential form of information processing in general are studied in the third section dealing with STM and related performance - also in an experiment - by deaf subjects. Finally, the last section presents an analysis of perceptual and recall errors which attempts to differentiate between the effective cues of audition and articulation in these tasks.

2.1 The acoustic similarity effect in short-term memory

The coding response performed by S as a perceptual response to verbal material has, as a result of its articulatory-acoustic properties, a greater stability than the decaying sensory registration. But the articulatory-acoustic cues may also lead to an increase in recall errors, namely if the material presented contains items for which the verbal codes are 'acoustically' similar. Items with similar codes tend to be confused with each other and sequences containing many similar items are generally hard to recall.

2.1.1 Acoustic confusions in STM experiments

By means of a qualitative analysis of errors in STM recall data Conrad (1962) demonstrated a systematic type of change occurring in verbal STM, which he assumed to be subject to spontaneous decay with time. He presented a total of 300 Ss with 40 different visual sequences of six letters made up from the following set of ten letters: B,C,F,M,N,P,S,T,V,X. The letters of each sequence appeared one at a time at a rate of .75 per letter. The Ss' task was to write down, immediately after the sixth letter had been presented, the whole sequence in the order of presentation. For the error analysis only those cases were considered where a single letter in an otherwise correct sequence had been substituted by an incorrect letter chosen from the set of ten letters used in the experiment. Most of the errors, therefore, had to be ignored but 1583 remained for examination.

Letters recalled	Letters presented	
	BCPTV	FMNSX
BCPTV	622	202
FMNSX	171	588

Table 2.1 Conrad's Matrix of confusions in recall (Conrad 1962, Table 1).

Although the presentation of the sequences was visual, the results displayed a largely acoustic effect. The analysis showed a very marked tendency for confusions to occur within acoustically defined subsets, as is obvious from Conrad's (1962) Table 1, which gives a condensed view[1] of his results (Table 2.1).

---

[1] In the present chapter several of these 'compressed matrices' derived from complete confusion matrices will be presented. For quick reference they are presented without headings. The columns represent the stimulus letters separated into two distinct sets of consonants with high 'acoustic' intraset similarity. The rows represent the corresponding two sets of incorrect responses. Systematic confusions are reflected by higher frequencies on the main diagonal. Since these matrices contain only errors, the expected frequencies for intraset confusions are lower than for the interset confusions. Multiplication of the cells on the main diagonal by n/n-1 (where n is the size of appropriate set) is the simplest correction for a more accurate comparison.

Moreover, a speech intelligibility study was performed of the ten test letters masked by varying amounts of white noise. There were two speakers and 50 listeners. The resulting auditory confusion matrix, containing 1900 perceptual errors, was closely related to the confusion matrix of the recall data. The rank correlation between the 90 cells of the two 10x10 matrices was highly significant ($\underline{r}_s = .56$; $t = 6.33$; $p<.001$).

A more detailed account of similar findings, obtained from an even larger group of subjects and supplemented with several refinements, is given in a later article by Conrad (1964) which has for some time been the most frequently quoted reference in the field. The 387 Ss of this experiment each recalled 40 visual sequences, made up from the same ten-letter vocabulary on a random basis with some restrictions to control for letter frequency, practice and serial position effect. Presentation and recall conditions and also the rules regarding selection of errors for analysis were the same as in the previous experiment. In order to obtain the maximum number of confusions, subjects were encouraged to guess rather than to leave blank spaces; during the test they had the vocabulary before them. The results, reduced to the same format as those of the earlier experiment, show exactly the same relationship (Table 2.2).

Letters recalled	Letters presented	
	BCPTV	FMNSX
BCPTV	790	270
FMNSX	233	757

Table 2.2 Synopsis of Conrad's 10x10 Recall confusion matrix (After Conrad 1964, Table 2).

The auditory confusability test was this time performed with 10 speakers and 300 listeners, while all the 26 letters of the alphabet were used. The relevant listening confusions were then placed in a 10x10 auditory confusion matrix (see Table 2.3) and a rank correlation coefficient of $\underline{r}_s = .64$ ($t = 7.68$; $p < .0001$) between the 90 cells of the two matrices was obtained.

It is stressed in the discussion of these findings that errors in verbal STM are not perceptual errors but errors resulting from partial decay of the memory trace. With respect to the close similarity between the two types of confusions, Conrad notes that "what is interesting is not that the memory distortions correspond to the sound of the original stimuli verbalized, but that they correspond so closely to the distortions occurring when speech sounds are partially masked by white noise" (Conrad 1964, p. 79). Finally, the important suggestion is made by Conrad that memory span is a function not of the size of the vocabulary (i.e. information per item), as has been predicted by information theory (e.g. Crossman 1960) but of the number of acoustic confusions to be expected from the vocabulary used.

Letters reported	Letters presented	
	BCPTV	FMNSX
BCPTV	2680	193
FMNSX	165	2922

Table 2.3 Synopsis of Conrad's 10x10 Listening confusion matrix derived from his 26x26 matrix of Listening Errors (After Conrad 1964, Table 1.).

These results, substantiated mainly by Wickelgren (1965a, 1965b, 1965c, 1965d, 1966b) have, in fact, led to the postulation of the 'speech' code in the model of Chapter 1. Conrad's suggestion, mentioned above, that there is equivalence between listening under noisy conditons and scanning of a memory trace in neural noise, as well as the term 'acoustic' confusions in STM, place the emphasis on a retrieval or sensory aspect of a part of STM. Of course this must correspond with a storage or input aspect in that part of STM. However, if this is not 'really' auditory, but at most the result of auditory imagination evoked by the identification response, then this response must itself provide a better explanation of the observed effect than the results of the scanning process. The response, as far as verbal items are concerned, is a speech response and as such it involves an articulatory-motor process. In the course of this chapter the implications of this will be repeatedly encountered. Although up to now there has thus been but half a reason (only retrieval, not storage) to speak of 'acoustic coding', we shall nevertheless maintain the term for the time being.

2.1.2 Acoustic similarity and interference in STM

There has been a serious dispute concerning the distinction between STM and LTM as two different systems in memory. The arguments were centred around autonomous decay as a characteristic of STM, in contrast to interference principles as the explanation of forgetting in LTM. The distinction between STM and LTM has been put forward mainly by investigators of information and communication theory background. Representatives of this group are Brown (1954, 1958, 1959), Broadbent (1958, 1963) and Conrad (1958, 1960, 1967; Conrad and Hille, 1958; Conrad and Hull, 1966). They all argued that recall decrement after a short delay should be attributed to decay with time rather than to interference from the activity interpolated between presentation and recall in order to prevent rehearsal. A continuous interpretation has been supported by theorists from the field of verbal learning such as Murdock (1961, 1967), Keppel and Underwood (1962), Melton (1963), Wickens, Born and Allen (1963), Postman (1964), Waugh and Norman (1965), and Wickelgren (1965a, 1965b, 1965c, 1966b, 1966c). Their argument is that STM, just as LTM is associative, and interference is the cause of forgetting in STM as well as in LTM.

One way of studying interference as a cause of forgetting is to present S with an original list (OL) to be learned, followed by an interpolated list (IL) also to be learned and, finally, to test S on OL. This is called a retroactive interference (RI) design. The general finding is that the negative effect of IL on the recall of OL is a function of the similarity between OL an IL. In LTM experiments with repeated trials on both lists an RI effect in function of semantic similarity has often been reported (e.g. McGeoch and McDonald 1931; Osgood 1949, 1953; Baddeley and Dale 1966). In a number of STM studies an RI design has also been followed, but no substantial effect of semantic similarity has been reported (e.g. Baddeley and Dale 1966; Baddeley 1966a; Dale and Gregory 1966; Dale 1967). On the other hand, if similarity is varied along the acoustic dimension, whose significance was demonstrated by Conrad (1962, 1964), RI does show up in STM. Dale (1964) obtained more errors in the recall of three-letter lists presented auditorily if the interpolated task was copying six acoustically similar letters than if the IL letters were dissimilar. More detailed results were reported by Wickelgren (1965a). He presented his Ss with auditory four-letter OL sequences for recall and eight IL letters to be copied at presentation. Increased acoustic similarity between OL and IL led to decreased recall performance, mainly due to the appearance of acoustically similar extralist intrusions in the OL recall. In two further experiments dealing with the recall and with the recognition of single letters in an RI design (Wickelgren 1966b; 1966d) the same adverse effect of acoustic similarity between the single original item and the interpolated items was observed. In the recognition experiment, moreover, there was a sharp increase of false recognitions of the single test letter if it was acoustically similar to the original letter. It should be noted that in all the RI experiments presentation of both OL and IL material was auditory.

Interference in STM can also be studied by varying the similarity between the items within the same sequence. The suggestion made by Conrad (1964) that acoustically homogeneous sequences will be recalled less well than acoustically heterogeneous sequences has already been mentioned. The results of experiments using visually presented lists of words and lists of consonants (Conrad 1963; Conrad and Hull 1964; Conrad, Freeman and Hull 1965) confirmed this suggestion. The effect of within-list acoustic similarity on the recall of mixed sequences of consonants and digits and on letter and digram sequences has also been established by Wickelgren (1965d; 1965c). Heterogeneous sequences result in better ordered recall and the errors tend to be acoustically similar intrusions. Wickelgren used auditory presentation, but he also introduced a control for perceptual errors. Moreover, in an elegant set of experiments by Baddeley (1964c 1966a) a large acoustic effect was demonstrated, also with visual presentation, on the recall of word sequences. Comparatively small, if any, effects of formal or semantic similarity were found.

The dispute has now been settled in the sense that interference is accepted for all traces in memory once the stimulus material has been adequately perceived and coded. The coding dimension - and thus the dimension along which similarity determines the amount of interference - may be different , but similar general principles seem to apply in the two mechanisms which may thus be regarded as two subsystems within the same single memory system. The acoustic aspects of STM material whose role was noted in the previous section appear to be not merely one quality among many of the STM code; indeed, the acoustic code occupies a most important place. Within the scope of the experiments described in the literature it is not as essential for correct recall what the items mean as how they sound or would sound if spoken aloud. With reference to the model of Chapter 1, this means that the verbal code can be given without immediately involving the semantic and associative networks

2.1.3 The coding response in memory tasks

There is common agreement in the experimental literature on the favourable effect of implicit or overt verbalization on human learning and retention. Subjects use verbalization to such an extent that one may wonder whether the distinction between 'verbal' and 'non-verbal' material in memory experiments can be justified. Nearly all material that has been presented in these experiments allows some kind of verbalization. As Riley has pointed out (Postman 1962) the systematic changes observed in the recall and recognition of figures - which were reported by Gestalt psychologists, and interpreted as evidence for a tendency of the memory trace toward 'figure goodness' - are due to the verbal labels used rather than to any autonomous internal changes. Indeed, coloured patches, dot patterns, drawings, etc., especially when presented in sequences, are readily coded by their appropriate verbalizations and what is retained is often no more than this verbal code, no other than if the material had been presented in a 'verbal' manner, i.e. in words or numbers.

This fact, and also its validity outside the sphere of laboratory tasks, has been emphasized by many authors, e.g. by Miller (1956) from whom the following lines are quoted. "When there is a story or an idea that we want to remember, we make a verbal description of the event and then we remember our verbalization. Upon recall we create by secondary elaboration the details that seem consistent with the particular verbal recoding we happen to have made" (p. 95). This view on the process of recoding is, incidentally, closely related to the concept of 'schemata' originally used by Bartlett (1932) and later, in a more specific sense by Oldfield (1954) and Attneave (1954, 1957, 1959) as a means of efficient recoding in perception and memory. The role of verbalization is perhaps not very much stressed by these latter authors; on the other hand the assumed relation between language and 'schematic' perception has led to a strong hypothesis stating the dependence of perceived reality on available concepts in the language (Whorf 1956).

Even within the range of simple memory tasks, however, some differences exist with respect to the type - and perhaps level - of verbalization that is assumed to affect the form of behaviour studied. In the experiments on shifts in discrimination learning by Kendler and Kendler (1961, 1962, 1968) for example, learning in adult humans is assumed to be mediated by conceptual, symbolic responses, available to subjects with language. The verbal loop hypothesis by Glanzer and Clark (1962, 1963) predicts recall of a sequence of stimuli from the length of the verbal chain in which the sequence can be coded. Verbal labelling, furthermore, in the classic study by Carmichael, Hogan, and Walter (1932) is claimed to explain errors in the reproduction of drawings; deaf subjects were reported in another study (Clarke 1951) not to make such systematic errors. Pronounceability of nonsense trigrams is said to be a better predictor of learning and retention than trigram frequency (Underwood and Schulz 1960) because pronounceability is closer related to the units of overt verbalization in speech than is the frequency of all possible trigrams. Lastly, also incidental learning has been observed to increase sharply if implicit verbalization is introduced by requiring the subjects to rate the incidental items for pronounceability (Mechanic 1962).

The assumptions on the role of verbalization in the above publications range from the use of abstract verbal concepts to the mere virtual articulation of nonsense material; or in other words, from the use of speech as a system which conveys meaning to the use of speech as a system made up of articulation and sound. The studies all agree, however, in establishing the universal fact that verbalization of material which is presented for memorization is a natural process and that the verbal code, resulting from the active use of language, supports learning and retention. Some kinds of material will be more difficult to code verbally than others; the most natural code is, of course, evoked by printed words, trigrams, and numbers. But the only real distinction between 'verbal' and 'non-verbal' material in this respect is probably situated in the certainty with which the experimenter can know in advance the exact verbalizations that will be used by his subjects when he confronts them with words rather than with drawings.

The generality of verbalization in memory experiments using visual presentation is beyond any doubt. It must, moreover, be assumed that as a result of learning involved in the enormous amount of reading experience verbal labels are readily available if S is confronted with verbal items such as words, trigrams, letters, or digits. Thus, verbal items are not only coded with greater consistency but also with greater speed than non-verbal items. Fraisse (1969) reported that a circular pattern in a context of letters is correctly given its letter name 'o' much faster than it is named 'circle' in a context of simple geometrical figures. In terms of our model this might mean that the speed at which a match can be brought about by the scan of VSR is faster if the 'templates' with which matching is attempted are verbal rather than pictorial. Not all reactions to verbal items are equally fast, however. Letters take longer, for example, than digits, if the task is simply naming. But it is of great interest in this context that Mackworth (1963) found a clearcut relation between the time taken per item for its identification (naming) and the number of items that can be recalled from a single short exposure of an array. Apparently naming and coding are correlated to the extent that the performance in the latter can be predicted from the amount of time involved in the former.

The use of verbal codes is universal; it is not restricted to meaningful material, as we saw in 2.1.1 and 2.1.2. We also saw, however, that a verbal code can be given more quickly for verbal than for non-verbal items. Although the product is the same, the processes, as described in terms of the model, naturally depend on the type of material presented. For verbal items we may assume learning processes have smoothed the way for direct connections via recognition and verbal code; this probably also applies to some extent for very familiar graphical representations (e.g. triangle, square, circle); but for completely 'meaningless' figures repeated comparisons must be made with permanently stored visual representations in order to achieve recognition. The reason that these figures are called meaningless is of course that recognition categories with the associated verbal labels are not directly available.

2.1.4 Rehearsal in STM experiments

A number of references to rehearsal processes have been made in the first chapter without much further comment. The assumed processes have been ascribed a central position in a number of models including the one we presented above. In the present section the concept of rehearsal will be discussed in some detail. In principle rehearsal may be regarded to denote all processes which operate on the stimulus item after termination of its presentation. Some authors (e.g. Waugh and Norman 1965; Sperling 1967) thus use the term for the coding response as well as for its cyclic repetition. Generally, however, the term rehearsal is used to refer to the circulation of perceived and coded stimulus information. Prolonged selective attention to non-verbal material, such as the position on a line (Posner 1966) or a tone (Wickelgren 1968) is thus also called rehearsal and it is assumed that a secondary task interpolated during the retention interval, which occupies the central processing capacity, precludes rehearsal. More specifically, the term is used for the more or less deliberate repetition of verbal items once they have been coded. We shall use the term in this sense.

A distinction must furthermore be made between overt and covert rehearsal. The former is very seldom allowed in STM experiments; the latter is supposed to be under control if a secondary task is used. Either controlled or not, covert rehearsal is assumed by many to differ from overt vocalization only in amplitude (Landauer 1962; Blackwood and Link 1968). Some authors suggest that overt rehearsal is more beneficial to short-term retention (Lewis and Teichner 1964) while others contend that subvocal rehearsal is superior because of its greater speed and flexibility (Corballis and Loveless 1967). In any case, rehearsal may be regarded as a form of implicit verbal activity in which articulatory processes are involved. The effectiveness of rehearsal preventing tasks is also related to the extent to which such tasks keep the articulatory mechanism engaged. In the next section we shall return to this issue.

It should be noted that rehearsal as a simple re-cycling of verbal items is largely determined by the experimental procedure and materials commonly employed in STM experiments. Different operations resulting in different codes are likely to be adopted by the subject with less stringent requirements on the one hand and more meaningful material on the other. These subject controlled curves, is somewhat less favourable than written recall to the last items the STM conditions, relatively straightforward rehearsal may be assumed. Its positive effect on retention has been established in a number of ways. If more time is available for rehearsal, the material is more resistent to forgetting (Sanders 1961; Hellyer 1962; Pollack 1963), especially if the opportunity for rehearsal is provided immediately after presentation (Pylyshyn 1965). If the presented stimuli allow short verbal codes and if over a certain interval more rehearsals can therefore be completed, there is better recall than if long, multisyllabic codes apply to the stimuli (Laugherry, Lachman, and Danserean 1965).

The manner in which spontaneous rehearsal is performed during the presentation in STM experiments may vary between Ss as has been shown in the recordings made by Corballis (1969) but obviously some kind of grouping occurs in all Ss. Grouping increases the span and decreases the difference between good and poor performers. In experiments that study the optimal size of the rehearsal group by means of varying instructions or presentation modes, a size of three items has consistently been found to be superior, virtually irrespective of the sequence length and over a wide range of presentation and recall conditions (Severin and Rigby 1963; Wickelgren 1964; Thorpe and Rowland 1965; Conrad and Hull 1967). Most likely, this rehearsal group size reflects on the one hand the number of items that can be rehearsed with ease in one single chunk and on the other hand a subdivision of the sequence into a manageable number of chunks for assembly at recall.

Apart from grouping, several different ways in which rehearsal brings about its positive effect on retention have been proposed. One notion (Brown 1958) is that rehearsal causes a delay of the onset of the decay process, whose rate is itself, however, unaffected by rehearsal. Other interpretations which are more generally supported by the data are that the trace itself is strengthened by the establishment of associations with increasing strength. It is common practice to consider each rehearsal cycle as equivalent to a repeated presentation and thus as constituting an 'internal reinforcement' of the inter-item associations (Adams 1967). The fact that rehearsal increases the probability of correct recall of the order of the items (Wickelgren 1964, 1967) supports this consideration. Moreover, there is also a greater chance of long-term semantic codes occurring with prolonged rehearsal. Such codes include natural language mediators (NLMs) which have been studied in this context by Groninger (1966).

The effect of rehearsal thus bears on the retention of the items, on the interitem relationships (i.e. on their order), and on their integration into existing semantic and associative networks in LTM. The relative weight of these factors will of course depend, as we saw above, on the material and on the amount of time available between presentation and recall and between the indivual items at presentation. With meaningless verbal items such as consonants or digits, presented sequentially at rates of approximately 1 per sec, rehearsal will probably not have the latter effect. If, moreover, a limited amount of items are used repeatedly in different sequences, so that the items themselves are easy to 'learn', the strongest beneficial effect of rehearsal will clearly be on the order of the items. Order errors will, correspondingly, receive special attention in several of the succeeding sections of this chapter. Rehearsal as a sequential process shows here its analogy to speech, where of course order of the elements is a primary source of information.

2.1.5 Rehearsal prevention by interpolated activity

If rehearsal is permitted a limited amount of information may be retained over long intervals. There is evidence that over increasing retention intervals up to 10 sec recall may even improve under these conditions (Crawford, Hunt, and Peak 1966). Several tasks have been devised in order to prevent rehearsal and to study the 'pure' course of forgetting in STM. Since such interpolated activity is, however, also a potential source of interference in the sense discussed above (2.1.2), care is usually taken to choose material from a class that is known to cause little associative or competitive interference with the original material (e.g. digit reading to suppress rehearsal of words). The most widely used technique is paced counting backwards in threes from a three-digit number spoken after the presentation of the items to be retained (Peterson and Peterson 1959). Further tasks are the generation or reading of random numbers, saying the alphabet in either direction, naming geometrical shapes, and many others. Generally they are tasks requiring considerable attention and paced overt verbal responses.

It has never been made explicit to what extent the 'attention' mechanisms and to what extent 'verbalization' is responsible for rehearsal suppression. Indeed, it may be impossible to separate these aspects where verbal material is concerned. The reason is not so much that the complexity of the attention requiring task cannot be adquately estimated, but rather that it is difficult to determine how much verbalization goes on in a task of some complexity requiring only a limited number of overt responses. An example may clarify this. Saying the alphabet is less rehearsal prohibiting than saying it backwards. Is this difference solely due to an increased load on central processing mechanisms under the latter task or does this task actually also involve more covert articulation? The latter is very likely if it is assumed that for every overt backward response from the alphabet a short covert run in forward direction is necessary. Such an assumption, however, need not even be made if the critical locus of incompatibility between rehearsal and interpolated activity is situated at a higher level where the programmes for articulation are made up. Simply one instruction might suffice to recite the alphabet in forward direction, leaving the execution to lower articulatory mechanisms. For backward responding, however, separate articulatory instructions will have to be set up.

This line of reasoning may apply, with some modifications, to a large variety of interpolated tasks in which complexity is varied. (e.g. Crowder 1965; Loess and McBurney 1965; Posner and Rossman 1965; Bruning, Schappe, and O'Malley 1966). It is also supported by the relative efficacy of saying "the-the-the" (Murray 1967) which puts a very low demand on the central processing mechanisms, but which is a highly uncommon sequence and therefore requires repeated articulatory instructions. In our opinion it is the compatibility of instructions that matters rather than the execution of the instructions or the rate of information processing as such.

If rehearsal is prevented by interpolated verbal processing, the interpolated items will become available as responses and may remain available for some time, so that at recall they may compete with the original items to be recalled. If, on the basis of their acoustic properties or their class membership, there is not much evidence to decide against such intrusions, items from the interpolated list will appear in the recall of the original list. This is a 'competition' interpretation of interference in STM as studied in RI designs (2.1.2); it contrasts with an associative interpretation in that no specific associations between original and interpolated items are presumed to be formed. A method to differentiate between these two interpretations is presented in the following paragraph (2.1.6.3).

2.1.6 Experiments on acoustic factors in STM

In the present section four experiments will be discussed. The first two intend a replication of the acoustic coding principle, and its effect on the recall of verbal items and their order, which has thus far not been established in the Dutch language. These experiments demonstrate that for various presentation rates and for different sequence characteristics the acoustic confusability of visually presented letter sequences determines the probability of errors. The third experiment investigates the role of modality similarity of interpolated activity on interference in the recall of short letter sequences. The fourth experiment tests a hypothesis concerning the manner of coding performed on the material of an interpolated task during the retention of varying numbers of verbal items.

2.1.6.1 Experiment 1 Item recall and acoustic coding

A straightforward manner to establish the relationship between coding in verbal STM and acoustic perception is to present Ss with visual letter sequences to be recalled under silent conditions, constructed such that, acoustically, they can be called homogeneous and heterogeneous and to compare their performance in the two types of lists. The Dutch alphabet when spelled allows a distinction quite similar to that applied by Conrad and Wickelgren to the English and American alphabets respectively. In the Dutch alphabet 9 letters are consonants followed by /e/ and 6 letters are consonants preceded by /ɛ/. If the letter Z, which is pronounced /zɛt/, is added to the latter group, and if two letters from the former group are omitted, two relatively homogeneous sets of equal size are obtained, one consisting of the seven consonants B,C,D,G,P,T,V which are pronounced with /e/ and the other consisting of the seven consonants F,L,M,N,R,S,Z which are pronounced with /ɛ/. These sets will be referred to as the B-set and the F-set respectively. Letter sequences made up from any one set only are called homogeneous and letter sequences made up from both sets are called heterogeneous. The prediction that more errors will occur in the recall of homogeneous seven-letter sequences after visual presentation was tested in Experiment 1. (The purpose of this experiment was also to collect confusion data and to test some further hypotheses concerning order error, serial position effect, and predictability; these issues will be considered later in this and in the following chapter).

Method. The material consisted of 56 seven-letter sequences in which from 0 to 7 letters were chosen from the B-set, the remaining from the F-set. The resulting 8 conditons, numbered from 0 to 7 correspondingly, each occurred 7 times, once in each block of 8 sequences. Thus there were 14 homogeneous sequences (conditions 0 and 7) and 42 heterogeneous sequences (conditions 1-6). Over all the sequences the probabilities that a letter was followed by one from the same set (within-set pairs) and from the other set (between-set pairs) were equal. All 14 letters occurred with equal frequency over conditions, blocks, and serial positions. An 18 mm film was made containing these 56 sequences preceded by an extra block for practice. As warning signals, red flashes (single frames) anounced every sequence; the last sequence item was always followed by a green flash to indicate the recall interval. Presentation rate was 0.6 sec per letter. Recall intervals were 12.5 sec. Longer intervals separated blocks and a pause was given after 4 blocks. The size of the letters on the screen was 29 mm. Recall was ordered in boxes printed on response sheets which also contained the 14 letter vocabulary to aid guessing. Ss were 80 grammar school pupils (61 male, 19 female) whose age ranged from 15 to 20. The Ss were run in four groups of approximately equal size.

Results and discussion. The main results are presented in Figure 2.1 which shows that acoustic similarity results in more errors than acoustic dissimilarity over all serial positions. Error percentages are 46.9 and 39.4 respectively. Tested by means of the Wilcoxon matched-pairs signed-ranks test this difference is significant well beyond the .001 level ($z= 4.2$). The proportion of errorless sequences was 12.8 under homogeneous and 18.0 under heterogeneous conditions ($z= 3.8$; $p < .001$). The effect of acoustic similarity, and thus the occurrence of acoustic coding, is clearly established for the recall of visually presented items. In the analysis separate attention is given to order errors. The expectation is that differences are still greater here. Pure order errors being hard to isolate, the analysis is restricted to 'paired transpositions' which are easily identified. They contain two adjacent letters recalled correctly but in the reverse order. Table 2.4 presents the distribution of these errors. Frequencies are corrected for the relative opportunity for their occurrence (numbers in brackets).

Condition	within	between	total
homogeneous	246 (1)	-	246 (1)
heterogeneous	229 (1)	80.5 (2)	130 (3)

Table 2.4. Paired transpositions involving within-set and between-set pairs in the recall of homogeneous and heterogeneous lists (Exp. 1).

The 636 paired transpositions (involving 1272 letters or 9,8 per cent of the error total) are distributed very unevenly indeed. Heterogeneous sequences are much less liable to order errors than homogeneous sequences, but the difference is entirely due to the decreased probability of transposition in *between-set pairs.*

Two further findings, concerned with the relation to auditory perception, may be reported. Firstly, the 182 cells of the 14x14 matrix containing all 11,662 substitution errors within the vocabulary show a significant correlation with the cells of the acoustic matrix (to be discussed in Section 2.4): $\underline{r}_s$=.69 ($p < .01$). This in nice agreement with the correlations reported by Conrad (1962, 1964) who was, however, careful in selecting his STM confusion errors (See 2.1.1). Secondly, over the 56 sequences there is a significant correlation between acoustic confusability (as expressed by an index based on the relative confusability of the six pairs of adjacent items per sequence) and recall errors: $\underline{r}_s$=.44 ($p < .01$).

Conclusion. *Acoustic coding has been established in the immediate ordered* recall of seven-letter sequences after visual presentation. Acoustically heterogeneous lists are less likely to produce errors than homogeneous lists. Over the sequences of the experiment there is a significant correlation

between acoustic confusability and errors in recall Moreover, STM recall confusions are closely related to confusions in auditory perception.
Paired transpositions show that order errors are less frequent in heterogeneous than in homogeneous sequences, the difference being greater than between the overall error percentages. Decreased order error in heterogeneous sequences must be ascribed to a decreased probability of transpositions between members of different acoustically defined sets.

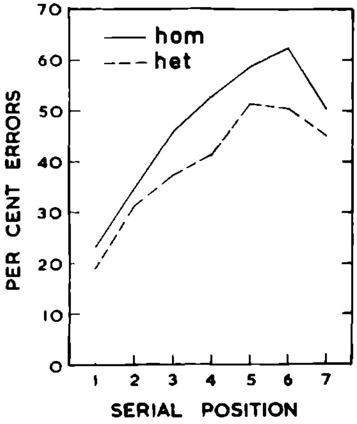


Figure 2.1

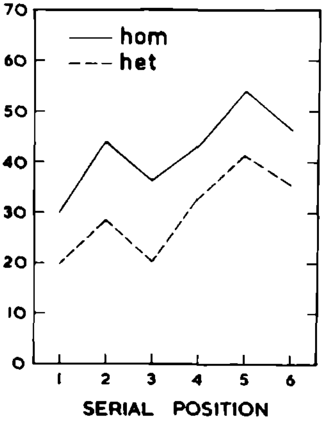


Figure 2.2

Figure 2.1. Errors in the STM recall of visually presented seven-letter sequences made up from the members of either B-set or F-set (homogeneous) or from both sets (heterogeneous) in various ratios (Exp. 1).

Figure 2.2. Errors in the STM recall of visually presented six-letter sequences yielding order errors exclusively. Error percentages are corrected for chance (Exp. 2).

2.1.6.2 Experiment 2. Order recall and acoustic coding

The present experiment is concerned with the recall of order in homogeneous and heterogeneous lists Even if S knows what letters to reproduce, he may be uncertain as to their correct order. The fact that order errors are especially likely to occur in homogeneous sequences has been reported by Wickelgren (1965c). The procedure followed to study order error is normally to select and count letters recalled correctly but in the wrong serial position. The present experiment employs a procedure in which the probability of recalling wrong letters is reduced to zero, so that all errors made may be ascribed to the imperfect retention of their order The expectation is that the effect of acoustic similarity observed in Experiment 1 is again found, but that it will be much greater and that the error ratio between homogeneous and heterogeneous sequences will tend towards that of the paired transposition errors of Experiment 1 rather than towards that of the overall error rate in Experiment 1.

Method. A ten-letter vocabulary was used to make up 360 six-letter sequences. There were two sets, viz. B,D,G,P,T, and F,L,N,R,S In each sequence only three different letters (trios) occurred in two different consecutive permutations. A trio was either wholly from one set (homogeneous) or from both sets (heterogeneous). Twelve patterns were selected to construct the 12 sequences from each trio. Twenty groups of 5 Ss were presented with 36 sequences. These were 12 heterogeneous sequences and 12 homogeneous sequences from each set. The heterogeneous trios were different for all groups, homogeneous trios were identical for every four groups. These arrangements allowed complete balancing of the 10 letters and their combinations. The 12 sequences made up of one trio were presented blockwise. They were followed by two more blocks of 12 with identical patterns but made up from different trios. Presentation order of conditions was balanced. Ss were informed of the 3+3 structure of the sequences and the trio used was printed on their response sheet for the appropriate block. Slides of single letters were presented by an automatic slide projector. The height of the letters on the screen was 7 cm. Presentation rate was 1.3 sec per letter. Recall intervals were 12 sec, terminated by a click of the projector anouncing the next sequence. The 100 Ss were mainly teacher training students (18 male, 82 female) with a modal age of 18.

Results and discussion The 5491 errors contained only 12 responses outside the appropriate trio, which does indicate that the task merely involved the retention of order rather than that of items. The principal data are shown in Figure 2.2 where again acoustic similarity appears to result in poorer recall over all serial positions. The error percentages corrected for 1/3 chance are 42.3 and 29.5 for the homogeneous and heterogenous sequences respectively. This difference is very significant as tested by the Wilcoxon matched-pairs signed-ranks test ($z= 3.61$; $p < .001$). The corresponding percentages of sequences recalled without errors are 42.8 and 57.2 ($z= 3.35$; $p< .001$). Thus, the effect on the retention of the correct order of known items is established, and as expected the effect is bigger than that in Experiment 1. Whereas the proportional increase for items plus their order from the heterogeneous to the homogeneous condition is 19.0 per cent (Exp. 1), the increase for order only is 43.4 per cent (Exp. 2). Comparison of the sets of serial position curves (Figures 2.1 and 2.2) shows a tendency for this differential effect to be greater in the early positions of the sequence.

Confusions were also analyzed. If the confusions involving within-set and between-set pairs are taken separately, it is evident that the low error rate under the heterogeneous condition must be ascribed solely to the lower probability of between-set confusions (See Table 2.5).

Condition	within	between	total
homogeneous	28.2	-	28.2
heterogeneous	28.7	15.6	19.9

Table 2.5. Order errors involving within-set and between-set confusions as percentages of the number of opportunities for such confusions (Exp.2).

This result accurately replicates that of the paired transpositions in Experiment 1 (Table 2.4) and thus provides strong evidence that the retention of order is impaired as a function of the number of opportunities for acoustically similar items to change places.

Conclusion. The effect of acoustic similarity is very great indeed if only order is to be retained. This confirms the suggestion based on the findings of Experiment 1. The other finding that order information in adjacent pairs strongly depends on the acoustic similarity of the members of the pair has further been substandtiated for groups larger than pairs. The two experiments taken together provide a firm basis for the assumption that acoustic coding is universal, and therefore for the speech-motor aspects of the model presented in Chapter 1. The question to what extent all verbal processing may be assumed to involve articulation will be dealt with in Experiment 3.

2.1.6.3 Experiment 3: Modality similarity and interference

Interference in STM has been called associative in the sense that similar items in OL and IL form similar associations and confuse with one another in recall on the basis of the high degree of similarity between these associations. The occurrence of intrusions from IL in OL recall has been taken as a measure of such interference e.g. by Wickelgren (1965a, 1965b, 1965c, 1966b, 1966c, 1966d, 1969b) who advocates an associative interpretation of forgetting in STM. A different, non-associative interpretation of intrusions from IL in OL recall is also possible, however. This interpretation, which is in agreement with the model of Chapter 1, is that IL items once they are coded in an articulatory manner remain available as responses some time. If there is no evidence to decide against such items on the basis of their class membership or of acoustic differentiation, they may appear as overt intrusions in OL recall. This possibility was mentioned above (2.1.5) in the context of discussing methods to prevent rehearsal. In the latter type of task, however, care is usually taken to choose IL items clearly distinguishable from OL items, so that overt IL responses in OL recall are unlikely, but apart from this easier rejection, the mechanism remains the same.

It is not possible to decide between the two interpretations if the relative availability of OL and IL items is kept constant within an experiment because then acoustic similarity affects the similarity of the assumed associations in a manner exactly the same as the similarity of competing responses between which S has to decide. This is the case with all experiments quoted above (2.1.2) where auditory presentation was always used for both OL and IL. The model suggests a greater availability of items spoken by S than of items written by S because the auditory consequence of their overt articulation is fed back through ASR which is essentially a repeated presentation; moreover, the greater persistence and more direct coding of ASR information than of VSR information add to the greater availability of spoken items. If OL performance consists of silently reading the items and silently writing them down at recall, a similar IL condition would require S also to write his IL recall silently, whereas a different IL condition would be to instruct S to speak his IL items. On an associative theory, therefore, a greater effect of acoustic similarity will occur with written IL, while the competition interpretation predicts a greater effect of spoken IL on intrusions in written OL recall. This prediction is tested in the following experiment.

Method. From the complete 14-letter vocabulary a series of 64 four-letter sequences was drawn up, 32 from the B-set and 32 from the F-set. Each of these quadrigrams was used as an original list (OL) in a retroactive interference (RI) design which implied that its recall would be required after the recall of another interpolated list (IL) whose presentation immediately followed that of OL. The ILs were trigrams with 3, 2, 1, or 0 letters from the same acoustically defined sets as the corresponding OLs.

Recall	modality	per cent list similarity			
of IL	similarity	100	67	33	0
written	high	89.7	88.5	87.6	81.2
spoken	low	91.8	83.8	77.3	59.0

Table 2.6. The effect of list similarity on within-set intrusions in the recall of OLs as a function of modality similarity between OL and IL (Exp. 3).

These constitute the four list similarity conditions 100, 67, 33, and 0. Their order was random within each of the 8 blocks of 8 trials. The same letters never occurred both in an OL and in the IL paired with it. Presentation was on a memory drum at a rate of 2 sec per list. First an OL appeared in one frame, then an IL in another frame, and finally there was an 8 sec interval for the recall of IL and OL. Recall of the OLs was always written. Interpolated recall of the ILs was written in Condition 1 and spoken in Condition 2. These constitute the two modality similarity conditons. The Ss were teacher training students with an age range of 17 to 22. They were assigned in approximately equal numbers ($n_1=22$, $n_2=20$) to either of the two conditions.Recall of OLs was in boxes on a response sheet. Ss were tested individually.

Results and discussion. The overall OL recall was not significantly different for the two IL modalities. Per cent errors was 48.7 after written IL and 52.8 after spoken IL. The main results, however, concern the type of the intrusion errors made under both modality similarity conditions. These are presented in Table 2.6. On an associative theory a greater effect of high modality similarity is expected, whereas the model of Chapter 1 predicts a greater effect of spoken IL, irrespective of its low modality similarity to OL. The results are clearly in agreement with the latter interpretation. Of the intrusions in OL recall after written IL the vast majority belongs to the (correct) set from which OL was composed. There is a gradual decrease over the four degrees of list similarity, but tested by means of a Kruskal-Wallis one-way analysis of variance these values (expressed as percentages in the top row of Table 2.6) do not differ significantly ($H=2.83$). After spoken IL, however, the acoustic class (set) of the intrusions in OL recall is increasingly determined by that of the IL items, i.e. there are fewer and fewer within-set intrusions in OL recall. This decline is significant ($H=14.21$; $df=3$; $p < .01$).

The observed interaction demonstrates that spoken IL items are more effective in interfering with written OL items than are written IL items. Interference may therefore have been through competition rather than through association, the largest effect being found for the low modality similarity condition. This confirms the suggestion stated above that the auditory information of IL recall, fed back through ASR, makes spoken IL items more available than written IL items and therefore more likely to intrude in OL recall.There may also be another interpretation, however. In terms of the model of Chapter 1 the result might mean that a 'lower' form of written IL processing may be possible in which articulation is suppressed or avoided altogether. This suggestion was expressed by three Ss in the written condition who remarked on different modes of coding for the two lists (all Ss were asked to report any points of interest on the back of the last response sheet). One of them phrased it as follows. "If I quickly write down the last letters (IL) that were in the frame (of the memory drum) I can still 'hear' the earlier ones and write these next." It is possible to decide from the 99 errors that were made in performing the written IL task that coding was not primarily acoustic. The confusions among written IL items, presented in a compressed matrix, show a tendency opposite to that known for written STM material. The high frequencies of some confusions (such as 38 in total between the pairs F-T, B-R, P-R) suggest visual or motor errors rather than acoustic coding.

31	33
23	12

These latter results suggest that indeed the OL material is coded in an articulatory-acoustic manner and that Ss attempt rehearsal according to the model. But if a short list is interpolated for immediate written reproduction, the acoustic code is avoided so as not to disturb ongoing rehearsal. Of course if spoken reproduction is required, the acoustic code cannot be avoided That adequate written IL performance is still possible may be ascribed to the effectiveness of VSR immediately after the IL has disappeared VSR may then be scanned quickly (resulting in visual errors) and be processed at a very high rate along the lines of the model, e.g in between two rehearsal cycles, or there may be a coding of VSR information 'directly' into motor patterns for copying, bypassing the articulatory mechanism. The latter possibility of course requires a separate channel in the model of Chapter 1. This channel would enable S to 'trace' letters from an array without causing much interference to simultaneously rehearsed items. In the present experiment the array would then be formed by the VSR registration of the trigram. In the next experiment we shall further investigate this possibility.

Conclusion. The larger effect of spoken IL on intrusions in OL recall, in spite of its modality difference from OL, has been established. If this can be further substantiated there is reason to doubt whether enough evidence remains for the assumption that STM is associative and therefore not different from LTM (2.1.2) However, a finding which was not anticipated on the basis of the model is that IL processing, when written, did not show any signs of acoustic coding. This might explain the results wholly in terms of more and less successful attempts by S to keep OL and IL items apart. This explanation has two consequences. First acoustic coding is essential for the (longer) retention of OL items to the extent that S apparently cannot keep them apart from spoken IL items. This is in agreement with the model. Second, the (immediate) recall of an IL trigram can be performed in a non-acoustic manner This evidently requires elaboration of the model by a direct route between recognition and output. We shall now investigate this suggestion further.

2.1.6.4 Experiment 4 Memory load and interpolated performance

The present experiment studies the mode of processing verbal items in an interpolated task (IL) as a function of the memory load involved in performance in the original task (OL). Fxperiment 3 suggested the possibility of a 'lower' form of processing characterized by non-acoustic coding for the immediate report of short messages, presented visually, when their acoustic coding would cause interference with ongoing rehearsal of other material. The question now is whether such low-level IL processing can be replicated and shown to be a function of the amount of OL rehearsal that must be supposed to be going on concurrently. If this is established, it may be assumed that such low-level processing is not the normal manner of dealing with verbal items but only forced upon S when his articulatory mechanism is fully engaged. If there is no relation between memory load and the mode of processing of interpolated material it must be assumed that S may select any code.

In the present experiment digit material was chosen for the construction of OLs. Copying of permanently displayed letter arrays is chosen as the interpolated task. The reason for selecting letters is - apart from their discriminability from digits - the opportunity they provide for establishing acoustic coding. The reasons for choosing a copying task rather than a recall task as in Experiment 3 is the following. Copying is a type of task which would evoke low-level processing even if there is only a slight chance of overloading the articulatory mechanism. Furthermore, in Experiment 3 it was suggested that S traces VSR information in a manner similar to copying a visual array. The present task corresponds to this analogy. Finally, the total amount of OL and IL items which can be retained with any chance of success is not much greater than the 7 items used in Experiment 3. Interpolated recall of ILs would require short ILs and allow only minimal variation of the length of OLs. Incidentally, short-term retention is of course also involved in copying, even over very short distances; this has been verified experimentally by Conrad and Hull (1967).

Method. Sixty-four digit sequences were drawn up from the digits 2-9. Composition of the sequences was random with restrictions concerning repetitions and ascending and descending series in order to avoid large variations in difficulty. Sequence lengths were 2,4,6, and 8; 16 sequences of each length. These constituted the original lists (OLs). Interpolated lists (ILs) were 64 seven-letter sequences, always composed of both B-set and F-set items. Their construction was random with some restrictions concering repetitions. Every sequence was used once in each condition. The OLs were printed on cards in a row of 6 cm high digits. The ILs were printed on the response sheets in clearly spaced typewriter (pica) capitals beside the row of seven boxes in which they had to be copied before recall of the OLs. The response sheets also contained boxes in rows of appropriate length for the recall of OLs. Exposures corresponded to a presentation rate of 1 digit per sec, followed by an interval allowing 1 sec per box to be filled plus 1 sec per presentation. The 'stop' signal anounced every next sequence. Timing was achieved by a taperecorded schedule. Short intervals separated blocks of 16 trials. Ss were 44 paid housewives from the APU subject panel. Their age range was 26 to 67. They were tested in three groups of approximately equal size. In the instruction the need of starting on the IL task immediately after the OL presentation was stressed. OL was, however, introduced to the Ss as the primary task.

Results and discussion. The percentage of sequences containing errors in OL recall and in IL copying performance over the four memory load conditions are presented in Table 2.7. OL recall shows a remarkably high level of performance. With a .50 criterion for errorless sequences the average value of digit span would be 6-7, which is high in view of the copying task interpolated between its presentation and recall. The increase in copying errors is somewhat smaller than anticipated: no two adjacent IL values in Table 2.7 are significantly different. Tested over the 16 sequences that were used once in every condition, the increase from condition 2 to 8 is however very significant (Wilcoxon matched-pairs signed-ranks test, $N=13$; $T=2.5$; $p < .005$). The error analysis of the interpolated task provided some difficulties because of the relatively small number of errors and the complex nature of many of them. The 327 errors made in the copying task contained as many as 91 omissions, providing no information on the mode of coding. Of the remaining 236 errors 69 were involved in complex error patterns and only 167 were 'simple' in the sense that a single box in a sequence contained a single wrong letter. Because of their small number it was decided to pool the IL data from OL sequence length 2 and 4 into a 'low' and the remaining into a 'high' memory load condition. The OL data of Table 2.7 justify such pooling. The 'simple' errors could now be entered into a 2x2 contingency table (Table 2.8), the criterion of acoustic confusion being the common possession by a wrongly copied item and the item presented of a vowel or a consonant (Wickelgren, 1965d). The table shows a tendency of non-acoustic errors to increase with high memory load. This tendency is just significant

Task	OL sequence length			
	2	4	6	8
OL (digits)	11.4	11.2	36.8	80.1
IL (letters)	5.7	7.1	9.2	9.7

Table 2.7. Percentage sequences in OL and IL tasks containing errors (Exp. 4).

at the 5 per cent level ($\chi^2=3.96$; $p < .05$). A number of scoring rules were designed to deal with the 69 'complex' errors such as anticipations, omissions leading to completions, and repetitions (Thomassen, 1966a). The purpose of these scoring rules is to determine which presented letter has acted as the stimulus for the specific response that is wrong for its position. Thus, an acceptable S-R relationship could be established for 60 of the 69 errors. Of these, 31 belonged to the acoustic and 29 to the non-acoustic class. For these errors no indication whatsoever was therefore found of any acoustic factor. Subdivision of these errors according to high or low memory load was not undertaken.

Conclusion. The result of Experiment 3 has been replicated: verbal material interpolated between the presentation and recall of verbal STM material is less likely to receive an articulatory code than the STM material itself. IL performance is nevertheless possible to a reasonable standard, which necessitates the addition to the model of Chapter 1 of a separate channel for low-level processing of verbal items. This channels bypasses the articulatory mechanism and feeds recognized items directly into a motor mechanism for writing. The fact that a high STM performance level is possible in spite of interpolated copying is further support for the relative separation of articulation (of rehearsal) and copying. The significant decline of acoustic errors in IL with increasing OL length indicates that the articulatory mechanism 'accepts' more IL items when it is not fully engaged (2-4 digits) than when it is rehearsing 6-8 items. The acoustic code may probably be considered the natural code even in copying if there is no concurrent rehearsal going on.

Memory load	Type of copying error	
	acoustic	non-acoustic
low	45	17
high	60	45

Table 2.8. Distribution of acoustic and non-acoustic confusions in letter copying presented as an interpolated task (Exp. 4).

2.2 The advantage of acoustic over visual presentation

2.2.1 Input modality in verbal STM experiments

Once the stimulus items of a sequence are coded no great effect of the modality of their presentation ought to be expected Indeed, large differences are never observed between the recall of auditory and visual sequences, but a small advantage of auditory over visual presentation systematically crops up in verbal STM experiments, especially with respect to the last items presented. Murray (1965a, 1966) obtained better recall of eight-digit sequences with increased overt vocalization. Inspection of his serial position curves reveals, however, that this difference is substantial only at the last serial positions. Posner (1967) reports similar results especially at higher speeds an advantage is observed of an auditory presentation. Corballis (1966) and Conrad and Hull (1968) report a clearcut positive effect of vocalizing on the last serial positions in the recall of visually presented digit sequences. Murdock (1967) obtained lower error rates when testing later paired associates PAs) from a list of six which was presented once if presentation was auditory. The same effect of modality is implicit in data of Atkinson and Shiffrin (1968) on continuous PA performance These data reveal better retention with short lags (only 1 to 3 items intervening the presentation and the test of a PA) under auditory conditions List of words (Woodhead 1966) and the last section of a prose passage (Poulton and Brown 1967) also show better retention if the material is vocalized at presentation.

There are several problems involved in the above modality effect in verbal STM One is that there is no control of perceptual differences occurring at presentation, so that e.g a visual sequence, especially if presented at a fast rate, may be inadequately perceived as compared to an auditory sequence presented at the same rate What should be the exposure duration of individual items to match their discriminability when presented auditorily? If Ss are required, as in most of the experiments mentioned above, to vocalize at some and not to vocalize at other trials, is the effect then attributable to enhanced articualtion or to auditory feedback? The first question is answered, in part, by results obtained by Conrad and Hull (1968) which show that for a .5 sec rate the recall of seven-digit sequences is not influenced by the on-off ratio if this varied between 50/450 and 450/50.

Arguments that the superiority of recall of the last items after vocalization is due to acoustic rather than to articulatory factors may be found in the fact that acoustic presentation without instructions to vocalize result in a similar effect, and that the advantage may be wiped out by delayed recall (Cooley an McNulty 1967) as well as by further acoustic stimulation. The latter form of stimulation may be either from later items presented (which may be the reason that the effect is limited to the last serial positions), from noise during presentation (Murray 1965b), or from the spoken recall of earlier items which, as may be seen from Murray's (1966) serial position curves, is somewhat less favourable than written recall to the last items of a sequence when this is vocalized at presentation.

The question of how the modality effect operates exactly cannot yet be answered. It may be true that the coding response for visual material requires some extra time because it is less 'direct' than the coding response for auditorily presented material. Indeed, the maximum presentation rate for adequate STM performance is reached sooner for visual than for auditory material (Posner 1967). Also in agreement with this notion is the fact that there is a faster decline in visual than in auditory STM performance with increasing age (McGhie, Chapman, and Lawson 1965). But as we have seen, the modality effect cannot be due only to the acoustic coding advantage, because it is only temporary and liable to interference by subsequent acoustic stimulation.

Among the factors involved in the described modality effect may be better temporal 'tagging' of auditory items. It has been suggested by Murray (1967) who found less order errors after auditory than after visual presentation, that the auditory mode provides more cues for item discrimination on the basis of temporal characteristics. This seems plausible on intuitive grounds and, furthermore, it is in agreement with such findings as those by Pollack, Johnson, and Knaff (1959). These authors report a decrease in the overall performance and a change in the serial positon curve when Ss recall the last items of long sequences with unknown length (running memory span). The discrepancies from short fixed-length performance are greater when auditory than when visual conditons are compared. It may be assumed that serial positon cues are generously avialable in short sequences with a fixed number of items. The 'normal' shape of the serial positon curve will be hard to explain without reference to the fact that the initial and terminal items in such sequences stand out and that they can both be anticipated. If auditory presentation facilitates the associations between these extreme and interposed items, thus providing serial position cues, the loss of these cues with long lists of uncertain length will be most conspicuous after auditory presentation.

There is at least one other factor involved in the modality effect. This is related to the characteristics of the auditory sensory register (ASR) of the model presented in the first chapter. Auditory information is longer available in uncoded form. Even after S has completed his recall of coded verbal items, this information may still be present as a kind of echo and support the recall of the last items. This is precisely what we observed in Experiment 3 where spoken IL recall appeared to interfere with subsequent written OL recall more than written IL recall did. We shall have more to say about ASR storage in the next chapter. Here it may suffice to indicate the superiority of ASR compared to VSR and to relate it to the modality effect discussed. In the next section, moreover, it will be related to performance in the assimilation of non-verbal sequential information.

2.2.2 The assimilation of sequential information

In the last section differences between visually and auditorily presented sequences of verbal items were discussed. Quantitatively no great differences were found, and qualitatively the effect of modality appears to be restricted to the last few items which are recalled better after auditory presentation. But we have also stressed the fact that verbal items when presented visually are readily coded by means of a set of verbal labels into a sequence with speech characteristics. Clearly, the coding efficiency described is dependent on learning, and the amount of learning that has been achieved in a normal S is probably great enough to obscure much greater differences which may exist between the assimilation by eye and by ear of sequential information. Sequences of discrete items that have no verbal labels closely attached to them will probably show much greater recall differences between modalities.

Auditory stimulation, speech and non-speech, is structured mainly along a temporal dimension, whereas visual information is primarily spatial. Sequential information will thus, on a priori grounds, be more adequately assimilated by the auditory system, supplied with storage or retention devices to integrate information over time. As we have discussed earlier, auditory speech perception is greatly aided by preperceptual storage in ASR which enables the listener to delay his perceptual decisions. It seems likely that the auditory processing of sequential messages other than speech will also rely on ASR - at high rates of presentation to delay perception, at lower rates to integrate the separate items into a single temporal sequence. VSR stores information spatially and apart from its considerably shorter duration each new stimulation tends to wipe out the earlier registration. Therefore, temporal integration of visual sequential items that are not (yet) verbally coded will be inferior to temporal integration of auditory information. Indeed, rhythmic patterns are reproduced much more accurately if the items are tones than if they are light flashes (e.g. Rosenbush and Gardner, 1968). Morse code is an equally obvious example of the priority of audition for dealing with sequential information. For well practised operators the optimal and maximal transmitting speeds are roughly three times as high for the auditory as for the visual mode.

Such differences are traditionally ascribed to "the poorer temporal resolving power of the eye as compared with that of the ear" (Taubman 1950; Woodworth and Schlossberg 1954; Eriksen and Collins 1967), but it is not really clear where exactly lies the cause of this big difference. The eye is probably most inferior to the ear at the higher levels of integration, and not so much at lower levels. Allowing for some generalizations this may be inferred from the following. Flicker fusion occurs at frequencies (e.g. 18 cps; Landis 1957) not much lower than those at which acoustic pulses cause a tonal sensation (Von Békésy 1960). Furthermore, visual flash rates are not discriminated unequivocally worse than auditory pulse rates of comparable frequencies (Pollack 1952, 1953; Mowbray, Gebhard and Byham 1956; Mowbray and Gebhard 1960). Lastly, temporal numerosity curves portraying the judged number of items as a function of the number presented at varying rates for visual, auditory, and tactile stimulation seem to be subject to quite similar limitations (White and Cheatham 1959). It should be noted that in these tasks no registration is required: S only has to report on certain quantitative characteristics of (redundant) stimulation. Differences from auditory processing apparently become great only when patterns must be assimilated for recognition, identification or recall.

This is obviously the case in visual Morse code transmission, where, incidentally, maximal flash rates (5-6/sec) are clearly below the range of fusion frequency and below most of the other frequencies studied in the above tasks. The 'natural' way in which auditory elements group together into the codes learned is illustrated by the fact that below a certain transmission speed the structures of the letter codes tend to disintegrate. On the other hand, in visual Morse the letter codes have to be separated by (also relatively) much longer intervals than in auditory signalling to prevent them from 'sticking' together. Other observations of importance are that even trained operators show a tendency to vocalize ('dih-dah-dah') the individual flashes at perception and their overt responses are in the order of letters rather than of words or groups as in auditory transmission.

2.2.3 Experiment 5: The effect of modality on STM for non-verbal items

We will now describe an experiment dealing with modality effects on STM for sequences of non-verbal items. The reason for selecting these items (tones and flashes) is to study 'pure' sequential performance under reduced opportunities for mediation by speech. With such items a much bigger effect is expected than the small discrepancies discussed above with respect to verbal material. It may even be questioned whether the visual system allows purely sequential processing at all of discrete items such as flashes. The individual signals, rapidly vanishing from VSR, probably require coding responses of some kind to integrate them into sequences. If the more obvious sequential code is in an auditory form, some mode or other of articulatory coding is still a possibility even with non-verbal items. This possibility was also suggested above by reference to the tendency to code visual Morse codes acoustically. It finds further support in the relative ease with which visual signalling can be interfered with by incompatible acoustic signals (e.g. Erkens 1968). The suggestion to be tested in the present experiment is that, although visual successive (VSU) presentation of non-verbal signals will result in much poorer performance than auditory (AUD) presentation, both will be similar qualitatively because the sequential processing that is going on VSU will be achieved primarily on the basis of acoustic coding. The shape of the serial position curve will be regarded as one index of similarity. Further comparisons will be made by means of correlations. Two more conditions will be included in our comparison. The first, visual simultaneous (VSI) presentation is adopted to differentiate between the visual system's capability to deal with spatially and temporarily structured information. Large differences are expected qualitatively, and to the extent that such comparisons can be made, also quantitatively. Secondly, a combined auditory and visual (CAV) condition is introduced. If visual processing (like auditory) is autonomous, without deliberate recoding, some effect on auditory processing must be expected. This would be either favourable because of enhanced information derived from the two separate channels, such as has been reported by Corcoran and Weening (1969) for single synthetic stimuli, or detrimental because of interference of some sort, in spite of the two channels providing identical information. If no effect is found, this is evidence for the ease with which visual information (unlike auditory) can be disregarded, indicating that it has no autonomous admission to higher levels of processing.

Method A series of 20 sequences was prepared, each containing 3 long and 3 short items. They were recorded on tape as sequences of 1500 cps tones with 330 and 50 msec durations. The interval between the onset of the tones in a sequence was constant at 540 msec. The recall interval succeeding each sequence was 21 sec, terminated by the anouncement (recorded on the other track of the tape) of the next sequence, which followed 2 sec later. The tape recording constituted the auditory (AUD) presentation condition. The same tape was used for operating a relay system responsible for switching on and off two sets of 40 W light bulbs. The relay was rapid (35 msec) and so were the rising and falling characteristics of the lamps (90 per cent within 10 msec). The flash sequences, which constituted the visual successive (VSU) condition, thus had near-identical time relationships compared to the AUD sequences. The flash sets (not visible themselves) projected their signals onto a screen, where they appeared as dim but clearly discriminable flashes. This arrangement was chosen to reduce the need of fixation and the effect of afterimages, thus increasing the similarity between AUD and VSU The relay system, enclosed in a sound absorbent box, was placed in a room at the other end of the building so that no acoustic cues were present under VSU. Combined presentation of auditory and visual signals (CAV) was achieved by using both AUD and VSU synchronously. Visual simultaneous (VSI) presentation was by slides, each containing a sequence in a horizontal block pattern, white on a black background. Slides were presented by an automatic projection tachistoscope set at a rate of 20 sec. Exposures were 540 msec, on the screen the patterns appeared 62 cm wide and 3.8 cm high. The same 20 sequences were used in the same order in all conditions. Two groups of Ss were tested. Group 1 was given AUD, VSU, and VSI, Group 2 received AUD, VSU, and CAV. The Ss of Group 1 were 120 recruits (age 19-22) tested in 6 subgroups of 20. The Ss of Group 2 were 103 grammar school pupils (age 16-18) tested in 6 subgroups of approximately 17. Order of conditions was balanced within Groups. Recall was required on response sheets in dots and dashes, ambient light was sufficient for this task New conditions were preceded by practice trials

Results and discussion. The primary results in terms of error proportions over the serial positions are shown in Figure 2.3. The two groups appear to perform at considerably different levels, both with respect to the proportion of item errors and to the proportion of sequences with one or more errors under conditions AUD and VSU which were identical for both groups (Table 2.9). Nevertheless, the ratio between these conditions and the shapes of the serial position curves are similar between the groups.

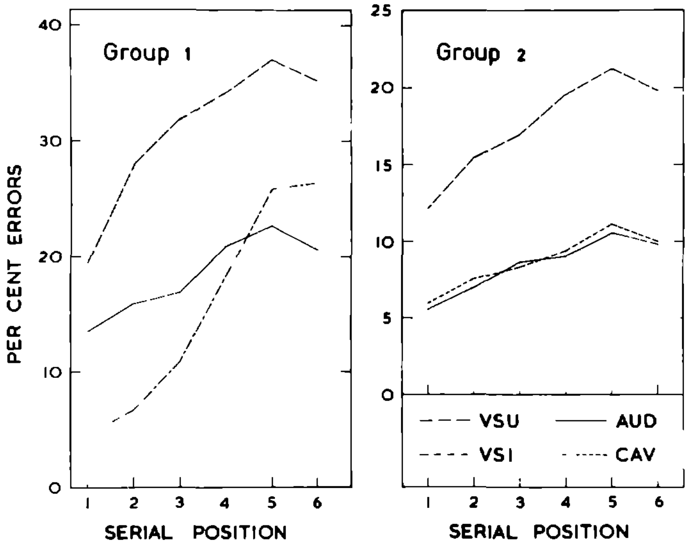


Figure 2.3. Performance in the STM recall of non-verbal sequences of six binary (short-long) items after auditory (AUD) and visual successive (VSU) presentation. For Group 1 a third condition was visual simultaneous (VSI); for Group 2 this was a combined auditory and visual successive (CAV) presentation (Exp. 5).

Within both groups there is much better AUD recall than VSU, and in both groups the serial position curves for AUD and VSU have highly comparable shapes. In Group 1 the quantitative difference between AUD and VSI is only very small and opposite for items and sequences. The VSI curve, however, represents the most salient feature of Figure 2.3. It is very different from all the other (sequential) curves. In Group 2 there appears to be no difference whatsoever between the AUD and CAV conditions. Correlations are given in Table 2.10 which also presents the significance levels of the differences discussed. It appears from Table 2.10 that both over Ss and over sequences the correlation coefficients are all comparable in size, with the only exception for those involving comparison with VSI.

Group	condition	items	sequences
1	AUD	18.2	44.6
	VSU	30.7	70.4
	VSI	15.4	45.5
2	AUD	8.5	21.4
	VSU	17.6	42.6
	CAV	8.8	22.6

Table 2.9. Proportion of errors in the recall of non-verbal sequences (AUD = auditory; VSU = visual successive; VSI = visual simultaneous; CAV = combined auditory and visual successive presentation). Fifty per cent wrong items equals chance performance (Exp. 5).

Group	conditions compared	significance of difference	rank correlation coefficient r_s	
			subjects	sequences
1	AUD/VSU	$p < .001$	.61	.94
	AUD/VSI	$p < .05$	.28	.32
	VSU/VSI	$p < .001$	.31	.40
2	AUD/VSU	$p < .001$	.59	.86
	AUD/CAV	n.s.	.62	.93
	VSU/CAV	$p < .001$	.56	.88

Table 2.10. Comparisons among the conditions of Table 2.9. Differences were tested by means of the Wilcoxon matched-pairs signed-ranks test. Correlations (r_s) were run over Ss (n_1= 120; n_2= 103) and over sequences (n= 20). (Exp. 5).

Correlations over the sequences are quite high between all conditions involving sequential presentations. This suggests a high degree of equivalence in spite of the large difference in AUD and VSU performance level. These results strongly support the acoustic recoding suggestion made above. The identity between AUD and CAV is confirmed by the highest correlation coefficients obtained for Group 2. Without specific coding responses, therefore, visual sequences are not assimilated. The shape of the VSI curve portrays a typical result of the VSR scan of the model of Chapter 1. The briefly presented array is available for some time in VSR. Its read-out is from left to right, but not rapid enough for the latter items to be coded, so that chances of incorrect recall increase sharply with serial position. Incomplete scanning is, however, not the only way in which VSI and the sequential conditions are different, because that would predict lower correlation over sequences only. The lowered correlation coefficients over Ss, however, strongly suggest that different processes are involved altogether.

Conclusion. The results are in agreement with the interpretation of acoustic coding even of non-verbal items, if these are successively presented discrete visual signals. This conclusion is based on qualitative similarities in spite of quantitative differences, on the obvious need for a specific code at presentation when visual material is sequential, and on its dissimilarity from processing simultaneous visual material. With reference to the model the finding may be considered as evidence for the superiority of ASR when ordered sequential information is involved and for the generality of coding activity to transform sequential information into a mode which somehow corresponds to this auditory superiority. The findings thus support the role of auditory images for sequential processing in general. It may correspondingly be claimed that the amount of auditory experience will determine the efficiency of sequential processing. In the next section we shall turn to the deaf, a group in which auditory experience - also of non-speech events - is minimal.

2.3 Short-term memory in deaf subjects

In the preceding section the superiority of the auditory system was indicated for the assimilation of sequentially presented information. The assimilation of sequential visual information was, moreover, related to the coding possibilities available in the sphere of the auditory system. This relation was seen as a result of learning processes involving audition which has 'biological priority' of some sort for temporally structured stimulation. The principal form of such stimulation is of course speech. If we now look at a group which is deprived of the auditory experiences that directly (through recoding) and indirectly (through learning) mediate the assimilation of visual information and which, moreover, has a limited experience with speech, a definite impairment of the capacity to process sequential presentation should be found. Such a group is formed by the deaf. In the next few pages we shall review some of the literature on perceptual and short-term memory processing by the deaf. This review will lead to several questions, some of which may be answered by the experiment to be reported in this section.

2.3.1 Observations in the literature on the deaf

There is a wide variety in the experimental literature on immediate memory in the deaf. This is mainly caused on the one hand by large variations in the criteria for deafness, in the type of education received by the particular group of deaf, and in the control groups chosen for comparison; on the other hand widely differing tests, presentation procedures, and methods of scoring are employed to establish immediate memory performance. There is, however, sufficient agreement to conclude the following. Immediate memory for spatial locations is not inferior to that of hearing Ss (Morsh 1936; Blair 1957; Doehring 1960). Likewise, movements in space and motor patterns are often more accurately reproduced by the deaf than by their hearing controls (Frisina 1955; Hiskey 1956; Blair 1957; Costello 1957; Fuller 1959). These findings are interpreted as evidence for compensation with respect to the assimilation of visual, visuo-motor, and kinesthetic information, due to their relatively greater practice.

Also visual perception and immediate reproduction of simultaneously presented simple and more complex patterns is at least as accurate by deaf as by hearing Ss (Hofmarksrichter 1931; Clarke 1951; Hiskey 1956; Blair 1957; Naffin 1959; Kilpatrick 1963). Delayed recognition or recall of such patterns is, however, often below the hearing standard (Hiskey 1956; Goetzinger and Huber 1964). These results have been related also to a compensatory superiority due to increased practice of the visual system, while in the deaf, moreover, perceptions are supposed not to be distorted by verbalizations which apply meaningful labels to meaningless or ambiguous patterns. On the other hand, prolonged retention is not supported, as it is in hearing Ss, by subvocal rehearsal. Similarly, successive presentation of pictures and dot patterns results in poorer performance by the deaf. This has been explained by a less well developed 'abstract capacity for the mental integration of the elements' of such sequences (Blair 1957).

For verbal items the results indicate an unambiguously inferior performance by the deaf. This applies to simultaneous as well as successive visual presentation of materials such as sentences (Brill and Orman 1953), digits and letters (Pintner and Paterson 1917; Hiskey 1956; Blair 1957; Olsson 1963; Conrad and Rush 1965) but also to sequences of domino patterns

(Blair 1957) and of colours (Hiskey 1956) which, as we have discussed above (2.1.3) are readily coded and thus effectively also verbal items.

In summary, immediate memory in the deaf is not by definition less efficient than in hearing Ss. There are three conditions in which the deaf show inferior STM performance. These are sequential presentation, delayed recall, and the use of verbal items or of items that are readily coded by hearing subjects into verbal items. It should be noted that the language defect of the deaf presents itself in all three conditions, not primarily as a defect in the use of language as a semantic system, but rather as a defect in the active use of language as a sequential skill. It is this skill which enables hearing Ss to assimilate successively presented verbal information, and to rehearse over the retention interval when recall is delayed. It was discussed above that STM is not primarily based on meaning (2.1.2); correspondingly, the STM defects in STM performance by the deaf are not in the first place seen by us as conceptual handicaps but rather as inadequate sequentializing as a direct (and indirect) result of a poor command of speech as a sequential skill.

2.3.2 Experiment 6: STM in deaf and hearing subjects

Some of the implications of the defect just described will now be tested in an experiment. First, it is of importance to replicate the differences between deaf and hearing Ss reported in the literature above with the kind of verbal items used in our earlier experiments. Second, just as in Experiments 1 and 2 acoustic coding was confirmed by a very significant difference between the recall of acoustically similar and dissimilar items, it will be tested whether this difference is absent in the deaf group. Third, even if this is confirmed and 'acoustic' coding therefore unlikely, it will be attempted to test the presence of articulatory coding in the deaf. If this form of 'speech' processing is found, its role in STM recall will be determined and the extent to which the model presented in Chapter 1 also applies to deaf Ss will thus be established. Fourth, and most specifically in connection with the present context, performance in the sequential aspects of the STM task, viz. order errors, will be isolated and their relative frequency tested.

Method. Twenty-six severely deaf pupils of the Instituut voor Doven, St.-Michielsgestel, were selected as Ss. Their ages ranged from 11-15. They were individually matched on sex (15 male, 11 female) and on non-verbal intelligence by 26 hearing children, whose age range was 11-13, of a local school. Ss were tested individually by an E with experience in testing deaf children. The material consisted of 34 cards with 21 mm high letters, clearly spaced horizontally. Sequence lengths were 2-7 letters. These were either all from the B-set or F-set (homogeneous) or from both (heterogeneous). All 14 letters occurred 6 times in a homogeneous and 6 times in a heterogeneous sequence. The cards were presented randomly but in ascending length, starting with length 2, which was exposed 2.5 sec, and ending with length 7, exposed 6.0 sec. Exposures were preceded by a short interval in which E gave S a warning signal by touch. Recall intervals ranged from 5 to 12 sec. Timing was monitored from a tape through E's earphones. Guessing was encouraged: the vocabulary was printed on the top of the form on which immediate ordered recall was required. Ten training sequences preceded the 34 experimental sequences.

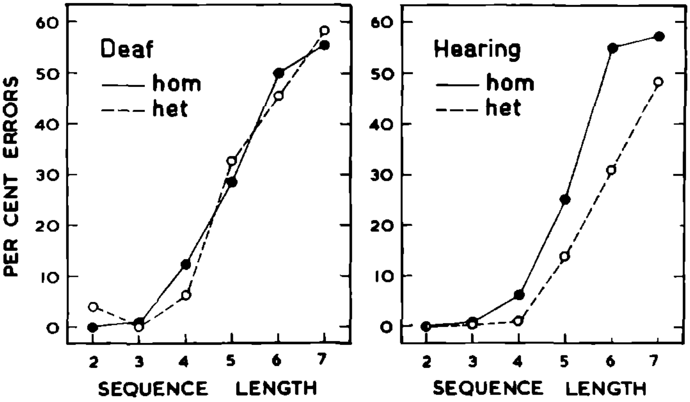


Figure 2.4. Percentage errors in the STM recall by deaf and hearing Ss of homogeneous and heterogeneous letter sequences of different length (Exp. 6).

Results and discussion. The overall performance of the hearing group is superior ($p < .025$ over the sequence criterion and $p < .015$ over the item criterion) but this difference is entirely due to better heterogeneous recall by the hearing Ss. Recall error percentages are shown in Figure 2.4. It is clear from the figure that for the deaf no differences are found between homogeneous and heterogeneous lists, whereas the hearing Ss show the significant effect of acoustic similarity also reported in Experiments 1 and 2. The differential effect is shown to hold over all sequence lengths with a tendency to increase up to length 7. A perfect match as found between the two groups in the homogeneous lists was not expected but it provides a nice opportunity for comparison. Besides suggesting that acoustic dissimilarity is the only type of cue on which is based the better hearing performance, it shows that identification of verbal items as such is certainly not worse by deaf than by hearing Ss. The compressed matrices of the heterogeneous sequences (with sequence length 7 omitted because of its high error rate) are presented below for comparison of the amount of acoustic coding observed.

74	71
104	73

deaf

57	29
34	54

hearing

Obviously there is little acoustic coding to be concluded from the deaf recall errors, at least as defined by the probability of intraset confusions. On the following pages evidence for an articulatory type of coding will be further discussed; also errors in the recall order will be given special attention.

The role of articulation is difficult to test if it is to be differentiated from that of auditory coding. However, since there are no more within-set than between-set confusions in recall by the deaf it must be assumed that the articulatory cues implied in the accompanying vowels, or in their position in relation to the consonants, result in no effect. Among the consonants themselves, six pairs can be formed with equivalent articulation (B-P, D-T, S-Z, F-V, S-C, Z-C). If the fair performance by the deaf is to be attributed to their enhanced articulation (which one is inclined to conclude when observing the articulatory movements in most deaf Ss engaged in the STM task) the proportion of such confusions over all other confusions should be greater in the deaf than in the hearing group. Indeed, this is the case: 12.9 vs. 9.8 per cent ($X^2 = 5.22$; $p < .025$). It also of interest to note here that within the BDPT-subset both groups show more confusions between the pairs with common place of articulation (B-P, D-T) than with common voicing (B-D, P-T), but that this preference tends to be greater among the deaf Ss. The difference, however, falls just short of the .05 level of significance ($X^2 = 3.51$; $p < .10$). To distinctions of this kind we shall return in the next section.

These finding do suggest that enhanced articulatory coding goes on in verbal STM performance by the deaf. The evidence seems too slight, however, to account for the relatively high score obtained by this group. And if articulation is used, why are heterogeneous sequences not better recalled? Our suggestion is that even if coding is articulatory by (some) deaf Ss, there is too little rehearsal to provide the cues for the items as well as for their sequential relationships - and it is principally in a sequential respect that homogeneous and heterogeneous sequences are different. In two very recent publications Conrad (1970a, 1970b) also reports articulatory coding in a certain proportion of his deaf Ss (now of a population different from the American deaf school which provided the Conrad and Rush (1965) data without any evidence of acoustic coding). But the Ss with acoustic confusions were not better than those without. Only when overt reading at presentation was required did the non-articulators recall less well. These findings confirm our suggestions in two respects. First, articulation cannot explain all STM coding by the deaf; second, if it occurs, it is not necessarily as advantageous as it is to hearing Ss whose subsequent rehearsal will also result in a better retention of the order between the items.

Two further sources of evidence are available in our data. The first is the low correlations that were found between STM and scores on the spelling test used to screen our deaf Ss (only normal spellers participated). Spelling and overall STM scores showed no correlation ($\underline{r}_{\underline{s}}$= .051), which parallels Conrad's finding. But spelling did not predict either whether a S would have relative difficulty with the homogeneous lists ($\underline{r}_{\underline{s}}$= .017). Thus, a good articulator (speller) among the deaf group is neither a good STM performer nor even a S who uses articulation to such an extent that it affects his ordered recall of acoustically similar and dissimilar sequences differentially. The second source of evidence for the minor role of articulation cues by the deaf is the occurrence of order errors, which will now be considered. In correspondence with the assumed order-strengthening function of rehearsal in the model of Chapter 1 it is suggested that order will be affected more severely in the deaf. The above passage underlines this suggestion by noting that even if individual items are given a 'speech' code by (some) deaf Ss, this code is not rehearsed sufficiently to provide interitem cues at recall.

Order recall was studied by considering only those sequences of which all letters were recalled correctly but in the incorrect order. The deaf group showed 56 such sequences (or 11.7 per cent of all sequences containing errors) while in the hearing group there were only 29 sequences with solely order errors (or 7.0 per cent of all sequences containing errors). The total numbers of letters involved were 254 for the deaf and 151 in the hearing group. This difference was present in 18 of the 26 matched pairs, with two pairs showing no difference (Sign test: $N= 24$; $x= 6$, $p= .011$). A more valid comparison is made by testing over homogeneous sequences only, because here total performance was identical for the two groups. The difference remains significant, however, in spite of the reduced numbers of errors, viz. 153 v. 90 (Sign test $N= 20$; $x= 5$; $p= .021$). When testing over heterogeneous sequences a distinction can be made between order errors involving within-set and between-set confusions (i.e. letters recalled in the position of a member of the same or of the other acoustic set). Both types of order error are more frequent in the deaf, the difference being much greater for between-set confusions ($X^2= 7.49$; $p < .01$) mainly because of a sharp decline from within-set to between-set order errors in the hearing group. These results thus strongly support our suggestion that decreased 'acoustic' (auditory or articulatory) processing has a relatively strong effect on the retention of order in STM.

Conclusion. The impaired verbal STM performance by the deaf, known in the literature, has been replicated. The differences with the hearing controls are, however, not great and they exist only with respect to those sequences which for the hearing have the advantage of acoustically defined heterogeneity. Acoustically homogeneous sequences were recalled equally well by hearing as by deaf Ss which indicates at least an adequate identification of verbal items by the deaf. The fact that in the deaf no differences occurred between acoustically homogeneous and heterogeneous sequences and the fact that no more acoustically similar (within-set) confusions were made indicates that the 'normal' acoustic coding as defined by vowel characteristics is absent. Articulatory equivalent items, however, are confused more by the deaf than by the hearing Ss. A role of articulation, therefore, seems to exist, but it is certainly not a large role. It is probably best described as one of supporting perception. The articulation effect, moreover, is only small. There were no correlations between total recall scores and scores obtained by the deaf Ss on an overt articulation task (spelling) involving the same type of items. Nor is there an effect of their ability to articulate (spell) to the degree to which homogeneous material causes interference. These facts indicate that articulation is present in (some) deaf Ss, but that the deaf cannot use the cues optimally. The same conclusion may be drawn from the order errors. These are more frequent in the deaf group, which confirms the assumed defect in sequential processing. The articulatory cues apparently are inadequate or used to an inadequate degree for them to provide order information. Normally rehearsal is responsible for firmly establishing the order between items, but rehearsal is probably either practised little by the deaf or the deaf have a reduced opportunity for cue utilization.

2.4 Acoustic or articulatory representation

2.4.1 Errors in auditory perception and in STM

In previous sections of the present chapter (2.1 and 2.3) repeated comparisons were made between STM for acoustically homogeneous and acoustically heterogeneous sequences and the number and type of errors were attributed to acoustic similarity between the correct letter presented and the letter incorrectly recalled in its place. Acoustic similarity was assumed on an a priori basis between consonants which, if pronounced such as in reading the alphabet, share a common vowel phoneme. This resulted in the distinction between the B-set and the F-set. Apparently the common vowel phoneme is responsible for the amount and the pattern of errors in STM: if a letter is partially forgotten, its vowel phoneme may be retained independently of its consonant phoneme so that intrusion errors are not random but tend to share the vowel phoneme with the correct letter. But also within the sets with common vowels certain confusions are more probable than others; this is the main cause of the correlation reported by Conrad (1962, 1964) between errors in STM and in auditory perception (2.1.1). In other words, there are elements smaller than phonemes which are retained or forgotten relatively independently in STM.

It has been shown that the auditory perception of vowels and consonants in constant contexts is achieved according to the distinctive features which describe the items phonetically (Peterson and Barney 1952; Miller 1956b; Miller and Nicely 1955). The question regarding the observed correspondence between errors in STM and in auditory perception may thus be posed: is the representation of verbal items in STM also in terms of distinctive features and do the same features describe the errors of both matrices equally well? With respect to the letters used in our STM experiments this would mean that confusions not only tend to remain within the vowel-defined sets but also to occur more frequently where presented and recalled letters have more consonant features in common. If it appears that the phonetic distinctive features may thus be regarded as the elements of the internal representation of verbal items in STM a comparison can be made with respect to the distinctive features that describe auditory perception. A complete agreement between the two would lend support to the notion that both STM and auditory perception are mediated - at some level - by one single 'auditory' mechanism, as suggested by those who refer to STM as an 'acoustic' system. If, however, differences are found, these may lead to conclusions regarding a specific type of representation in STM.

Attempts to determine whether English vowel and consonant phonemes are coded in STM as sets of distinctive features were made by Wickelgren (1965e; 1966e). In his vowel experiment six-item sequences were read at a rate of 2 sec per item. Each item (CVC) in a sequence had a different vowel in a constant consonant environment. Analysis of the items incorrectly recalled included a test of relative confusion frequencies against binary predictions based on existing distinctive feature systems. Predictions on the basis of a system consisting of the two dimensions 'place of articulation' and 'openness of the vocal tract' appeared to be very precise.In the consonant experiments (Wickelgren 1966e) six and nine-item lists were read at a rate of 1 sec per item. The items (CVs) were a consonant, chosen from a set of 16 or 23, followed by the vowel /a/ in all cases. Predictions based on a distinctive feature system comprising voicing, nasality, openness of the vocal tract, and place of articulation as dimensions were reasonably accurate when tested against the confusion data. Thus it may be concluded that in the case of auditory presentation both vowels and consonants are stored in STM as sets of distinctive features in a manner similar to that which had been observed for auditory perception of vowels (Peterson and Barney 1952; Miller 1956b) and consonants (Miller and Nicely 1955).

The succeeding pages of the present section will be devoted to the collection and ordering of error data obtained in STM experiments and in an auditory perception task. The main purpose is to compare the confusion probabilities of the items used in both tasks. These probabilities are first laid down in confusion matrices, then they are subjected to a cluster analysis, and the causes of clustering are subsequently described in terms of distinctive features . Another set of data is obtained in visual perception tasks using the same letters (with identical shapes) as in the STM tasks where presentation was also visual. The main difference between the two types of tasks is that in STM there are several successive items to be recalled, each presented clearly, whereas in the visual perception tasks there is only one single letter in each trial, but presentation is such that identification cannot be perfect. The purpose of presenting these data is not an analysis in terms of the distinctive features used to describe the other types of errors but rather to demonstrate that imperfect visual identification or the mere impairment of a visual sensory trace is not a factor of importance in the STM data. The visual perception errors also provide an opportunity for the verification of the 'visual coding' hypothesis suggested in Experiments 3 and 4.

Our approach is different from that by Wickelgren in the following respects. First, the STM data are obtained with visual rather than auditory presentation. This seems essential for our purpose since, if acoustic coding is involved as a specific STM factor, one should avoid any contamination with acoustic factors at the input side, i.e. auditory presentation. Second, the items used in our experments are verbal items which are readily coded in the sense described above (2.1.3), and not 'artificial' items such as 'leck' or 'zha' as used by Wickelgren. Third, not only STM data are analyzed but also auditory perception data obtained with the same item vocabulary. Where Wickelgren makes comparisons with auditory perception he builds upon data collected by others under widely different conditions and with different item vocabularies. Fourth, our analysis is not concerned so much with deciding among different feature systems as it is with comparing and contrasting STM and auditory perception errors and the dimensions which determine their relative frequencies when one representative feature system is used. Our results will, therefore, not be presented as percentages correct predictions derived from various feature systems but rather as degrees of correspondence between the two matrices with respect to the relative weight of the dimensions constituting the feature system adopted for our analysis.

2.4.2 Confusion matrices

The following pages describe the manner of collecting data for comparisons between coding in STM and in auditory perception. Since our interest is mainly in the 'acoustic' coding of visually presented items, confusions in the visual perception of these items are also discussed. The general method of collecting the confusion data practised under all conditions was to present a large number of Ss with a large number of printed or spoken verbal items, always from the same 14 letter vocabulary. The items were arranged in a quasi-random order allowing for equal frequency and even distribution over the entire material in order to minimize differences in a priori probability of errors. Ss were instructed to guess if uncertain and the 14 letters, printed in the same type as used for presentation, were permanently displayed in an alphabetical order. The within-vocabulary errors are entered into confusion matrices in which the rows represent the incorrect responses to the stimuli at the top of each column. For easy identification of intra-set confusions the B-set and the F-set are grouped in the matrices. Hierarchical clustering schemes of the confusion data will be presented in the next paragraph, where also the pattern of visual confusions is discussed. The discussion of the STM and auditory perception confusions will be postponed till 2.4.4.

2.4.2.1 Matrix I: Short-term memory (STM)

The STM data collected for entry into a confusion matrix were collected in two separate experiments both using visual presentation. The first has already been reported (Exp. 1; 2.1.6.1); the second will be reported in Chapter 3 (Exp. 8; 3.4.2). Both these experiments satisfy the conditions mentioned above. In the error analysis all letters which were not recalled in the correct serial position were scored as errors. If these errors were omissions or letters from outside the vocabulary they were not included in the STM matrix proper. In Experiment 1 a total of 80 Ss were given 56 seven-letter sequences presented visually at a rate of 0.6 sec per letter. Error rate was 41.3 per cent with 4.1 per cent omissions and responses outside the vocabulary, leaving 37.2 per cent or 11,662 errors for analysis. In Experiment 8 a total of 120 Ss were given 56 six-letter sequences presented visually at a rate of 1.3 per letter. Error rate was 25.3 per cent with 0.5 per cent errors and confusions outside the vocabulary, leaving 24.8 per cent or 9,979 errors for analysis. The resulting 21,641 errors were entered into a single STM confusions matrix (Appendix 1).

There are two variables not under control in these experiments. Even if all letters occurred with equal frequency in all serial positions, frequency of one letter occurring adjacent to another letter was random and therefore unequal. Since paired transpositions make a relatively large contribution to confusion data this factor may be disturbing. Furthermore, each experiment served other purposes as well. These required certain structures (in terms of B-set/F-set ordering) to be present in the sequences. To the extent that the Ss learned these structures (especially in Experiment 8) this may have had an effect on recall in favour of intra-set confusions.

2.4.2.2 Matrix 2: Auditory perception (AUP)

On a two-channel recorder (Philips EL 3569) a tape was prepared of five male speakers each speaking five consecutive blocks of 14 letters chosen from the set B,C,D,G,P,T,V,F,L,M,N,R,S,Z. The letters were spoken at a rate of one every 4 sec. Each letter occurred twice in every two blocks of 14 and over the whole experiment every letter occurred twice in each serial position of a block. The 350 experimental letters were preceded on the tape by an instruction and by three blocks of practice letters. On the second track of the tape white noise was recorded. Its level was monitored electronically so that playback would result in a preset signal/noise (S/N) ratio at the loudspeakers. For the first practice block S/N was -4db, for the second -7db, and for the third practice block as well as for all the experimental letters S/N was -10db. Listeners were 95 male pupils of a grammar school. Their age ranged between 15 and 17. They were tested in two groups of approximately equal size in a room with no special acoustic provisions.Two 20 Watt loudspeakers (S and N) were placed in a vertical arrangement in the front centre of the room. The Ss wrote every perceived letter in the appropriate box on a response sheet. There were 14 boxes in a line; a short extra pause and a warning on the tape preceded the first letter of every line. The 14 letters of the vocabulary used were printed on the top of the response sheets. Ss were encouraged to guess rather than to leave blank spaces. Including a short rest in the middle of the test, the whole procedure lasted approximately 40 min for each group of Ss.

Results. Of the 33,250 responses 27.6 per cent were correct. There were 4.4 per cent omissions and responses outside the vocabulary, so that 68.0 per cent or 22,619 incorrect responses remained for analysis. Since there were no obvious differences between the five speakers, all these errors were entered into a single auditory (AUP) confusion matrix (Appendix 2).

2.4.2.3 Matrix 3: Visual perception (VP)

For the sake of comparison, confusion data for visual perception were also collected with the same letters and of the same type as used in the STM experiments. The purpose of collecting these data was solely to check whether the STM confusions, also obtained with visual presentation, were indeed different from perceptual confusions. There are a number of ways in which such perceptual errors can be evoked. One is to present letters at a sufficiently long distance for errors to occur, a second procedure is tachistoscopic presentation, and a third method is to present letters on a TV screen and to mask them by a large amount of visual noise. Because there is a great deal of arbitrariness in any of these methods and since there is no means of deciding a priori which type of presentation will result in errors similar to those that might occur in visual STM tasks, all three procedures were in fact employed.

Distance errors were obtained by presenting normally typed, well spaced, and clearly marked letters at a distance of 1.5 to 2 m, the exact distance being determined by advance testing of eyesight. Twenty Ss were tested individually. Their task was self-paced reading and their responses were recorded by E. After some practice, 392 letters were read by all Ss in a balanced order. Tachistoscopic confusion data were collected by projecting automatically one letter every 5 sec with an exposure duration of .013 sec. The height of a projected letter on the screen was 29 mm. Ss were seated at a distance of 2.25 m. Forty Ss participated; they were tested in groups of 4 or less. Following a block of practice trials 112 letters were exposed in a balanced order. Report was written. Noise errors were obtained by presenting one letter every 8 sec on a TV screen. The letter height was 135 mm. This signal was mixed with the output from a random noise generator, the S/N ratio being .04. The mixed stimuli were recorded on video tape Exposures were 5 sec, followed by an acoustically marked response interval of 3 sec. Ss were tested in two groups of 30 and 24 respectively. Their seats were 3 to 9 m. from the screen. Practice trials in which noise levels were increased were followed by 168 experimental letters in a balanced order. Report was written.

In each of the three methods the main confusion pattern between letters became clear with only relatively few observations. There was a much greater tendency for overcrowding in certain cells (G-C, P-F, B-S, Z-L) than in the other confusion matrices. There were 1909 intravocabulary confusions obtained by the distance method when it was decided to stop collecting further data. With tachistoscopic presentation this decision was taken when only 408 such errors were made. The noise condition, which involved testing in larger groups, yielded 960 errors within the vocabulary. The error percentages were 25.1, 12.0, and 13.1 respectively. Although the matrices were not identical, especially not with respect to the amount of symmetry and scatter, they were considered to be of sufficient similarity for their pooling into a single analysis for the present purpose. The 3277 errors, thus collected under largely varying conditions, are presented in the visual perception (VP) confusion matrix (Appendix 3). The letter shapes

(Pica capitals) used in the present visual perception tasks as well as in all verbal STM experiments with visual presentation are shown in Appendix 4.

2.4.3 Hierarchical cluster analysis

In order to compare the structure underlying the confusion frequencies observed under auditory perception (AUP) and short-term memory (STM) conditions the confusion matrices were subjected to a hierarchical cluster analysis (Johnson 1967). This analysis was adopted because of its minimal requirements with respect to the nature of the similarity data, its clear and comprehensive representation, and its usefulness for the description of confusions in consonant perception as demonstrated by Johnson in his examples for application. A hierarchical clustering scheme (HCS) provides a hierarchical system of clustering representations among a series of objects, ranging from one in which each of n objects is represented as a separate cluster (level 0) to one in which all objects are grouped together in a single cluster (level n - 1). The order of clustering is determined by similarity, the closer objects clustering earlier than the more remote objects. Smaller clusters are incorporated into larger clusters, thus giving rise to a hierarchical scheme.

The analysis, which was performed using computer programme HICLU in Fortran IV G/H for IBM 360/50 OS (Roskam and Brandsma 1969) first requires transformation of the confusion matrix into a symmetric similarity matrix. This was achieved in accordance with Johnson's suggestion by the formula

$$s(x,y) = \frac{f(x,y)}{f(x,x)} + \frac{f(y,x)}{f(y,y)}$$

where s(x,y) is the similarity measure, f(x,y) the confusion frequency between the presented letter x and the reported letter y; (x,x) the frequency of correctly reported letters x, etc. Subsequent steps in the analysis are concerned with the recursive selection of the highest cell frequency in the matrix and the combination of the corresponding objects into clusters. Each level of clustering reduces the number of objects by one. The same steps are repeated in the smaller matrix, and so on. The similarity of a newly formed cluster to any other object (or cluster) is defined as either the smaller (Minimum method) or the greater (Maximum method) of the two distances between that object (or cluster) and the two objects (or clusters) involved in the new clustering. This procedure implies that only a fraction of the original number of the matrix cells contributes to the clustering scheme.

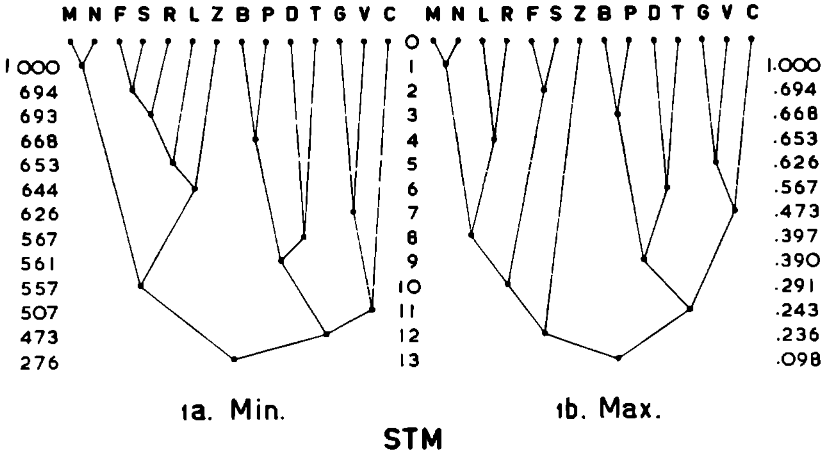


Figure 2.5. Hierarchical clustering schemes (HCSs) of confusions in the recall of visually presented letter sequences (STM).

An implication of the Minimum method is that new objects are added relatively easily to existing clusters, which results in long, internally 'connected' clusters, whereas the Maximum method tends to postpone the extension of existing clusters, which results in small 'compact' clusters.

The HCSs presented in Figures 2.5, 2.6, and 2.7 are drawn as trees. The order of clustering is downward from 0 to 13. The points where two branches join (into clusters) have a corresponding similarity value, derived from the transformed confusion matrix in which the largest cell frequency is given a value of 1 and the other cell entries are multiplied correspondingly. These

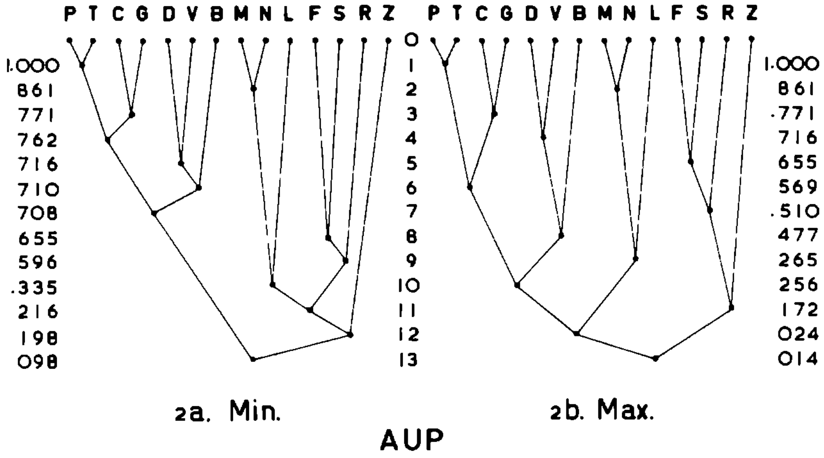


Figure 2.6. Hierarchical clustering schemes (HCSs) of confusions in the auditory perception of single letters (AUP).

values have no more than an order relation, which is expressed by equal vertical distances between the clustering levels in the trees. Although our main interest is in the compact Maximum solutions, there is no reason for omitting the Minimum solutions. Therefore, for all matrices both representations are presented and discussed.

As anounced above, we shall now discuss the visual perception (VP) HCSs and leave the STM and AUP HCSs till later in this section (2.4.4). If the discussion of VP confusions is limited to the main features of the HCSs 3a and 3b of Figure 2.7 the following descriptions may be given.

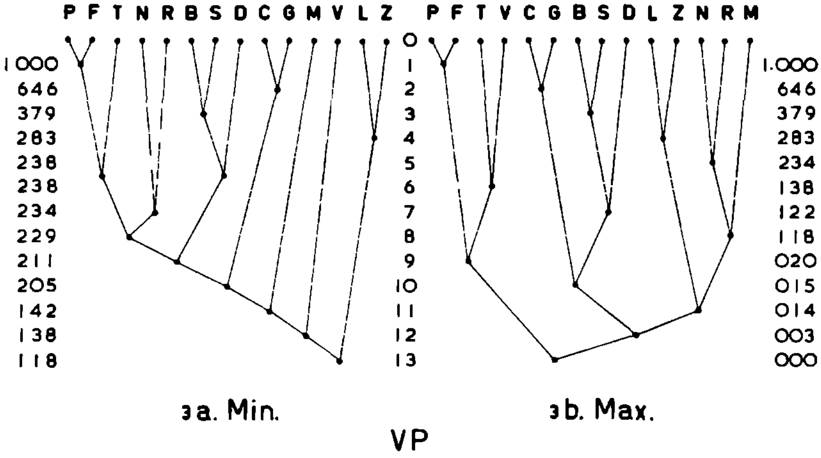


Figure 2.7. Hierarchical clustering schemes (HCSs) of confusions in the visual perception of single letters (VP).

Visual perception (VP): Minimum method (Fig. 2.7. a).

The larger clusters are formed between the single-legged letters (P,F,T) and the letters with right-hand bottom curves (B,S,D). Smaller clusters exist between letters with right-hand down strokes (N,R), left-hand curves (C,G), and flat bases (L,Z).

Visual perception (VP): Maximum method (Fig. 2.7. b).

The three strong clusters may be named (1) Narrow base, composed of an asymmetric (P,F) and a symmetric (T,V) cluster; (2) Round base, composed of left-hand curves (C,G) and right-hand bottom curves (B,S,D); and (3) Wide base, composed of flat (L,Z) and multi-legged (N,R,M) bases.

The most striking characteristic of the visual perception HCSs in the present context is that they do not resemble the STM and AUP HCSs in any respect. The obvious clustering within the 'acoustic' B and F-set in the latter HCSs is completely absent in VP. On the contrary, the HCSs for VP are readily described in terms of morphological features. An explanation of the fact that the lower half of the letters seems to be a stronger determinant of clustering than the upper half is not attempted here.

2.4.4 Distinctive feature analysis

Various systems for the classification of Dutch consonants have been proposed (e.g. Eijkman 1955; Roorda, as quoted in Kruisinga 1955; Cohen, Ebeling, Fokkema, and Van Holk 1961; De Groot 1963; Van Dongen 1962). These classifications differ mainly with respect to the number of dimensions used to describe all the consonants (some of which are not present in our letter-sets) and with respect to the number of values on each dimension. For the present purpose a classification which is in general agreement with the principal subdivisions in the above feature systems seems desirable. The feature system of Table 2.11 was chosen to describe the consonant phonemes of the letters used in the STM and listening tests.

Dimension	value	category	letters
1. voicing	0	voiceless	C G P T - F S
	1	voiced	B D V - L M N R Z
2. place of articulation	0	bilabials	B P - M
	1	labiodentals	V - F
	2	dentals	C D T - L N S Z
	3	velars	G - R
3. nasality	0	orals	B C D G P T V - F L R S Z
	1	nasals	- M N
4. affrication	0	plosives	B D P T -
	1	fricatives	C G V - F S Z
	2	others	- L M N R

Table 2.11. Distinctive feature system used in the analysis of confusions in STM and in Auditory perception.

The fact that the similarity among the letters both in STM and in AUP may be described in terms of the system of distinctive features adopted is illustrated in Figure 2.8, where the median number of confusions obtained in the two matrices is represented as a function of the number of dimensions on which the letters involved in such confusions differ. In this analysis the confusions within the B-set and within the F-set are combined. With one minor exception (which may be due to the very small number of observations at distance 4) both curves show a monotonous relation between probability of confusion and the number of dimensions on which a different value occurs. The latter may thus be regarded as a measure of the distance between the items (Hamming distance).

A comprehensive description of the HCSs presented above will now be attempted in terms of distinctive features. It is convenient to consider clustering within the B-set and the F-set separately. This is in agreement with the most striking aspect of the HCSs, namely that clustering between the elements of these two sets occurs only at the very highest levels.

Short-term memory (STM): Minimum method (Fig. 2.5. a).

B-set. At level 8 there are clusters of bilabial plosives, and labiodental plosives. These cluster at level 9 and to the resulting cluster are added at level 12 the fricatives. These do not form a cluster on the basis of voicing, since only V is voiced.

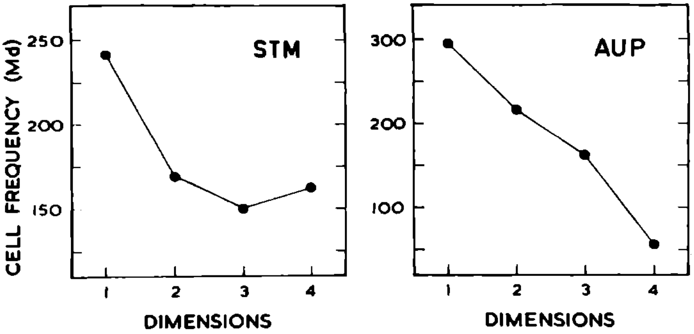


Figure 2.8. Median cell frequencies in the short-term memory (STM) and auditory perception (AUP) confusion matrices as a function of the number of dimensions on which the letters involved in confusions have a different value (Hamming distance).

F-set. At level 2 there are clusters of nasals and of voiceless fricatives. To the latter are added at subsequent levels the liquids L and R and the voiced fricative Z. At level 10 they combine with the nasals.

Short-term memory (STM): Maximum method (Fig. 2.5. b).

B-set. Exactly the same clusters are found as with the Minimum method, viz. bilabial plosives, labiodental plosives, and fricatives. The only difference is, perhaps, the relatively early clustering of the latter elements.

F-set. At level 4 there are again clusters of nasals and voiceless fricatives. The liquids now cluster before combining with the nasals and before the voiceless fricatives and the individual voiced fricative Z do so.

Auditory perception (AUP): Minimum method (Fig. 2.6. a).

B-set. The three main clusters are voiceless plosives, voiceless fricatives, and voiced plosive plus the voiced fricative V. These clusters are formed at the levels 1, 3, and 6 respectively. The voiced cluster is not primarily a cluster of plosives, because the voiced plosive D combines with the voiced fricative V earlier than with the other voiced plosive B.

F-set. At level 8 there are clusters of nasals and of voiceless fricatives. The liquids L and R are joined to these clusters at levels 10 and 9 respectively. Only after these two enlarged clusters are combined at level 11, is the individual voiced fricative Z added to the resulting cluster.

Auditory perception (AUP): Maximum method (Fig. 2.6. b).

B-set. Exactly the same clusters are obtained as with the Minimum method. Again there are voiceless plosives, voiceless fricatives, and voiced plosives plus the voiced fricative V. The clusters are now formed at the levels 1, 3, and 8 respectively.

F-set. At level 5 there are again clusters of nasals and of voiceless fricatives. The liquids L and R again combine with these clusters, now at levels 9 and 7 respectively. The voiced fricative Z joins with the latter cluster at level 10 and only after the nasals plus L cluster is combined with the whole of the B-set, at level 12, are the elements of the F-set all joined in one cluster. This is the only exception where clustering between sets occurs before all the elements within each set are completely clustered.

To summarize, there is a large correspondence between the hierarchical clustering schemes (HCSs) for confusions in the immediate recall of visually presented letter sequences (STM) on the one hand and in the perception of these letters when spoken individually against a background of white noise (AUP) on the other. Not only do both types of confusions show a stronger tendency of elements to cluster within than between the B and F-sets (which are defined in terms of the vowels accompanying the consonants) but there is also obvious agreement that clustering in all HCSs similarly occurs between certain plosives, fricatives, and nasals. The relative isolation of the letter Z (with its unique pronunciation) is also represented in all four HCSs. However, there are consistent differences between the STM and AUP clustering schemes. Firstly, the F-set as a whole is characterized in STM by a greater similarity between the elements than the B-set, while in AUP there is a greater tendency for B-set letters to cluster. Secondly, the liquids L and R show a greater similarity in STM than in AUP, where they cluster with nasals and fricatives respectively. Thirdly, in the B-set plosives and fricatives form exclusive clusters, whereas this is not the case in AUP: here voiceless and voiced clusters appear. Fourthly, in STM the plosives cluster on the basis of common place of articulation, while in AUP they join the basis of common voicing.

The fact that verbal coding and rehearsal determine the type of storage in STM implies that recall is based on the traces of these verbal operations in which, according to the model of Chapter 1, articulation is involved to a large extent. Thus the model requires that in the STM confusion matrix articulatory features are present. The problem is, however, that articulatory and acoustic features are difficult to differentiate because of the causal relationship between them. Comparison among certain dimensions of the feature system may nevertheless be considered as a means of disentangling the two. The voicing dimension has been shown by Miller and Nicely (1955) to be of great importance in the auditory perception of consonants. The authors found over a wide range of S/N ratios that this dimension carries a considerably larger amount of information than the place of articulation dimension. Moreover, for initial plosives the difference appears to be very great indeed: consonants of the set /p,t,k,b,d,g/ are confused in auditory perception much rather within the voiced and voiceless subsets /p, t, k/ and /b, d, g/ than within the subsets /p, b/, /t, d/, and /k, g/ with a common place of articulation.

But if voicing is an important feature in the perception of normal, vocalized speech and in the auditory images which correspond to it, this is not necessarily the case in forms of speech that - although quite intelligible - are characterized by incomplete vocalization. One such form is whispering. If confusions between plosives under normal speech and under whispered speech are compared, as in Table 2.12, it appears that the importance of the voicing dimension is sharply reduced in favour of the place dimension. If coding and rehearsal are considered as other forms of speech related to whispering perhaps as much as to normal speech, it is feasible that in these kinds of subvocal processing too, place of articulation provides more important cues than in normal listening and perhaps even relatively more than voicing. Without wishing to imply that the cues for STM recall are solely or primarily provided by kinestetic feedback from the peripheral speech organs the following observation may clarify the matter. Consonants may be 'formed' subvocally in such a manner that the kinesthetic and proprioceptive feedback from the speech musculature has informative significance; it is however much less likely that cues related to voicing can be derived from articulatory movements if these are not accompanied by actual vibrations of the vocal chords which are the main source of the difference between voiced and voiceless consonants as produced by a speaker.

Speech mode	Errors within classes with common voicing	place
	(b-d) and (p-t)	(b-p) and (d-t)
Normal	174	42
Whispered	12	559

Table 2.12. Confusions among plosives in a listening task. Phonetically balanced (PB) lists of Dutch words were read in a normal conversational manner or whispered. Forty per cent of the words were perceived correctly under both conditions (Compiled from Kruisinga 1955, Diagrams 1 and 2).

The above discussion clearly attributes a larger role to the place dimension in STM and to the voicing dimension in AUP. If we return to the raw matrices it appears that, indeed in STM relatively small frequencies occur in cells that represent confusions between letters with a different value on the place dimension, and that in AUP the same is true for confusions between letters with a different value on the voicing dimension. See Table 2.13. Testing by means of a Sign Test over the cells involved in the two matrices shows that the differences for place ($n= 32$) and for voicing ($n= 11$) are both significant at the .05 level. Table 2.13 also shows, however, that the differences are much greater in the B-set than in the F-set. STM confusions between letters of the B-set are relatively low if the letters involved have a different place of articulation; AUP confusions are relatively infrequent when different values on the voicing dimension are involved. A comparison within the set of plosives (B,D,P,T) analogous to that of Table 2.12 in which total confusion frequencies are given, confirms the picture of Table 2.13. In the AUP matrix 63 per cent of the confusions involving similarity on either the voicing or the place dimension is found in cells with a common voicing value (1665 out of 2640), whereas in the STM matrix this is only 45 per cent (826 out of 1836). This difference is significant ($\chi^2 = 6.52$; $df = 1$; $p < .025$).

The relatively greater weight of the place dimension in STM as compared to AUP can further be tested by determining the predictive power of the ordered relation between the four values on this feature dimension. On the basis of these values the distance between B and G is larger than between B and T, etc., so that more B-T confusions are predicted than B-G confusions. Within the B-set the complete predicted order of confusions with the letter B is then B-P > B-V > B-C = B-D = B-T > B-G. In this order 12 binary predictions are contained in terms of greater or smaller confusion frequencies with the letter B. The B-set as a whole allows 76 such predictions and the F-set 67. The predictions can be tested both over the rows and over the columns of the matrices, holding constant response bias and item difficulty respectively.

Matrix	place	p + v	voicing
STM	171	159	220
B-set	167	164	253
F-set	193	157	184
AUP	314	191	145
B-set	376	196	233
F-set	177	115	101

Table 2.13. Relative weight of place and voicing dimensions on STM and AUP confusions. Median cell frequencies were determined for confusions between letters with a different value on the place dimension, on the voicing dimension, and both (p + v). Rows 1 and 4 represent the medians of the within-set confusions in the B and F sets pooled.

The accuracy of these predictions may further be expressed as the proportion of the total number of predictions which is confirmed in each of the two matrices to be compared. These results are given in Table 2.14. They indicate that the order relation is more clearly present in STM than in AUP. The number of accurate predictions in each row of the STM matrix was paired with that of the same row of the AUP matrix. Over the 14 rows of the two matrices a Wilcoxon matched-pairs signed-ranks test was run; it established the significance of this difference ($N= 12$; $T= 13$; $p < .025$). Testing over columns yielded similar results ($N= 13$; $T= 16$; $p < .025$). Again, the difference between STM and AUP is considerably larger for the B-set than for the F-set.

Our suggestion that articulation plays a greater role in STM than in auditory perception is thus confirmed by the comparisons made between the place and voicing dimensions in the STM and AUP matrix. In summary, the evidence is the following. Raw cell frequencies indicate that place differences significantly reduce the probability of confusion in STM conditions, whereas voicing differences have a similar effect on auditory perception. An unexpected finding was that this effect is larger within the B-set than within the F-set. Total confusion frequencies among the plosives are, moreover, in complete agreement with these results. Furthermore, the four values on the place dimension are ordered to a significantly greater extent in STM than in auditory perception. Because in the other comparisons all 'different' values, irrespective of the size of the difference, were combined into a single class, the latter comparison may be considered to provide independent evidence. The evidence is also reasonably strong in view of the fact that the place values themselves cannot be expected to correlate more than loosely with the actual position of maximum constriction during articulation.

Matrix	Rows			Columns		
	B-set	F-set	B+F	B-set	F-set	B+F
STM	57	46	52	57	42	50
AUP	40	34	37	39	37	38

Table 2.14. Percentage of confirmed predictions based on an order relation between the four values on the feature dimension 'Place of articulation'.

2.4.5 On distinguishing auditory from articulatory representation

The arguments quoted on the preceding pages in support of the suggestion that in STM articulatory cues are a major source of information imply that STM is characterized by a specific manner of coding which is different from acoustic coding in auditory perception. The difference, moreover, is in agreement with the articulation involved in the coding and rehearsal processes as postulated in the model of Chapter 1. There are also other sources of evidence in agreement with our own data obtained in STM experiments with English and American Ss. A Kruskal analysis performed on the confusion matrices reported by Conrad (1964) indicates that there are differences with respect to voicing and place exactly parallel to those discussed above (Thomassen 1966b). From confusion data collected by Hintzman (1967) who studied recall of the initial plosive consonants in visually presented CVC trigrams, it also appears that place of articulation is a greater determinant of confusions in STM than in Miller and Nicely's (1955) perception data. Finally there are the results of the consonant experiments by Wickelgren (1966e) already mentioned. If the results over his two experiments are combined (as in his Table VII) it may be concluded that, although presentation was auditory, the place dimension gives a somewhat better description of STM confusions than the voicing dimension (76 vs. 64 per cent accurate binary predictions) while the reverse is the case when the same predictions are tested in Miller and Nicely's (1955) perception data (65 vs. 100 per cent).

The contrasts observed in the former paragraph must be considered with some reservation, however. Firstly there is the use of noise as a means of collecting perceptual confusions. A relatively low S/N ratio is required to attain sufficiently large numbers of observations in all the cells of the confusion matrix. This is better achieved by adding noise than by using very low intensity levels. The relative amount of noise, as well as its frequency characteristics may, however, determine the type of confusions observed. The results of Miller and Nicely (1955, Table XXI and Figures 1, 3, 5) may serve as examples. As stated above, the amount of transmitted information for place is below that for voicing at all S/N ratios but the rate of increase with increasing S/N ratios is very different for the two dimensions. The greatest gain increase for voicing occurs between -18 db and -6 db, whereas the greatest gain increase for place seems to occur above +6 db. These relations are reflected by maximal differences between voicing and place confusions at -6 db and 0 db. Miller and Nicely's filtering data give an impression of what non-white noise would do to the type of perceptual confusions. Low-pass filtering (which would have an effect similar to high-frequency noise) is far less detrimental to voicing than it is to place. High-pass filtering (or low-frequency noise) has, however, an almost identical effect on both the voicing and the place dimension.

These considerations certainly moderate the significance of quantitave comparisons such as those made above; the exact ratio of voicing over place confusions within an auditory perception matrix depend on amount and type of noise and if the frequency of either of these is used in comparison with STM confusions it must be realized that they do not represent 'the' auditory perception characteristics. On the other hand, these considerations do not invalidate the general finding that over a very large range of conditions voicing is a greater source of information than place in the auditory perception of consonants. This is shown by all the available data, especially those of Miller and Nicely (1955) but also by those of Conrad (1964) when subjected to a Kruskal analysis (Thomassen 1966b) and by the AUP data reported above. In contrast, a greater weight of place of articulation is present in the STM data after visual presentation as reported above, as well as in the confusions obtained by Hintzman (1967) and after auditory presentation as collected by Wickelgren (1966e).

A second point to be made is that comparison between errors in STM and in auditory perception involves a specific problem with respect to the Ss providing the data. In the model of Chapter 1 it is suggested that S's recall after coding or rehearsal is based on scanning the coding or rehearsal output, i.e. on a 'sensory' process concerned with the feedback of S's own verbal activity. The best comparison would thus be achieved is S's perceptual errors are compared with his own STM errors, and - which is equally essential - if the perception task involves S's identification of verbal items spoken by himself. This degree of perfection has a great procedural drawback, however, because of the large number of observations required from each S in individual sessions. In the preceding paragraphs only group comparisons were made and although large differences between the speakers and listeners in the auditory perception task and the Ss in the STM experiments were avoided, the groups were not identical. This may have influenced the results in the sense that, owing to regional differences certain values in the feature system adopted give a better description of the speech habits of one group (e.g. listeners) than of another (e.g. STM Ss). An example may be the letter R which is known to have as many as five different pronunciations in Dutch (Damsteegt 1969). Our feature system describes its consonant phoneme as a velar /R/, which is probably valid for most of the listeners, but for the major part of the Ss in the STM experiments a dental /r/ may have been a better description. The HCSs do show a closer similarity between the (dental) L and R in the case of STM than in AUP. In this case our discussion of the STM data would, therefore, have underestimated the place factor.

A third, and more fundamental problem related to the distinction between cues in STM and in auditory perception is concerned with articulation as a factor involved in the perception of speech. If it was found on the one hand that articulation plays a larger role in STM than in auditory perception, but if it also became clear on the other hand that the cues of 'the' auditory perception cannot be studied completely by using noise, the possibility that under different perceptual conditions the auditory coding system is equivalent to the STM coding system is still not ruled out. In that case the articulation cues involved in perception would be relatively more impaired than the 'pure' acoustic cues, and all the perception data discussed so far would reflect this asymmetry. In fact there is a certain class of evidence in the literature on speech perception which attributes to articulation a necessary role in the perception of auditorily presented verbal items. Thus, the perception of consonants, whose acoustic correlates vary largely with the vowel environment, is described more easily in articulatory than in acoustic terms (Liberman, Cooper , Shankweiler, and Studdert-Kennedy 1967). This approach has actually referred to place of articulation as an important cue in the perception of consonants (Liberman, Delattre, and Cooper 1958). We shall discuss the 'active' interpretation of speech perception in the first section of Chapter 3. We may now restrict ourselves to the observation that the extent to which the active interpretation is correct will determine the difficulty of discriminating between the cues involved in auditory perception and in STM. If indeed articulation mediates all perception of verbal items, findings which merely indicate the role of articulation in STM have no other value than to demonstrate that the same rather than different coding occurs in STM and in auditory perception.

CHAPTER THREE

Chapter 3

Towards a model for the representation of verbal items in short-term memory

After having discussed in the preceding chapter the articulatory and sequential characteristics of verbal STM performance, as dealing with units of speech, we shall now look at verbal STM against the background of verbal processing as such. It has been made clear that verbal STM belongs to the sphere of sensory and motor speech behaviour rather than to the region of short-term retention of non-verbal sensory and motor forms of behaviour. The present chapter will, correspondingly, start with a discussion of speech perception models in which a fundamental function is ascribed to articulation. These models may, perhaps, indicate an approach leading to the integration of the study of immediate memory with the articulatory processes involved and the study of perceptual and motor events occurring respectively upon presentation of verbal information and at the production of speech. It is not at all clear, at present, what the exact relation is between speech perception, speech production, and short-term retention of verbal information. Of course, they all should be accounted for in a final model of verbal processing. The role of STM in such a model would be to describe the storage involved in delayed perceptual decisions, as discussed in Chapter 1, of verbal sequences with largely varying size, but also of specific STM behaviour as described in the last chapter, and, finally, also the storage involved in the delay between anticipated verbal response by a speaker and the ultimate production of an overt response. The problems related to the study of verbal processing - apart from the meaning components involved - may all be related to the single question: How are verbal items internally represented? Secondary questions would then be concerned with how these representations are activated during perception and production of speech, and how during retention in STM tasks their level of activation is maintained over a period of seconds. It seems that we are still far from such an approach. The present chapter, by dicussing evidence for articulation in speech perception and relating it to that in STM, and by presenting and testing an extended version of the preliminary model of Chapter 1, in which to some extent perceptual factors are incorporated, may constitute a modest step in the intended direction.

3.1 Auditory and articulatory processes in speech perception

In the preceding chapter arguments were forwarded for articulation in short-term retention. In the last section (2.4) these arguments were derived from comparisons with auditory perception. The section was closed (2.4.5) however, with some notes indicating the difficulty of tracing the cues of 'pure' auditory perception. The possibility that the representation of verbal material in STM is auditory cannot be excluded completely as long as there is no decisive method to distinguish between the cues of articulation and of audition. Moreover, the superiority of STM for auditory patterns as discussed in the previous chapter (2.2) seems to show that acoustic representation would fulfil the requirements of an adequate STM performance. However, this acoustic superiority must not be unduly generalized, firstly because with verbal items differences from visual presentation only appeared to be moderate and confined to the last items of the sequence; secondly, and more importantly, because large differences exist between speech and non-speech material. There is an impressive amount of evidence, also neurologically, that auditory speech perception is mediated by mechanisms different from those for tones and non-speech in general. The former are situated in the speech-dominant (usually left) hemisphere, the latter in the other hemisphere. The predominantly contralateral connections between the ear and the cortical speech areas are generally responsible for superior speech perception and immediate recall if presentation is to the right ear; listening through the left ear is usually better with non-verbal stimulation (Kimura 1961a; 1961b; 1964; Broadbent and Gregory 1964; Bartz, Satz, and Fennel 1967; Webster and Chaney 1967; Bryden 1969). These results make the drawing of parallels between processing of speech and non-speech very dubious. In fact they indicate that auditory perception of verbal material belongs to the region of speech. It is thus related to articulation as much as to auditory processing. The existence of a certain class of evidence supporting an articulatory interpretation of speech perception has also been mentioned in the last chapter (2.4.5). In the present section we shall review some of this evidence. If it appears that articulation of auditorily presented material does not only occur at coding and rehearsal as proposed in the model of Chapter 1, but also beforehand, preceding recognition as its necessary requirement, the role of articulation would be different from and more fundamental than that proposed.

3.1.1 Active interpretation of speech perception

If it is assumed that phonemes are the segments of speech as it is spoken and perceived, various difficulties arise if one also wishes to assume that the auditory perceptual system follows the acoustic signal segment-by-segment in tasks which involve the perception of speech at normal conversational speed. In the first place, there is the problem that, on the one hand, a single phoneme may result in a variety of different acoustic signals, while, on the other hand, any one acoustic signal may represent several phonemes. This has been shown especially with respect to the influence exerted upon the acoustic cues of a phoneme by the presence of a certain other phoneme following it (Liberman, Delattre, and Cooper 1952; Liberman 1957). In other words, "... for many of the important consonants there is no way to define the acoustic cues so as to have, except in a small number of phonetic contexts, an invariance between *acoustic cue and phonemic perception*" (*Liberman, Cooper, Harris, MacNeilage,* and Studdert-Kennedy 1967, p. 76). The second problem is, that in speech perception the listener can deal with some 10 to 20 phonemes per second, which is much faster than the maximum rate established for the auditory discrimination and perception of non-speech signals, even if these are highly distinct. Furthermore, the difficulties in machine recognition of speech and in speech synthesis, if these are based on segments of phonemic length (Harris 1953; Liberman, Ingemann, Lisker, Delattre, and Cooper 1959) cast serious doubt on the possibility that phonemes are the segments on which speech perception is based.

Liberman and his associates suggest that the successive phonemes of a string of e.g. word length have overlapping acoustic cues and are encoded into units of approximately syllabic size. Although this suggestion copes with the above problems, it leaves open the question of how individual phonemes are perceived in the acoustic stream of speech in which they are not present as segments. It is this question which Liberman's motor theory attempts to answer. The core of this theory is, always somewhat vaguely formulated, that 'the sounds of speech are somehow perceived by reference to the way they are generated' (Liberman et al. 1952, Liberman 1957, Liberman, Cooper, Harris, and MacNeilage, 1962; Liberman et al. 1967). It is assumed that in the speaker's CNS each phoneme of his language is represented by a specific neural event. At this level there is a one-to-one relation between the phonemes of the language and the speaker's neural activity. However, when a string of phonemes is going to be spoken, the commands from the CNS to the articulatory muscles will overlap in time, and there will be interactions between the activated muscles. The result is the complex relation which is observed between the acoustic signals and the spoken phonemes they represent. The listener, on his part, is as a potential speaker assumed to possess the same mechanism as is used by the speaker for 'putting the segments through the successive recodings that result eventually in the acoustic signal'. The recovery by the listener of the simple one-to-one relation between acoustic signal and phoneme is made possible by this mechanism now used as a decoder. Thus, speech perception would be based on an articulatory motor mechanism rather than on a sensory system processing the acoustic signals in a number of auditory steps.

The evidence quoted in support of Liberman's motor theory of speech perception may be summarized as follows. First, the perception of speech signals seems to be different form that of non-speech signals. Not only is there the ear asymmetry to which we referred above, but there is also a dichotomy in the listener's perception when a synthetic acoustic signal is continuously varied in such a way that it moves across the boundary between speech and non-speech The acoustic continuum apparently does not have its counterpart in perception. A signal heard as speech is, moreover, easier discriminated from other signals and it is better learned in a PA task (House, Stevens, Sandel, and Arnold 1962). Second, the perception of synthetic speechlike signals often seems to be categorial. If such signals are continuously varied along one dimension (e g. on the common dimension for /b,d,g/ as voiced stop consonants), the subject tends to hear the signals as either /b/ or /d/ or /g/, rather than as intermediate sounds somewhere in between any two of the consonant phonemes. At the category (phoneme) boundaries, perception seems to make a 'quantal jump'. Similarly, two signals that are separated by a certain distance on one dimension are much easier to discriminate if they are located on either side of a phoneme boundary than if they are both located at the same side of such a boundary. Such increased discriminability at certain positions on relevant dimensions has repeatedly been shown, especially in the case of voiced and voiceless stop consonants (Liberman, Harris, Hoffman, and Griffith 1957; Bastian, Eimas, and Liberman 1961; Liberman, Harris, Kinney, and Lane 1961; Liberman et al. 1962; Eimas 1963). No such peaks in discriminability have been found between non-speech acoustic signals if these are similarly varied along the same dimension as the one that produces a 'quantal jump' in speech signals (Bastian et al. 1961).

Evidence thus far is in support of two different modes of perception, one within and one outside the range of signals that are perceived as speech. In the case of vowels, however, perception is again continuous like that of non-speech signals. There are no 'quantal jumps' in vowel perception, and the subject can discriminate many more vowellike signals than he can identify (Fry, Abramson, Eimas, and Liberman 1962, Stevens, Öhman, Studdert-Kennedy, and Liberman 1964). These results seem to suggest that isolated vowels have non-speech properties, which is probably in agreement with their tonal characteristics on the one hand, and with their articulatory characteristics on the other. different vowel phonemes are mediated by groups of related muscles which operate in different degrees, whereas different muscle groups are involved in the case of different consonants (Liberman, Cooper, Harris, MacNeilage, and Studdert-Kennedy 1967). It is of great interest to note here that the ear asymmetry with respect to speech perception has not been established for vowel perception (Shankweiler and Studdert-Kennedy 1967). A third source of evidence is provided by EMG recordings. The complex relation between the spoken phonemes and the actual acoustic signals produced by the speaker formed the background for the research reviewed above. Liberman's model, however, requires that, at some level, there is an invariant relation. Several attempts to demonstrate such invariance at the level of motor commands to the speaker's articulatory muscles have indeed been reported. Certain EMG tracings are specific for certain phonemes, irrespective of their phonetic context formed by other, temporally and spatially overlapping phonemes (Harris, Schvey, and Lysaught 1962, Harris 1963, MacNeilage 1963, 1964).

A model of speech perception which is essentially based on the same principles as Liberman's motor theory of speech perception - and which is supported by the same type of evidence - has been proposed by Halle and Stevens (Halle and Stevens 1959, 1962, Stevens 1960; Stevens and Halle 1967). Their analysis-by-synthesis model describes in some detail one possible way in which speech perception may be achieved through its relation to speech production The main assumptions are that a preliminary analysis of the acoustic signal enables the listener to make a hypothesis about the speech units uttered by the speaker. The hypothesized unit sequence is then actively produced by the listener and the corresponding (virtual) auditory pattern is tested against the received acoustic signal. If there is a match, the signal is perceived and processed further; if not, a new hypothesis is generated, and so on.

In the terminology used by Stevens and Halle (1967), both speaker and listener use the same representations of the lexical items (e g words) of the language, which are stored in memory, and the same rules for transformation. The representations $\underline{P}$ are abstract, in terms of segments (e g phonemes) which themselves are complexes of distinctive features. Phonological rules transform the abstract representations P into sets of instructions $\underline{V}$ for articulation. The resulting acoustic output S during speech production does not directly reflect the abstract representations, because S is also a function of the phonological rules and of the dynamics of the vocal tract In speech perception the same rules are used by the listener to decode the acoustic signal $\underline{S}'$ into the corresponding abstract representation $\underline{P}$. The acoustic signal $\underline{S}'$ is first transformed into an auditory pattern $\underline{A}$, which is entered into the analysis-by-synthesis mechanism Here, $\underline{A}$ is subject to a preliminary analysis resulting in a hypothesis concerning the abstract representation $\underline{P}_{\underline{trial}}$. The phonological rules operate on $\underline{P}_{\underline{trial}}$ to yield an articulatory pattern $\underline{V}_{\underline{trial}}$ During speech perception, actual articulation is, however, suppressed and, instead, an equivalent auditory pattern $\underline{A}_{\underline{trial}}$ is derived from $\underline{V}_{\underline{trial}}$. $\underline{A}_{\underline{trial}}$ is then fed into a comparator and matched against the original auditory pattern A, which had been temporarily stored for later comparison. If, by some criterion, $\underline{A}_{\underline{trial}}$ and $\underline{A}$ are in agreement, the hypothesized speech unit was correct, and $\underline{P}_{\underline{trial}}$ is established as $\underline{P}$. If there is no agreement between the patterns, the control component generates new hypothetical sequences until one is accepted by the comparator as a match.

Although in Liberman's interpretation the perception criterion is articulatory, while in the latter model the ultimate comparison is on acoustic characteristics, the same classes of evidence are regarded affirmative with respect to the analysis-by-synthesis model as the ones that are claimed to support the motor theory of speech perception. These were, in short, (a) perception of speech and non-speech are different, (b) the perception of consonant phonemes is categorial, (c) motor innervation of speech shows greater invariance in relation to spoken phonemes than the acoustic signal does. Regarding speech production, Stevens and Halle (1967) also point to the fact that cineradiographic data (X-ray motion pictures) of voci-motor behaviour demonstrate that "in the actual speech events the discreteness of phonetic segments and features is blurred or totally obliterated". If this observation is taken together with the findings related to $\underline{c}$, the evidence seems to suggest, in accordance with the model, a decreasing invariance between spoken phonemes and the following order of events: (1) the neural commands to the articulatory muscles, (2) the articulatory movements produced by these muscles: (3) the acoustic signal resulting from articulation. Extrapolation to a zero-stage seems to have some plausibility if this decreasing invariance is not merely due to the methodological aspects of measuring output. The zero-stage, then, would be the 'highest', abstract, level of internal representation where there is perfect invariance with respect to the spoken phonemes.

On the speech perception side, there is also reference to the capacity of understanding a wide range of different dialects by anyone who speaks just one dialect. If it is assumed that the utterances are identified by means of classification of the successive segments of the acoustic signal, this capacity would remain unexplained because the dialectal differences often are precisely differences in the inventory of the speech sounds. "If, on the other hand, we assume that dialectal differences in the sounds are due to the fact that a given abstract representation of a speech event is actualized in accordance with different phonological rules, then the performance of the normal speaker becomes at once understandable. Having listened to a relatively small sample of utterances in a dialect different from his own, the speaker of a language is evidently able to determine modifications of a few phonological rules of the dialect as compared with those in his own dialect. He is then able to utilize these rules to identify correctly combinations of elements or words he has never heard before in that dialect" (Stevens and Halle 1967, pp. 94-95).

3.1.2 Critical notes on the active interpretation

To the above 'active' interpretation of speech perception a number of serious, and partly obvious, objections may be made. Firstly, if the preliminary auditory analysis provides all the necessary information for the articulatory mechanism to generate the articulatory or auditory signal which is then recognized or matched, the auditory system may just as well be capable of 'passively' decoding the auditory input directly, without reference to articulation. If the first hypothesis results in a bad match, the error signal fed back to the control component must contain sufficient information for a considerable improvement at the next trial. It is unlikely that more than a few hypotheses can thus be tested before the stored raw input has decayed too far for its use by the comparator. This limiting factor will apply most strongly to verbal items spoken in isolation because no context is given to reduce the number of alternatives. In normal speech the context provides ample opportunity for alternative reduction but here the limiting factor is formed by speed requirements set by the task of understanding normal or rapid speech.

Another objection is that Ss must be assumed to rely on learning with respect both to the verbal elements they listen for and to the responses they have available for the discrimination between or for the production of the perceived verbal units Even if the instruction is to listen for sound qualities Ss will tend to listen for familiar speech units, i.e. phonemes Similarly, in discriminating between (synthetic) speech sounds each member of the class of stimuli to which a certain phoneme applies is identified by that response and it is differentiated from the members of other classes more easily than from other members of its own class. This interpretation is in terms of 'codability' rather than perception, and analogous to the finding by Brown and Lenneberg (1954) who reported better recognition of colours that were named more readily and consistently by the Ss, the efficiency of coding being determined by the S's experience in his culture (Lenneberg 1961) This point is clearly related to our discussion on the coding response (2.1.3).

A third argument against analysis-by-synthesis in speech perception is concerned with findings in the experimental work on attention. If S is presented two speech messages simultaneously and instructed to attend to only one of them, he will generally succeed in this task But occasionally words from the discarded message (the unattended channel) will be perceived (e.g. Treisman 1964) If there is active articulatory matching of the attended message it is unlikely that a parallel matching programme is set up for the neglected message, such dual processing is extremely unlikely in view of the limitations discussed above.

To these notes, which in some form or other have been made in various contexts in the literature the comments by Fant (1967) should be added. His argument is that, though it is difficult to prove the motor theory either right or wrong, it is inappropriate to abandon prematurely the search for acoustic cues in auditory processing. Fant proposes a model in which motor and sensory mechanisms are separated at first as auditory patterns on the one hand and motor patterns as subsidiary to these on the other. But they become more and more involved as processing proceeds from more peripheral to more central stages (from smaller segments to larger wholes). The message as conceived would involve mechanisms related both to the perception and to the production of language. Analysis-by-synthesis is considered unlikely by Fant at levels lower than phonemes, only at the word level and higher, he assumes a 'running prediction' of the most probable continuation of the message followed by a check against the sensory input. At this level the role of active predictions, mediated by motor patterns, is, however, merely to ensure "that the listener hears what he expects to hear; but this is not the same as stating that one hears only what one can say" (p. 115). At the lower levels distinctive features are assumed on the sensory and on the motor side. But not mechanisms that distinguish in an identical manner e.g. between the voiceless /p/ and voiced /b/ as between the voiceless /s/ and voiced /z/ at the arrival of the consonant phoneme segment. On the contrary, voicing of a consonant phoneme is reflected by " a relatively greater amount of voicing in the sound substance within as much as two connected syllables" (p. 118). Similarly, vowels preceding (and also following) a nasal segment appear to be nasalized also. Fant thus draws attention to the dynamic aspects of auditory speech perception and he points at the temporal contrasts present in the acoustic signal at transitions between phonemes rather than at the cues provided by the individual phonemic segments. Finally, Fant (1967) presents a number of spectograms demonstrating the fact that formant patterns of the acoustic signal of a stop consonant is indeed not invariant (Liberman 1952) but affected by the following vowel. The relationship is, however, not so specific that only an articulation interpretation would be capable of explaining invariance in the perception of the consonants. The formant pattern of spectral energy of the succeeding vowel simply seems to assimilate that of the preceding consonant, so that 'knowledge' of such rules and their specific generalizations by the auditory system does not seem unfeasible; such simple rules would suffice to perceive 'passively' the consonants with constancy in a number of different vowel contexts, e.g. /g/ in /gu/ and /gi/, and similarly /k/ in /ku/ and /ki/.

Summing up the evidence on articulation in speech perception does not yield, in our opinion, sufficient cause to presume that it plays an indispensable role as a prerequisite for perception. That articulation occurs as an accompanying effect with a supporting function, depending on the task requirements, will have to be accepted most probably in those cases where unusual performance is required of S, such as indicating or reproducing a perceived phoneme, or recalling several verbal items such as in STM experiments. But even in these cases articulation merely seems to serve a supporting function, which in fact will be hard to differentiate from the coding function we ascribed to articulation in the model of Chapter 1. This support then comes from articulation as the efferent aspect of a speech processing mechanism which provides the perceptual categories in terms of perceptual responses. A system with 'abstract' representations of verbal codes for speech segments of varying size may be considered feasible. Such codes may then be activated after specific sensory events have occurred at the level (segmental size) scanned. This level is dependent on the task. Activation of verbal codes may always imply articualtion, but its relative weight in supporting 'perception' or higher processing may equally vary with the task

It is thus an open question whether the perception of individual phonemes (which is possible through context, as Fant indicated, rather than in spite of context) may be counted as a normal activity of the speech perception mechanism. We saw that the process of speech perception is characterized by a great flexibility in the sense that there are ample opportunities for preperceptual storage and delayed decisions In this flexibility valuable possibilities must be assumed for perception in much larger units than phonemes, larger even than words. If and when articulatory mechanisms are involved in coding at perception (such as in STM tasks) and if perception is concerned with larger units than phonemes, it must be considered unlikely that (the same) articulatory mechanisms also deal with individual phonemes etc. at lower (and therefore preperceptual) levels

3.2 Morton's interpretation of verbal processing

Although it became clear in the preceding section that the perception of speech and further processing of speech stimuli, or of verbal items in general are not identical, it was also indicated that they cannot really be studied independently The model of Chapter 1, accordingly, includes in STM the sensory registration which forms the basis for recognition. However, to our knowledge there are no models available in the literature in which both perception and further STM processing and recall of verbal items are explicitly accounted for. Two models that to some extent indicate the conditions for perception while also describing a possible relation between perception and (re)production from immediate memory have been advanced by Norman (1968) and by Morton (1968a). The former model is in fact a theory on attention and the latter a model on language behaviour also covering the understanding and spontaneous production of grammatical sentences. But both models contain a section which is closely related to our present question. what are the properties of the representations of verbal items in perception and (re)production of speech? Both models, too, are similar in assuming a permanent storage system containing the representations of verbal items which both by sensory information and by context ('pertinence') information may be activated. Although the models differ in some respects (e.g the relation between different compartments in memory), these differences "are almost trivial and certainly reconcilable" (Morton 1969a, p. 177). We shall briefly discuss Morton's model below because it is most explicitly concerned with verbal items and because it has been applied to experimental conditions comparable to those of some of the experiments to be reported in a later section in the present chapter.

The functional model proposed by Morton (1964a, 1964b, 1964c, 1968a, 1969a; Morton and Broadbent 1967) has as its basic characteristic a set of 'logogens'. Each logogen corresponds to a certain verbal response, e.g a word or a letter name. Sensory (auditory or visual) and contextual information contribute cumulatively to the level of excitation of a logogen If a threshold value is attained in a logogen the appropriate response becomes 'available' in the sense that it may be uttered without further sensory or contextual evidence The logogen is thus defined (Morton 1968a) as the unit at which all information relevant to a certain verbal response converges, regardless of the source of the information and from which the response is made available. A logogen contains no information as to the kind of evidence causing its excitation. The availability of a response implies the presence also of the logogen's semantic meaning. Sensory contribution to the excitation level decays very rapidly, contextual evidence is somewhat more stable. Once a response has been made available, a temporary decline of the threshold of the logogen is the result. Logogens corresponding to high-frequency words thus permanently have lower thresholds. Both the presence of contextual information and of a lowered threshold have the effect that less stimulus information is required for a response to become available. Available responses are fed into an output buffer which may or may not lead to an overt response. The 'perceptual' system involved in the accumulation of evidence in one logogen rather than in another is not specified in Morton's model. It is claimed, however, that the system is a passive one based on filtering rather than on active patterning as in the interpretations discussed above. Each logogen has its own present level of excitation and its own threshold value so that decisions are only made within logogens. The presentation of one single stimulus may result in the availability of several different responses, but since the exit from the logogen system into the output buffer is single channel, the first response to become available will be entered first.

Morton's model has not been devised in the first place to account for

STM phenomena, but its relation to STM has been briefly indicated e.g. in the 1968a and 1969a papers. Available responses entering the output buffer are coded in a form related to a series of instructions for the articulation of the word. On exit from the output buffer the articulatory code is translated into a response code. This code may again be entered into the logogen system, which constitutes a rehearsal cycle. The output buffer is related by Morton (1969a) to primary memory (a term used by William James and later adopted by Waugh and Norman (1965) to denote the immediate form of memory for events which since their occurrence have not left S's attention and which need not be retrieved in the proper sense). Events in the logogen system, both the increasing excitation of the logogens and the effects upon the threshold values, are regarded by Morton as two sources of information for secondary memory (i.e. the psychological past) following primary memory. Morton further states that the output buffer is reflected by the eye-voice span and the ear-voice span and he suggests that the material within it is coded in terms of articulation parameters.

These relationships between Morton's model and STM raise some difficulties, however. Firstly, as we saw above, it is not clear how a certain logogen is informed that appropriate information has arrived for its excitation level to increase. Secondly, accumulating evidence preceding the availability of a response is entered into the logogen system in secondary memory before it could ever reach the output buffer which is in Norman's and other models the entrance into secondary memory. The problem is that one system stores both current (preperceptual) sensory information and the (postperceptual) aftereffects of earlier activations of the logogen concerned. These kinds of information, however, have a different input and must be kept apart. Thirdly, only to the extent that there is no perceptual delay can eye-voice and ear-voice spans be assumed to reflect the output buffer. The model as a whole, however, is similar to the one we presented in the first chapter. The most important discrepancies are that in Morton's model the single notion 'logogen' is used for the set of separate boxes in our model which are labeled 'verbal code, semantic networks, and articulatory patterns', and that in Morton's model entrance into STM is passive, i.e. without even an active scanning process as in our model. Of course, context variables and threshold lowering may lead to the same selective result as scanning, but not necessarily, especially not if verbal items are presented for 'non-verbal' processing.

In a series of experiments reported recently Morton (1969b) applies details of his model to card sorting. The cards employed contained different numbers of symbols which could - with varying degrees of effort - be verbalized. The general finding was that sorting by the number of symbols is interfered with by the verbal nature of the symbols themselves. It is suggested that the number sorting response is verbally mediated and that there is interference at the exit of the logogen system where the verbal number response is set up, because the verbal response to the symbol is also set up in the logogen system in an autonomous fashion immediately upon perception by S. It is also claimed by Morton that there is 'equivalence of channels' which implies that the appropriate number logogen is activated autonomously, just as the symbol name is activated autonomously. The evidence is based upon the fact that the presence of more identical symbols on a card resulted in a decreased rate of sorting by the symbol (but better tachistoscopic identification of the symbol), as compared to a single symbol (Exp. IX) and that, similarly, sorting by the number of rows on a card was interfered with by the (irrelevant) number of columns on the same card (Exp. XI). These findings may indeed support the notion of interference due to the near-simultaneous transfer of two responses to the output buffer, but the experiments do not look into the specific characteristic of the logogens, namely that they may be activated -simultaneously- by different sources of information. In fact, the above results on interference may also be explained by 'increased verbal stimulation' unrelated to logogens and their critical level of activation. If there is equivalence of channels there should be a faster build-up of evidence within a logogen if both channels (dimensions) carry information to the same logogen. Under these conditions facilitation should occur. This might be achieved e.g. by introducing 'compatible' cards with numbers of digits corresponding to the digits themselves. A second problem is that if entrance into the output buffer is serial, a minimal difference in time between the availability of two responses is sufficient to avoid interference. Thus, if e.g. one digit out of a number of identical digits is recognized only slightly before the number of digits present on the card is perceived with sufficient accuracy for a numeric response to become available, there should, on Morton's model, be no output interference. To these problems we shall return in the experimental sections of the present chapter. They are all concerned with the properties of logogens and their excitation i.e. with the internal representations of verbal items and the processes leading to their (overt) production or recall. As indicated above, the two stores, one permanent logogen system and one temporary output buffer, roughly correspond to the verbal code store and STS respectively in our model as presented in Chapter 1, so that the general outline of the models is sufficiently similar for a closer comparison with respect to the issues mentioned.

3.3 A model for the representation of verbal items in STM

Before making such a comparison we shall first look into some of the characteristics of the model as presented in Chapter 1 and be more explicit on several of its aspects, also in view of the findings reported in Chapter 2. There is reason to refine the preliminary model in several aspects. Thus, it is not clear in the model as presented why exactly homogeneous sequences (Exp. 1) and how especially a better order retention is achieved under heterogeneous conditions (Exp. 2). Also, it is not possible according to the model that visually presented items are processed under certain conditions along non-acoustic lines (Exp. 3; Exp. 4). The model is not explicit on the assimilation of non-verbal information (Exp. 5) but the data seem to show that the principle of recoding applies to successively presented sequences of flashes as much as to the other kinds of 'non-verbal' items discussed in the first section of Chapter 2 (2.1.3). Finally, it appeared that articulation provides some information to the deaf, though not to the same degree as to the hearing; a question that remains is why especially order cues seem to be inadequate (Exp. 6). On the following pages an adapted version of the model of Chapter 1 will be presented in which these points will receive special attention. Certain details of the 'extended' model will be tested in the experimental sections of the present chapter.

3.3.1 Characteristics of the model

The basic outline of the flow diagram presented in Figure 3.1 is the model of Chapter 1 (Figure 1.1). Where no alterations are discussed, the properties of this preliminary model also apply to the present model. A number of additions have been adopted and several further specifications will be given in order to describe the data from the previous experiments more accurately and to provide a framework for further experiments which, in part, intend to differentiate between the model by Morton as discussed above and our own conception. Firstly, a scanning mechanism and a template store

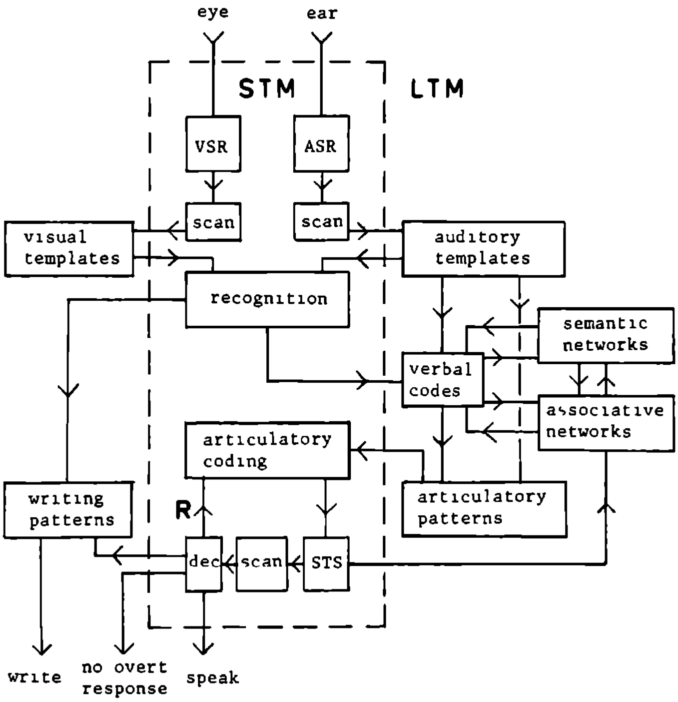


Figure 3.1. Model for the representation of verbal items in STM

have now been added to each of the 'peripherous' channels preceding recognition. It was implicit in the preliminary model already that scanning of certain forms of information from the total amount of stimulation present had to be assumed as an analogy of 'attention' or 'filtering' such as have been demonstrated e.g. by Treisman (1964). Scanning implies a directed activity which in visual perception is selective with respect to certain parts of the visual field or with respect to items belonging to a certain class (e.g. digits rather than letters). Similarly, in auditory perception there may be selective scanning of items e g. with certain physical properties or with respect to items having e.g. the characteristics of words or phonemes, that are then separated out from the acoustic context which, physically, may hardly allow their differentiation. These selection processes require the presence of models or templates which, on the one hand monitor the scan and which, on the other hand form the basis for recognition. These templates have purely sensory characteristics, the output of the scan, which is also sensory, is matched against these templates. By assuming (visual and) auditory template stores we have given preference to a speech perception model in which the criterion is auditory-sensory rather than articulatory-motor. It is further assumed that different template stores exist for different classes of stimuli including non-verbal classes. Thus there are visual template stores e.g. for letters, words, colours, and various familiar shapes. Auditory template stores exist e.g. for phonemes, words, and phrases in the mother tongue whilst other stores for these verbal items may be present for a language learned at a later age. There are also templates for non-verbal sounds. Different template stores have different accessibility. Non-verbal items for which no templates exist will be recoded, certainly if ordered recall of a sequence of such items is required.

Some evidence for separating the sensory and motor aspects of verbal items, i.e. distinguishing between auditory templates and articulatory patterns is provided by the neurological distinction between sensory and motor aphasia. The distinction, although not perfect in any of the patients described by Penfield and Roberts (1959, pp. 220-234) has led these authors to conclude that "this strongly suggests that the motor units for words and phrases are separated somehow, spatially, from the sensory units. But it is also clear that they are both located in the general region of the cortico-thalamic speech areas of the left side, where they are closely interrelated in function" (p. 247). In their tentative model, Penfield and Roberts themselves distinguish between sound unit, verbal unit and speech-motor unit.

A match between the output of VSR scan or ASR scan and the contents of one of the template stores leads to recognition. Recognition will normally directly result in the active application of a verbal code via articulatory patterns which are available for execution by an articulatory coding mechanism. Normally, therefore, recognition and coding can be lodged into one single box. But recognition may also be achieved if no verbal code can be applied directly (e.g. to the shape of a forgotten letter of the Greek alphabet). It is also possible that articulatory patterns are activated without verbal codes being applied, e.g. in shadowing material in an unknown language. It is assumed that the verbal codes form the abstract long-term representation of verbal items of which the sensory parallels are situated in the auditory templates and of which the motor parallels are similarly stored separately.

The latter two represent the modes along which verbal items are operated. But the verbal items also have a permanent existence without such operation, i.e. in an abstract manner. This may be illustrated by the fact that one may have to spend considerable effort to 'find' the word one intends to speak, i.e. to activate the correct articulatory patterns for the verbal code that is already active. Such states have been described by Brown and McNeill (1966) as the tip of the tongue (TOT) phenomenon. It is assumed that the 'abstract' form of the item (verbal code) is connected with semantic networks, also stored in an abstract manner, and that there is two-way traffic so that not only the verbal code may arouse semantic networks, but the verbal code may itself be aroused by semantic or context information. Closely related to the sematic networks are the associative networks. They are located in a separate box to indicate that not every semantic relation is associative and not all long-term associations have semantic components. It is essential to the model that the activation of the sensory or motor aspect of a verbal code does not necessarily involve its semantic or associative connections. This point was stressed already in the discussion of the model of Chapter 1; it finds further support in the lack of evidence of semantic coding in a large number of STM experiments (see 2.1.2).

The abstract representation of a verbal item in terms of a verbal code might be named, differently, a concept. This term is not satisfactory, however. Firstly because a concept is defined semantically (L.B.J. and president Johnson are the same concept but different verbal codes). Secondly, because concepts do not include all gramatical word forms and classes, nor verbal units smaller than words. An (abstract) verbal code can perhaps be best described as representing a certain category of possible speech events with common auditory and articulatory properties.

Thus, our verbal codes differ in at least two respects from the logogens in Morton's model. Firstly, they are 'abstract' with separate auditory and articulatory aspects; secondly, they do not directly contain semantic and associative information. A third aspect in which verbal codes differ from logogens is their activation. As we saw above, in Morton's (1969b) model it is claimed that logogens have autonomous entry and even that there is equivalence of channels, e.g. of numerosity and digits. In our present model, on the contrary, selective scanning is required for recognition and the subsequent application of a verbal code to a sensory event.

For verbal items to which a verbal code has not (yet) been applied there is, in our model, no way into semantic and associative networks. Single coding will normally not result in permanent associations, but repeated coding at presentation or rehearsal will. Thus, effects of language redundancy and other sequential dependencies only work through STS. As specified in Chapter 1 verbal codes are realized through an articulation mechanism (which does not necessarily involve speech musculature). There is reason to believe (2.1.5) that 'higher' rather than 'lower' levels of articulatory activity are involved in this process.

STS information is scanned, the scan being based on articulatory traces. Correspondingly differentiation between items with varied articulation (heterogeneous) is better than between items similar articulation (homogeneous). Retrieval of the items in their order is most likely also dependent on articulation cues We shall assume that the realization in STS by the articulatory mechanism of a verbal item in close temporal proximity to another item is affected by this environment as much as it is in overt speech. This would imply that, especially as a result of rapid processing during rehearsal, there is allophonic coding rather than phonemic coding. By 'allophone' here is meant a phoneme specified by its phoneme context on either side, e.g. /o/ in /top/ as /${}_{t}o_{p}$/) A related suggestion, although in a different context, has been made by Wickelgren (1969b). If individual allophones are retrieved correctly their ordering is correct with a much greater probability than if individual phonemes are correctly scanned. This interpretation of ordered recall contrasts with an associative interpretation in that it assumes specific sequential cues in the form of 'context-sensitive' representations of the individual items rather than abstract connections (associations) formed by contiguity between items which are themselves unaffected.

A decision mechanism must moreover be assumed to decide on further rehearsal, on recall (spoken or written) or, perhaps, on the suppression of overt vocalization. If rehearsal or recall are overt, an extra loop is formed, re-entering the information into ASR. For written recall writing patterns are available. To account for the results of Exp. 3 and Exp. 4 it must be assumed that they can also be activated directly via the recognition state if there is a minimal delay between visual presentation and written recall or copying and if, simultaneously, the articulation mechanism is engaged in rehearsal.

It must be stressed that the model as presented and discussed does not give a complete description of how the presented stimulus is perceived. This would demand a complete theory of visual pattern recognition and auditory speech perception. As regards the auditory channel, it should be noted that where the term 'matching' is used it does not - as in some speech perception theories - indicate active generation of internal speech representations until a match occurs, but merely a comparative pairing of the scanned sensory information and the templates of the most likely item(s), permanently stored as sensory acoustic representations As has been said, the model is in this sense passive. It is not specified which are the attributes judged in connection with the 'most-likely' decision and what criteria are employed for a match It must be noticed in this context that in the STM tasks as studied by us it is customary to work with relatively small vocabularies only, which may render the first stage of perception considerably less complicated. On the other hand, the perceptual process in these tasks is not supported by redundancy such as in the perception of normal speech. With respect to the model this means that a box labelled 'linguistic rules' may be omitted

If we now return to the beginning of this section which summed up some questions arising from the data of the experiments in Chapter 2 we can state that the consequences of coding and articulation as worked out in the above pages may be referred to as providing a tentative explanation in all cases. Heterogeneous sequences provide more varied articulatory cues, especially if allophonic coding prevails in STM there is considerably more differentiation between the items from a heterogeneous than from a homogeneous sequence. If, for example, the quadrigram BFDI is to be retained, the supposed context-sensitive ('allophonic') segments would be $/{}_{*}be_{\varepsilon}, {}_{e}\varepsilon f_{d}, {}_{f}de_{l}, {}_{e}\varepsilon l_{*}/$ which is quite varied as compared to the segments for BTDV that would be $/{}_{*}be_{t}, {}_{e}te_{d}, {}_{e}de_{v}, {}_{e}ve_{*}/$ This explanation, though tentative and open to further testing, can deal with the questions regarding item and especially order recall in Experiments 1, 2, and 6 mentioned at the beginning of this section. The data of Experiments 3, 4, and 5 are clarified by reference to special coding principles in the former two experiments Ss may be assumed to make use of the low-level loop directly to 'writing patterns', bypassing the verbal code store In Experiment 5 there may have been sequential articulatory coding of scanned and recognized non-verbal items under the VSU condition.

The discrepancies between our present model and that proposed by Morton (e.g. 1968a) may be summarized as follows. Firstly, the logogen system does not differentiate between the sensory and semantic aspects of a verbal item. It is specified in Morton's model that these are attributes directly associated with logogens. The association is in two directions: semantic information can activate the logogen, and upon activation of the logogen its semantic attributes are also activated. In our model the verbal code e.g. corresponding to a 'word' response will under normal speech perception conditions also arouse the appropriate semantic networks, but not necessarily. It is possible to perceive words accurately, and also to repeat them without a meaning attached to them. For the semantic networks to be attached to the message an extra activity on the part of S may be required. Such forms of 'minimal' verbal processing demonstrate that the verbal codes may be considered a separate entity. Reference has already been made to the large number of STM experiments in which no or few 'semantic' effects were observed (2.1.2).

The second difference is related to the activity leading to the activation of a logogen. We have assumed a scanning mechanism to represent selective processes preceding perception. Morton suggests that a logogen is activated autonomously in the sense that any sensory impression as well as 'context' contributes towards making the response available through the logogen system. The idea of a logogen also being activated by context again stresses the semantic features of the logogen system. In our model these context aspects have an effect on the scanning mechanism which begins an (active) search for certain words, given a certain context. It is of course difficult to differentiate (experimentally) between these two interpretations, but a certain kind of experimental condition seems to be available to distinguish between the two. We shall look into these problems in the next section.

Thirdly, our model, in contrast to that by Morton, indicates how the activation of a verbal code occurs. Templates, as the permanently stored sensory representations of e.g. words, serve as a reference for ('passive') recognition in both sensory channels. Template stores can be more or less readily accessible depending on context and also on the amount of experience with the verbal items whose sensory correlates are represented in the template store concerned. With context and amount of sensory stimulation constant, recognition may be delayed if the templates are not readily accessible, and independently, coding may be delayed if the articulatory patterns are not readily available. Morton would suggest a higher threshold of the logogen due to a relatively low frequency of its activation. A test that would decide between these two interpretations would be one in which auditory and articulatory cues are varied independently. Further objections to Morton's interpretation are concerned with the relationships between 'primary' and 'secondary' storage implying a lasting form of storage without coding; with the interpretation of the eye-voice and the ear-voice span as 'reflecting' the output buffer storage with the implication that all material is coded in an articulatory manner immediately upon presentation; with the incomplete investigation of the capacity of different sources of information adding independently to a logogen's excitation level; and, finally, with the implications which are related to the last point made, of parallel processing involved in 'equivalent channels'. If a response is made available from the logogen system upon criterion excitation of the logogen, and if transfer into the output buffer is serial in the order of availability, there should be no interference occurring at this transition except for those instances where two (different) responses become available at exactly the same moment.

Finally, a special characteristic of recall after auditory presentation, which has incidentally been noted several times earlier (1.4.2, 2.2.1) is that the last items of a sequence may be recalled somewhat better than after visual presentation because ASR information on these items is still available. This phenomenon may be found in the literature in so many different forms that it must be assumed to constitute an important variable and a powerful tool for differentiating between recall data in which auditory-sensory information is involved to a considerable extent (terminal items) and recall data relying mainly on articulatory codes (initial items). The evidence will be summed up below.

Conrad (1958) observed a large decrement in recall performance when his subjects were required to respond with a redundant 'zero' prefix before recalling eight-digit numbers. This effect has repeatedly been replicated, e.g. by Conrad (1960a), Dallett (1964), and Crowder (1967). In these studies interpretations have been attempted in terms of autonomous decay, memory load, and interference through formal similarity. A more recent interpretation is in terms of modality or phoneme similarity between the prefix item and the list presented for recall (Crowder 1969). In general, the prefix effect is an increase in errors spread out evenly over all serial positions with, perhaps, a slight tendency toward the earlier items.

Another procedure of interpolating redundant items between presentation and recall is to add such items at the end of the lists to be recalled. Thus, Dallett (1965) found impaired recall of seven-digit lists if they were followed by a redundant 'zero' which was not to be recalled. The proportion of correctly recalled sequences under this condition was as low as under a condition where the added (eighth) digit was not redundant and recalled. This suffix effect, which at first sight seems identical to the prefix effects just described, results, however, in a different distribution of the error increase. In contrast to the response prefix which causes a relatively nonselective error increase, the stimulus suffix works mainly on the terminal items of auditorily presented sequences. Crowder (1967) reports a complete loss of the typical recency effect after auditory presentation when he suffixed - but not when he prefixed - a redundant 'zero' to an eight-digit list. In the suffix condition, errors on the last digit increased with a factor 3. Morton (1968b) observed a large terminal effect in a slightly different condition entailing the recall of six-letter lists in which each letter was followed by an irrelevant digit. A smaller and evenly distributed effect was obtained when the digits preceded the letters to be recalled.

An interpretation in close agreement with our description of ASR, together with more evidence, has been provided recently by Crowder and Morton (1969): Precategorical acoustic storage (PAS) provides extra read-out time for identification. In general, better recall after auditory than after visual presentation may therefore be expected. Because, however, only the last items are not overwritten by new auditory information, the advantage of auditory over visual sequences is limited to their terminal items. This explains the recency effect which is a typical recall characteristic of acoustically presented lists (see 2.2.1). Redundant suffix items, if auditory, will wipe out this recency effect. A visual suffix will have no such effect on an auditory sequence; conversely, an auditory suffix will not have the effect on a visual sequence because no acoustic traces were set up in the first place. The prefix effect is not located in this precategorical but rather in a categorical stage of processing: subjects provide a prefix at recall, i.e. after having coded all the perceived items (in an articulatory mode). The prefix effect, therefore, will work on all the items in the sequence, and be much less dependent on modality or rate. These predictions were borne out in two experiments (Crowder and Morton, 1969). In an auditory running memory task using digits, the expected interactions were observed between recall conditions (prefix, suffix, control) and serial positions, and between recall conditions and rate (4, 2, 0.5 digits/sec) of presentation (Exp. 1). Visual presentation of nine-digit sequences resulted in parallel serial position curves, showing no selective effect of either an acoustic suffix or an acoustic prefix spoken by the subject (Exp. 2). Also other results that have since been obtained (Morton 1970; Morton and Holloway 1970) confirm the modality specificity of the suffix effect on the recall of auditorily presented lists.

3.3.2 Aspects suitable for experimental testing

The functional model as described above does not allow testing in any single experiment. Dependent on the elements and their functional interrelationships to be investigated, various types of experiments will be required. Within the framework of the present study, which deals with the ordered STM recall of sequences of about span size, it is of primary importance to test the descriptive value of the model for such relatively simple STM situations. Given the framework of the present chapter, it is of equal importance, however, to test the feasibility of the model for the description of other simple forms of processing verbal items, outside the sphere of STM. Of course, the two types of experimental conditions will have overlapping elements in terms of the model. The aspects and characteristics of the model studied in the experiments to be reported are discussed below.

With respect to the STM situations mentioned, the model explicitly stresses the presence of two relatively independent types of information available at recall after auditory presentation. The first, which is also present after visual presentation (as we have seen throughout Chapter 2), reflects the 'normal' STM representation of verbal items, stored in STS following articulatory coding and characterized by the possibility of rehearsal. The second, which is typical after auditory presentation only, directly reflects the characteristics of the auditory stimulation in a non-coded form stored temporarily in ASR (see 1.4.2, 2.2.1, and the discussion above). Opportunities for distinguishing between recall data based on STS on the one hand and ASR on the other, are present in the distribution of errors over the serial positions of recalled sequences. The most obvious example is the absence of a 'recency' effect (improved recall of the last items of a sequence) after visual presentation (see 2.2.1). The characteristics of overwriting at presentation and the accessibility of auditory templates are mainly related to ASR; effects of these variables will appear later in the recalled sequence. The availability of articulatory patterns as a requirement for a fast execution of articulatory coding and especially for interpolated rehearsal, are typical STS characteristics, reflected throughout the sequence if presentation is visual, and mainly early in the sequence after auditory presentation. In general, rehearsal is known to have a stronger effect on the early items in the sequence (e.g. Howe 1965). There is, thus, an indirect opportunity for studying the two forms of storage and the accompanying auditory and articulatory systems relatively independently if the shape of the serial position curve is taken into account.

One of the characteristics of the model that may, tentatively, be studied in this manner is the principle of separate systems for the auditory and articulatory forms of realization of verbal codes, i.e. auditory templates on the one hand and articulatory patterns on the other. It may be assumed that well-learned and long established verbal codes are characterized by a large degree of equivalence between accessibility of auditory templates and articulatory patterns. With newly learned and not so well established verbal codes there is a greater likelihood of a difference between the efficiency with which references to the auditory templates and to the articulatory patterns are possible - this difference most likely favouring auditory templates. Now, if a differential effect on the serial position curve is found when STM performance for sequences of 'old' and sequences of 'new' verbal items are compared, this will confirm the principle of auditory templates and articulatory patterns as two separate components of verbal codes.

A second aspect of the model concerns the use of articulatory cues for coding not only of individual items but also of their order. Starting from the principle of context-sensitive coding, especially resulting from rehearsal, and from the articulatory properties of the individual verbal items used, predictions may be attempted on the probability of correct ordered recall of visually presented sequences. Finally, with respect to STM differences may be studied in sensitivity to stimulus variables such as rate of presentation and articulatory differentiation. The specifically defined characteristics of ASR (as sensitive to overwriting but not to articulatory differentiation) and of STS (as sensitive to articulatory differences but not to rate) may thus be tested. By employing a 'non-associative' probe technique, moreover, the exact temporal relationship and the possibility of retention independent of item-to-item associations may be studied. Since in STM as described in the model no true associations are formed, the recall of an item from STM would, apart from the consequences of context-sensitive codes, be dependent on temporal cues rather than on direct item-to-item associations.

With respect to aspects of the model concerned with verbal processing independent of short-term retention, the principal problems are related to the question of how verbal responses are made available. One interpretation, provided by Morton, is that the process follows 'verbal stimulation' in an autonomous manner. In our model verbal codes are assumed to be involved only after recognition following specific VSR or ASR scanning. A method of assessing the presence of verbal activity is to use a secondary verbal task and to look for interference on the primary task. Tasks involving verbal processing will show a greater and more specific effect of such interference. In analogy to our discussion on rehearsal prevention (2.1.5) it may be stated that simultaneous verbal processing of two verbal messages gives rise to interference. Subsidiary verbal tasks to be performed simultaneously with primary tasks of which the verbal components are to be determined, therefore provide a useful tool in experimentation Likewise, dual tasks with verbal codes attached to more than one attribute (dimension) of the verbal items involved (e.g. words printed in a colour to which thus a secondary verbal code is attached) may be similarly employed.

By means of these tasks a number of variables may be manipulated so that also other aspects than the mere presence or absence of verbal processing may be studied. One, which is related to autonomous processing, is the feasibility of the notion of equivalence of channels proposed by Morton. This notion implies that sensory and context information concerning a verbal item arriving from different and independent sources cumulatively add to the level of activation of the specific logogen corresponding to that verbal item. This interpretation of activating a verbal code (which with its auditory, articulatory and semantic-associative components attached is equivalent to Morton's logogen) may be studied by comparing conditions in which the channels provide compatible and incompatible information. Facilitation should occur under the former condition and interference under the latter.

Another aspect is concerned with higher-order associations formed between items after their repeated contiguous presentation. Research on the lasting effects of repeated presentations indicates that these effects are dependent on coding responses made to the items by the Ss (Cohen and Johansson 1967). If long-term effects are found, reflecting sequential dependencies among items, it may therefore be concluded that coding responses have occurred Verbal items(such as colour words) to which two different verbal responses can be made, may be sequentially dependent on either dimension.If an effect of sequential dependency is found on a dimension, articulatory coding may, accordingly, be assumed for that dimension. Comparison of the strength of the effect on the two dimensions may, moreover, decide which of the two dimensions was given priority in coding.

Thirdly, the relative weight of, different dimensions (semantic, articulatory) in the course of processing verbal items may be assessed by varying the similarity between the relevant and irrelevant dimensions. If semantic components are involved in the type of processint at hand, a coulour name will be given as a response with greater ease if the word printed in that colour is semantically related to the colour than if the word is unrelated. Likewise, articulatory similarity may be varied. The model would predict a larger effect on the articulation dimension.

3.4 Testing short-term memory aspects of the model

We shall now describe three experiments dealing with short-term memory aspects of the model specifically. The three experiments have in common that the main effects studied are on the distribution of errors over the serial positions. Two experiments, both involving auditory presentation, are concerned with differences between ASR and STS information. The former (Experiment 7) attempts to differentiate between auditory templates and articulatory patterns as distinct forms of realization of verbal codes. The latter (Experiment 9) tests the temporal characteristics of ASR in contrast to the articulatory properties of STS. The experiment using visual presentation (Experiment 8) deals with specific order cues provided by rapid articulation during rehearsal which is known to have a stronger effect on the first sequence half. The order cues are supposedly based on the phonemic characteristics of the items used.

3.4.1 Experiment 7: Auditory templates and articulatory patterns

The above discussion on the prefix and suffix effect thus is in close agreement with the distinction we made in our model between ASR and STS as sources of information for the recall of auditorily presented material. According to the model, the early items in an auditory sequence are dependent on STS and rehearsal in articulatory terms for almost the total duration between presentation and recall, whereas for the last items of such a sequence, if not overwritten, ASR provides 'purely' acoustic subsidiary cues for the recall of the 'recency' items. Thus there is a possibility of differentiating between the accessibility of auditory template stores on the one hand and the availability of articulatory patterns on the other. Auditory templates must have accessibility for recognition and subsequent coding of all items in the sequence, but if ASR is to have its supporting effect, the last items in a sequence must once more be scanned after the recall of the early items. ASR information then must again be subject to a sensory match in the template store before its decay has proceeded beyond recognition. Especially this second scan requires rapid access to the auditory templates. Since the recall of the later items depends less on rehearsal, it will reflect the auditory aspect of the verbal codes rather than their articulatory aspect.

Conversely, for coding and especially for rehearsal as a rapid articulatory activity the undelayed availability of articulatory patterns is required. Paced articulatory coding and interpolated rehearsal are requirements entailed in the recall of the early items. Performance on these items will therefore mainly reflect articulatory aspects of the verbal codes. Of course, all items, also those whose recall is supported by ASR information, will have to be coded before recall, but coding of individual items at presentation need not be as rapid as during interpolated rehearsal. Moreover, coding of the last items of an auditory sequence may be delayed to some extent so that coding is still adequate even if it is slightly slower than presentation.

Now, if it is assumed that in the course of learning a foreign language auditory templates are made accessible at the same rate as the articulatory patterns, there will be parallel serial position curves in all groups of different progress in that language. But if in the course of learning auditory templates become accessible before articulatory patterns are made available, non-parallel serial position curves will be found. A very early stage of practice would result in curves parallel to STM performance in the mother tongue, an intermediate stage would favour later items, and a final stage would again result in serial position curves parallel with - but now at near-identical levels to - those obtained for STM in the mother tongue.

The latter suggestion will be investigated in the present experiment which is concerned with the recall of Dutch and English digit sequences by pupils with different levels of schooling in English. For testing the relative independence of auditory templates and articulatory patterns, which are assumed to function as separable aspects of verbal codes in the model, different stages in the course of learning a foreign language were chosen because they provide an occasion for studying relatively well-defined phases in the development of new verbal codes. In older, completely established verbal codes any differentiation between auditory and articulatory aspects may be more difficult, if not impossible.

By studying independently the rate of copying at presentation it was thought possible to compare the groups' read-out capacity directly from ASR, unmediated by rehearsal. The expectation is that copying differences will, correspondingly, reflect differences in the second halves of the recalled sequences rather than in the first halves. The copying task in the mother tongue would, moreover, provide an estimate of maximum output rates by the different groups.

Method. Eighty 6-digit sequences for recall were made up of the digits 1 through 9 with no digit occurring more than once in a sequence A tape recording was made with a male voice speaking blocks of 10 sequences in Dutch, alternated by blocks of 10 sequences in English spoken by a female voice Presentation rate was 5 sec per digit. Recall intervals were 12 sec. Every new block was anounced in the appropriate voice. Practice trials were given in both languages. Written ordered recall was required in boxes on response sheets. Another 64 ten-digit sequences for copying at presentation were recorded in 4 eight-sequence blocks in each language This dictation task was given following the recall task after a 5 min pause. Presentation rates, increasing over blocks, were 700, 600, 500, and 400 msec per digit. Intervals between sequences were 3 sec. Again languages alternated between blocks while - as in the memory task - a block in Dutch always preceded a block in English. The instruction for the copying task was to attempt keeping the pace of presentation. Ss were informed that scoring would involve only the length of the string of correct digits preceding the first error in each sequence. This instruction served to discourage Ss from relying on rehearsal and memory (resulting in poorer performance in the middle of the sequence and improved performance on the very last items). Also the rapid succession of sequences aimed at obtaining coding performance at presentation rather than recall from memory. Six groups of approximately 20-25 Ss (136 female grammar school pupils in total) were tested in a normal classroom during an English lesson. There were two subgroups for each of three standards of English. first, third, and fifth year of formal English lessons. The modal ages for these three groups were 13, 15, anc 17.

Results and discussion. Performance levels in both STM tasks were unexpectedly high. There were only 8 to 10 per cent errors in Dutch and 10 to 15 per cent errors in English. Since, moreover, the differences in raw scores as well as ratios English/Dutch between the two higher forms were only minimal, it was decided to analyze the data of the latter groups into one single 'Advanced' group (n=84). Correspondingly the first form will be denoted as the 'Beginners' group (n=52). The results of the copying task will be presented first. They are given in Table 3.1. It appears from the table that the Ss of the two groups differ slightly with respect to the speed of copying Dutch items. A significant difference, however, only occurs at the 400 msec rate (Mann-Whitney, z=2.61; p<.01). At a rate of 500 msec or slower both groups appear to be capable of dealing with at least 7 digits in Dutch. The English dictation caused more difficulties for both groups, but now the Beginners group lags behind at all rates. The smallest z (=1.98; p<.025) was obtained for the 700 msec rate. The other differences were all significant beyond the .01 level. The Advanced group is still capable of handling 7 items at the 500 msec rate, whereas for the Beginners group less than 6 English items are copied at this rate.

Group	Dutch				English			
	700	600	500	400	700	600	500	400
Beginners	9.2	8.2	7.7	4.6	7.9	6.3	5.7	3.8
Advanced	9.6	8.7	8.2	6.3	9.1	8.2	7.1	5.6

Table 3.1. Mean number of successive digits correctly copied during the presentation of 10-digit sequences presented at various rates (in msec per digit). (Exp. 7).

The results of the recall task are presented in Figure 3.2. The two groups are hardly distinguishable with respect to the Dutch sequences. This confirms the finding that for Dutch no significant differences in copying (and other output variables) were found at the 500 msec and slower dictation rates and agrees with the expectation that no differences in auditory templates or articulatory patterns for Dutch digit names would be present. The English sequences, however, are recalled less well by the Beginners than by the Advanced, the difference being mainly in the latter half of the sequence.

On the pooled positions 1-3 there is no significant difference; on the positions 4-6 the difference is highly significant (Mann-Whitney; $n_1=52$; $n_2=84$; $z=2.66$; $p<.005$). Table 3.2 shows that, in proportion to the Dutch sequences, the early English items are recalled equally poorly by both groups (61 per cent error increase). The later items are recalled relatively well by the Advanced but not by the Beginners. In terms of the three stages of language mastery sketched above (which we hoped to find in the three original different standard groups) this would mean that the Beginners are still at the 'very early stage of practice', the English/Dutch ratios for this group being almost equal (indicating both inadequate rehearsal and auditory template access). The Advanced would be in an 'intermediate stage', their 1-3 ratios being relatively high (indicating still inadequate rehearsal but increased accessibility of auditory templates).[1]

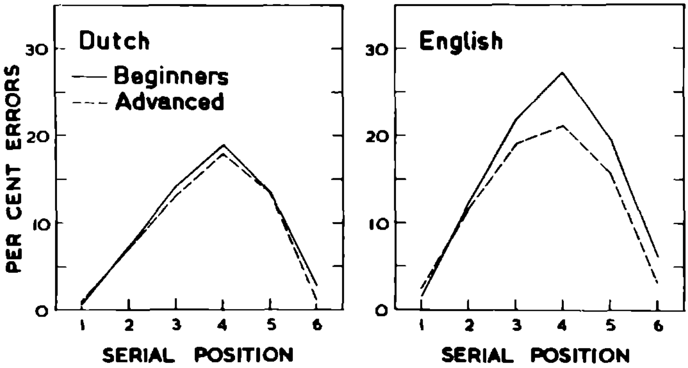


Figure 3.2. STM recall performance in auditory 6-digit sequences presented in Dutch and in English to groups of Dutch Ss, beginners and advanced in the study of English (Exp. 7).

---

[1] These findings are closely related to data obtained by Dornic in numerosity experiments. Bilingual Ss were requested to count the number of rapidly presented pulses (flashes, tones, vibration) either in their mother tongue (Swedish) or in their second language (English). Judged number was lower under the English condition than under the Swedish condition, the difference increasing with the number of pulses presented. The author's interpretation is in terms of decreased pronounceability and inadequate coding capacity for the high speed required. This interpretation is similar to the above discussion on the availability of articulatory patterns. (Dornic, S. Verbal factor in number perception. Rep. Psychol. Lab., Univ. of Stockholm, Feb. 1968. No. 244).

Group	serial positions	
	1-3	4-6
Beginners	1.61	1.51
Advanced	1.61	1.23

Table 3.2. Mean ratio of errors (English/Dutch) in the STM recall of 6-digit sequences for both sequence halves separately (Exp.7).

On the whole the results support the suggestion made above that ASR provides information partly independent of rehearsed articulatory activity in STS. But the recency effect in both serial position curves indicates that also Beginners use ASR, be it to a significantly smaller degree. Special weight should be given to the fact that Beginners and Advanced have similar ratios over the first sequence half. This finding seems to rule out an interpretation in terms of initial coding difficulties by the Beginners, resulting in accumulation in ASR so that this information cannot be used. Another finding implicit in Table 2.3 is that for the Advanced group relatively more errors occur early in the English sequences than in the Dutch sequences. This differentiates the present results from the effect of delayed coding resulting from fast presentation or resulting from noise (see Aaronson 1967, reporting on data from Conrad and Hille 1958, and from Pollack and Rubenstein 1963).

There is a possibility that the results are due to a faster output rate by the Advanced group. This possibility is implicit in the difference between the groups for copying Dutch digits at 400 msec. But results from the same experiment by Conrad and Hille (1958) show that in constrast to input rate no effect on the shape of the serial position curve is obtained by varying output rate. To the extent that recall from ASR was found in the present experiment to vary independently from recall from STS rehearsal the data support the notion of separate stages in the learning of a second language in which auditory templates are made accessible through practice sooner than the articulatory patterns. As such the results constitute evidence for separate mechanisms dealing with the verbal items: one auditory for recognition and one articulatory for coding and - at higher rates - for rehearsal.

3.4.2 Experiment 8: Articulatory cues for ordered recall

Visual presentation of a verbal sequence will result in an ordered coding of the individual items into a single chain or several chains, depending e.g. on the amount presented. After recognition and coding of the first few items interpolated rehearsal will occur, followed immediately by coding of subsequent items. Irrespective of the mode of later rehearsal it may in general be assumed that the earlier items in a visual sequence have a greater likelihood of rehearsal in a small group than later items. Given the optimal size three of a rehearsal group (2.1.4) it may be considered probable that in a 6-item sequence interpolated rehearsal will occur following the presentation and coding of the first three items. The effect of rehearsal will therefore be stronger on the first half than on the last sequence half. In the last chapter (2.1.4) we also saw that a major effect of rehearsal is strengthening the correct order of the items. In discussing the order cues effective in the model presented in the present chapter we suggested (3.3) that individual items have context-sensitive codes, which implies that, as a result of sequential rehearsal, they are supplied with order cues analogous to 'allophones' which contain implicit information on the phoneme preceding the first and on the phoneme following the last phoneme of an item itself. It is now possible to anticipate on the basis of the relation between consonant and vowel phonemes contained in the verbal items of our B-set and F-set the relative probability of correct ordered recall of these items in the first and second halves of a 6-letter sequence respectively. In B-set items the more informative consonant phoneme always precedes a long vowel. The item itself is terminated by a redundant vowel of relatively long duration. F-set items (except Z) are consonants preceded by a short vowel. The vowel, moreover, is subject to the influence of - and therefore carries information concerning - the consonant following it (for Dutch nasals and liquids in postvocalic positions this has recently been pointed out e.g. by 't Hart 1968). Termination of F-set items is (except for Z) by the non-redundant consonant phoneme. Thus, there are three types of difference between B-set and F-set items. First, in F-set items articulatory consonant cues are present for a relatively long time; at slower rates of presentation they may even be extended over time by S. Second, within the F-set also the vowel contains some information.[1] Third, B-set items have a redundant vowel phoneme and F-set items have a non-redundant consonant phoneme in their terminal positions. The former two differences will result in a stronger representation of F-set items than of B-set items in general. The third difference may be related in the following manner to order cues in a sequence. At fast processing rates such as are assumed to prevail during rehearsal, the articulatory code of an item preceding a B-set item is affected by the following non-redundant initial B-set consonant. Especially if terminal vowels (of other B-set items) precede, these contain context cues supporting ordered recall (e.g. BVDP as $/{}_{*}\mathrm{be}_{\mathrm{v}},{}_{\mathrm{e}}\mathrm{ve}_{\mathrm{d}},{}_{\mathrm{e}}\mathrm{de}_{\mathrm{p}},{}_{\mathrm{e}}\mathrm{pe}_{*}/$). An item preceding an F-set item, however, only contains information about the redundant /ɛ/ (e.g. FSRL as $/{}_{*}\varepsilon\mathrm{f}_{\varepsilon},{}_{\mathrm{f}}\varepsilon\mathrm{s}_{\varepsilon},{}_{\mathrm{s}}\varepsilon\mathrm{r}_{\varepsilon},{}_{\mathrm{r}}\varepsilon\mathrm{l}_{*}/$). A requirement for these cues to be made available is that the rate of processing is fast; at moderate rates of presentation they will only occur during interpolated rehearsal. Given this rate, there will thus be better order cues for B-set items, especially when they occur in 'clusters' than for F-set items. F-set items, irrespective of their serial position in the sequence, however, have as we saw above a stronger representation.

This line of reasoning leads to a strong hypothesis concerning the ordered recall of B-set and F-set items from the first and second half of a visual 6-letter sequence presented at a moderate rate. On an ordered recall criterion (well-rehearsed) B-set items in the first sequence half will be recalled better than (equally but less successfully rehearsed) F-set items in the first sequence half. Conversely, (less well rehearsed) B-set items in the second sequence half will be recalled below the level of (equally poorly rehearsed, but stronger represented) F-set items in the second sequence half. Support for the stronger representation of F-set items, irrespective of their order, comes from the literature and from our Experiment 3. Wickelgren (1966b) required recall of a single letter under varying interference conditions. He obtained 81.2 per cent correct F-set items versus 75.1 per cent correct B-set items. Similar differences are present in the results of other experiments including recognition situations reported by Wickelgren (1965a, 1966d). Wickelgren, however, used smaller F-sets than B-sets and his presentation mode was auditory. But with visual presentation and with equal size sets the results are similar: in our

---

[1] At the acoustic level this is also implicit in the recordings by Chistovich (quoted by Fant 1967) showing that in fast shadowing of VCVs the articulation of the consonant may be initiated by S even before he has received its main part. - The relatively long and virtually identical vowel phonemes among the items of the B-set, in contrast to the shorter and relatively discriminate vowel phonemes among the items of the F-set, may be illustrated by the spectograms of these items presented in Appendix 5.

Experiment 3 reported above (2.1.6.3) both groups recalled more letters correct from the F-set than from the B-set under all conditions. From the data of one experiment by Conrad (1967) it may moreover be concluded that F-set items are recalled as well as B-set items if recall starts 2.4 sec after the presentation, but somewhat worse if recall is delayed another 5 sec. Extrapolation to a zero interval suggests better representation of F-set items immediately after presentation. This is in agreement with the above findings, but it also indicates steeper forgetting of F-set items during rehearsal if the requirement is recall in the correct position. With respect to order errors, in Experiment 1 (2.1.6.1) the homogeneous F-sequences gave rise to more order errors than homogeneous B-sequences. Also in Experiment 2 (2.1.6.2) heterogeneous sequences composed of a trio containing one B-set and two F-set items were significantly more subject to order errors than sequences made up of a trio of two B-set items and one F-set item (23.0 vs. 16.9 per cent errors; $p<.01$). Similarly, Wickelgren (1965a), using independent criteria for item recall and for position recall, noted that F-set items are retained better but that their order is weaker than of B-set items. Especially the latter result, summarized by Wickelgren as follows, "Apparently it is easier to remember items from the ĕ confusion class (F-set) than items from the ē confusion class (B-set) ($p<.001$), but it is easier to remember the position of items from the ē confusion class ($p<.01$)..." (p. 58) lends support to our reasoning regarding a differential effect of rehearsal on the items from the B-set and F-set. The latter class of items, though better equipped with cues for item recall, profits less from rehearsal than B-set items because the effect of rehearsal is mainly on order, and order cues are weaker in F-set items. Now, since the two homogeneous conditions of Experiment 2 fulfil the requirements of list structure and presentation rate, support for the stated hypothesis will be present in the slopes of the serial position curves. As appears from Figure 3.3 the B-sequences indeed lead to fewer errors at the first three serial positions where the F-sequences are more difficult, and vice versa. This cross-over effect is very significant when tested over the 40 homogeneous sequences used ($U=97.5$, $n_1=n_2=20$, $p<.005$. The following experiment further tests the hypothesis, now with sequences requiring not only order recall, as in Experiment 2, but requiring ordered recall of 6-letter sequences with different letters in every sequence. Furthermore, in the present experiment heterogeneous sequences were used, constructed to yield better and poorer recall merely on the basis of different distributions of B-set and F-set items.

It is possible that the anticipated interaction between vocabulary and serial position must be ascribed to a difference in the ease of coding and rehearsal between the B-set and the F-set rather than to a differential effect on order cues provided by rehearsal. Since such a difference would show up also in vocalization tasks involving overt coding and rehearsal, the experiment to be reported includes data on the rate of reading aloud and rehearsing aloud letters from the two sets.

Method. From the complete 14-letter vocabulary 56 six-letter sequences were drawn up, each letter occurring with equal frequency in each serial position. Half the sequences consisted of three B-set letters followed by three F-set letters, the other half of three F-set letters followed by three B-set letters. These sequences constituted the BF and FB conditions respectively. They were presented alternately in four blocks of 14, each block being preceded by three practice sequences. A short pause was interpolated after the second block. The sequences were presented by an automatic slide projector at a rate of 1.3 sec per letter. A click of the projector anounced every next sequence. Recall intervals were 10 sec. Recall was written in the correct order on a form containing blocks of boxes and the 14-letter vocabulary to assist guessing. Ss were not informed of the structure of the sequences. Twelve groups of 10 Ss were run under identical conditions, but with a balanced order of blocks. The 120 Ss were female teacher training students with an age range of 18-21.

The rate of verbalization in reading aloud and in overt rehearsal was tested in a separate study in which 118 Ss of the main experiment participated. The individual tests were carried out approximately two months later. The reading task consisted of reading aloud 10 blocks of 75 well-spaced letters at maximum speed. The letters making up each block were randomly chosen either from the B-set or from the F-set. Immediate repetitions of a letter were never admitted. Ss were given a stopwatch to time themselves on each block. The order of presentation was balanced. In the rehearsal task sequences of six letters were once read aloud and immediately repeated four more times at maximum speed. E measured the total time thus needed for each sequence. There were seven B-set and seven F-set sequences, and seven (mixed) sequences in which B-set and F-set items alternated. A sequence never contained a letter twice; otherwise they were random. The order of presentation was balanced.

Results and discussion. The error proportions for both conditions over the six serial positions are given in Table 3.3. The observed differences between BF and FB sequences are very small indeed, but they are all in the predicted direction and present at all serial positions. The mean percentages of wrongly recalled letters are 24.2 and 26.4 for BF and FB respectively. The percentages of sequences with errors are 57.2 and 61.1. Both these differences are significant at the .0001 level (t=3.92 for letters; t=3.70 for sequences; df=119). For a comprehensive comparison with the relevant data from Experiment 2, the error percentages for the first and second sequence halves in both experiments are shown in Figure 3.3. It can be seen from the figure that the two homogeneous conditions of Experiment 2 have crossing serial position curves and that, correspondingly, the two heterogeneous conditions of the present experiment have parallel curves, running at different levels.

Condition	serial position					
	1	2	3	4	5	6
BF	10.6	17.7	21.2	23.9	35.8	36.2
FB	12.4	21.0	22.1	25.6	36.6	40.4

Table 3.3. Error percentages over the serial positions of 6-letter sequences with different distributions of B-set and F-set items (Exp. 8).

A special source of evidence is, furthermore, present in the recall probability of the letter Z. This verbal item does not correspond to the other letters of the F-set in that its (relevant) consonant is not preceded but followed by the vowel. The increased difficulty that is assumed to be present for the members of the F-set when occurring early in a sequence is therefore not expected for Z. Because of its unique articulatory properties, Z is usually responsible for less than 1/7 of the overall numbers of errors made to the letters of the F-set, but its relative contribution to the error total of the F-set should be relatively low over the first and relatively high over the last serial positions. The observed percentages confirm this reasoning. The percentages are 9.3 and 12.2 for the first and second sequence half respectively. Finally, a specific comparison may be made between C and S which, having identical consonant phonemes, are different only in the vowel aspect. Of all the errors made to C, 25.7 per cent were made when C occurred in the first half; for S this percentage was 33.2. The corresponding error distribution significantly confirms the main effect of the present experiment (χ^2=10,50; df=1; p<.005).

Estimates of the ease of coding were obtained by dividing the total time spent on the overt reading task by the number of letter (44,250) read aloud from each of the sets. For the B-set the mean latency was 355 msec, for the F-set 376 msec per letter. There is thus a small difference in favour of the B-set items. On a two-tailed Wilcoxon matched-pairs signed -ranks test this difference is significant (z=6.68; p<.001). Estimates of the rate of rehearsal were similarly obtained on the basis of the total time needed for the vocalization of all the letters per condition (24,780). For the B-set the mean latency was 233 msec, for the F-set 238 msec per letter. Rehearsal rate thus appears to be near-identical for the two sets. Mixed sequences were rehearsed considerably faster (210 msec).

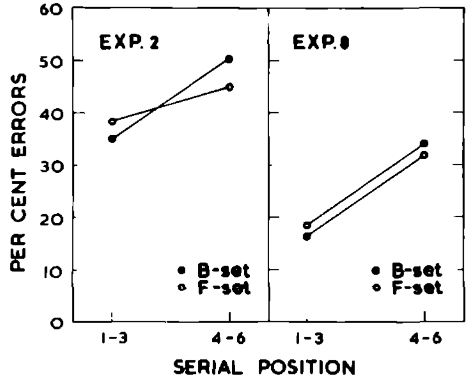


Figure 3.3. Errors in the immediate recall of visual 6-letter sequences with different distributions over the first and second sequence halves of items from the B-set and the F-set. The left-hand figure portrays the two types of homogeneous sequences of Experiment 2 which are combined in the top curve of Figure 2.2 (Exp. 8).

The observed difference in recall performance, although very significant with the large number of paired observations, is really very small: there is only a 9.1 per cent error increase from the 'easy' BF to the 'difficult' FB sequence. The interpretation in terms of a differential effect of rehearsal on order cues is, therefore, at least incomplete. The evidence, however, is not invalidated by the data on overt rehearsal rate, showing equal performance. Furthermore, the fact that the difference obtained a posteriori from the data of Experiment 2 is larger than the effect in the present experiment, also agrees with our interpretation since the conditions in Experiment 2 were designed to obtain errors in the retention of order specifically. The observed difference in reading rate, as an index of coding performance independent of rehearsal, will be left undiscussed.

These results thus favour the interpretation of the data obtained in the present experiment in terms of serial position cues being attached also to the earlier items at presentation. This interpretation is, moreover, in agreement with the phenomenon of serial position intrusion. items from a certain serial position in one sequence are reproduced above chance as intrusions in a later series in that same serial position (Conrad 1960b). Also between successive rehearsal groups within one single sequence such an effect has been reported (Wickelgren 1967). Our interpretation thus points to another factor in the ordered recall of verbal items, which, as we saw in Experiment 8, is partly but insufficiently explained by context-sensitive coding.

3 4.3 Experiment 9· Differential effects on STS and ASR

The model requires that recall of auditorily presented sequences - if fast - may in part rely on ASR information, which favours the most recent items specifically. With respect to these last items two consequences follow from this. One is that retention of these items will be possible without being dependent on rehearsal and extended storage in STS. The other consequence is that the recall of these items will reflect input variables that determine the probability of successful ASR scanning. The present experiment will investigate these consequences once more, this time not by studying the accessibility of template stores and the availability of articulation patterns, as in Experiment 7, but by varying stimulus characteristics whose roles in these two types of storage are established. In Chapter 2 the role of articulation in STM coding was extensively discussed. Correspondingly, articulation patterns were adopted in the model as the input into STS. There is no further dependence of stimulus variables in STS once the articulatory code has been performed. Material that received an articulatory code (or that was rehearsed) some time before is subject to errors made in scanning STS, where an important parameter is 'acoustic' similarity. Measures of retention of verbal items presented more than a few seconds ago will, accordingly, reflect differences between homogeneous and heterogeneous sequences as illustrated by several of the earlier experiments, regardless of conditions of presentation and testing.

More recent items may be retained with greater ease if presentation is auditory because of the enduring characteristics of ASR. In Experiment 7 the assumption was confirmed that with increasing practice on the verbal items used Ss increasingly employ ASR as a subsidiary source of information. An important ASR feature, discussed in the last section (3.3.2) and in the first chapter (1.4.3), is that ASR is subject to overwriting. If presentation is fast, there is a smaller probability of successful ASR scanning after a minimal delay than if presentation is slow. Items scanned from ASR for immediate use are, moreover, less subject to the effect of the articulatory principles in STS. If, therefore, Ss are tested on the retention of individual items in STM immediately after the auditory presentation of a sequence, their performance will reflect the articulatory criteria of STS mainly on the earlier items and the properties of ASR overwriting mainly in the later items. In other words, the effect of homogeneous vs. heterogeneous sequences will appear principally on the early serial positions and the effect of presentation rate will be typical for the later serial positions.

The preceding two experiments also studied differential effects on serial positions. In these experiments recall was used as a measure of STM performance after auditory and after visual presentation respectively. It appeared possible to obtain the predicted effects in recall in spite of the fact that recall causes extra delay and interference especially with respect to the last items. The present experiment, which is most specifically concerned with temporal characteristics of ASR will use a 'probe' technique which tests only one item from each sequence immediately after the presentation of the last item, varying the posititons to be tested over successive sequences. This allows a more accurate assessment of the retention of an item, unaffected by output variables, as a function of its serial position at presentation. Another feature of the experiment is that it attempts to control rehearsal. Interpolated rehearsal during the presentation of a sequence has been established as a universal phenomenon in many STM situations (see 2.1.4). Rehearsal has, moreover, been assumed to affect recall performance in a specific manner in the two previous experiments. In the present experiment, however, rehearsal would constitute an uncontrolled variable because it would quantitatively counteract the recency effect and qualitatively add extra articulatory cues to the earlier items. The procedure adopted was to require S to provide his own auditory presentation, which involved the vocalization of each item at its appearance. This articulatory activity was supposed to suppress interpolated rehearsal directly, but also indirectly by pacing S's coding, and thus leaving little opportunity for rehearsal, which requires flexibility at input. Finally, it is typical for the present experiment that few sequential factors are involved. On the one hand the sequences contain repeated letters, which minimizes the role or interitem connections (see Wickelgren 1966c); on the other a sequential output mode (such as in recall or in techniques probing for the 'next' item) is absent.

Method. From the letters C,G,P,T - F,L,N,S eight-letter sequences were made up such that a sequence always contained only four different letters, each repeated once. No more than one adjacent repetition was allowed in a sequence. The sequences were homogeneous B-set, homogeneous F-set, or heterogeneous combinations of two letters from each set. Sequences were presented at a rate of 400, 600, or 800 msec per letter. Presentation was by means of a Nixie tube mounted in a display unit at eye level 60 cm from S. The Ss were to indicate the serial position of the earlier occurrence of the letter whose repetition constituted the eighth (and last) letter of the sequence. This last letter remained visible until after S's response. Effectively, the task thus involved a search of that letter among the seven earlier serial positions. The display unit was fed by punch tape. Nine different tapes were prepared, one for each of the 3x3 conditions. Thirty-six Ss were tested individually. During a single session 9 blocks of 28 sequences were presented. In a block each serial position was tested 4 times; within blocks the 4-letter vocabulary and the presentation rate remained constant. All possible heterogeneous vocabularies were used, one for each S. This was achieved by reconnecting tape reader and display unit. The design allowed balancing of the order of conditions and serial positions tested and the use of letters, letter combinations, and sequence structures. The Ss were first-year psychology students, 25 male and 11 female, paid volunteers. They were instructed to read aloud all letters at presentation. Responses were made as crosses in boxes corresponding to the seven serial positions. Following a response S started the next sequence himself. Knowledge of results in the form of the number of hits was given after each block.

Results and discussion. The mean percentages of errors, pooled over rates, were 44.8 and 45.8 respectively for the homogeneous B-set and F-set sequences. The heterogeneous sequences resulted in only 36.2 per cent errors. These error percentages are corrected for the a posteriori probability of indication of each of the seven serial positions. Figure 3.4 depicts the serial position curves of these corrected error data collapsed over rates and over vocabularies. The expected effects on the serial position curve are clearly

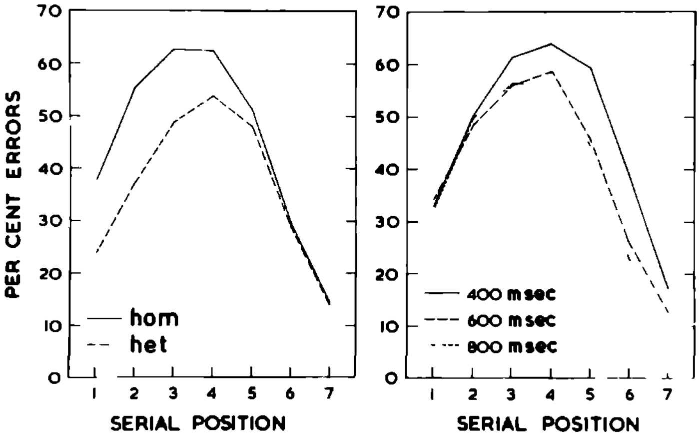


Figure 3.4. Error data, corrected for response frequency, obtained in a position indication task. Mean values are pooled over rates and over vocabularies in the left and right-hand figures respectively (Exp. 9).

present in the figure. Compared to heterogeneous sequences, homogeneous sequences result in more errors especially in the first serial positions, whereas the effect of rate is strongest in the latter half of the curve. The difference between the 600 and 800 rates is considerably smaller than between the 400 and 600 rates. Separate analyses of variance (Programme VARIAN/01; Kwaaitaal en Roskam 1969) were run for the seven serial positions. Subjects (A) were a significant main factor ($p<.01$) in all analyses. Rate (B) was significant ($p<.01$) only at positions 5, 6, and 7. Vocabularies (C) reached the .01 level of significance at positions 1, 2, and 4 and very closely ($p=.0124$) approached this level at position 3. The only significant interaction was AxB at positions 6 and 7. These results thus confirm the graphic representations and they are in agreement with the expectations formulated above.

The effect of acoustic differentiation on early serial positions and of presentation rate on later positions has thus been demonstrated under conditions of controlled rehearsal, not requiring recall. Apart from demonstrating the different characteristics of STS and ASR the present experiment shows that fairly accurate performance is possible under conditions which do not provide or require cues for sequential processing. This feature of the experiment implies that STM performance may in part be based on 'temporal' cues, resulting in accurate position indication independent of the connections (either associative or allophonic) between the items This is most probably the case with information scanned from ASR immediately after presentation of the sequence It is feasible that these later items -still audibly available- are directly judged on their 'age' resulting in no differentiation between homogeneous and heterogeneous items, and 'age' being less accurately determined when rate of presentation is high The question whether the earlier items are indicated also directly on the basis of their age (or, equivalently, their serial position) or on the basis of item-to-item associations seems to be more difficult to answer If performance decreases with homogeneous sequences this implies that items are not retained independently of one another But even then the cause may be similar item-to-item associations or allophonic connections, or similar separate items competing individually for retrieval at a certain serial position (see our discussion in 2.1.6.3)

In terms of inter-item associations, a dissimilar item in a sequence would introduce two dissimilar direct associations, one on either side of the item If only first-order associations in forward direction are considered, this would reduce the error rate of the dissimilar item and of the item following it If, however, connections are mainly between an item and its serial position, the error rate of the dissimilar item itself would be reduced but not of the item following it The present experiment does not provide sequential data for deciding between these alternatives, but the results of ordered recall in Experiment 1 may be analyzed to answer the question.

It may be seen from the data of Experiment 1 that the position at which an acoustically defined 'isolated item' occurs is characterized by a lowered error frequency (Figure 3.5). This is especially clear in sequences with only one item from one set in the middle of a sequence which for the rest is made up of items from the other set (Conditions 1 and 6). The single isolated item (left-hand figure) has a definite effect on the serial position curve, characterized by a sudden drop in the error rate. The item following the isolated item (serial position 5) is indeed recalled better than the item in the corresponding serial position in homogeneous sequences, but it is back on the smooth serial position curve that would be formed if the dip at position 4 were omitted from the graph. Similarly, when two or three items from one set are included in a sequence whose majority is from

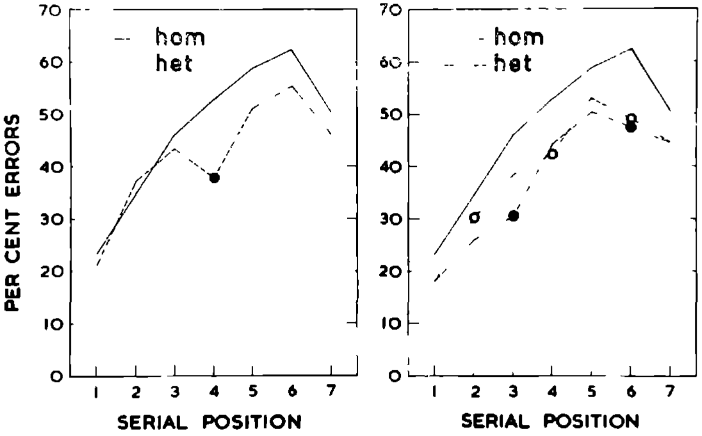


Figure 3.5. Serial position curves of STM recall of visual 7-letter sequences with acoustically 'isolated' items (filled and open circles). Data from Experiment 1.

the other set (right-hand figure) there is a general lowering of the level of the curve, but the effect of the isolated items is locally restricted to the serial positions at which they occur. Neither compared to the shape of the serial position curve for homogeneous sequences nor in relation to the appropriate heterogeneous curve itself is there an exceptionally low error rate or any serial position following that of an isolated item.

3.4.4 Discussion of evidence on the STM aspects

The results obtained in the three experiments described are all in support of the investigated aspects of the proposed model. In summary, the findings suggest, in accordance with the model, that auditory templates are relatively independent of articulatory patterns in the sense that for newly learned verbal codes the articulatory patterns are made available at a somewhat slower rate than the rate at which the auditory templates are made available (Experiment 7). Moreover, they suggest that verbal items in STM are not only represented by the corresponding individual articulatory codes, but also by the articulatory ('allophonic') cues involved in the sequential characteristics of a string of verbal items; these characteristics are, moreover, mainly due to rapid processing in rehearsal (Experiment 8). Lastly, differential aspects of presentation rate on the one hand and articulatory discriminability on the other, demonstrate the specific characteristics of ASR and STS as two distinct storage mechanisms in auditory STM tasks, one characterized by sensory and temporal attributes and the other by articulatory codes, relatively, independent of input variables (Experiment 9).

It should be noted, however, that especially in Experiments 7 and 8 the differences between the groups and conditions respectively are small enough to suggest that the model only partly accounts for the factors involved. The fundamental problem here is that evidence is largely indirect, reflecting many stages of processing in addition to the ones of primary interest. Firstly, there is unobserved and uncontrolled rehearsal, assumed to be involved in the differential effects obtained in Experiments 7 and 8. Secondly there is recall as the mode of responding by Ss. Recall does have the advantage of most directly reflecting sequential processing as it is going on, and in sequences of span length as used throughout the experiment reported thusfar it is generally spontaneously chosen by Ss as the 'natural' order of responding (e.g. Corballis 1969). One of the merits of Experiment 9 which used position probes is, therefore, that it indicates that the general principles of STS and ASR also hold under non-recall conditions. The probability of artifacts in either of the methods is thus reduced. With respect to Experiment 7 it must again be stressed that the observed error percentages were very low. Differences between the two languages would have been considerably larger if 7-item sequences had been used. Our expectation is that, compared to performance in the mother tongue the relative superiority of the Advanced group on the later items in the foreign sequences will show up more clearly, whereas the group performance levels on the early items will remain equal. With respect to the results of Experiment 8 which used a moderately slow presentation rate, an improved method would involve faster presentation of the later items. This would maintain the opportunity for differential rehearsal effects early in the sequence but it would cause recall of later items to depend more exclusively on unrehearsed codes.

A separate possibility for testing the relatively greater role of articulatory patterns in the first half of a visually presented letter sequence is to determine independently the correlation between pronounceability and recall performance for the two sequence halves. A higher correlation for the first sequence half would be expected because of the greater demands on the articulatory mechanism for rehearsal of the first few items. The data of Experiment 1 were analyzed accordingly. Since, however, pronounceability ratings would require a separate session involving the cooperation of a large number of raters, it was decided to use digram frequency instead of pronounceability. This decision seemed justified for our comparison because of the consistent correlation between the two variables (Underwood and Schulz 1960). For each of the 56 seven-letter sequences of Experiment 1 two digram frequency values were obtained, based on the statistical properties of newspaper Dutch (Van Berckel at al. 1965).'First half' values included the four digram frequencies contained in the transitions from 'space' to serial position 1, from 1 to 2, from 2 to 3, and from 3 to 4. 'Second half' values included the four digram frequencies contained in the transitions from serial position 4 to 5, from 5 to 6, from 6 to 7, and from 7 to 'space'. Spearman rank correlations between these values and the number of recall errors made to the serial position 1-4 and 5-7 respectively, were $\underline{r}_{\underline{s}}=-.292$, and $\underline{r}_{\underline{s}}=-.241$. These findings are in the predicted direction, but the difference between the correlation coefficients is not significant ($p>.05$). It may still be that predictability as expressed by digram frequency is operative along two different lines in the two sequence halves. In the first half highly predictable letter sequences might be rehearsed with greater efficiency, whereas in the latter half, where the sequences are always subject to more forgetting as shown by the serial position curves, highly predictable sequences might profit more from guessing. A possible approach to differentiate between these two modes of operation would include pronounceability ratings after all. This approach, however, has not been pursued.

3.5 Testing verbal processing aspects of the model

The four experiments to be reported in the present section are concerned with verbal interference, either from a subsidiary verbal task to be performed simultaneously with the primary verbal task, or from two conflicting verbal codes implicit in different dimensions of single verbal items. In Experiment 10 letter cancelling tasks are employed. The purpose of this experiment is to study the relation between VSR scanning and verbal codes, i.e. to determine whether all verbal items are processed in a verbal articulatory manner, as is suggested in Morton's model, and to study interference as a function of the criterion used for cancelling on the one hand and the type of verbalization required of S in a subsidiary task on the other. The remaining experiments employ colour-word tasks. The first of these (Experiment 11) studies the feasibility of information accumulating in a facilitatory manner, in parallel channels also suggested by Morton's model, before an overt verbal response is made. Experiment 12 investigates the build-up of effects of sequential dependency among verbal codes that are either vocalized or not, i.e. the articulatory processing stage reached by each of the two interfering dimensions of colour-word items. Finally, in Experiment 13 the relative weights of verbal interference are determined for semantic and articulatory dimensions respectively. In terms of the model, semantic networks are not only separable from verbal codes but also less directly involved in STM than articulatory patterns are. Accordingly, assessment of the greater weight of the articulatory dimension in verbal interference is attempted.

3.5.1 Experiment 10: Articulatory interference in verbal processing

Visually presented letters as verbal items will normally pass the articulatory coding mechanism. The consequences of this have been shown e.g. in most of the preceding experiments. Letters, however, also have non-verbal properties, such as their shape characteristics (e.g. capital vs lower case; straight vs curved). If letters are to be cancelled on a 'shape' criterion, there would be VSR scanning merely e.g. for curves rather than for letters. Even though a decision on the shape criterion would take longer than naming the item (see Fraisse 1969; 2.1.3), no recognition, and therefore no articulating mechanism, will be involved. Morton's model, on the other hand, would anticipate articulation due to autonomous verbal processing.

In contrast, if the item is to be classified according to a verbal category (e.g. alphabetical position) articulatory coding will occur following recognition. Such articulatory activity, furthermore, may itself be made to provide the criterion for cancelling (e.g. by requiring S to cancel letters with a long e-vowel when pronounced). The former type of classification would involve articulation as an accompanying code, followed by a search in the LTM associative alphabet list. The latter would involve scanning in STS for the output of the coding mechanism. There would, thus, be prolonged articulatory activity under this last condition, characterized by an 'articulation' scan.

The presence of articulation in the three types of tasks mentioned may be assessed by their sensitivity to interference by a subsidiary covert verbal task (e.g. silent counting) to be performed simultaneously. Such a task would occupy the articulatory coding mechanism to the detriment of the processing rate of letters. The prediction is that 'shape' tasks will show little if any effect of interference, in contrast to the other type tasks. The subsidiary task, moreover, would, on our model, cause less interference to the 'alphabet' task than to the 'vowel' task because both counting and the 'vowel' task depend on STS scanning. If counting in the subsidiary task is overt, the same general relationships will be found, with the exception that scan in STS will differentiate counting responses more easily from letter coding responses because the former have accompanying ASR information. This would be analogous to findings on dichotic listening tasks which show that one of two simultaneous messages is perceived with greater efficiency even if the loudness of the relevant message is far below that of the message to be ignored (see Broadbent 1958, pp. 18-35).

The main experiment in which these predictions were tested was preceded by a pilot study with a separate group of Ss. The pilot study served two functions. First, it was intended to provide baseline performance data on the six different tasks. Any pairs of tasks performed on a different criterion but with similar baseline rates would, moreover, provide a valuable extra opportunity for testing the differential effect of the counting conditions to be introduced in the main experiment. Second, it aimed at determining whether different sets of letters would be cancelled, also after prolonged practice, on the criterion provided by the instruction. When Ss become familiar with a certain set of letters, there is a possibility that further cancelling responses are made increasingly on the basis of membership of this learned set rather than on the characteristics that define the set (e.g. A, B, C are such a familiar set, also independent of their alphabetical position). In that case one would expect a differential speed increase tending towards a common latency plateau for all tasks. Further evidence for the criterion of cancelling throughout the experimental session would come from an error analysis. A specific error pattern for each task, remaining constant over the two presentations would support the lasting effect of the instruction.

Method. Booklets were prepared containig twelve identical pages with clearly spaced capitals printed in 20 lines of 28 letters. Each of the letters B C D F G L M N P R S T V Z randomly occurred twice in every line.

Each of these 12 pages was preceded by an instruction page to introduce the next cancelling task and to provide a practice set containing all 14 letters. In the pilot study every subject had to work through the following tasks twice in a balanced order. Cancelling of the letters
(1) with curved lines (B C D G P R S)
(2) with only straight lines (F L M N T V Z)
(3) belonging to the first half of the alphabet (B C D F G L M)
(4) belonging to the second half of the alphabet (N P R S T V Z)
(5) with a long /e/ vowel when pronounced (B C D G P T V)
(6) with a short /ε/ vowel when pronounced (F L M N R S Z)
After an extensive group instruction the Ss were given time to work through the first practice set which, naturally, always contained seven letters to be cancelled and seven not. Only after this was completed by every member of the group the 'ready-go' signal was given. The Ss then immediately turned over the page and worked at maximum speed for 3 min. After every 30 sec, as indicated by E, the subjects put an extra dash behind the last letter checked, thus providing information on the rate of progress. The 3 min period was concluded by E's 'stop' signal, upon which the Ss turned to the next instruction page, and so on. After the completion of the sixth 3 min period, when the Ss had worked on each task once (first presentation), a short rest period was inserted. This pause was followed by a second presentation of the same six tasks in the same order, which was, however, different for every subject. There were 36 Ss, tested in three groups of 12. Ss were trainee social workers.

In the main experiment 24 Ss were tested individually in a balanced order. They were each given 24 one-minute trials of letter cancelling. Each of the six tasks used in the pilot study was presented once under three different conditions. Moreover, at the beginning as well as at the end of the session there were control trials under these conditions. The three conditions comprised a subsidiary counting task to be performed simultaneously with the cancelling task. They were as follows:
(a) without counting
(b) counting silently
(c) counting aloud
The a condition was identical to that used in the pilot study. The only exception was that the lines on the response sheets now contained 20 letters. The last 8 letters of every line were omitted in this experiment so as to introduce some uncertainty in the total number of

letters to be cancelled (and counted under the b and c conditions) per line. In the pilot study this number was always 14; by the above procedure a range from 7 to 13 was introduced. The b condition required no other overt activity from S than a statement at the end of every line of the number of letters cancelled on that line. This number was noted down by the experimenter. At every new line counting started at 1. The c condition required the subject to count aloud every letter cancelled. The line totals were again noted by the experimenter and, as under the b condition, the subject started counting again at the beginning of every line. The 18 experimental trials were preceded and followed by three one-minute control trials consisting of crossing out every other letter under the a, b, and c conditions. Ss were first and second year students.

Results. The data from the pilot study will be presented first. Table 3.4 presents the mean latencies for the six tasks, averaged over complementary pairs to yield Shape (tasks 1 and 2), Alphabet (tasks 3 and 4), and Vowel (tasks 5 and 6) scores for each of the two presentations. RT estimates were obtained by dividing total time by the number of letters considered for cancellation.

3 min Period	Shape	Alphabet	Vowel
first presentation	471	604	665
second presentation	415	551	598

Table 3.4. Mean latencies (msec) for cancelling letters on the basis of their shape, alphabetical position, and vowel characteristics. Latencies are expressed as RT per letter checked (Pilot study of Exp. 10).

The difference between Shape and Alphabet as tested by means of the Wilcoxon matched-pairs signed-ranks test are significant beyond the.01 level for both presentations. Vowel is only just significantly slower than Alphabet ($p<.04$ and $p<.06$ for the two presentation periods respectively). Almost identical results were obtained in both presentations of 'second half of the alphabet' (614 and 564 msec for the two presentations) and 'letters with a long e-vowel' (623 and 563 msec).

The increase in speed at the second presentation is in the order of about 60 msec for each of the three tasks. This indicates parallel learning slopes. Performance data after every 30 sec (not presented here) unambiguously confirm this parallel progress in the course of each of the two 3 min presentation periods for the different tasks. The error analysis showed that error frequencies were extremely low, never exceeding the 0.3 per cent level. Omissions, however, were more frequent (3 to 4 per cent). The specificity and consistency of the pattern of the omissions may be concluded from Table 3.5. The consistency of the pattern of omissions over the 14 letters was determined by means of test-retest reliability coefficients over the two presentations per task. The obtained values were much higher than the correlations between the patterns of the different tasks. This specificity is also illustrated by the fact that in the Alphabet task there was a steep and smooth gradient with maximum probability of omission at the cenral letters L, M, N and near perfect performance at both ends of the alphabet B, C, D and T, V, Z. In the Shape task, however, L, M, N were hardly ever overlooked, and in the Vowel task Z was the letter yielding by far the most omissions.

	Shape	Alphabet	Vowel
Shape	(.94)	-.25	.34
Alphabet		(.88)	-.25
Sound			(.98)

Table 3.5. Spearman rank correlation matrix showing specificity and consistency of the pattern of omissions in three letter cancelling tasks. The values on the diagonal represent reliability coefficients (Pilot study of Exp. 10).

In the results of the main experiment the data from the 24 trials per S have again been avaraged to yield Shape, Alphabet, Vowel, and Control latency scores for each of the three conditions. These are presented in Table 3.6. The Control task is obviously not subject to interference from the counting conditions; the three mean latency scores under the a, b and c conditions are all very close to 340 msec. The same lack of an effect of counting is found for Shape where, however, the latencies are in the order of 470 msec, which nicely replicates the RTs of the first presentation in the pilot study. Alphabet shows a 40 msec delay under the silent counting condition b as compared to the a and c conditions which are virtually identical at about 580 msec. The differences among these Alphabet scores are, however, not significant when tested by means of a Kruskal-Wallis analysis of variance. By the same analysis the differences among the Vowel latency scores showed up as very significant ($p<.01$). Compared to the a condition, the silent counting condition b gives rise to an extra delay of 109 msec ($p<.01$) and the counting aloud condition c to an increase of 45 msec ($p<.025$). The difference between b and c is also significant at the .025 level. The latter differences were tested by means of the Wilcoxon matched-pairs signed-ranks test.

Condition	Shape	Alphabet	Vowel	Control
a. no counting	467	584	690	338
b. silent counting	476	621	799	347
c. counting aloud	461	578	735	336

Table 3.6. Mean latencies (msec) for cancelling letters on the basis of different classification principles under three conditions of verbal interference. Latencies are expressed as RT in msec per letter checked (Exp. 10).

The latencies of condition a are generally close to those of the first presentation of the pilot study. The Shape RTs, however, are much closer than the Alphabet RTs or the Vowel RTs. This difference between the two experiments - which may be due to a difference in the Ss, as well as to a difference in the procedure - prevents us from making a closer comparison between the two conditions which led to near-identical results in the pilot study. The latencies under condition a were 574 msec for the 'Second half of the alphabet' and 658 msec for 'Letters with a long e-vowel'. These scores are more than 80 msec apart, so that the convenient basis for comparison is lost.
This is the reason that the results, which in general confirm our predictions, must be regarded with caution. Now there is still a possibility that the more 'difficult' tasks (with long RTs) are more subject to interference by a subsidiary task, irrespective of its articulatory characteristics. However, this would not explain why covert articulatory activity is more interfering than overt articulation. Finally, it may have been noted that the Shape (a) RTs are 50 to 100 msec longer than the latencies found for naming the same letters in the reading task of Experiment 8. This implies, in terms of the interpretation by Morton, that there may have been ample opportunity for stimulus information to build up for a logogen to reach criterion excitation

and to make an articulatory response available. Our data show, however, that no articulation whatsoever is involved in the Shape task, which supports the VSR scan of our model. To some extent ASR scan may be supposed to be supported indirectly by the fact that auditory-plus-articulatory feedback from counting aloud is less disturbing than articulatory feedback only from silent counting. In terms of our model ASR scan would be reasonably successful in not selecting auditory information under the c condition.

3.5.2 Experiment 11: Inhibition vs facilitation of recognition

Colour-word (CW) performance is concerned with naming the colours in which words are printed. If the word is the name of a colour different from the colour in which the word is printed, CW performance is slowed down considerably. A task of this type, used as a diagnostic tool for testing sensitivity to interference, is known as the Stroop colour-word test (see Jensen and Rohwer 1966 for a review). In terms of our model CW performance may be regarded as illustrating interference at the level of the articulatory mechanism where one set of articulatory patterns is executed but blocked from vocalization by the decision mechanism ('no overt response') and another set is let through for responding ('speak'). In general there will be easier access to word templates than to colour templates, leading to earlier recognition and faster coding of the word. This is, however, the response to be discarded, so that the articulatory and decision mechanisms are always engaged with the execution of a different articulatory pattern of a different response at the time of the arrival of the articulatory patterns for the correct colour name response. This, as a result, will be delayed. This interpretation holds for the 'normal' CW items in which colour and word have different verbal codes. If CW items are incidentally printed in the correct colour, there should still be blockage by the decision mechanism until the colour name instructions have arrived for the decision mechanism to decide that all is clear for vocalization.

The interpretation of CW performance of Morton's model would be that two logogens are both activated and that there is interference on the exit from the logogen system into the output buffer. But if compatible items occur, the same logogen's excitation level should be increased by two different, equivalent channels. This would lead to more rapid transfer to the output buffer. Since there would, moreover, be no output interference, the response should be faster than a word reading response made to a colour name printed in black. In other words, the logogen interpretation would predict facilitation on compatible items, whereas on our model the normal latency for naming the colour of a patch would be expected because the articulatory patterns originating from colour recognition must first arrive independently and at the normal rate for colour-matching in the template store. No further interference may be expected because the articulatory patterns are identical so that vocalization can occur undelayed. The present experiment attempts to differentiate between these two interpretations. Most specifically the 'equivalence of channels' notion by Morton is subject to experimental testing. In the present experiment the probability of compatible items in the list of one hundred CW items is varied over 5 levels, including all-incompatible (CW 0) and the all-compatible (CW 100). The former would provide 'standard' CW scores; the latter would provide pure compatibility measures. The three remaining conditions would provide an opportunity for comparison at intermediate probabilities.

Material. Three different types of lists, all containing 100 items in a 10 x 10 format, are used for the tasks Word-Reading (WR), Colour-Naming (CN), and Colour-Word performance (CW). The WR list contained the (Dutch equivalents of) the words 'green', 'brown', 'red', and 'blue' equally often in random order with the restriction that two identical words were never adjacent -horizontally or vertically- in the list and that each word was followed in the list by every other word with approximately equal probability. The WR words were printed in black lower case letters. The CN list contained 100 rectangular patches in the colours green, brown, red, and blue. The distribution of the patches was subject to the same restrictions as the distribution of the words in the WR list. The five CW lists made up for this experiment varied colour-word compatibility over five levels. A compatible CW item is a colour word printed in the colour corresponding to the word, e.g. 'green' printed in green. An incompatible item is a colour word printed in one of the three colours that do not correspond to the word, e.g. 'green' printed in blue. The probability of an item being compatible was 0, 25, 50, 75, and 100 per cent in the five CW lists. The compatible and incompatible items were distributed evenly throughout the lists in their appropriate frequencies. For example, in the 50 per cent compatibility CW list (CW 50) there were 13 green, 12 brown, 12 red, and 13 blue compatible CW items and the remaining 50 item positions were occupied in approximately equal frequencies by the 12 incompatible CW items.

Subjects. Twenty Ss, 18 male and 2 female, participated in this experiment. Their cooperation was obtained during breaks in the procedure of their individual testing programmes set up for selective purposes. The age range was 19 to 43 with the mean at 31.8 years.

Procedure. Ss were tested individually, seated at a table on which was placed a sloping desk, put at a convenient angle and at an adjusted distance for ease of reading. After the instructions for the next list, this list was placed on the reading desk for a short training trial. When this was satisfactory, S began his reading or naming task, working from left to right down the 10 lines making up the list. All Ss started with two successive trials on the WR list, followed by two trials on the CN list. The five CW lists were each presented four times. In each block of four trials every list was presented once in an order that was different for each S and for each block. The total duration of a session was approximately 45 min including short rest periods that were interposed between trials. Maximal speed and accuracy were stressed in the instruction. E monitored S's responses and signalled any errors to him. S was instructed to correct his error before continuing. E noted down the total number of errors per list and the total time required to read or name the 100 items correctly.

Results. The main results are presented in the bottom row of Table 3.7. The five conditions differ very significantly (F=207.80; df=4; p<. 001). It is clear that the CW 100 items are named as fast as the colour words are read, but not faster. This indicates a complete absence of facilitation. At the

Presentation	Word reading	Colour naming	Per cent compatible CW items				
			0	25	50	75	100
1	433	687	1122	1061	1007	904	488
2	408	655	973	954	859	798	436
3	-	-	920	888	817	747	409
4	-	-	892	854	809	748	398
Mean	421	671	977	940	873	799	433

Table 3.7. Mean latencies in msec per item in WR, CN, and CW performance under various compatibility conditions. Latency scores are obtained by dividing the total time by 100 (Exp. 11).

intermediate levels 25, 50, and 75 there is a possibility for testing predictions based on the interpretation of responses to compatible CW items being made at WR rates or at CN rates. The former type of predictions would be obtained by (p-1)(CW 0)+p(WR), where p is the probability of a compatible item. The latter type predictions would, similarly, be obtained by (p-1)(CW 0)+p(CN). For example, if compatible items have latencies equal to the Word reading latencies, the mean latency for the first presentation of the CW 25 list is predicted as .75(1122)+.25(433)=950. The observed value is 1061. Differences between observed mean latencies and the two types of predictions on both lists are presented in Table 3.8 for all presentations. Extrapolated values of 400 and 400 for WR and 635 and 620 for CN were employed as estimates for the third and fourth presentation of the lists respectively. It is clear from the table that all predictions underestimate the mean latencies, but that those based on CN performance are much closer to the observed values. Moreover, with more compatible items in a list the discrepancy between observed and predicted values increases sharply in spite of the overall decrease of latencies. With an exception for the first presentaion, there are no striking changes over the four presentations in Table 3.8, although the presentation x condition interaction was very significant (F=5.13; df=12; p<.001).

Presentation	WR			CN		
	25	50	75	25	50	75
1	111	230	299	48	102	108
2	122	168	231	60	45	63
3	98	157	217	39	39	41
4	85	163	225	30	53	60

Table 3.8. Difference between observed latencies and those predicted on the assumption that latencies to compatible items equal those to WR and those to CN items respectively (Exp. 11).

Discussion. The results do not support the 'equivalence of channels' notion as a characteristic of Morton's logogens system. Instead, the finding that latencies are predicted much closer if it is assumed that naming compatible items is equivalent to colour naming, lends support to our interpretation that CW performance effectively remains a colour naming task in spite of the presence of many compatible items. The CW 100 list, of course, provides an exception; this list was soon recognized by Ss as 'the easy one', and from the second trial onwards a different scanning strategy must be assumed. Trials 2 and 3 are remarkably close to trials 1 and 2 on WR. Accurate matches between WR and CW 100 were found for all 20 Ss. The systematic underestimation of latencies for compatible items will be left without discussion.

Experiment 12: Sequential dependency among verbal codes

An STM sequence that is repeated over and over again, embedded in a long list of similar sequences, may gradually be recalled better. This effect has been named after Hebb who reported it in 1961. Since then a number of variables have been investigated that were supposed to be involved (e.g. Melton 1963). It has become clear that long-term retention of material presented under typical STM conditions does not occur unless a coding response by S is made (e.g. Cohen and Johansson 1967). The statement may also be reversed to indicate that if a persisting effect occurs, there must have been STM coding. With colour-word items used for the CW described in Experiment 11, it is possible to introduce sequential dependency on one dimension irrespective of the other.Correspondingly, the presence of an effect of sequential redundancy on any of the dimensions will be taken as evidence that coding has occurred. In terms of the model, the word's interference power results form the fact that it is recognized. It must therefore have received a code. The model thus predicts a persistent effect of sequential dependency on both dimensions. The present experiment tests this hypothesis. It may be pointed out that an effect of sequential dependency will affect CW performance in opposite directions depending whether the effect is on colours or on words. If colours have a high transition probability the colour-naming response will be made with greater efficiency, resulting in higher speed. On the other hand, if a word response is made available by a high transition probability, the ready availability of the word will cause extra interference and slow down the rate of CW performance.

Material. As in the preceding experiment three different types of lists (WR, CN, CW) were used. They were made up in the same format and according to the same general rules. Transition probabilities, however, were either low (.33) or high (.80) in the different lists. These values indicate the probability of occurrence of a certain word (or a certain colour) following the occurrence of a certain other word (or colour). Thus there were two WR lists, one as in the previous experiment (where all transition probabilities were .33) and another list in which every word was followed in 80 per cent of the cases by a specific other word and in approximately 10 per cent by either of the two remaining words.

Type of list	Code	Transition probability between: Words, Low	Words, High	Colours, Low	Colours, High
Word reading	WR 33	x	-	-	-
	WR 80	-	x	-	-
Colour naming	CN 33	-	-	x	-
	CN 80	-	-	-	x
Colour-word performance	CW 33	x	-	x	-
	CW 80W	-	x	x	-
	CW 80C	x	-	-	x

Table 3.9. Word reading, colour naming, and colour-word lists used in Experiment 12.

For example, the (Dutch equivalent of the) word 'green' occurred 25 times; it was followed 20 times by 'brown' but only three times by 'blue' and twice by 'red'. Similarly, there were two CN lists, prepared on the same principles, except for the fact that the colours involved in the higher transition probabilities did not correspond to the words in the WR list. Thus, whereas in the WR list the word 'green' was followed most often by the word 'brown', the colour green was followed most often by the colour blue in the CN list. Lastly, there were three CW lists, one standard list as used in the preceding experiment (with a transition probability of .33 with respect to colours as well as words) and two high transition probability lists, one with respect to words, and one with respect to colours. If there was high transition probability between words, the transition probability between colours was kept at .33 and vice versa. The word transitions were the same in the WR and CW lists; a completely redundant word sequence in either list would have been 'green'-'brown'-'blue'-'red'-'green'- etc. Likewise, colour transitions were the same in the CN and CW lists; a completely redundant colour sequence would have been green-blue-red-brown-green- etc.

To sum up, the seven[1] experimental lists varied transition probability with respect to words and colours as in Table 3.9.

Subjects. Twenty undergraduate psychology students served as subjects in this experiment. During the course of the experiment an altered presentation order of the lists was adopted. Thus, the 20 Ss fall into two groups, viz. Group 1 (n=11) that received random presentation, and Group 2 (n=9) that was given blocked presentation.

Procedure. The general procedure was identical to that of the preceding experiment. The main difference is in the lists and in the order of their presentation. All 11 lists were read twice in succesion by all the Ss: For Group 1 a random order of the lists was determined; each of the subjects started with a different list but the order was kept constant, so that over the whole group each list was presented once in each serial position. For Group 2 systematic sequences (blocks) with increasing transition probability were chosen for words as well as colours. The 'word' block of lists, was WR 33, WR 58, WR 80, CW 80W, CW 58W, CW 33, and for 'colours' it was CN 33, CN 58, CN 80, CW 80C, CW 58C, CW 33. The standard list CW 33 was thus presented twice to the members of Group 2, once in the 'word' context, and once in the 'colour' context. The reason for changing the presentation order was that the chances of a redundancy effect building up were very low indeed for the first group. These Ss learned in one list e.g. that the word 'green' was followed with a certain high probability by the word 'brown', while in the next list this probability was again reduced to chance level. The original expectation was that the effect could build up within each list, but since the first 11 subjects failed to show any sign of an effect, the blocked presentation order was adopted. This implied that the transition probabilities between words and between colours were 'learned' successively in two successive blocks of lists. In Group 2 the odd Ss started by the 'word' block, followed by the 'colour' block; the even Ss were presented with colours first and then with words.

---

[1] In addition to the 7 lists described there were 4 more lists with an intermediate transition probability of about .58 each. They were a WR list, a CN list, a CW list with word redundancy, and a CW list with colour redundancy. Performance on these lists was not analyzed because there was an error in the design of the scheme used to draw up the .58 lists. The resulting transition probabilities were inaccurate and uneven. The latter lists were, therefore, only of importance for the procedure.

Results. As is clear from Table 3.10 there is no effect of sequential dependency in Group 1. The mixed presentation order of the lists apparently counteracted a build-up of the effect of the high transition probabilities. Even the WR lists show only a minimal difference of 17 msec in favour of the highly redundant WR 80 list. The results of Group 2 are more distinct. There is an average speed increase of 42 and 31 msec for the redundant WR and CN lists respectively. The results of the CW lists are both larger and in the predicted opposite directions. With words redundant, the CW task (CW 80W) is 71 msec slower than the CW control task, whereas with colours redundant, the CW task (CW 80C) is 67 msec faster than the CW control task. Whereas the differences between the WR and CN lists failed to reach significance, both differences on the CW lists are significant as tested by the Wilcoxon matched-pairs signed-ranks test (CW 80W: $N=9$, $T=7.5$, $p<.05$; CW 80C; $N=9$, $T=6.0$, $p<.025$). An unanticipated finding is that there is some 40 msec difference between the *two* trials on the *same* CW 33 *list*. It is apparently easier to perform on this list after training on high transition probability CW lists with colours redundant than with words redundant. This difference is not disturbing, however, because it is counter to confirming our prediction.

Group	Word reading		Colour naming		Colour-word performance			
					Words		Colours	
	33	80	33	80	33	80	33	80
1	381	364	704	715	941	943	941	951
2	393	351	688	657	918	989	879	812

Table 3.10. Mean latencies (msec) per item read or named in WR, CN, and CW lists with varying transition probability between items. Group 1 received the lists in random order; Group 2 was given blocked presentation. (Exp. 12).

Discussion. The effect of sequential dependency has thus been established for both dimensions. Articulatory coding may accordingly be assumed to be confirmed for the colour response vocalized as well as for the word response suppressed. The fact that coding of items to be vocalized and of non-vocalized items seems to proceed through the same stages is an argument for the postulated decision mechanism in the model to decide at a 'late' stage of processing that a response cannot be made. These findings, even more than the results on colour-word performance in general, illustrate the imperfect activity of VSR scan: the irrelevant word aspect of a CW item is scanned, recognized and processed as much as the irrelevant colour aspect.

3.5.4 Experiment 13: Semantic networks vs articulatory patterns

In the present experiment the relative weights are determined of semantic -associative factors on the one hand and of articulatory factors on the other in CW performance. In our discussin of these factors in STM, which were noted in Chapter 2, ample attention was paid to evidence showing that acoustic -articulatory rather than semantic factors are involved in this type of performance (2.1.2). Also in discussing definitions of STM we laid emphasis on the fact that in the early stages of processing verbal items semantic components may be altogether absent. This was the reason e.g. for introducing a separate box 'semantic and associative networks' in the model of Chapter 1. There are other interpretations also, however. Morton's model, as discussed above (3.2) attaches semantic components directly to logogens, so that if these are activated, their meaning is, somehow, also present - and vice versa: if a certain semantic context is given, the logogen's activity level may rise.

With respect to colour-word tasks there is an important study by Klein (1964) showing that interference is maximal under the condition used in our Standard CW task with words and colour names belonging to the same set. Klein showed that using different colour words interfered less; still less interference was obtained with unrelated words. Nouns with strong associative connections with colours (e.g. lemon) are more effective in producing interference than 'neutral' words. In fact, Klein obtained a gradient from more to less interfering verbal items, which reflected, to some extent semantic similarity to the colours used. Given the primarily articulatory representation of verbal items in STM and other simple verbal processing, it seems that articulatory interference will be predominant, in contrast to semantic interference, whose role may, however, be present, as shown by Klein. This prediction is tested in the present experiment. The experiment employs four groups of Ss, two for a high and low value on the 'semantic' dimension, and two for a high and low value on the 'phonemic' similarity condition respectively. Each group received its own experimental list.

Material. Again three different types of lists (WR, CN, CW) were used. The standard word reading (WR) and colour naming (CN) lists were identical to those used in the previous experiments. The standard colour word (CW) list, however, was complemented by another CW list in which a certain colour word was printed only in one incompatible colour, viz. 'green' only in brown, 'red' only in blue, 'brown' only in red, and 'blue' only in green. The latter CW list provided a control for the following CW lists in which semantic and phonemic similarity was varied. In the CW lists with high and low semantic similarity the words were not colours but monosyllabic nouns with a dominant colour associated to each. These nouns (and the colours associated to them) were the Dutch equivalents of 'field' (green), 'blood' (red), 'wood' (brown), and 'sky' (blue). In the high semantic similarity (SHS) list all words were printed in the appropriate colour, e.g. 'field' always in green, etc. In the low semantic similarity (SLS) list there were only the following four inappropriate word-colour combinations: 'sky' in green, 'field' in red, 'blood' in brown, and 'wood' in blue. The CW lists with high and low phonemic similarity also contained four monosyllabic nouns, this time each word being phonemically similar to one of the colour words. These nouns (and the colour words related to them) were the Dutch equivalents of 'shoe' (green), 'plume' (brown), 'lead' (red), and 'rope' (blue). High phonemic similarity (PHS) items were these nouns printed in the colours given in brackets which are phonemically related to them. Low phonemic similarity (PLS) items were: 'plume' in green, 'lead' in brown, 'rope' in red, and 'shoe' in blue. It was assumed that in the semantic similarity lists the nouns chosen were relatively neutral phonemically, and in the phonemic similarity lists the nouns were relatively neutral semantically. Table 3.11 gives an overview of the stimulus words and their colours used in the different conditions, and in Table 3.12 a phonetic transcription and translation of the words may be found.

		Words used in each list				
Colour of word	Response	CW control list	Semantic		Phonemic	
			high	low	high	low
green	groen	blauw	veld	lucht	schoen	pluim
brown	bruin	groen	hout	bloed	pluim	lood
red	rood	bruin	bloed	veld	lood	touw
blue	blauw	rood	lucht	hout	touw	schoen

Table 3.11. Stimulus words and colours combined with these words used in the four experimental and control conditions of Experiment 13.

Phonetic transcription			English translation		
Response	Semantic high	Phonemic high	Response	Semantic high	Phonemic high
/χru:n/	/velt/	/sχu:n/	green	field	shoe
/brœyn/	/hɔut/	/plœym/	brown	wood	plume
/ro·t/	/blut/	/lo·t/	red	blood	lead
/blɔu/	/lœχt/	/tɔu/	blue	sky	rope

Table 3.12. Phonetic transcription and English translation of stimulus words and colour name responses combined with these words under the high similarity conditions of Experiment 13.

The nouns chosen for the semantic similarity dimension were obtained by an informal procedure in which 10 people were asked to give five associations to each of the colour words. From these associations a selection had to be made because the most frequent association was in one case a two-syllable word ('aarde' to brown) and in another case it shared the two initial phonemes with the colour word serving as its stimulus ('gras' to green). The final set of four words was again presented to 12 Ss with the request to associate with colour names. Out of the 48 first responses 45 were the 'intended' colour associations. The words adopted for the phonemic similarity lists each had one consonant and one vowel phoneme in common with one colour name. Shared phonemes were not accepted in initial positions because it was supposed that that might induce S to read the first letter(s) of the words and then to complete the required colour name starting from the common letters read. An exception to the rule that two phonemes are shared is the noun 'pluim' paired with the colour 'bruin'. The different phonemes are, however, closely related, so that the total phonemic similarity between 'pluim' and 'bruin' is at least as great as that between the members of the other pairs. The colour associations to these four words, provided by the same 12 Ss, were quite varied and based exclusively on semantic properties of the words. There were no associations 'groen' to 'schoen', etc.

Subjects. Forty Ss, 32 male and 8 female, participated in the experiment. Their cooperation was obtained during breaks in the procedure of their individual testing programmes set up for selective purposes. Their age range was 17 to 48 with the mean at 26.7.

Procedure. Upon entrance, S was assigned to one of the four conditions SHS, SLS, PHS, PLS. He was then given two successive trials on five lists, always starting with the WR and CN lists, followed by the standard CW list, the experimental list (SHS, SLS, PHS, or PLS), and lastly, the CW control list (CWC). Ten Ss served in each of the four groups. Especially in this experiment, E stressed the requirement that S should clearly pronounce the colour names. Accordingly, E signalled any imperfect responses for correction by S. This measure was of essential importance in the PHS condition, where a relaxed criterion of pronunciation would even accept unclear utterances of the printed words.

Experimental group	Word reading (WR)	Colour naming (CN)	Colour-word performance (CW)		
			Standard	Control	Experimental
Semantic High SHS	418	678	1039	903	784
Semantic Low SLS	418	651	1063	861	802
Phonemic High PHS	409	666	1066	879	716
Phonemic Low PLS	433	666	1080	912	829

Table 3.13. Mean latencies (msec) per item read or named by the four experimental groups in Experiment 13.

Results. Table 3.13 presents the mean latencies per item in each of the five lists for the four experimental groups. It is obvious from the WR, CN, standard and control CW performance columns that the groups are all very similar. It is of interest to note that the control CW latencies are on average over 150 msec shorter than the standard CW latencies. This justifies the adoption of the extra CW control list. Apparently it is easier to name the colour of an incompatible colour word if this colour is always the same for each of the colour words. If not the raw scores are compared but only the Interference Scores (which are obtained by subtraction of the CN score from the CW score; Klein 1964), the difference between standard and control CW performance turns out to be in the order of 50 to 60 per cent. The data of main interest in Table 3.13 are the Experimental CW latencies in the last column. Both the low similarity scores are higher than the high similarity scores, the difference between PHS and PLS being relatively great. Better comparisons are, however, made within groups by expressing the experimental CW interference per group in terms of its standard or control interference. This has been done in Table 3.14. From this table it is clear that experimental interference is maximal when the CW items are of low phonemic similarity (PLS) and of low semantic similarity (SLS). An intermediate position is occupied by SHS, and by far the smallest interference is found for high phonemic similarity (PHS) items.

Experimental group	Interference score			Exp. Interf. in % of	
	Standard	Control	Experimental	Standard	Control
Semantic High SHS	361	225	106	29.4	47.1
Semantic Low SLS	412	210	151	36.6	71.9
Phonemic High PHS	400	213	50	12.5	23.5
Phonemic Low PLS	415	246	163	39.3	66.3

Table 3.14. Interference scores and experimental CW interference in proportion of standard and control CW performance (Exp. 13).

A Kruskall-Wallis one-way analysis of variance was applied to the standard interference scores, the group means of which are in the first column of Table 3.14. This analysis confirmed our assumption that the four groups were equal with respect to their sensitivity to colour-word interference ($H=1.89$; $df=3$; $p=.60$). One-tailed Mann-Whitney U-tests were applied to test differences between the 'semantic' and between the 'phonemic' interference scores in proportion to the individual standard interference scores, whose group means are in the third column of Table 3.14, and between the same interference scores in proportion to the individual standard interference scores, whose group means appear in column 4 of Table 3.14. In both cases the same results were found. The difference between SHS and SLS fails to reach the 5 per cent level of significance ($U=32$ for the interference scores; $U=34$ for the proportional scores; $n_1=n_2=10$), while the difference between PHS and PLS is significant ($U=15$ for the interference scores; $U=12$ for the proportional scores; $n_1=n_2=10$; $p<.01$ in both cases). There is also a significant tendency for PHS interference scores to be smaller than SHS interference scores ($U=28$ for the interference scores; $n_1=n_2=10$; $p=.05$; $U=25$ for the proportional scores; $n_1=n_2=10$; $p<.05$).

Discussion. The results are in agreement with our prediction; high phonemic similarity naming is considerably faster than low phonemic dissimilarity, and also faster than high semantic similarity. The small semantic effect obtained, which replicates Klein's results, was moreover not significant. Another feature of the results is that Standard and Control CW lists show such a big difference. An explanation which is in complete agreement with the results of Experiment 12 is that both word and colour dimensions of CW items are coded by S, so that after repeated presentation they become learned as a PA, resulting in a facilitation of the colour response upon the preceding recognition of the word acting as a stimulus. Such PA-learning would have a neutralizing effect on the conditions. But if (long-term) associative elements are present in the task, why would (short-term) phonemic effects then be bigger than semantic-associative aspects? The answer may be that the presence of overlapping articulatory elements of the two competing responses in STS may allow rapid assembly of the required colour response.

3.5.5 Discussion of evidence on the verbal processing aspects

The data from the last four experiments provide supporting as well as conflicting evidence towards an answer to the questions implicitly posed above (3.3.1, 3.3.2). Regarding the question whether a verbal code, or logogen in Morton's terminology, includes its semantic meaning, we only found a very small effect of SHS as compared to PHS in Experiment 13. With more observations, however, the SHS-SLS difference might show up as significant. Thus semantic networks are aroused to a smaller extent than articulatory patterns, but their effectiveness is not ruled out. Similarly, if articulatory patterns were completely involved, regardless of the meaning of the word, perhaps a greater interfering effect of PLS might have been expected: it should really be considerably larger than the interfering effect of SLS.

Secondly, the question whether VSR scanning is required in a model of the type at hand has not been answered yet. In the letter cancelling task of Experiment 11 there was no sign of verbal processing when letters were cancelled on a Shape criterion. However, words are continuously scanned in the CW tasks, and, as it appears from Experiment 12, words are processed as much as colour names. Is scanning at all present? The answer seems to be that context, items and overt responses all work against the effectiveness of scan. Merely the fact that CW interference is much smaller and scanning more effective when the words are non-colour words indicates that under less unfavourable conditions scanning is normally effective.

The third question, concerned with equivalence of channels, must be answered negatively. In all those cases where decreased latencies would be expected on the basis of different sources of information accumulating evidence independently, there is no matter of such facilitation, either on colour naming or on word reading.

3.6. Evaluation and perspective

The experiments concerning properties of the model, reported above, have in general yielded positive results. Major problems remain, however, with respect to the operationalizations used. The indirect approach of rehearsal was indicated above (3.4.4). But even more fundamentally: can auditory templates be distinguished more directly than has been attempted above, namely by reference to the process of learning new verbal codes? An equally indirect mode of testing was involved in the study of articulation cues for ordered recall. Here too, however, not many alternatives seem to be present because the structure of the sequences must remain constant. Moreover, the BF and FB conditions employed in Experiment 8 constitute the conditions where the difference to be obtained by our method should be maximal. The fact that the observed differnce appeared to be so small adds to the necessity of looking for a different procedure, and equally to studying other cues for order retention such as appeared e.g. in Experiment 9. It will remain difficult, however, to differentiate between associative and 'allophonic' cues if only sequential order is taken into account.

A consideration for further investigation, based on the notion of allophonic coding in articulatory cues for STM performance, and on the findings of context effects in speech perception is the following. It may be attempted to make use of the increasing correspondence - described above in our discussion of speech perception (3.1.1) - between phonemes on the one hand and on the other their realization at more and more central levels. Whereas the acoustic signal reflects e.g. the consonant's sensitivity to its vowel context, articulation patterns may or may not be equally affected. If, then, e.g. /ki/ and /ti/ are confused more easily in auditory perception than /ku/ and /tu/ this will also be found in STM if there is context-sensitive coding but there will be constancy in confusion between /k/ and /t/ in STM if the articulatory consonant patterns are less sensitive to their environment. In the case that auditory perception is more generous in providing context cues than articulation in STM is, STM confusion matrices should show independence of context, while auditory confusion matrices for consonants should in principle be different for every vowel context.

SUMMARY

Summary

The memory processes which take place during and immediately following the reception of a verbal message form the subject of the present study. They are denoted by the collective term 'short-term memory' (STM). When a person is presented with a short verbal message, he can 'retain' it for a short time in accordance with preliminary storage principles. These enable him to perceive the message in larger units at a time and also, after correct perception, e.g. to repeat this message. Two types of storage in STM have thus been indicated: on the one hand a preperceptual, raw storage; on the other hand a storage of the message in its perceived and coded form. Both types of STM storage are, however, of a temporary and unstable kind and they differ also qualitatively from storage in permanent memory structures.

In Chapter 1 observations and research from a wide range of sources are reported, from which it appears that the first stages of memory are similarly, more unstable than the later forms of memory storage. The integration with existing memory structures, required for permanent storage, has not yet taken place at that early stage. It seems feasible that verbal stimulation too has a very short, direct aftereffect and that the information is rapidly lost if integration with existing networks has not yet been achieved. Verbal material in its original form remains available only for a few seconds. If constant attention is given to the message a sort of internal repetition ('rehearsal') of the material can be effected so that the message may be longer preserved in its original form, which increases the probability of its permanent integration. There has since long been interest in the study of immediate memory and especially in its constant, limited capacity. Only during the past decade, however, has explicit attention been paid to the specific properties of the processes involved. This has, among other results, led to the construction of a number of functional and mathematical STM models. Partly on the grounds of a comprehensive discussion of these models, STM is defined as a processing mechanism in which verbal items are registered in a preliminary store from where they may be read into a categorization mechanism applying a verbal code which is stored somewhat longer and which

may be rehearsed until a response is required. The chapter is concluded by the presentation of a prelimnary model of verbal STM in which pre-perceptual storage, recognition, coding and rehearsal are included. This model serves as a framework for a discussion of findings from literature and of the experiments to be reported in the following chapter.

Chapter 2 discusses 'acoustic' factors in STM It appears that errors made in the recall of visually presented verbal material are not random but that specific confusions occur among items with common 'acoustic' properties. This reflects the characteristics of the coding response. It likewise appears that it is mainly 'acoustic' factors which determine correct recall from STM. Semantic factors play a much smaller part in STM errors. The properties and consequences of rehearsal, considered as a verbal coding process, are subsequently considered in more detail. In the next section of the chapter, four experiments are reported, all concerned with 'acoustic' coding. Experiment 1 shows that the immediate written recall of a visually presented series is significantly more difficult when this series is composed of items with equal vowel phonemes (homogeneous) than when the series is made up of items with different vowel phonemes (heterogeneous). The items are the letters B C D G P T V (B-set) and F L M N R S Z (F-set). Experiment 2 confirms the expectation that with an increased probability of order errors, especially homogeneous series will be affected. Experiment 3 studies the occurence of wrongly recalled items in a sequence (OL) intruding from an interpolated sequence (IL), as a function of the 'acoustic' similarity between IL and OL on the one hand and their similarity as regards modality of IL reproduction (spoken vs written) on the other. The results support the interpretation interms of the availability of wrong responses and not the associative interpretation in terms of a greater amount of confusion accompanying a greater degree of similarity. The suggestion that during great rehearsal activity of a difficult OL, written recall of a short IL can take place bypassing 'acoustic' coding was investigated in a succeeding experiment. Experiment 4 confirmed this hypothesis. Rehearsal of a short number sequence appears to be in agreement with verbal processing of interpolated items; rehearsal of a long number sequence appears to allow less secondary verbal activity. The findings of these experiments confirm the generally very important role of 'acoustic' coding. The question to what extent acoustic stimulus presentation is favourable for the assimilation of sequences which must be immediately recalled is now posed. Auditory presentation of verbal sequences appears to favour recall of the last items of these sequences. This is a typical effect of modality of presentation. For the assimilation of non-redundant sequences 'the ear' is equipped with a greater capacity for integration. In Experiment 5 visual and auditory presentation modes of non-verbal sequences are compared. The results underline the inferior capacity of 'the eye' for flash sequences and point in the direction of recoding into auditory sequences. This indicates the importance of auditory experience for non-auditory stimulation. A group for whom this auditory experience is strongly reduced are the deaf. From the literature it appears that, disregarding verbal tasks, they do indeed perform below the level of the hearing exclusively in tasks involving sequential presentation or delayed recall. In Experiment 6 verbal STM performance following simultaneous visual presentation is compared between deaf and hearing Ss. The mean level of performance is significantly lower for the deaf but there are no differences between the deaf and the hearing as regards homogeneous sequences. A minor role of articulatory cues in the recall data provided by the deaf could be indicated tentatively. As anticipated, the deaf make more order errors. The final section of Chapter 2 is concerned with distinguishing auditory ('acoustic') and articulatory cues during STM performance on visually presented verbal items. Confusions in STM are compared to those in auditory perception (AUP) of letters individually spoken, masked by loud white noise (Appendix 2). Confusion data for visual perception (VP) of the same letters as those used in STM tasks were also collected. Results were analyzed by means of Johnson's hierarchical cluster analysis. The STM and AUP clustering schemes display, in general, a great similarity to each other. They show remarkably little similarity to the VP clustering schemes. The discussion of the clustering schemes is in terms of a system of 'distinctive features', which as a whole fairly well predicts the cell frequencies in both STM and AUP matrices. Place of articulation was found in STM more often as a common feature in a cluster than it was found in AUP. The opposite was found for the voicing dimension. Taking all cells of the matrix into account, STM resulted in relatively low median cell frequencies for cells with different values on the dimension 'place of articulation'. Conversely, for AUP relatively low cell frequencies were found where different values occurred on the dimension 'voicing'. These data were taken to indicate the greater part of articulatory cues in the representation of verbal items in STM. The chapter is concluded by various comments on fundamental problems concerning the distinction between auditory and articulatory cues. One of them is related to the evidence for articulation as a necessary reference in the perception of speech.

In Chapter 3 this evidence is further discussed. The arguments for active speech perception are not considered sufficiently strong for regarding articulation as a necessary requirement for perception. The active interpretations (and the counter-argumentation), however, have indicated various types of cues that may also be effective in STM. One of them is the importance of context for the perception of individual phonemes. The chapter continues with a discussion of the interpretation by Morton of verbal processing, as one of the few models in the literature in which a place is given to the perception of verbal units as well as to STM. The discussion of this model leads to points of criticism, some of which are suitable for experimentation. Before undertaking this, the preliminary model of Chapter 1 is worked out in further detail. The refined model is then compared with Morton's interpretation. In the experimental sections of Chapter 3 seven experiments are reported, three of which involve STM situations, the remaining four being concerned with simple forms of processing verbal items. Of the experiments employing STM situations Experiment 7 investigates the distinction between auditory templates and articulatory patterns as two not completely overlapping opportunities for realization of verbal coding. The finding of a differential effect on the serial position curve when comparing well learned with less well established verbal codes, was taken to suggest a certain degree of independence. The succeeding experiment (Experiment 8) tests and confirms a hypothesis on the presence of order-cues. As a result of rehearsal context-sensitive articulation cues are postulated for the individual items of visually presented sequences. A prediction based on this postulate, anticipating the relative difficulty of one out of two types of sequences, was confirmed. The last STM experiment (Experiment 9) shows that the recall of the two parts of an auditorily presented sequence may be affected by two different variables. Recall of the first part depends on coding and rehearsal in the short-term store (STS); recall of the second part - especially immediately upon presentation - depends on auditory sensory registration (ASR). In general, the results confirm the postulated characteristics of the model. The experiments concerning non-STM situations have a bearing on simple processing tasks with verbal items. Common to these four experiments is that verbal interference is used. The first experiment (Experiment 10) tests the autonomous entry, postulated by Morton, of verbal items into a stage of verbal processing. For some tasks such autonomous entry must be denied. The second experiment (Experiment 11) studies consequences of Morton's model. From the data it must be concluded that the verbal units ('logogens') postulated by Morton do not correspond to the characteristics ascribed to them if facilitation rather than inhibition is intended. The third experiment (Experiment 12) assumes coding as a condition for the establishment of a permanent effect of sequential dependency. In colour-word items it appears that such an effect occurs both on the relevant (colour) and on the dimension to be discarded (word). Verbal processing of the dimension to be discarded indicates an imperfectly functioning scanning mechanism. The final experiment (Experiment 13) determines the relative importance of articulatory patterns on the one hand and semantic-associative relations on the other for simple verbal tasks. In correspondence with the model a significantly greater effect is found for articulatory patterns. In an evaluating paragraph the data are discussed.

SAMENVATTING

Samenvatting

De geheugenprocessen die zich voltrekken tijdens en onmiddellijk na het ontvangen van een verbale boodschap vormen het onderwerp van deze studie. Zij worden aangduid met het verzamelbegrip 'short-term memory' (STM). Wanneer mensen een korte verbale boodschap wordt aangeboden, kunnen zij die enige tijd 'vasthouden' volgens voorlopige opslagprincipes. Deze stellen hen in staat de boodschap in grotere eenheden tegelijk waar te nemen en ook om deze boodschap na juiste waarneming, bv. na te spreken. Hiermee zijn twee vormen opslag in STM gegeven, enerzijds een preperceptuele, ruwe opslag, anderzijds een opslag van de boodschap in zijn waargenomen, gecodeerde vorm. Beide vormen van STM opslag zijn echter tijdelijk en labiel van aard en verschillen ook kwalitatief van opslag in permanente geheugenstructuren.

In Hoofdstuk 1 worden observaties en onderzoekingen van zeer uiteenlopende aard vermeld waaruit blijkt dat de eerste stadia van het onthouden eveneens veel labieler zijn dan latere vormen van geheugenopslag. De voor een blijvende opslag vereiste integratie met bestaande geheugenstructuren is dan nog niet tot stand gekomen. Het lijkt aannemelijk dat ook verbale prikkeling een zeer korte directe nawerking heeft en dat de informatie snel verloren gaat als geen integratie met bestaande netwerken heeft plaats gehad. In zijn oorspronkelijke vorm kan verbaal materiaal slechts enkele seconden beschikbaar blijven. Als voortdurende aandacht aan de boodschap wordt geschonken kan een soort interne repetitie ('rehearsal') van het materiaal plaatsvinden waardoor de boodschap langer in zijn oorspronkelijke vorm behouden kan blijven en waardoor de kans vergroot dat er blijvende integratie plaats vindt. De belangstelling voor de studie van het onmiddellijk geheugen en wel voornamelijk voor zijn constante, beperkte omvang is al oud. Pas gedurende de laatste tien jaar is expliciet aandacht besteed aan de specifieke eigenschappen van de betrokken processen. Dit heeft onder meer geresulteerd in een aantal funktionele en enkele mathematische STM modellen. Mede op grond van een samenvattende bespreking van deze modellen wordt het verbale STM omschreven als een verwerkingsmechanisme dat een voor-

lopige opslag omvat waarin verbale elementen ('items') worden geregistreerd, en vanwaaruit ze worden ingelezen in een classificatie mechanisme dat een verbale code toepast die enige tijd langer opgeslagen blijft en die door rehearsal behouden kan blijven tot een reactie is vereist. Het hoofdstuk wordt besloten met een voorlopig model van het verbale STM waarin pre-perceptuele opslag, herkenning, codering, en rehearsal een plaats krijgen. Dit model dient als leidraad voor de bespreking van literatuur en experimenteel onderzoek in het volgende hoofdstuk.

Hoofdstuk 2 bespreekt 'akoestische' factoren in STM. Het blijkt dat fouten gemaakt in de reproductie van visueel aangeboden verbaal materiaal niet toevallig zijn maar dat er specifieke verwarringen voorkomen in termen van items met gemeenschappelijke 'akoestische' eigenschappen. Dit weerspiegelt de eigenschappen van de coderingsrespons. Het blijken dan ook voornamelijk 'akoestische' factoren te zijn die de kans op juiste reproduktie uit STM bepalen. Semantische factoren spelen daarbij een veel geringere rol. De eigenschappen en consequenties van rehearsal - opgevat als verbaal coderingsproces - worden nader besproken. In het hoofdstuk worden vier experimenten vermeld die 'akoestische' codering tot onderwerp hebben. Experiment 1 toont aan dat de onmiddellijke schriftelijke reproductie van een visueel aangeboden reeks significant moeilijker is wanneer deze reeks (homogeen) is samengesteld uit items met gelijke klinkerfonemen dan wanneer de reeks bestaat uit items met verschillende klinkerfonemen (heterogeen). De items zijn de letters B C D G P T V (B-set) en F L M N R S Z (F-set). Experiment 2 bevestigt het vermoeden dat bij een toenemende kans op volgorde fouten vooral homogene reeksen getroffen worden. Experiment 3 bestudeert het optreden van fout gereproduceerde items in een reeks (OL) die afkomstig zijn uit een geinterpoleerde reeks (IL) en dit in funktie van de 'akoestische' gelijkenis tussen IL en OL enerzijds en hun onderlinge gelijkenis voor wat betreft de modaliteit van IL reproductie (gesproken vs geschreven) anderzijds. De resultaten steunen de interpretatie in termen van het beschikbaar komen van verkeerde responses en niet de associatieve interpretatie in termen van een grotere mate van verwarring bij een grotere mate van gelijkenis. De suggestie dat tijdens grote rehearsal activiteit van een moeilijke OL schriftelijke reproductie van korte IL kan plaatsvinden buiten 'akoestische' codering om werd in een volgend experiment onderzocht. Experiment 4 bevestigde deze hypothese. Rehearsal van een korte cijferreeks blijkt verenigbaar met verbale verwerking van geinterpoleerde items; rehearsal van een lange cijferreeks blijkt minder secundaire verbale activiteit toe te staan. De bevindingen van deze experimenten bevestigen de over het algemeen zeer grote rol van 'akoestische' codering. De vraag in hoeverre akoestische stimuluspresentatie gunstig is voor het opnemen van sequenties die onmiddellijk gereproduceerd moeten worden komt vervolgens aan de orde. Auditieve presentatie van verbale sequenties blijkt de reproductie van de laatste items van de reeks te begunstigen. Dit is een typisch effect van de stimulusmodaliteit. Voor het opnemen van niet-redundante sequenties beschikt 'het oor' over een beter integratievermogen. In Experiment 5 worden visuele en auditieve presentatiewijzen van niet-verbale reeksen vergeleken. De resultaten onderstrepen de geringere geschiktheid van het oog voor flitssequenties en wijzen in de richting van visuele hercodering tot auditieve sequenties. Daarmee wordt het belang aangegeven van auditieve ervaring voor niet-auditieve stimulatie. Een groep waarbij deze auditieve ervaringen sterk zijn gereduceerd vormen de doven. Uit de literatuur blijkt dat zij, afgezien van verbale taken, inderdaad met name na sequentiële presentatie en bij vertraagde reproductie tot slechtere prestaties komen dan horenden. In Experiment 6 worden verbale STM prestaties bij simultane visuele aanbieding tussen doven en horenden vergeleken. Het gemiddelde prestatie niveau is significant verschillend maar er zijn geen verschillen tussen doven en horenden voor wat betreft homogene reeksen. Doven maken, zoals verwacht, meer volgorde-fouten. Een kleine rol van articulatiekenmerken (cues) bij reproductie uit STM van verbale items door doven kon tentatief worden aangehoord. De laatste sectie van het tweede hoofdstuk betreft het onderscheid tussen auditieve ('akoestische') en articulatie -cues tijdens het uitvoeren van STM taken met visueel gepresenteerde verbale items. Verwarringen in STM (Appendix 1) worden vergeleken met die in de auditieve perceptie (AUP) van individuele letters gesproken onder een grote hoeveelheid ruis (Appendix 2). Ook werden verwarringsdata verzameld voor visuele perceptie (VP) van dezelfde letters als die welke in de STM taken werden gebruikt (Appendix 3). De resultaten werden geanaliseerd met behulp van Johnson's hierarchische clusteranalyse. De STM en AUP clusterschema's vertonen in het algemeen onderling grote overeenkomst. Opvallend weinig overeenkomst wordt gevonden met de VP cluster-schema's. Gepoogd wordt de representatie van de verbale items te beschrijven in termen van distinctive features. De bespreking van de clusterschema's geschiedt in termen van een systeem van 'distinctive features', dat als geheel een redelijk goede beschrijving geeft van de celfrequenties in de STM en AUP matrices. Plaats van articulatie werd in STM vaker als gemeenschappelijke eigenschap in een cluster aangetroffen dan in AUP. Voor stemhebbendheid werd het omgekeerde gevonden. Gerekend over alle cellen van de matrices werden voor STM relatief lage frequenties gevonden bij verschillende waarden op de 'plaats van articulatie' dimensie. Omgekeerd werden voor AUP relatief lage frequenties gevonden bij verschillende waarden op de 'stemhebbendheid' dimensie. Deze gegevens worden opgevat als indicatie voor de relatief grote rol van articulatie-cues in de representatie van verbale items in STM. Het hoofdstuk wordt besloten met enkele opmerkingen over fundamentele problemen met betrekking tot het onderscheid tussen auditieve en articulatie-cues. Een daarvan hangt samen met de evidentie voor de rol van articulatie bij de perceptie van de spraak.

In Hoofdstuk 3 wordt deze evidentie nader besproken. De argumenten voor de actieve spraakperceptie worden niet van zodanig gewicht geacht dat op grond daarvan articulatie als noodzakelijke voorwaarde voor perceptie moet worden beschouwd. De actieve interpretaties (en de contra-argumentatie) hebben echter de aandacht gevestigd op enkele typen van cues die ook als cues in STM werkzaam kunnen zijn. Een ervan is het gewicht dat de context heeft voor het waarnemen van individuele fonemen. Het hoofdstuk wordt vervolgd met de bespreking van Morton's interpretatie van verbale verwerkingsprocessen als een van de weinige modellen uit de literatuur waarin zoveel aan perceptie van verbale eenheden als aan STM een plaats wordt toegekend. De bespreking van dit model leidt tot punten van kritiek, waarvan er enkele voor experimenteel onderzoek in aanmerking komen. Alvorens hiertoe over te gaan wordt het voorlopige model van Hoofdstuk 1 nader uitgewerkt. Het uitgewerkte model wordt vergeleken met Morton's interpretatie. In het experimentele deel van het derde hoofdstuk worden zeven experimenten gerapporteerd, waarvan er drie betrekking hebben op STM situaties en vier op andere vormen van eenvoudige verwerking van verbale items. Van de experimenten betreffende STM heeft Experiment 7 betrekking op het onderscheid tussen auditieve- en articulatie-patronen als twee niet geheel samenvallende realisatie-mogelijkheden van verbale codes. Door een differentieel effect op de seriële positiecurve bij vergelijking van goed geleerde met minder goed gevestigde verbale codes wordt tot een zekere mate van onafhankelijkheid besloten. Het volgende experiment (Expeiment 8) toetst en bevestigt een hypothese over de aanwezigheid van volgorde-cues. Ten gevolge van rehearsal worden contextgevoelige articulatie codes gepostuleerd voor de individuele items van visueel gepresenteerde sequenties. Een hierop gebaseerde voorspelling over de relatieve moeilijkheid van twee typen reeksen werd bevestigd. Het laatste STM experiment (Experiment 9) geeft aan dat de reproductie van de twee delen van een auditief gepresenteerde reeks door verschillende variabelen te beinvloeden zijn. De reproductie van het eerste deel berust op codering en rehearsal in de 'short-term store' (STS); de reproductie van het tweede deel - zeker direct na aanbieding - berust op auditieve sensorische registratie (ASR). In het algemeen leveren de resultaten een bevestiging op voor de gepostuleerde eigenschappen van het model.

De experimenten betreffende niet-STM situaties hebben eenvoudige verwerkingsprocessen van verbale items tot onderwerp. Gemeenschappelijk is aan deze experimenten dat verbale interferentie als middel tot bestudering wordt gehanteerd. Het eerste experiment (Experiment 10) toetst onder meer de door Morton gepostuleerde autonome toegang van verbale items tot een verbaal verwerkingsniveau. Voor sommige taken moet deze autonome toegang worden ontkend. Het tweede experiment (Experiment 11) bespreekt consequenties van Morton's model. Uit de data moet worden geconcludeerd dat de door Morton gepostuleerde verbale eenheden ('logogens') niet beantwoorden aan de toegeschreven eigenschappen wanneer facilitatie in plaats van inhibitie wordt beoogd. Het derde experiment (Experiment 12) gaat uit van codering als voorwaarde voor de vestiging van een permanent effect van sequentiële afhankelijkheid. Bij kleur-woord items blijkt dat zulk een effect zowel op de relevante (kleur) als op de te verwaarlozen (woord) dimensie optreedt. De verbale verwerking van de te verwaarlozen dimensie duidt op een onvolledig funktionerend 'scanning' mechanisme. Het laatste experiment (Experiment 13) weegt voor eenvoudige verbale taken het belang af van articulatie-patronen enerzijds en semantisch-associatieve relaties anderzijds. In overeenstemming met het model wordt voor articulatie-patronen een significant groter effect gevonden. De data worden in een evaluerende paragraaf kritisch besproken.

APPENDICES

	Stimulus													
Response	B	C	D	G	P	T	V	F	L	M	N	R	S	Z
B	-	99	211	138	256	164	180	54	40	36	29	54	31	51
C	96	-	94	158	104	152	202	59	72	32	29	31	76	65
D	243	140	-	192	178	220	129	45	53	42	61	67	50	49
G	217	255	199	-	186	113	292	80	61	63	83	74	72	62
P	281	169	186	105	-	148	143	65	45	58	65	100	63	43
T	159	204	253	165	224	-	157	142	83	92	64	79	58	64
V	169	177	161	219	155	116	-	109	51	57	46	55	65	72
F	71	72	39	61	51	72	98	-	157	189	119	243	274	105
L	60	85	66	77	58	76	45	244	-	170	256	291	187	107
M	59	66	84	78	70	58	61	246	203	-	482	171	156	113
N	60	62	74	68	83	63	69	168	192	331	-	188	118	137
R	60	53	66	63	91	62	64	286	240	154	197	-	155	138
S	52	122	46	82	85	85	87	260	181	116	114	145	-	236
Z	63	89	50	47	40	45	75	113	105	93	103	135	304	-
Other/no response	115	122	116	103	113	109	114	110	96	92	104	101	105	77

Appendix 1. STM (short-term memory) confusion matrix of errors in the recall of visual six and seven-letter sequences presented sequentially. All column totals, including correct responses, are 5120 (see 2.4.2.1).

		Stimulus													
		B	C	D	G	P	T	V	F	L	M	N	R	S	Z
Response	B	-	84	280	106	145	103	236	13	9	39	31	15	13	19
	C	101	-	82	129	144	151	89	7	8	12	22	6	4	11
	D	391	175	-	246	266	233	378	11	14	54	60	13	11	20
	G	380	426	332	-	372	388	350	18	17	46	55	11	16	25
	P	346	410	298	442	-	621	347	19	32	45	49	22	22	31
	T	247	333	251	379	373	-	247	28	17	29	41	12	10	24
	V	266	194	265	196	191	149	-	11	28	47	59	9	11	17
	F	14	34	17	32	29	23	35	-	138	88	55	531	677	334
	L	17	47	31	42	34	27	55	92	-	380	445	222	112	55
	M	23	40	28	23	33	26	23	49	87	-	437	113	54	30
	N	40	50	43	46	44	40	51	58	145	795	-	142	90	24
	R	31	50	33	58	48	29	36	596	310	138	114	-	501	317
	S	12	30	9	23	25	16	18	380	70	71	45	315	-	187
	Z	31	58	27	42	37	25	47	155	30	26	16	122	156	-
	Other/no response	131	163	129	127	159	138	141	102	58	72	72	45	60	62

Appendix 2. AUP (auditory perception) confusion matrix of errors in the perception of auditorily presented individual letters masked by white noise (S/N = -10 db). All column totals, including correct responses, are 2375 (see 2.4.2.2).

		Stimulus													
		B	C	D	G	P	T	V	F	L	M	N	R	S	Z
Response	B	-	6	16	5	23	3	2	25	58	4	20	73	148	6
	C	2	-	2	155	2	4		2	4		1	4	24	2
	D	111	14	-	21	15	10	3	11	62	2	15	21	57	5
	G	3	192	3	-	1	3		1	1	3	7	2	83	3
	P	25	6	12	6	-	24	2	215	5	2	4	70	5	3
	T	5	3	3		36	-	48	76	4		3	3	3	25
	V	1	1			8	34	-	52	3	4	4		2	4
	F	8	6	2	2	239	35	8	-	13	5	15	23	3	9
	L	7	5	1	1	4	13		8	-	6	2	12	7	145
	M	16	4		2	1	3	12	3	2	-	44	40	6	5
	N	30	1	1	3	5	2	3	8	7	38	-	77	15	10
	R	31	2	8	7	40	7		44	8	21	42	-	10	11
	S	34	9	1	16	1	6	1	1	4	3	11	6	-	13
	Z	4		1	1	2	13	8	16	7	6	6	7	11	-
	Other/no response	15	30	6	11	15	20	9	39	41	30	24	44	33	16

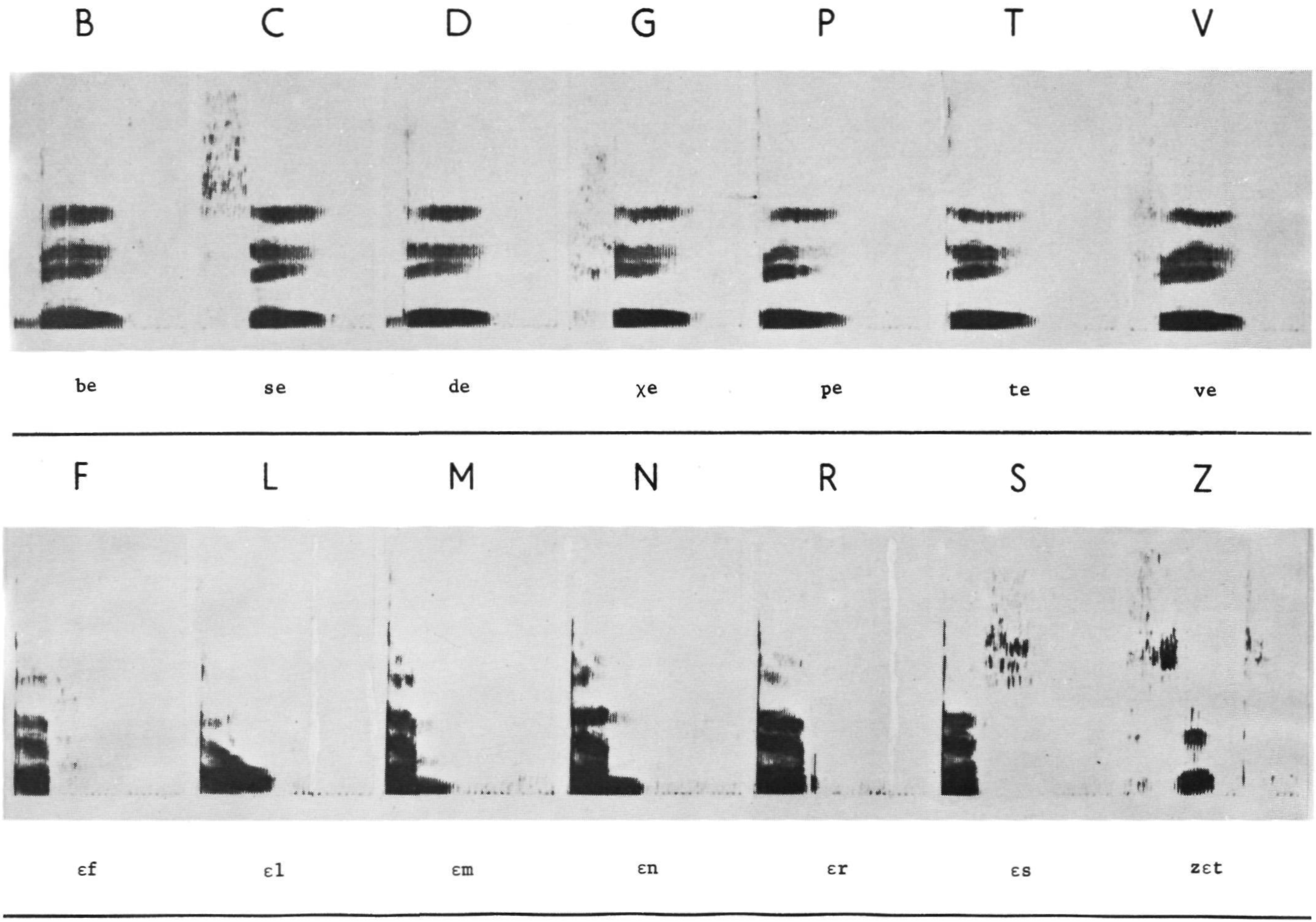
Appendix 3. VP (visual perception) confusion matrix of errors in the perception of visually presented individual letters under reduced viewing conditions. All column totals, including correct responses, are 1480 (see 2.4.2.3).

Appendix 4

B C D G P T V

F L M N R S Z

Appendix 4. Letter types (Pica capitals) employed in the verbal STM experiments using visual presentation and in the collection of the visual perception (VP) confusion data contained in the matrix of Appendix 3. Also the IL items of Experiment 3 and the letters presented for interpolated copying in Experiment 4 were printed in this type.



<u>Appendix 5.</u> Spectograms (Kay sonagrams) of the spoken 14-letter vocabulary. This material was prepared for me in 1967 at the IPO, Eindhoven, by Mr. J. 't Hart, whose kindness is hereby acknowledged.

REFERENCES

Aaronson, D. (1967). Temporal factors in perception and short-term memory. *Psychol. Bull.*, 67, 130-144.

Aborn, M., and Rubenstein, H. (1952). Information theory and immediate recall. *J. exp. Psychol.* 44, 260-266.

Adams, J.A. (1967). *Human memory*. New York. McGraw-Hill.

Atkinson, R.C., and Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence and J.T. Spence (Eds.) *The psychology of learning and motivation: Advances in research and theory*, Vol.2. New York: Academic Press

Attneave, F. (1954). Some informational aspects of visual perception. *Psychol. Rev.*, 61, 183-193.

Attneave, F. (1957) Transfer of experience with a class schema to identification learning of patterns and shapes. *J. exp. Psychol.*, 54, 81-82.

Attneave, F. (1959). *Applications of information theory to psychology*. New York: Holt, Rinehart, Winston.

Averbach, E., and Coriell, A.S. (1961). Short-term memory in vision. *Bell syst. tech. J.*, 40, 309-328

Averbach, E , and Sperling, G.A (1961). Short-term storage of information in vision. In E.C. Cherry (Ed.) *Information theory: Proceedings of the fourth London symposium*. London Butterworth.

Baddeley, A.D. (1964a). Immediate memory and the "perception" of letter sequences. *Quart. J. exp. Psychol.*, 16, 364-367.

Baddeley, A.D. (1964b) Language habits, S-R compatibility, and verbal learning. *Amer. J. Psychol.*, 77, 463-468

Baddeley, A.D. (1964c). Semantic and acoustic similarity in short-term memory. *Nature*, 204, 1116-1117.

Baddeley, A.D (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quart. J. exp. Psychol.*, 18, 362-365.

Baddeley, A.D., Conrad, R., and Hull, A J. (1965). Predictability and immediate memory for consonant sequences *Quart. J exp. Psychol.*, 175-177.

Baddeley, A.D., and Dale, H.C.A. (1966). The effect of semantic similarity on retroactive interference in long- and short-term memory. *J. verb. Learn. verb. Behav.*, 5, 417-420.

Baddeley, A.D., and Warrington, E.K (1969). Amnesia and the distinction between long- and short-term memory. *Paper read at 19th. Int Congr. Psychol., London*.

Barbizet, J. (1963). Defect of memorizing of hippocampal-mammillary origin: a review. *J.Neurol.Neurosurg.Psychiat.*, 26, 127-135.

Bartlett, F.C. (1932). *Remembering A study in experimental and social psychology*. London. Cambridge Univ. Press.

Bartz, W.H., Satz, P., and Fennell, E. (1967). Grouping strategies in dichotic listening: The effects of instructions, rate, and ear asymmetry. *J. exp. Psychol.*, 74, 132-136.

Bastian, J., Eimas, P., and Liberman, A.M. (1961). Identification and discrimination of a phonemic contrast induced by silent interval. *J. acoust. Soc. Amer.*, 33, 842 (Abstract)

Binet, A. (1894). Les origines et les modalités de la mémoire. *Rev. Philos.*, 21, 235-242.

Blackwood, R.O., and Link, J.M. (1968). A method for observing rehearsal and natural language mediation in short-term memory. *Psychon. Sci.*, 10, 287-288.

Blair, F.X. (1957). A study of the visual memory of deaf and hearing children. *Amer. Ann. Deaf*, 102, 254-263.

Blankenship, A.B. (1938). Memory span: A review of the literature. *Psychol. Bull.*, 35, 1-25.

Booth, D.A. (1967). Vertebrate brain ribonucleic acids and memory retention. *Psychol. Bull.*, 68, 149-177.

Bower, G.H. (1967). A multicomponent theory of the memory trace. In K.W. Spence and J.T. Spence (Eds.) *The psychology of learning and motivation: Advances in research and theory*, Vol. 1. New York: Academic Press, Pp. 229-325.

Brener, R. (1940). An experimental investigation of memory span. *J. exp. Psychol.*, 26, 467-482.

Brill, R.G. and Orman, J.N. (1953). An experiment in the training of deaf children in memory for sentences. *Amer. Ann. Deaf*, 98, 270-279.

Broadbent, D.E. (1956). Successive responses to simultaneous stimuli. *Quart. J. exp. Psychol*., 8, 145-152.

Broadbent, D.E. (1957). A mechanical model for human attention and immediate memory. *Psychol. Rev.*, 64, 205-215.

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

Broadbent, D.E. (1963). Flow of information within the organism. *J. verb. Learn.verb.Behav*., 2, 34-39.

Broadbent D.E., and Gregory, M. (1964). Accuracy of recognition for speech presented to the right and left ears. *Quart. J. exp. Psychol.*, 16, 359-360.

Broadbent, D.E., and Gregory, M. (1965). Some confirmatory results on age differences in memory for simultaneous stimulation. *Brit. J. Psychol.*, 56, 77-80.

Broadbent, D.E., and Heron, A. (1962). Effects of a subsidiary task on performance involving immediate memory by younger and older men. *Brit. J. Psychol.*, 53, 189-198.

Brown, J. (1954). The nature of set-to-learn and of intra-material interference in immediate memory. *Quart. J. exp. Psychol.*, 6, 141-148.

Brown, J. (1955). *Immediate memory*. Doctoral thesis. Univ. of Cambridge.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quart. J. exp. Psychol.*, 10, 12-21.

Brown, J. (1959). Information, redundancy and decay of the memory trace. In: *The mechanisation of thought processes. Nat. Phys.Lab. Sympos.* No. 10. H.M.S.O.

Brown, J. (1964). Short-term memory. *Brit. med. Bull.*, 20, 8-11.

Brown, R.W., and Lenneberg, E.H. (1954). A study in language and cognition. *J. abnorm. soc. Psychol.*, 49, 454-462.

Brown, R., and McNeill, D. (1966). The "tip of the tongue" phenomenon. J. verb. Learn. verb. Behav., 5, 325-337.

Bryden, M.P. (1969). Binaural competition and division of attention as determinants of the laterality effect in dichotic listening. Canad. J. Psychol., 23, 101-113.

Bruning, J.L., Schappe, R.H., and O'Malley, J.J. (1966). Active vs passive interpolated activity in STM. Psychol. Rep., 19, 126.

Carmichael, L., Hogan, H.P., and Walter, A.A. (1932). An experimental study of the effect of language on the reproduction of visually perceived form. J. exp. Psychol., 15, 73-86.

Cherry, E.C. (1953). Some experiments on the recognition of speech with one and with two ears. J. acoust. Soc. Amer., 25, 975-979.

Cherry, E.C., and Taylor, W.K. (1954). Some further experiments on the recognition of speech, with one and with two ears. J. acoust. Soc. Amer., 26, 554-559.

Clarke, L.M. (1951). An investigation into the effect of verbalization on the trace. Unpublished B. Ed. thesis, Queens University, Belfast.

Cohen, A., Ebeling, C.L., Fokkema, K., van Holk, A.G.F. (1961). Fonologie van het Nederlands en het Fries. Second Ed. The Hague: Nijhoff.

Conrad, R. (1958). Accuracy of recall using keyset and telephone dial, and the effect of a prefix digit. J. appl. Psychol., 42, 285-288.

Conrad, R. (1960a). Very brief delay of immediate recall. Quart. J. exp. Psychol., 12, 45-47.

Conrad, R. (1960b). Serial order intrusions in immediate memory. Brit. J. Psychol., 51, 45-48.

Conrad, R. (1962). An association between memory errors and errors due to acoustic masking of speech. Nature, 193, 1314-1315.

Conrad, R. (1963). Acoustic confusions and memory span for words. Nature, 197, 1029-1030.

Conrad, R. (1964). Acoustic confusions in immediate memory. Brit. J. Psychol., 55, 75-84.

Conrad, R. (1967). Interference or decay over short retention intervals? J. verb. Learn. verb Behav., 6, 49-54.

Conrad, R. (1970a). How do the deaf remember words? Teacher Deaf, 68, 3-12.

Conrad, R. (1970b). Short-term memory processes in the deaf. Brit. J. Psychol. 61. In press.

Conrad, R., Freeman, P.R., and Hull, A.J. (1965). Acoustic factors versus language factors in short-term memory. Psychon. Sci., 3, 57-58.

Conrad, R., and Hille, B.A. (1958). The decay theory of immediate memory and the paced recall. Canad. J. Psychol., 12, 1-6.

Conrad, R., and Hull, A.J. (1964). Information, acoustic confusion and memory span. Brit. J. Psychol., 55, 429-432.

Conrad, R., and Hull, A.J. (1968). Input modality and the serial position curve in short-term memory. Psychon. Sci., 10, 135-136.

Conrad, R., and Hull, A.J. (1967). Copying alpha and numeric codes by hand: An experimental study. J. appl. Psychol., 51, 444-448.

Conrad, R., and Hull, A.J. (1966). The role of the interpolated task in short-term retention. *Quart. J. exp. Psychol.*, 18, 266-269.

Conrad, R., and Rush, M.L. (1965). On the nature of short-term memory encoding by the deaf. *J. Speech Hear. Dis.*, 30, 336-343.

Cooley, R.K., and McNulty, J.A (1967). Recall of individual CCC trigrams over short intervals of time as a function of mode of presentation. *Psychon. Sci.*, 9, 543-544.

Corballis, M.C. (1966). Rehearsal and decay in immediate recall of visually and auditorily presented items. *Canad. J. Psychol.*, 20, 43-51.

Corballis, M.C. (1969). Patterns of rehearsal in immediate memory. *Brit. J. Psychol.*, 60, 41-49.

Corballis, M.C., and Loveless, T. (1967). The effect of input modality on short-term serial recall. *Psychon. Sci.*, 7, 275-276.

Corcoran, D.W.J., and Weening, D.L. (1969). On the combination of evidence from the eye and ear. *Ergonomics*, 12, 383-394.

Costello, M.R. (1957). A study of speech reading as a developing language process in deaf and in hard of hearing children. *Unpublished Ph. D. Thesis*, Northwestern University, Evanston.

Craik, F.I.M., and Masani, P.A. (1967). Age differences in the temporal integration of language. *Brit. J. Psychol.*, 58, 291-299.

Crawford, J., Hunt, E., and Peak, G. (1966). Inverse forgetting in short-term memory. *J. exp. Psychol.*, 72, 415-422.

Crossman, E.R.F.W. (1961). Information and serial order in human immediate memory. In E.C. Cherry (Ed) Information theory. *Proceedings of the fourth London symposium*. London: Butterworth.

Crowder, R.G. (1965). Verbal short-term memory as a function of degree of learning on a perceptual-motor interpolated activity. *Diss. Abstr.*, 26, 2877.

Crowder, R.G. (1967). Prefix effects in immediate memory. *Canad. J. Psychol.*, 21, 450-461.

Crowder, R.G. (1969). Phonic interference and the prefix effect. *J. verb. Learn. verb. Behav.*, 8, 302-304.

Crowder, R.G., and Morton, J. (1969). Precategorical acoustic storage (PAS). *Percept. Psychophys.*, 5, 365-373.

Dale, H.C.A. (1964). Retroactive interference in short-term memory. *Nature*, 203, 1408.

Dale, H.C.A. (1967). Semantic similarity and the A-B, C-D paradigm in STM. *Psychon. Sci.*, 9, 79-80.

Dale, H.C.A., and Gregory, M. (1966). Evidence of semantic coding in short-term memory. *Psychon. Sci.*, 5, 75-76.

Dallett, K.M. (1964). Effects of a redundant prefix on immediate recall. *J. exp. Psychol.*, 67, 296-298.

Dallett, K.M. (1965). "Primary memory". The effects of redundancy upon digit repetition. *Psychon. Sci.*, 3, 237-238.

Damsteegt, B.C. (1969). Over relevante kenmerken van fonemen en de nederlandse *R*. *Public lecture*. Leiden. Univ. Pers.

De Groot, A.W. (1964). *Inleiding tot de algemene taalwetenschap*. Second Ed. Groningen: Wolters.

Delattre, P.C., Liberman, A.M., and Cooper, F.S. (1955). Acoustic loci and transitional cues for consonants. J. acoust. Soc. Am., 27, 769-773.

DiMascio, A. (1959). Learning characteristics of nonsense syllables: A function of letter frequency. Psychol. Rep., 5, 585-591.

Doehring, D.G. (1960). Visual spatial memory in aphasic children. J. Speech Hear Res., 3, 138- 149.

Doppelt, J.E., and Wallace, W.L. (1955). Standardization of the Wechsler Adult Intelligence Scale for older persons. J. abnorm. (soc.) Psychol., 5, 312-330.

Duncan, C.P. (1949). The retroactive effect of electroshock on learning. J. comp. physiol. Psychol., 42, 32-44.

Eijkman, L.P.H. (1955). Phonetiek van het Nederlands. Second Ed. Haarlem: Bohn.

Eimas, P. (1963). The relation between identification and discrimination along speech and non-speech continua. Language and Speech, 6, 206-217.

Eriksen, C.W., and Collins, J.F. (1967). Some temporal characteristics of visual pattern perception J. exp. Psychol., 74, 476- 484.

Erkens, L.J.M.G. (1968). De verwerking van temporeel gestructureerde stimuli. Unpubl. Thesis, Psychol. Lab., Univ. of Nijmegen

Fant, G. (1967). Auditory patterns of speech. In W. Wathen-Dunn (Ed.) Models for the perception of speech and visual form. Cambridge, Mass.. M.I.T. Press.

Fraisse, P. (1969). Why is naming longer than reading? In W.G. Koster (Ed.) Attention and Performance II: Proceedings of the Donders centenary symposium on reaction time. Amsterdam: North Holland Publ. Co.

Frisina, D.R. (1955). A psychological study of the mentally retarded deaf child. Unpublished Ph. D. Thesis, North western University, Evanston.

Fry, D., Abramson, A.S., Eimas, P., and Liberman, A.M. (1962). The identification and discrimination of synthetic vowels. Language and Speech, 5, 171-189.

Fuller, C. (1959). A study of the growth and organization of certain mental abilities in young deaf children. Unpublished Ph. D. Thesis, Northwestern University, Evanston.

Gerard, R.W. (1953). What is memory? Scient. Amer., 189, (3), 118-136.

Gilbert, J.G. (1941). Memory loss in senescence. J. abnorm. (soc.) Psychol., 36, 73-86.

Glanzer, M., and Clark, W.H. (1962). Accuracy of perceptual recall: An analysis of organization. J. verb. Learn. verb. Behav., 1, 289-299.

Glanzer, M., and Clark, W.H. (1963). The verbal loop hypothesis. Binary numbers. J. verb. Learn. verb. Behav., 2, 301-309.

Goetzinger, C.P. and Huber, T.G. (1964) A study of immediate and delayed visual retention with deaf and hearing adolescents. Amer. Ann. Deaf, 109, 297-305.

Green, D.M., and Swets, J.A. (1966). Signal detection and psychophysics. New York: Wiley.

Groninger, L.D. (1966). Natural language mediation and covert rehearsal in short-term memory. Psychon. Sci., 5, 135-136.

Grossman, S.P. (1967). Textbook of Physiological Psychology. New York: Wiley.

Gurowitz, E.M. (1969). The molecular basis of memory. Englewood Cliffs, N.J.: Prentice Hall.

Halle, M. and Stevens, K.N. (1959). Analysis by synthesis. In W. Wathen-Dunn and L.E. Woods (Eds.) Proceedings of the seminar on speech compression and processing. Vol. 2. AFCRC-TR-59-198, USAF Camb. Res. Ctr., Paper D7.

Halle, M., and Stevens, K.N. (1962). Speech recognition: A model and a program for research. IRE Trans. PGIT, IT-8, 155-159.

Harris, C.M. (1953). A study of the building blocks of speech. J. acoust. Soc. Amer., 25, 962-969.

Harris, K.S. (1963). Behavior of the tongue in the production of some alveolar consonants. J. acoust. Soc. Amer. 35, 784 (Abstract).

Harris, K.S., Schrey, M.M., and Lysaught, G.F. (1962). Component gestures in the production of oral and nasal labial stops. J. acoust. Soc. Amer., 34, 743 (Abstract).

Hayes, J.R.M. (1952). Memory span for several vocabularies as a function of vocabulary size. In: Quart, Progr. Rep., Cambridge, Mass.: Acoustics Lab., M.I.T., Jan.-June.

Hebb, D.O. (1949). The organization of behaviour: A neuropsychological theory. New York: Wiley.

Hellyer, S. (1962). Supplementary report: Frequency of stimulus presentation and short-term decrement in recall J. exp Psychol., 64, 650.

Heriot, J.T., and Coleman, P.D (1962). The effect of electroconvulsive shock on retention of a modified "one-trial" conditioned avoidance. J. comp. physiol. Psychol., 55, 1082-1084.

Heron, A., and Craik, F.I.M. (1964). Age differences in cumulative learning of meaningful and meaningless meterial. Scand. J. Psychol., 5, 209-217.

Hintzman, D.L. (1967). Articulatory coding in short-term memory. J. verb. Learn. verb. Behav., 6, 312-316.

Hiskey, M.S. (1956). A study of the intelligence of deaf and hearing children. Amer. Ann. Deaf, 101, 329-339.

Hofmarksrichter, K. (1931). Visuelle Kompensation und Eidetik bei Taubstummen Vergleichende Untersuchungen an Taubstummen und Vollsinnigen. Arch. ges. Psychol., 82, 329-402.

House, A., Stevens, K.N., Sandel, T., and Arnold, J. (1962). On the learning of speechlike vocabularies. J. verb. Learn. verb. Behav., 1, 133-143.

Inglis, J., and Caird, W.K. (1963). Age differences in successive responses to simultaneous stimulation. Canad. J. Psychol., 17, 98-105.

Jensen, A.R., and Rowher, W.D. (1966). The Stroop color-word test: a review. Acta Psychol., 25, 36-93.

Johnson, S.C. (1967). Hierarchical clustering schemes. Psychometrika, 32, 241-254.

Katzenberger, R. (1965). Dimensionen des Gedachtnisses. Zeitschr. exp. angew. Psychol., 12, 451-492.

Kay, H.E. (1959). Theories of learning and ageing. In J.E. Birren (Ed.) Handbook of ageing and the individual. Chicago: Univ. of Chicago Press.

Kelley (1954). On "span memory", quoted in Ross, J. (1961) A factor test of a memory model Ed Test. Service, Res. Bull., 61, 14

Kempen, G. (1970). Memory for word and sentence meanings. A set-feature model. Doctoral dissertation, Univ of Nijmegen

Kendler, H.H., and Kendler, T.S. (1961). The effect of verbalization on reversal shifts in children. Science, 134, 1619-1620.

Kendler, H.H. and Kendler, T.S. (1962). Vertical and horizontal processes in problem solving. Psychol. Rev., 69, 1-16.

Kendler, H.H., and Kendler, T.S. (1968) Mediation and conceptual behavior. In K.W. Spence and J.T. Spence (Eds.) The Psychology of learning and Motivation. Vol. 2, New York: Academic Press.

Keppel, G , and Underwood, B.J (1962) Proactive inhibition in short-term retention of single items. J verb. Learn. verb. Behav., 1, 153-161.

Kilpatrick, J.M.L. (1963). The effect of verbalization on the reproduction of visual forms. Teacher Deaf, 61, 147-155.

Kimura, D. (1961a). Some effects of temporal-lobe damage on auditory perception. Canad. J. Psychol., 15, 156-165.

Kimura, D. (1961b). Cerebral dominance and the perception of verbal stimuli. Canad. J. Psychol., 15, 166-171.

Kimura, D. (1964). Left-right differences in the perception of melodies. Quart. J. exp. Psychol , 16, 355-358.

Klein, G.S. (1964) Semantic power measures through the interference of words with color naming. Amer J. Psychol , 57, 576-588.

Kral, V.A. (1958). Neuro-psychiatric observations in an old people's home. Studies of memory dysfunction in senescence. J. Geront., 13, 169-176.

Kruisinga, R.J.H. (1956). Slechthorendheid en het verstaan van spraak. Ph. D. Thesis. Univ. of Groningen

Kwaaitaal, T., and Roskam, E. (1968). Analysis of variance. Programme VARIAN/01 Psychol. Lab., Univ. of Nijmegen. May 1968.

Ladefoged, P., and Broadbent, D.E. (1960) Perception of sequence in auditory events. Quart. J. exp. Psychol., 12, 162-170.

Landauer, T.K. (1962). Rate of implicit speech. Percept. mot. Skills, 15, 646.

Landis, C. (1957). Determinants of flicker. In L.M.N. Bach (Ed) Tulane Symposium on Flicker, Tulane Univ. and U.S. Army E.R.D.L.

Laugherry, K.R., Lachman, R., and Danserean, D. (1965). Short-term memory: Effects of item pronunciation time. Proceed. 73 Ann. Conv. A.P.A., 83-84.

Lenneberg, E.H. (1961). Color naming, color recognition, color discrimination a reappraisal. Percept. Mot. Skills, 12, 375-382.

Leukel, F.A. (1957). A comparison of the effects of E.C.S. and anesthesia on acquisition of the maze habit. J. comp. physiol. Psychol., 50, 300-306.

Lewis, M.E., and Teichner, W.H. (1964). The effect of various modes of rehearsal on short-term recall. Navtradevcen Tech. Rep., 1303-1

Liberman, A.M. (1957). Some results of research on speech perception. J. acoustic Soc. Amer., 29, 117-123.

Liberman, A.M., Cooper, F.S., Harris, K.S., and MacNeilage, P.F. (1962). A motor theory of speech perception. In Proceedings of the speech communication seminar. Vol. 2. Stockholm: Royal Inst. of Technol.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., and Studdert-Kennedy, M. (1967). Perception of the speech code. Psychol. Rev., 74, 431-461.

Liberman, A.M., Cooper, F.S., Harris, K.S., MacNeilage, P.F., and Studdert-Kennedy, M. (1967).Some observations on a model for speech perception. In W. Wathen-Dunn (Ed.) Models for the perception of speech and Visual Form:Proceedings of a Symposium. Cambridge, Mass.: M.I.T. Press.

Liberman, A.M., Delattre, P.C., and Cooper, F.S. (1952). The role of selected stimulus-variables in the perception of the unvoiced stop consonants. Amer. J. Psychol., 65, 497-516.

Liberman, A.M., Delattre, P.C., and Cooper, F.S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. Language and Speech, 1, 153-167.

Liberman, A.M., Harris, K.S., Hoffman, H. and Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. J. exp. Psychol., 54, 358-368.

Liberman, A.M., Harris, K.S. Kinney, J., and Lane, H. (1961). The discrimination of relative onset-time of the components of certain speech and non-speech patterns. J. exp. Psychol., 61, 379-388.

Liberman, A.M., Ingemann, F., Lisker, L., Delattre, P., and Cooper, F.S. (1959). Minimal rules for speech synthesis. J. acoust. Soc. Amer., 31, 1490-1499.

Loess, H., and McBurney, J. (1965). Short-term memory and retention-interval activity. Proceed. 73 Ann. Conv. A.P.A., 85-86.

MacKay, H.A., and Inglis, J. (1963). The effect of age on short-term auditory storage. Gerontologia, 8, 193-200.

Mackworth, J.F. (1963). The relation between the visual image and post-perceptual immediate memory. J. verb. Learn. verb. Behav., 2, 75-85.

MacNeilage, P.F. (1963). Electromyographic and acoustic study of the production of certain final clusters. J. acoust. Soc. Amer., 35, 461-463.

MacNeilage, P.F. (1964). Electromyographic study of the coarticulation of vowels and consonants. J. acoust. Soc. Amer., 36, 1989 (Abstract).

Madsen, M.C., and McGaugh, J.L. (1961). The effect of E.C.S. on one-trial avoidance learning. J. comp. physiol. Psychol., 54, 522-523.

Mandler, G. (1968). Association and organization: Facts, fancies, and theories. In T.R. Dixon and D.L. Horton (Eds.) Verbal behavior and general behavior theory. Englewood Cliffs, N.J.: Prentice Hall.

Martin P.R., and Fernberger, S.W. (1929). Improvement in memory span. Amer. J. Psychol., 41, 94-97.

McGaugh, J.L. (1961). Facilitative and disruptive effects of strychnine sulphate on maze-learning. Psychol. Rep., 8, 99-104.

McGeoch, J.A., and McDonald, W.T. (1931). Meaningful relation and retroactive inhibition. Amer. J. Psychol., 43, 579-588.

McGhie, A.J., Chapman, J., and Lawson, J.S. (1965). Changes in immediate memory with age. Brit. J. Psychol., 56, 69-75.

Mechanic, A. (1962). Effects of orienting task, practice, and incentive on simultaneous incidental and intentional learning. J. exp. Psychol., 64, 393-399.

Melton, A.W. (1963). Implications of short-term memory for a general theory of memory. J. verb. Learn. verb. Behav., 2, 1-21.

Miller, G.A. (1956a). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychol. Rev., 63, 81-97.

Miller, G.A. (1956b). The perception of speech. In M. Halle (Ed.) For Roman Jakobson. Pp. 353-359. The Hague: Mouton.

Miller, G.A. (1958). Free recall of redundant strings of letters. J. exp. Psychol., 56, 485-491.

Miller, G.A. (1962). Decision units in the perception of speech. IRE Transactions on information theory, IT-8, 81-83.

Miller, G.A., Bruner, J.S., and Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. J. gen. Psychol., 50, 129-139.

Miller, G.A., and Nicely, P.E. (1955). An analysis of perceptual confusions among some English consonants. J. acoust. Soc. Amer., 27, 338-352.

Miller, G.A., and Selfridge, J.A. (1950). Verbal context and the recall of meaningful material. Amer. J. Psychol., 63, 176-185.

Milner, B. (1959). The memory defect in bilaterial hippocampal lesions. Psychiat. Res. Rep. 11. 43-52.

Milner, B. (1966). Amnesia following operation on the temporal lobes. In C.W.M. Whitty and O.L. Zangwill (Eds.) Amnesia. London: Butterworth, Pp. 109-133.

Milner, B. (1968). Neuropsychological evidence for differing memory processes. Abstract for the symposium on short-term and long-term memory. Proceedings of the 18th International congress of Psychology, Moscow, 1966. Amsterdam: North-Holland Publishing Co.

Milner, B., and Penfield, W. (1955). The effects of hippocampal lesions on recent memory. Trans. Amer. Neurol. Ass., 80, 42-48.

Moray, N., and Taylor, A. (1958). The effect of redundancy in shadowing one of two dichotic messages. Language and Speech, 1, 102-109.

Morsh, J.E. (1936). Motor performance of the deaf. Comp. Psychol. Monogr., 66.

Morton, J. (1964a). A model for continuous language behavior. Language and Speech, 7, 40-70.

Morton, J. (1964b). A preliminary functional model for language behavior. Int. Audiol., 3, 216-225.

Morton, J. (1964c). The effects of context upon speed of reading, eye-movements, and eye-voice span. Quart. J. exp. Psychol., 16, 340-354.

Morton, J. (1964d). The effects of context on the visual duration threshold for words. Brit. J. Psychol., 55, 165-180.

Morton, J. (1968a). Considerations of grammar and computation in language behavior. Studies in Language and Language Behavior, 6, 499-545.

Morton, J. (1968b). Selective interference in immediate recall. Psychon. Sci., 12, 75-76.

Morton, J. (1969a). Interaction of information in word recognition. Psychol. Rev., 76, 165-178.

Morton, J. (1969b). Categories of interference: Verbal mediation and conflict in card sorting. Brit. J. Psychol., 60, 329-346.

Morton, J. (1970). The modalitv of post-stimulus cueing of recall order in the memory span. In press.

Morton, J., and Broadbent, D.E (1967) Passive vs. active recognition models, or is your homunculus really necessary? In W. Wathen-Dunn (Ed.) Models for the perception of speech and visual form. Cambridge, Mass.. M.I.T. Press.

Morton, J. and Holloway, C.M. (1970). The absence of a cross-modal suffix effect in short-term memory. In press.

Mowbray, G.H., and Gebhard, J.W. (1960). Differential sensivity of peripheral retina to intermittent white light. Science, 132, 672-674.

Mowbray, G.H., Gebhard, J.W., and Byham, C.L. (1956). Sensivity to changes in the interruption rate of white noise. J. Acoust. Soc. Amer., 28, 106-110.

Murdock, B.B., Jr. (1961). The retention of individual items. J. exp. Psychol , 62, 618-625

Murdock, B.B., Jr. (1967a). Recent developments in short-term memory. Brit. J Psychol., 58, 421-433.

Murdock, B.B., Jr. (1967b). Theoretical note: A fixed-point model for short-term memory. J. math. Psychol., 4, 501-506.

Murdock, B.B., Jr. (1967c). Auditory and visual stores in short term memory In A.F. Sanders (Ed.) Attention and Performance. Amsterdam: North-Holland Publ. Co.

Murray, D.J. (1965a). Vocalization-at-presentation, and immediate recall. Nature, 207, 1011-1012.

Murray, D.J. (1965b). The effects of white noise upon the recall of vocalized lists. Canad. J. Psychol., 19, 333-345.

Murray, D.J. (1966). Vocalization-at-presentation and immediate recall, with varying recall methods. Quart. J. exp. Psychol., 18, 9-18.

Murray, D.J. (1967). The role of speech responses in short-term memory. Canad. J. Psychol., 21, 263-276.

Naffin, P. (1959). Die psychologischen Voraussetzungen der Erziehung des Taubstummen Kindes. In H. Hetzer (Ed.) Paedagogische Psychologie. Gottingen: Hogrefe.

Neisser, U. (1967). Cognitive psychology. New York: Appleton-Century-Crofts.

Norman, D.A. (1968). Attention and retention. Psychol. Rev. 76, 522-536.

Norman, D.A. (1969). Memory and attention: An introduction to human information processing. New York: Wiley.

Norman, D.A., and Wickelgren, W.A. (1965). Short-term recognition memory for single digits and pairs of digits. J. exp. Psychol., 70, 479-489.

Norman, D.A. and Wickelgren, W.A. (1969). Strength theory of decision rules and latency in retrieval from short-term memory. J math. Psychol., Vol. 6, No. 2, June.

Oldfield, R.C. (1954). Memory mechanisms and the theory of schemata. Brit. J. Psychol., 45, 14-23.

Olsson, J.E. (1963). The influence of language experience on visual memory span. Unpublished Master's thesis, Catholic University of America.

Osgood, C.E. (1949). The similarity paradox in human learning. Psychol. Rev. 56, 132-143.

Osgood, C.E. (1953). Method and theory in experimental psychology. New York: Oxford Univ. Press.

Penfield, W., and Milner, B. (1958). Memory deficit produced by bilateral lesions in the hippocampal zone. AMA Arch. Neurol. Psychiat., 79, 475-497.

Penfield, W., and Roberts, L. (1959). Speech and brain mechanisms. Princeton, N.J.: Princeton Univ. Press.

Peterson, L.R. (1963). Immediate memory: Data and theory. In C.N. Cofer and B.S. Musgrave (Eds.) Verbal behaviour and learning. Problems and Processes. New York: McGraw-Hill, Pp 336-353.

Peterson, G.E. and Barney, H.L. (1952). Control methods used in a study of the vowels. J. acoust. Soc. Amer , 24, 175-184.

Peterson, L.R., and Peterson, M.J. (1959). Short-term retention of individual verbal items. J. exp. Psychol., 58, 193-198.

Pinter, R., and Paterson, D.G. (1917). A comparison of deaf and hearing children in visual memory for digits. J. exp. Psychol., 2, 76-88.

Pollack, I. (1952). The assimilation of sequentially encoded information, HumARO memo. Rep., No. 25.

Pollack, I. (1953). Assimilation of sequentially encoded information. Amer. J. Psychol., 66, 421-435.

Pollack, I. (1963). Interference, rehearsal, and short-term retention of digits. Canad. J. Psychol., 17, 380-392.

Pollack, I., Johnson, L.B., and Knaff, P.R. (1959). Running memory span. J. exp. Psychol., 57,137-146.

Posner, M.I. (1963). Immediate memory in sequential tasks. Psychol. Bull., 60, 333-349.

Posner, M.I. (1966). Components of skilled performance. Science, 152, 1712-1718.

Posner, M.I. (1967). Short term memory systems in human information processing. Acta Psychol., 27, 267-284.

Posner, M.I., and Rossman, E. (1965). The effect of size and location of informational transforms upon short-term retention. J. exp. Psychol., 70, 496-505.

Postman, L. (Ed.) (1962). Psychology in the making: Histories of selected research problems. New York: Knopf.

Postman, L., (1964). Short-term memory and incidental learning. In: A.W. Melton (Ed.) Categories of human learning. New York: Academic Press, Pp. 145-201.

Poulton, E.C., and Brown, C.H. (1967). Memory after reading aloud and reading silently. Brit. J. Psychol., 58, 219-222.

Pribram, K.H. (1969). The neurophysiology of remembering. Scient. Amer., 220, (1), 73-86.

Pylyshyn, Z.W. (1965). The effect of a brief interpolated task on short-term retention. Canad. J. Psychol., 19, 280-287.

Richet, C. (1886). Les origines et les modalités de la mémoire. Rev. Philos., 21, 235-242.

Rosenbush, M.H., and Gardner, D.B. (1968). Reproduction of visual and auditory rhythm patterns by children.Percept. mot. Skills, 26, 1271-1276.

Rosenzweig, M.R. (1969). The effect of changes in the Central Nervous System on memory. Doc. Geigy, Basle, Memory, 5-6.

Roskam, E., and Brandsma, I. (1969). Programma HICLU. Progr. Bull. No. 13, Psychol. Lab., Afd. Mathem. Psychol., Univ. of Nijmegen.

Rubenstein, H., and Aborn, M. (1954). Immediate recall as a function of degree of organization and length of study period. J. exp. Psychol., 48, 146-152.

Russel, W.R., and Nathan, P.W. (1946). Traumatic amnesia. Brain, 69, 280-300.

Sanders, A.F. (1961). Rehearsal and recall in immediate memory. Ergonomics, 4, 29-34.

Sanders, A.F. (1964). Recent onderzoek over het onmiddellijk geheugen. Ned. Tijdschr. Psychol., 19, 165-185.

Sanders, A.F. (1968). Short-term memory for spatial positions. Ned. T. Psychol., 23, 1-15.

Severin, F.T., and Rigby, M.K. (1963). Influence of digit grouping on memory for telephone numbers. J. appl. Psychol., 47 (2).

Shankweiler, D., and Studdert-Kennedy, M. (1967). Identification of consonants and vowels presented to left and right ears. Quart. J. exp. Psychol., 19, 59-63.

Sperling, G.A. (1960). The information available in brief visual presentation. Psychol. Monogr., 74, Whole No. 498.

Sperling, G.A. (1962). Auditory interference with a visual memory task. Paper read at Eastern Psychological Assoc., Atlantic City, New Jersey.

Sperling, G.A. (1963). A model for visual memory tasks. Hum. Factors, 5, 19-31.

Sperling, G.A. (1967). Successive approximations to a model for short-term memory. In A.F. Sanders (Ed.) Attention and Performance. Amsterdam: North-Holland Publishing Co.

Stevens, K.N. (1960). Toward a model for speech recognition. J. acoust. Soc. Amer., 36, 1989 (Abstract).

Stevens, K.N. and Halle, M. (1967). Remarks on analysis by synthesis and distinctive features. In W. Wathen-Dunn (Ed.) Models for the perception of speech and visual form: Proceedings of a symposium. Cambridge, Mass.: M.I.T. Press.

Stevens, K.N., Öhman, S.E.G., Studdert-Kennedy, M., and Liberman, A.M. (1964). Crosslinguistic study of vowel discrimination. J. acoust. Soc. Amer., 36, 1989 (Abstract).

Süllwold, F. (1964). Das unmittelbare Behalten und seine denkpsychologische Bedeutung. Göttingen: Hogrefe.

Taubman, R.E. (1950). Studies in judged number. II. The judgement of visual number. J. gen. Psychol. 43, 195-219.

Terman, L.M. (1916). The measurement of intelligence. New York: Houghton Mifflin Co.

Thomassen, A.J.W.M. (1966a). Performance in an interpolated task during STM retention. Unpublished report, Psychol. Lab., Univ. of Nijmegen.

Thomassen, A.J.W.M. (1966b). Een nadere analyse van de verwarringsmatrijzen van Conrad (1964). Unpublished report, Psychol. Lab., Univ. of Nijmegen.

Thompson, R., and Dean, W. (1955). A further study of the retroactive effect of E.C.S., J. comp. physiol. Psychol., 48, 488-491.

Thorpe, C.E., and Rowland, G.E. (1965). The effect of 'natural' grouping numerals on short-term memory. Human Factors, 7, 38-44.

Thurstone, L.L., and Thurstone, T.G. (1941). Factorial studies of intelligence. Psychometric Monogr., No. 2. Chicago.

Treisman, A.M. (1964). Verbal cues, language and meaning in selective attention. Amer. J. Psychol., 77, 206-219.

Underwood, B.J., and Schulz, R.W. (1960). Meaningfulness and Verbal learning. New York: Lippincott.

Van Berckel, J., Brandt Corstius, H., Mokken, R., and Van Wijngaarden, A. (1965). Formal Properties of Newspaper Dutch. Amsterdam: Mathematisch Centrum.

Van Dongen, A. (1962). Fonetiek- en Artikulatiekursus. Third Ed. St. Michielsgestel: Inst. voor Doven.

Von Békésy, G. (1960). Experiments in Hearing. New York: Mc. Graw-Hill.

Wallace, J.G. (1956). Some studies of perception in relation to age. Brit. J. Psychol., 47, 283-297.

Waugh, N.C., and Norman, D.A. (1965). Primary memory. Psychol. Rev., 72, 89-104.

Webster, J.C., and Chaney, R.B. (1967). Information and complex signal perception. In W. Wathen-Dunn (Ed.) Models for the perception of speech and visual form. Cambridge, Mass.: M.I.T. Press.

Weiskrantz, L. (1964). Neurological Studies and animal behavior. Brit. med. Bull., 20, 49-53.

Welford, A.T. (1958). Ageing and human skill. London: Oxford Univ. Press.

White, C.T., and Cheatham, P.G. (1959). Temporal numerosity: IV. A comparison of the major senses. J. exp. Psychol., 58, 441-444.

Whorf, B.L. (1956). Language, thought, and reality. New York: Wiley.

Wickelgren, W.A. (1964). Size of rehearsal group and short-term memory. J. exp. Psychol., 68, 413-419.

Wickelgren, W.A. (1965a). Acoustic similarity and retroactive interference in short-term memory. J. verb. Learn. verb. Behav., 4, 53-61.

Wickelgren, W.A. (1965b). Short-term memory for repeated and non-repeated items. Quart. J. exp. Psychol., 17, 14-25.

Wickelgren, W.A. (1965c). Short-term memory for phonemically similar lists. Amer. J. Psychol., 78, 567-574.

Wickelgren, W.A. (1965d). Acoustic similarity and intrusion errors in short-term memory. J. exp. Psychol., 70, 102-108.

Wickelgren, W.A. (1965e). Distinctive features and errors in short-term memory for English vowels. J. acoust. Soc. Amer., 38, 583-588.

Wickelgren, W.A. (1966a). Consolidation and retroactive interference in short-term recognition memory for pitch. J. exp. Psychol., 72, 250-259.

Wickelgren, W.A. (1966b). Phonemic similarity and interference in short-term memory for single letters. J. exp. Psychol., 71, 396-404.

Wickelgren, W.A. (1966c). Associative intrusions in short-term recall. J. exp. Psychol., 72, 853-858.

Wickelgren, W.A. (1966d). Short-term recognition memory for single letters and phonemic similarity of retroactive interference. Quart. J. exp. Psychol., 18, 55-62.

Wickelgren, W.A. (1966e). Distinctive features and errors in short-term memory for English consonants. J. acoust. Soc. Amer., 39, 388-389.

Wickelgren, W.A. (1967). Rehearsal grouping and hierarchical organization of serial position cues in short-term memory. Quart. J. exp. Psychol., 19, 97-102.

Wickelgren, W.A. (1968). Sparing of short-term memory in an amnesic patient: Implications for strength theory of memory. Neuropsychol., 6, 235-244.

Wickelgren, W.A. (1969a). Associative strength theory of recognition memory for pitch. J. math. Psychol., 6, 13-61.

Wickelgren, W.A. (1969b). Context-sensitive coding, associative memory, and serial order in (speech) behavior. Psychol. Rev., 76, 1-15.

Wickelgren, W.A., and Norman, D.A. (1966). Strength models and serial position in short-term memory. J. math. Psychol., 3, 316-347.

Wickens, D.D., Born, D.G., and Allen, C.K. (1963). Proactive inhibition and item similarity in short-term memory. J. verb. Learn. verb. Behav., 2, 440-445.

Williams, M. (1950). Memory studies in electric convulsion therapy. J. Neurol. Neurosurg. Psychiat., 13, 30-35.

Williams, M., and Zangwill, O.L. (1952). Amnesia following concussional head injury. J. Neurol. Neurosurg. Psychiat., 15, 54-59.

Woodhead, M.M. (1966). Varying the number of alternatives in short-term recall. Brit. J. Psychol., 57, 45-52.

Woodrow, H. (1939). The common factors in fifty-two mental tests. Psychometrika, 4, 99-108.

Woodworth, R.S., and Schlossberg, H. (1954). Experimental Psychology. New York: Holt.

STELLINGEN

I

Het feit dat de spraakwaarneming geschiedt in termen van segmenten van zeer uiteenlopende omvang vraagt om een theoretisch kader waarin zowel preperceptuele als postperceptuele aspecten van het proces zijn opgenomen.

II

Dat in het onmiddellijk geheugen associatieve relaties tussen de elementen van een reeks worden gevormd is niet overtuigend aangetoond.

III

Bij het zoeken naar een verklaring van het modaliteitseffect op de seriële positiecurve bij de onmiddellijke reproductie van verbale items dient rekening te worden gehouden met de ruimere mogelijkheden voor 'rehearsal' bij visuele aanbieding.

IV

Integratie van linguistische en psychologische benaderingen van verbale gedragsvormen is, behalve op het niveau van de semantiek en van de grammatica, ook gewenst op het niveau van de fonetiek en de articulatie.

V

Tijdens het roepen van slogans zijn hogere mentale functies, zoals die bv. nodig zijn voor het beoordelen van de juistheid van een zin, aanzienlijk gereduceerd.

M. Hammerton, Interference between low information verbal output and a cognitive task. Nature, 1969, 222, (No. 5189), p. 196.

VI

Deconditionering van bepaalde vormen van automutilatief gedrag bij oligofrenen is in principe mogelijk.

VII

Voor het spraakonderwijs aan doven verdient het beproeven van methoden waarmee rechtstreeks articulatiepatronen kunnen worden overgebracht meer aandacht.

VIII

Een verklaring voor het feit dat er zoveel meer mannelijke dan vrouwelijke stotteraars voorkomen houdt wellicht verband met een verschillende rol die kinaesthetische en proprioceptieve terugkoppeling spelen bij de sexen.

IX

Afschaffen van 'stadslichten' als verlichtingsmogelijkheid op auto's is uit veiligheidsoverwegingen gewenst.

X

Dat weinig mensen in staat zijn zonder grove overschatting tegen een wand de hoogte aan te geven van de schrijftafel waaraan zij dagelijks hun werk verrichten, moet eerder worden verklaard in termen van een waarnemingstheorie dan in samenhang met het eigen oordeel over het aan die tafel geleverde werk.

Nijmegen, 23 juni 1970 A.J.W.M. Thomassen