

INFORMATION PROCESSING
IN SHORT-TERM MEMORY TASKS

Stephen Monsell

Wadham College

Thesis submitted for the degree of Doctor of Philosophy in the
University of Oxford, July 1973.

ABSTRACT

This thesis is about the nature of a man's immediate memory for a short sequence of verbal items. Two components of immediate memory may be distinguished: an 'active' memory component, which is progressively destroyed within a few seconds if, subsequent to presentation of the sequence, the man is required to perform a distracting task which prevents rehearsal of the items, and a residual 'inactive' component which is more permanent. My concern is with behavioural evidence relating to the question of how the attributes of a verbal sequence are represented in active memory subsequent to identification of the constituent items. It is possible selectively to probe the retrieval of different attributes of the sequence by the use of appropriate tasks. Evidence on the performance of normal adult subjects in various short-term memory tasks, from the literature and from my own experiments, is extensively reviewed. There emerges a fundamental contrast between two groups of tasks: on the one hand, tasks which require the subject to indicate whether a test item was present in the memorised sequence (item recognition) or to judge the relative recency with which items have occurred; on the other, tasks which require the subject to base his response on the order or precise position of items in the sequence. To account for the nature of the contrast I put forward a general hypothesis which is an amended and elaborated version of parts of a general model of memory described by Morton (1970). It is argued that subsequent to identification of an item, representations of it may temporarily be held in active memory by two distinct storage mechanisms. The first is as a decaying trace at the neural unit, called (following Morton) a 'logogen', responsible for the identification of that item.

This trace may be treated as possessing a unidimensional 'strength' whose magnitude is dependent on how recently the item has occurred. Item recognition and judgements of recency are held to be mediated by assessment of the test item's trace strength with respect to a decision criterion located on the strength continuum. Secondly, the item may also be represented as one of several maintained in a serially-organised limited-capacity storage mechanism called (following Morton) the 'response buffer'. This holds a small number of the verbal responses recently made available by the logogens, coded as a string of articulatory descriptions or commands. It has two crucial properties. Firstly, items represented in it as potential responses may be fed back in sequence to the logogens for re-identification (the process of sub-vocal rehearsal). Secondly, since the serial order in which items enter the response buffer is retained as an intrinsic property of its structure, retrieval from it provides a ready mode of access to the order or position of items in the memorised sequence, which item-traces at logogens do not.

The thesis falls into two parts. The first three chapters contain a theoretical review of the literature. In Ch.1, the active/inactive distinction is introduced with some reference to the historical background (1.1). Some points of technique are raised (1.2), the aims of the present work are outlined (1.3) and the two-store hypothesis is described (1.4).

In Ch.2, evidence is reviewed from short-term memory experiments in which accuracy is the main dependent variable, beginning with experiments in which either serial recall of the whole sequence or probed recall of a single item is required. The hypothesis that the order of items is retained in active memory as an intrinsic property

of the memory structure is contrasted with theories (e.g. Wickelgren, 1972) emphasising the formation of temporary associations between the representations of items adjacent in the list and/or between item and position representations. It is concluded that inter-item associations play no major role in active memory, and that item-position associations cannot account both for the partial independence of order and item errors and for the relationship between order errors and phonemic similarity. Conrad's (1965) model is introduced as a precursor of my own. It is concluded that items are recalled primarily from the (ordered) response buffer, but that traces at logogens influence the availability of responses as guesses when items are wholly or partly lost from the response buffer (2.11). Some evidence is then described which suggests that variables may be identified which differentially influence retention of information about the order of items and about their occurrence (2.12). Experiments comparing the effectiveness of position, context or both, as cues for recall are argued to imply that order and position are coded by the same mechanism, but that access to items retained by it may be more direct given a position rather than a contextual cue (2.13). We then turn to experiments on short-term recognition and judgements of recency, which are discussed in relation to the trace strength theory of Wickelgren & Norman (1966) (2.21). It is concluded that an item's recent identification is represented as an exponentially-decaying trace located at a unique permanent representation of that item, to which access is direct (2.22). But order-recognition experiments provide little evidence that associations between items are represented in active memory in the same way (2.23). The two-store hypothesis is then applied to some general problems of the functions of active memory: in speech comprehension (2.31), as an 'address register'

for inactive memory (cf. Broadbent, 1971) (2.32) and as a working memory. Finally, the logogen system is discussed in relation to Sperling's (1967) 'recognition buffer' (2.34), and evidence for articulatory as opposed to acoustic coding in immediate memory is reviewed (2.35).

In Ch.3, I introduce the experimental paradigms pioneered by Sternberg (1966), which involve presentation on each trial of a different sub-span list for memorisation, followed by a probe to which reaction-time (RT) is measured. Several models are outlined of the nature, dynamics and format of the representations mediating performance in Sternberg's item-recognition paradigm (IRn), in which the subject must indicate whether the probe was or was not a member of the list. Predictions are derived from Sternberg's two scanning models and two versions of a trace strength hypothesis and compared to data available in the literature. The evidence favours the hypothesis that subjects perform the IRn task by judging the strength of the trace at the amodal representation (logogen) to which the probe provides direct access (3.1). Evidence from a version of the IRn paradigm in which the memory set remains constant from trial to trial is then reviewed; it is argued that the results obtained are necessarily equivocal with respect to active memory; nevertheless, the concept of trace strength may be of use in this context also (3.2). Next, we see that Sternberg's own experiments on contextual recall (CR), in which the subject must respond with the item following the probed item in the list, yield a pattern of results very different from that obtained in IRn. His own explanation is in terms of different strategies of search within a single store. In view of the failure of the exhaustive scanning model of IRn, the differences may be better accounted for by the two-store model. Since these RT

paradigms provide a relatively pure and sensitive way of probing the retrieval of different attributes from active memory, a strategy is suggested of looking for factors which differentially influence performance in IRn, CR and related tasks. Evidence already available suggests that these factors may include serial position, phonemic similarity and the availability of a position cue.

The second part of the thesis describes my own experiments. These employ Sternberg's varied-set IRn and CR paradigms and a task in which the subject must indicate the location of the probe in the list presented on that trial (Ln). In Ch.4, I raise some methodological issues inherent in the use of reaction time as a measure and give details of the apparatus and general procedure common to the experiments.

In Ch.5 are described three experiments on the effects of phonemic similarity between items of the memory set on performance in the three tasks. Five-consonant lists were presented visually at a fast rate; the probe followed immediately and subjects were instructed not to rehearse; vocabulary and other factors were held constant across tasks. In Exp.I (CR) and Exp.III (Ln) phonemic similarity increased both RT and error frequency. The relationship between serial position and the effects of similarity in both tasks suggested a slow forward serial search through a temporally-ordered phonemically-coded store, save perhaps for terminal items of the sequence. The effects of similarity were in both tasks diminished when the probed item was phonemically distinct from the list items and, in CR, when the response item was similarly isolated. In Exp. II, no comparable effects upon IRn performance were found. Moreover, whereas marked increases with serial position probed were obtained in both CR and Ln, in the IRn task RT showed an accelerating decrease

with serial position probed. The contrast between the two patterns of results is argued to be incompatible both with Sternberg's account in terms of two retrieval strategies and with Wickelgren's associative theory.

The first experiment described in Ch.6 (Exp.IV) demonstrated that IRn was not facilitated by a cue presented along with the probe which informed the subject which half (in terms of ordinal position) of the four digit memory set was relevant to his decision. This result is compatible with the proposal that access to the active memory representations involved in IRn is direct, and does not require a search through an ordered store. Exp.V compared the effects upon IRn and CR of giving with the probe a cue telling the subject in which of four spatio-temporal positions the probed item had been presented (if, in IRn, it was a positive probe). Experimental parameters and subjects were the same for both tasks. Information provided by the position cue had a marked facilitatory effect on CR, none on IRn (as in Exp.IV). The difference is held to reflect the fact that location in the list of the probed item and retrieval of its successor, in CR, requires a serial search of the response buffer, a search which may however be abbreviated if the subject knows at what position to begin the search.

In Ch.7 I give an account of a single experiment (Exp.VI) intended to test predictions critically distinguishing the models of IRn outlined in Section 3.1. Memory sets of 1-4 consonants were visually presented at a fast rate; subjects were told not to rehearse. Two variables were manipulated: 1) the recency of occurrence, prior to the present trial, of items presented as negative probes; 2) the recency of items in the memory set. The latter was varied by presenting the probe either immediately or following a short delay during

which rehearsal was prevented by a vowel-naming task. It was found that the more recently a negative probe had occurred, the slower and less accurately it was rejected; the effect increased with set size. The introduction of a filled delay between list and probe reduced the slopes of the serial position and set-size functions. The former showed monotonic recency effects. These and other aspects of the data are incompatible with any scanning model and favour the trace strength inspection hypothesis.

In Ch.8, two early exploratory experiments are described in which the task was again IRn. Mixed memory sets of digits and letters were presented; the probe was randomly either a letter or a digit. The question was whether the subject could use the semantic class of the probe to restrict a search to the appropriate subset of the list; such a strategy would not be possible according to a pure trace strength hypothesis. While Exp. VII provided no evidence for such a 'directed search' strategy, the data from Exp.VIII, in which the conditions were changed so as to maximise the possibility of such a strategy, showed an increase in RT with the size of the subset of the same class as the probe. However, it transpires that given the conditions of the experiment, this result does not rule out an explanation in terms of the trace strength inspection model. Changes are suggested which would make the experiment more decisive.

The last chapter is devoted to concluding comments, speculations and plans. The necessity for exploring the adequacy of the two-store model in relation to other areas of research is acknowledged; these include speech production perception and comprehension, attention, non-verbal memory and the study of patients with memory disorders due to brain damage. Further experiments are suggested to compare the effects in the various short-term memory tasks of semantic similarity,

rate of presentation, and interpolated material. In conclusion it is suggested that a greater emphasis on the representation of motor as well as perceptual events does much to clarify the nature of active memory.

Acknowledgements

My first thanks are due to my supervisor, Pat Rabbitt, who has encouraged me in times of flagging morale, and who in spite of a hectic schedule, found time to grapple with a draft of the first six chapters even more unwieldy than the final version. Nancy Waugh read a draft of Chapter III with meticulous care. Crispin Hasler made some perceptive comments on the first three chapters. I am also grateful to Graham Hitch for lending me a copy of his thesis; his review of the literature was particularly influential in helping me to formulate the arguments of the first part of Chapter II. In addition, I have benefited from several conversations with Saul Sternberg.

This thesis could not have been completed without the help of the following: Serena Bacon, who typed the text in two weeks in the face of impossible odds, Judy Turner, who typed the bibliography, Jeremy Broad, who photographed the illustrations, Maggie Gee and David Nightingale, who read the proofs, and Susan Grey, who checked the bibliography, monitored my sanity and kept smiling. I am deeply grateful to them all.

Any errors and omissions are, of course, entirely my own.

During the last four years I have received invaluable assistance from the staff, students and technicians of the Department of Experimental Psychology at Oxford. I owe a particular debt to Geoff Cumming and Stephen Fearnley, who taught me to use the computer; the latter made available to me his Analysis of Variance programme and was always ready to help pursue elusive bugs in my own programmes. It is a particular pleasure to record my appreciation of the intellectual company and friendship of my colleagues Guy Claxton, Nigel Harvey and Caro Williams.

The work reported would not have been possible without the

loyal service of many experimental subjects; Margaret Blanchard in particular served as a pilot subject for many of the experiments. They endured several sessions of boring tasks, often at inconvenient times, with good humour, for very little financial reward.

Finally, I should like to thank the Fellows of Wadham College for their hospitality during my years as an undergraduate and graduate. Particular mention should be made of Ian Crombie, whose philosophy tutorials taught me how to think; he is not to be blamed if I have since forgotten.

For financial support I was dependent during the course of the work reported in this thesis on a Research Studentship from the Medical Research Council.

Preface

An abbreviated version of Chapter V of this thesis was read to a meeting of the Experimental Psychology Society at University College on January 4th, 1973 under the title, "The effect of acoustic similarity and the relation between immediate recall and recognition."

A decimal notation is used to number the sections of chapters to which cross-references are made. Thus 'Section 3.21' refers to the first sub-section of the second section of the third chapter. The tables and figures for each chapter are bound together at the end of that chapter, and numbered accordingly.

TABLE OF CONTENTS

<u>CHAPTER I. Short-term memory: background and aims.</u>	
Introduction.	<u>p.</u> 1
1.1 The background.	4
1.2 Techniques.	11
1.3 Aims and strategy.	15
1.4 A model.	18
<u>CHAPTER II. Short-term forgetting.</u>	
Introduction.	26
2.1 The representation of order in immediate memory.	27
2.2 Short-term recognition and the concept of trace strength.	53
2.3 Active memory: capacity, function and coding.	69
Summary and conclusions.	87
<u>CHAPTER III. Immediate memory and reaction times.</u>	
Introduction.	90
3.1 Immediate recognition: the varied-set paradigm.	92
3.2 The constant-set recognition paradigm.	136
3.3 The other 'memory-scanning' paradigms.	152
3.4 Is this active memory?	166
3.5 Conclusions.	169
<u>CHAPTER IV. Methodology and Method.</u>	
Introduction.	171
4.1 The use of reaction times.	171
4.2 The Experiments: Apparatus and General Procedure.	182

<u>CHAPTER IX. Concluding comments, speculation and plans.</u>	<u>P.</u>
Introduction.	285
9.1 Theory.	285
9.2 Experiments.	292
9.3 What is active memory?	296
<u>BIBLIOGRAPHY</u>	302

"This field of research is not, at the time of writing, in a state which lends itself to simple and tidy conclusions."

(Broadbent, 1971, p.398, in a review of 'Primary Memory')

"Essentially nothing is known about the relation between event and order memory."

(Wickelgren, 1970a, p.92)

PART I. Theoretical review.

CHAPTER I. Short-term memory: background and aims.

Introduction.

1.1 The background.

1.2 Techniques.

1.3 Aims and strategy.

1.4 A model.

CHAPTER I. Short-term memory: background and aims.

If a man listens to or watches a sequence of four or five words presented to him at a rate comparable to that of ordinary speech, the chances are that he will be able immediately to recall the words and their order without error. His memory for the sequence may be tested in a variety of ways. For instance, we can ask him to say or write down the words in their order of presentation, or to recall an indicated word, or to tell us if a given word was included in the list, or to tell us about the position or context of a word which he knows to have been presented. If between the sequence and the test we have him perform a distracting task lasting several seconds; we shall find that, on all these tests, his performance is severely impaired¹. It is evident that at least some part or component of the stored memory representation mediating the original perfect performance is labile and progressively destroyed or permitted to decay by any task which prevents the man from attending to or 'rehearsing' the words. The greater the complexity and duration of the distracting task, the less he will remember about the list, save that after a period of the order of 15-30 secs., the level of performance typically ceases to change much - suggesting a small residual component of memory for the list which is not radically influenced by the duration or amount of processing required by the distracting task. It is widely assumed that these two components of his memory for the list reflect two different mechanisms of retention in the brain. I shall call the first of the hypothesised mechanisms 'active memory', the second 'inactive memory' (following Sternberg, 1969a). The terms reflect the possible physiological

¹It is assumed that the subject has already had several such trials.

basis of the two mechanisms: it is probable that retention in active memory is mediated by dynamic neural activity which is transient, fading away unless refreshed in some manner, while retention in inactive memory depends on a more permanent structural change (in the "neural tissue"). However, neurophysiological evidence lies outside the ambit of this thesis. My intention is to consider only behavioural evidence from human subjects performing short-term memory tasks, in the hope of elucidating the nature of active memory for verbal material.

To give a strict behavioural definition of the distinction between active and inactive memory at this point is not easy. As is now generally realised, both mechanisms are involved in most 'short-term memory' tasks. But there exist means of separating the two mechanisms experimentally. These will be briefly reviewed at the end of Section 1.1. For the present we will make do with this informal definition: active verbal memory is that hypothetical mechanism from which the traces representing a sequence of recently-presented verbal items are lost in the absence of rehearsal, rehearsal being prevented by any task which takes up all the subject's capacity for verbal processing; information represented in it is in some sense readily available to the subject. I do not mean to imply that active memory is a unitary mechanism - I shall in fact argue that it is not - nor that there is no process of retrieval. Nor do I mean to suggest that inactive memory is not subject to interference from subsequent material; it clearly is (Melton & Irwin, 1940), but in this case some similarity between the learned and interfering material is probably essential, while it is not so for loss from

active memory. Some further definitions may be helpful. I shall follow recent practice and use the terms 'short-term memory' and 'long-term memory' in an operational sense, to refer not to hypothetical mechanisms but to experimental procedures. 'Short-term memory (recognition, recall)' will be used to refer to situations where memory for small sets of items is tested within seconds. The term 'immediate memory (recognition, recall)' will be used atheoretically and loosely to indicate the general human capacity tested in such situations. What I have called active and inactive memory are by some theorists called 'primary' and 'secondary' memory and I shall occasionally use these terms in the context of reviewing other people's work, though to me they have undesirable connotations concerning the sequential arrangement of the two stores. By 'verbal' material, I intend to indicate any events or objects with readily available names. Lastly, the term 'recognition' is widely used to indicate both the experimental paradigm in which subjects are asked to indicate whether items occurred in a previously presented set, and the process by which a man establishes that a particular concatenation of perceptual features is, say, his grandmother. I shall restrict my use of the term 'recognition' to the former sense, and reserve the term 'identification' for grandmother- (or word-) naming.

A sequence of verbal items may logically be analysed into component elements or attributes (Underwood, 1969), such as order and content, relations between item and item and between item and position, relative position and absolute position, some of which are alternatives in that various subsets of them will serve absolutely to specify the sequence. Moreover, individual verbal items

characteristically possess visual, acoustic, tactual, articulatory semantic and abstract (e.g. frequency of occurrence) attributes, and may in principle be represented by any or several of these in memory. A particular emphasis will be placed on which attributes of the sequence are represented in active memory, how they are represented and retrieved, and to what degree the encoding and forgetting of different attributes is independent. We should also bear in mind that active memory may be active in a sense distinct from that already discussed. The subject is not passive: he may exert control over what he stores and how he stores it; he may be able to base his response on one of several alternative representations; he may perform transformations upon the stored material during the course of a trial.

1.1. The background.

Nineteenth century ideas on the nature of active memory were limited to little more than awareness of the existence of a rapidly fleeting record of recent experience. This was usually conceived of as a fading image, specific to a sensory modality but distinct from a purely sensory after-image. The 'primary memory' described on the basis of introspection by William James (1890), who quotes also the descriptions of Exner and Richet, appears to be what we would now call 'iconic' memory for vision, or 'echoic' memory for audition (Neisser, 1967) - a trace lasting not longer than a second or two in a preattentive sensory buffer. In the classic experimental textbooks of Titchener (1905) and Myers (1911) we find descriptions of experiments on immediate recall (e.g. the method of 'letter squares') in the context of discussion of the modalities of memory; these are supposed to be fairly directly related to sensory modalities. We

may note that Myers anticipates Conrad (1962) in pointing out that errors in immediate recall, with visual presentation, are often auditory or articulatory confusions; he also suggests means of selectively facilitating or interfering with storage in the different modalities. Apart from brief mention of 'memory after-images' nothing is said of differences between immediate and delayed recall. Nor do later accounts of memory such as that of Woodworth (1938) emphasise the distinction, save to describe the standard digit-span experiment and note that "the fact that the memory span is limited presents something of an enigma". Woodworth was also impressed by the fact that while lists shorter than the span can be learned in a single presentation, when the span is exceeded, the time for memorising the list takes a sudden jump.

Modern experimental work on immediate memory may be said to have started with the work of Brown (1958) and Peterson & Peterson (1959). They introduced the distractor technique: the interpolation of a rehearsal-preventing task between the presentation on each trial of a list and the subject's attempt to recall it. The distracting task in neither case involved new learning; the rapid forgetting which resulted seemed therefore to be the consequence of the absence of rehearsal rather than the formation of competing associations. This discovery was the basis for a major shift of emphasis away from interference theory (Postman, 1961), which had hitherto dominated conceptions of memory (see Broadbent, 1971). This was expressed in Brown's suggestion that forgetting in this task was by decay from a buffer storage system to which information could be recirculated by a rehearsal process; this retention mechanism was limited in terms of informational capacity. Meanwhile, Broadbent's (1958) dichotic

digit span experiments led him also to accept the idea of a short-term store to which information could be recirculated via a limited capacity channel; this store, the 'S-system' in his model, was, like James' 'primary memory', sensory, preattentive and located prior to the process of identification or categorisation, though its decay time was of the order of several seconds.

Later developments have somewhat changed this picture. Firstly, interest in information theory as a metric for immediate memory has proved transitory, partly as a result of Miller's (1956) classic paper. Secondly, it soon became clear that while the existence of a short-lived preattentive visual memory (now called 'iconic' memory, thanks to Neisser, 1967) may indeed be demonstrated by the use of post-stimulus visual masking and partial report techniques (Sperling, 1960, 1963; Averbach & Corriell, 1961), its decay time, which appears to be of the order of less than a second, was not sufficient to account for the size of the digit span or the recency effect in free recall, for visual material. On the basis of more recent evidence (Massaro, 1970; Norman, 1969a; Crowder & Morton, 1969; Morton, 1970; Morton Crowder & Prussin, 1971) a similar conclusion may be reached for the auditory modality, though speech sounds may be stored in what Crowder and Morton call 'pre-categorical acoustic storage' (PAS) for perhaps as much as two seconds. However, this hypothesis is not yet universally accepted (see Murdock & Walker, 1969). Thirdly, the demonstration that even with visual presentation immediate memory errors tended to be 'acoustic' in nature (Conrad, 1962, 1964; Sperling, 1963) implied that recall was from a form of storage into which material was encoded subsequent to identification. Fourthly, that amount of decay was shown to be a

function primarily of number of subsequent items rather than time, for both probed recall (Waugh & Norman, 1965), and free recall (Glanzer, Gianutsos & Dubin, 1969), suggested that forgetting was a matter of displacement rather than pure decay. Recent theorists have tended therefore to propose an active memory whose capacity is limited in terms of the number of items rather than of 'bits', which is post-categorical, coded phonemically, and from which information is lost in the absence of rehearsal mainly through displacement by subsequent items, though some pure decay may also occur. This minimal description may be said to characterise the current 'consensus' view, as expressed in influential papers by Waugh & Norman (1965), who distinguish 'primary' and 'secondary' memory, and Atkinson & Shiffrin (1968) who posit a 'short-term store' separate from both the 'sensory register' and the 'long-term store', and also in recent reviews by Craik (1971), Baddeley & Patterson (1971) and Baddeley (1972). The consensus is far from universal; among the dissenters may be numbered: Murdock (1972) who still prefers the idea of short-term modality-specific sensory stores; Wickelgren (1972), who is even more old-fashioned, in a sense, in that while accepting the existence of non-associative iconic and echoic stores, he considers active memory to be merely a transient component in the network of associations constituting inactive memory; and Broadbent (1970, 1971) who to his short-term sensory buffer has added something called 'the address register' (see Section 2.32).

Early in this saga, the postulate of a separate kind of memory to account for short-term forgetting had provoked a reaction from those nurtured in the tradition of interference theory (Mellon, 1963; Postman, 1964), who argued that there were demonstrable similarities

and continuities between short and long-term forgetting. However, these arguments could on the whole be countered by Waugh & Norman's (1965) proposal that items might in the short term be represented in both primary and secondary memory, and that there was therefore a secondary memory component even in immediate recall. Also, interference theory as an account of long-term memory suffers from numerous internal weaknesses, and has in recent years been, if not disproved, outnumbered by other approaches informed by the same view of the brain as an information-processor as characterises the consensus model of active memory, (Kintsch, 1970; Norman, 1970; Tulving & Madigan, 1970; Tulving & Donaldson, 1972). Nevertheless, there remained a need for the positive demonstration of a clear dissociation between active and inactive memory in terms of the empirical variables which differentially affect them. This need has now largely been fulfilled. The evidence for dissociation is reviewed in detail elsewhere (Baddeley, 1972; Craik, 1971; Baddeley & Patterson, 1971; Broadbent, 1970, 1971) and may be summarised briefly under the following headings:

- 1) The recency effect in free recall. When a range of different list lengths is used, this, corrected for the 'secondary memory component', yields a reasonably constant estimate of primary memory capacity (Murdock, 1967). Only the recency portion of the serial position curve is eliminated when recall is delayed by a rehearsal-preventing task (Glanzer & Cunitz, 1966). On the other hand, a number of variables such as presentation rate, repetition and frequency have been shown to affect the level of the flat portion of the curve, while leaving the recency portion unaffected (Raymond, 1969).

2) Acoustic versus semantic coding. Baddeley (1966a,b) showed that immediate serial recall of words was hindered when they shared a common set of phonemes but not when they were near-synonyms, while the reverse was true for delayed recall. Kintsch & Buschke (1969) demonstrated the same dissociation using probed recall, and moreover showed that it was congruent in extent with the recency effect.

3) Two-component retention functions. If one prevents rehearsal and varies the time/number of items intervening between presentation of an item and its test by means of a probe technique, the retention function relating p (correct) to time/items elapsed shows two quite distinct components: one showing rapid decay within 5-20 secs., and a later asymptotic portion showing no or very slow decay. This has been shown for both probed recall (Waugh & Norman, 1965; Norman, 1966) and item recognition (Wickelgren & Berian, 1971).

4) Capacity limitation. Active memory appears to be limited in terms primarily of the number of items it may contain, inactive memory by the low rate at which it can accept new information. Baddeley et al. (1971) showed that the informational demands of a sorting task concurrent with input influenced the secondary but not the primary memory component of recall of a list of words. On the other hand, loss from active memory is a function of the informational demands of a task interpolated between presentation and test (Posner & Rossman, 1965), while there is no reason to believe that this is true of inactive memory.

5) Neurology. (Warrington, 1971; Warrington & Weiskrantz, 1973). Patients suffering from the classical amnesic syndrome typically perform normally in short-term memory tasks (digit span, free recall, Brown-Peterson paradigm). Recently, Warrington & Shallice (1969, 1972) have observed a patient K.F. who appears to show the inverse of these symptoms: that is, while his performance

in long-term memory tasks is relatively normal, that in short-term tasks, at least when presentation is auditory, is extremely poor. If it is certain that this syndrome cannot be otherwise interpreted (for example, in terms of aphasia) then this constitutes evidence for an anatomical dissociation between active and inactive memory.

Moreover, Weiskrantz (1970) has drawn attention to a further implication: that entry of an item into inactive memory need not require that it shall previously have been represented in active memory¹, as is supposed by both Waugh & Norman (1965) and Atkinson & Shiffrin (1968). This suggests that active memory¹ must be conceived of as essentially parallel to inactive memory, except in the sense that the identification of items prior to storage in active memory must require contact with permanent representations of those items. Weiskrantz is inclined to go further and view such a short-term memory mechanism as merely an overflow system for cases where 'the long-term system is too hard pressed'. This seems to me excessively iconoclastic. Some of the functions of active memory in human cognitive processing will be discussed in Section 2.3.

It may be concluded from this brief survey of the field that confidence in the existence of post-categorical active memory as a mechanism separable both from inactive memory and from purely sensory storage is now general. As this confidence has grown, so attention has been diverted to a more detailed consideration of the nature of active memory. Here the consensus breaks down, and one is confronted by a bewildering array of inconsistent evidence and conflicting theoretical formulations. Some theorists have responded by abandoning the idea of multistore models (Cermak, 1972; Craik & Lockhart,

¹or at least that component of active memory which K.F. lacks

1972). The theoretical construct they suggest is that of 'depth of processing', the idea that there is a virtual continuum of levels to which information presented to the senses may be analysed, and that there may remain temporary traces at all these many levels, with characteristics specific to the level. To me this formulation seems either excessively abstract or excessively complicated. In the remainder of this thesis I intend to attempt a rather different approach and see if some of the difficulties may not be resolved by the making of a single sub-division within post-categorical memory, a sub-division which, as it happens, has already been implicit in the literature for eight years.

1.2. Techniques.

Among the changes that have occurred since the early experiments on immediate memory, four important developments in technique may be discerned:

1) Probe techniques. Early experiments tended to require complete recall of the presented list. However, it was apparent that the time or activity involved in recalling the first items caused forgetting of items later in the output sequence. (Brown, 1958). This effect, 'output interference', has been clearly demonstrated for short-term paired-associates learning (Tulving & Arbuckle, 1963), serial recall (Anderson, 1960) and recognition with several test items (Norman & Waugh, 1968). The resultant confounding of the effects of input and output position may be avoided by testing memory in a manner which requires only one response on each trial: thus, we may cue recall of one item only in P-A learning (Murdock, 1963a) and

memory for serial lists (Waugh & Norman, 1965), we present a single recognition probe (Norman & Wickelgren, 1965). This procedure has an additional crucial advantage; it may enable us to test independently the retention of different attributes of a list, corresponding to the different relationships which may be specified between the probe and the response required. For instance, it is possible that the retention of an item is tested independently of that of its position in the sequence when we present an item as probe and ask for a recognition response. Similarly it may be possible to separate order retention from position retention by probing for recall of an item with either the preceding item or a position cue. That these different attributes are logically separable does not of course imply that their retention in active memory is independent, but this fundamental question may be put to the test by the comparison of performance with these different classes of probe. If the same variables had the same effect regardless of the probe type, then there would be no grounds for supposing that the different attributes are independently represented. Evidence to the contrary will occupy a large part of this thesis.

2) Recognition versus recall. A given variable, such as acoustic similarity, may have its effect at any of three phases of memory: acquisition, time in storage and retrieval. A 'consolidation' phase may also be required in long-term memory, but may be considered instantaneous for short-term memory (Wickelgren, 1970a, 1972). In attempting to separate storage from retrieval processes, many researchers have turned to the use of recognition paradigms on the grounds that recognition largely bypasses retrieval processes, and in the hope that the application of decision theory measures in recognition (see Section 2.2) may provide a pure estimate of amount of forgetting independently of the subject's response biases (e.g. Norman &

Wickelgren, 1965; Broadbent, 1971). We shall consider the assumption of direct access in recognition in detail in later chapters. It should be noted that a second assumption is commonly made: that the memory traces mediating item recognition are identical to those underlying recall. The foregoing discussion on the separability of attributes should have made it clear that this may not necessarily be so.

3) Control of rehearsal. The early experimenters tended to use fairly leisurely presentation rates and leave subjects to their own devices as regards what they thought about during presentation and testing. However, the heavy emphasis placed by Atkinson & Shiffrin (1968) on the distinction between 'memory structure' and 'control processes' reflected the general growing awareness of the crucial importance of subjects' ability to allocate the mechanisms, processes and stores available in their brains in a variety of different ways to the task in hand. In immediate memory experiments, with single trials, immediate test or filled delay, and presentation rates that permit little complex recoding or mnemonic activity, the main strategy open to control by the subject is rehearsal. The experimenter's problem may be seen as one of avoiding the possibility that the subject is covertly (re)presenting the material to himself at times and in an order not under experimental control. Three solutions are possible: making rehearsal observable, controlling it through instructions, or eliminating it altogether. Thus one can instruct subjects to rehearse aloud (Rundus, 1970), or to attend at any moment only to the last item presented (Waugh & Norman, 1965) or one can use immediate test and presentation rates which make rehearsal difficult, unnecessary or both. Though none of these lacks

disadvantages, there is no longer any excuse for leaving rehearsal entirely uncontrolled if one intends to draw conclusions concerning retention functions or the effects of serial position.

4) The use of reaction times. Until comparatively recently, latencies had not generally been measured in immediate memory experiments save as an occasional adjunct to accuracy measures (e.g. Peterson & Peterson, 1959). As a result, little was known about performance when sub-span lists are recalled immediately, since recall is effectively error-free. However Sternberg (1966, 1967a, 1969a) has combined the measurement of reaction times with the use of probe techniques in a series of pioneering studies on immediate memory for sub-span lists. This work is part of a general renaissance of interest in the use of reaction times as a tool for the separation and isolation of processing stages as first suggested by Donders (1969). This trend is due partly to increasing technological sophistication, partly to the influence of the computer sciences on psychology and partly to the theoretical work of Sternberg himself (1969b). As regards the immediate memory experiments, since performance is largely error-free, the main theoretical emphasis to date has been on retrieval processes. But it should be noted that the subject's ability to recall a sub-span list correctly does not imply that no decay, loss or transformation of information has occurred. Information may initially be represented redundantly, or with a 'strength' considerably in excess of what is necessary for error-free performance.

In the later chapters of the thesis, experimental work will be reported which employs reaction-time techniques due to Sternberg in

both item recognition and probed recall studies. A major concern will be whether these two classes of task do indeed tap the same form of active memory. Rehearsal is controlled with the use of fast presentation rates, immediate test and appropriate instructions.

1.3. Aims and strategy.

A glance at the journals that report research on human memory reveals that the experimental paradigms introduced by Sternberg (1966, 1967a,b, 1969a), particularly item recognition, have been taken up with some enthusiasm. One of the more bizarre features of this trend is the apparent lack of contact between the resulting studies and the somewhat older tradition initiated with the work of Brown, Broadbent, and the Petersons. Thus there may be distinguished two largely separate bodies of research on immediate memory: experiments on 'short-term forgetting', in which loss of information from memory is deliberately induced by the use either of distracting extraneous tasks or of super-span lists, and accuracy is the chief dependent variable; and 'memory-scanning'¹ experiments, in which forgetting is largely avoided by the use of short lists, immediate test and probe techniques, and reaction time is measured. There have to date been few signs that these two lines of research are converging on common theoretical conclusions about the nature of active memory. For instance, the exhaustive scanning model proposed by Sternberg (1966) for speeded item recognition bears little resemblance to the idea of direct access in recognition which we find implicit in the theory of Wickelgren & Norman (1966). Nor does the rate attributed

¹The term 'memory-scanning', though derived from Sternberg's model, is not here intended to have theoretical connotations.

by Sternberg to the scanning process have any analogue in models deriving from studies on short-term forgetting. This apparent schism is surely distressing. If one is to have any faith in the existence of active memory as a separate mechanism with describable properties, these properties cannot be allowed to depend on whether the experimenter has chosen to measure errors or reaction times, nor on whether he has chosen to ply his subjects with lists just shorter or just longer than the memory span. Of course, the different classes of paradigm may selectively emphasise different aspects of active memory; for instance, it may be the case that RT serial position curves will primarily reflect differences due to retrieval processes, while the corresponding accuracy curves may depend rather more on differences in storage. But by and large the same picture should emerge from both sorts of experiment. One of the motives behind the work reported in this thesis is the itch of a synthesist to reconcile these hitherto uneasy and distant partners, in the belief that each should more closely inform the other than is the case at present. In pursuit of this aim I shall pursue two partly independent lines of argument.

One is developed in the context of a comparison between the various models available to account for data from studies employing Sternberg's (1966) item recognition paradigm. It has recently been suggested (Baddeley & Ecob, 1970; Nickerson, 1972) that a satisfactory alternative to Sternberg's exhaustive scanning hypothesis may be found in a model employing concepts of 'trace strength', unique representations, and direct access without search. These postulates are also characteristic of a model derived from experiments on short-term forgetting by Wickelgren & Norman (1966), which I shall

review briefly in the second part of Chapter 2. The first two sections of Chapter 3 will be devoted to a critical review of the many studies employing Sternberg's item recognition paradigm; I shall elaborate two 'trace strength' models, and compare their predictions with those derived from other models in the light of data reported in the literature; experiments of my own germane to this issue will be reported in later chapters.

Another argument develops when we consider in some detail the relationship between the retention in active memory of order and content information. It will emerge that models of active memory which suppose a unitary form of storage for these different attributes are inadequate to account both for short-term forgetting (reviewed in Chapter 2) and for such data as we already have from memory-scanning studies, which are summarised in the third section of Chapter 3. A strategy for the investigation of the nature of the differences between item recognition and probed recall is there outlined and experiments guided by this strategy are described in following chapters. It will be argued that the differences which emerge may plausibly be accounted for by a hypothesis assuming two forms of active verbal memory. There already exists a general model of memory which includes the two elements we need and has already proved useful in a wide range of related areas of research. This is Morton's (1970) 'logogen' model. Since a long argument is easier to follow if its goal is known, I shall, in the last section of this chapter, anticipate my conclusion by describing briefly the relevant features of Morton's model and indicating the amendments I

shall find necessary.

In summary then, the two issues running through the discussion will be: the nature of the representations underlying short-term item recognition, and the relationship between these representations and those involved in the immediate recall of a verbal item from a short ordered sequence. The approach is eclectic; with the help of Morton's (1970) framework, I shall attempt to combine Wickelgren & Norman's (1966) concept of content-addressable representations possessing 'trace strength' with the idea of active memory as an ordered buffer or queue (Atkinson & Shiffrin, 1968; Reitman, 1970) in a manner similar to that first sketched by Conrad (1965). The experiments I shall describe employ reaction time as a measure, but are suggested by and intended to have consequences for theories of short-term forgetting.

1.4. A model.

Morton's 'logogen' model was originally developed as an account of word identification and other language behaviour (Morton, 1964a,b, 1969a; Morton & Broadbent, 1967). It and elements contained in it have since proved useful as an explanatory framework in accounting for the Stroop effect (Morton, 1969b), object naming and comparisons between printed words and pictures (Seymour, 1973), and inter-modal integration (Harvey, 1973). In his 1970 paper, Morton elaborates the logogen model as a general functional model of memory. From our point of view, it contains two crucial elements, the

'logogen system' and the 'response buffer'. I shall summarise Morton's view of each of these and indicate where I wish to differ from or add to his account, though the evidence for these changes must be left for later chapters. Fig. 1.1 shows the principal structural components of the model.

The logogen system: The starting point for the model is that when a verbal response such as 'chair' becomes available, the same final unit in the brain is assumed to produce that response regardless of whether the source of the information is visual (e.g. the actual object, or the written word 'chair'), acoustic (the spoken word 'chair') or semantic (as when one is asked to complete the sentence 'The table is next to the', or associate to the word 'table'). These units, of which there is one for every identifiable word, are termed 'logogens' and may be viewed as counting devices for the collection of evidence from both external (visual, auditory, etc.) and internal (semantic, articulatory) sources ; it may also be helpful to view them as translation devices mediating between representations of the different attributes (visual, acoustic, articulatory, semantic) of a word. Logogens are held to be unique word-specific terminal units in a passive identification system based on hierarchies of feature detectors; as a pattern-recognition system the model has affinities with the models of Selfridge (1959), Norman (1968) and Treisman (1969). As in these models, the effects of context and expectancy on thresholds for identification are taken into account. A logogen's threshold, which is the level of the 'count' (or amount of evidence collected, or degree of activation) necessary for the response to become available, is altered:

a) by context and expectancy, via the 'cognitive system' (the rest of inactive memory) which stores and mediates the effects of prior experience;

b) by frequent repetition, which permanently lowers the threshold;

c) by the operation of the logogen itself, which temporarily lowers its own threshold. Thus following identification of a word, the threshold of its logogen is lowered, subsequent to which it returns slowly (and non-linearly - Morton, 1968) to its base-line level. This is intended to account for the facilitative effect of the prior presentation of a word on its subsequent identification (Neisser, 1954). I propose that this 'lowering of threshold', while it lasts, serves as one form of post-categorical active memory representation of the occurrence of an item. This suggestion is very similar to Norman's (1968) conception of the nature of primary memory. I shall call this representation an 'item-trace'. Morton himself does not make much of the possibility that traces at logogens could have this kind of active memory role, though he mentions that "traces remaining in logogens, whether in the values of the count or the threshold could be regarded as two possible sources of information (with different time characteristics) for secondary memory" (Morton, 1969a), and in the 1970 paper points out that it is implicit in his (Morton 1968) results on judgements of recency (see Section 2.22) that information represented in a lowered threshold can be utilised in at least one task. We should note that the logogen itself is a permanent amodal or conceptual representation, and that it is content-addressable (Norman, 1968) - i.e. accessed directly, without search, given the appropriate physical or semantic

features.

I shall want to amend and qualify Morton's account of the way in which logogens function in several ways.

i) I shall assume that as well as for words, logogens exist for digits, letters and probably any familiar 'entities'. I shall nevertheless continue to refer to logogens, in spite of the etymology.

ii) Morton's account of the activation of a logogen is very complex. He proposes the following elements:

- the 'count', which reflects the number of relevant semantic and sensory attributes currently available, and is subject to rapid decay when these are no longer available,

- a semantic threshold, which is the level of the count a logogen must achieve before its activation can influence that of other logogens,

- a (higher) threshold for the word to become available as a response, both thresholds being internal to individual logogens,

- the lowering of the latter threshold following response availability (the 'priming' effect); this decays more slowly back to the baseline than does the count,

- a facility for lowering the thresholds at all logogens if no response becomes available.

Part of the complexity is due to Morton's desire to avoid positing an external decision mechanism. Yet in spite of his efforts, an external decision process still seems essential. Firstly, according to Morton's own exegesis of the nature of recency judgements (Morton, 1968, 1970) the subject must be able to compare the state of activ-

ation in different logogens (Section 2.22). Secondly, on Morton's account, it seems possible that a number of logogens could exceed their upper thresholds simultaneously, in which case there must be some means of choosing between the responses thus made available. Thirdly and most importantly, evidence from experiments on short-term item recognition (Section 2.22, Section 3.1, et seq.) suggests that subjects are able to compare the degree of activation of a logogen against an adjustable external decision criterion.

For our purpose, therefore, it seems simpler to propose that the decision criterion is external to logogens. Activation of a logogen will be said to take place along a single and continuous dimension of 'trace strength' (Wickelgren & Norman, 1966). The more semantic and sensory attributes relevant to a logogen are presented, the greater the increment in trace strength. Trace strength decays non-linearly. (The initial rapid component of this decay may be due in part to the rapid loss of activation by pre-logogen feature analysers - see discussion of the 'sensory component of trace strength' in Section 3.131). Given appropriate content cues for access to a logogen (Section 2.31) its trace strength may be compared against an adjustable decision criterion. Adjustments to such a criterion may also be responsible for ensuring that only one response becomes available, and may thus participate in the process of response selection (see below).

The response buffer: Once a verbal response has been made available by a logogen, Morton proposes that it remains available for a while in a form of storage called the response buffer; this

was originally introduced to account for the fact that in reading aloud (Morton, 1964c) or in shadowing (Treisman & Geffen, 1967) subjects may lag by as much as five words between what they are reading or hearing and what they are saying. The response buffer in many respects resembles Atkinson & Shiffrin's (1968) 'rehearsal buffer'. It has a limited capacity of items (about five); these are coded (purely) in terms of articulatory features. Items in the store are ordered temporally - a crucial property. Forgetting is by displacement; an item entering the buffer when it is full bumps out an old item. The process usually called silent rehearsal consists in the recirculation of items from the response buffer back to the logogen system.

Again some amendments:

(i) Morton assumes that forgetting from the response buffer is by displacement only. I shall want to assume that some decay may also occur in that phonemic features of items may be lost independently (see Ch. 5), though whether this loss is all-or-none I cannot say.

(ii) Morton assumes that when a verbal response is made available by the logogen system, it automatically 'captures' the response buffer, so that the contents of the response buffer are necessarily the last few responses made available, encoded in the order they become available (unless presentation is very fast). I consider it necessary to propose that there is some selection of items for entry to the response buffer (see Sections 2.11 and 2.32). This selection is by omission: responses must become available (i.e. exceed a strength criterion) before they can be entered into the

response buffer, but once available, they need not be entered.

(iii) The nature of access to items represented on the one hand as traces at logogens and on the other as potential responses in the response buffer is radically different. Logogens are content addressable, and require a content cue for access (see Section 2.31). The response buffer contains only a string of articulatory commands. To locate an item in the response buffer given only a content cue requires a process of search. The contents of the response buffer must be fed, in order, back to the logogen system, until the logogen of the item searched for is reactivated. This process of search, which is the same as subvocal rehearsal is held to be slow, not exceeding about eight letter names a second (Landauer, 1962). Though there is no direct access in terms of content, it is possible that the subject may control the position in the sequence at which he starts the search process, given a position cue (see Ch. 6).

As a model of active memory, Morton's model is almost entirely passive. The changes I have suggested largely involve the introduction of hypothetical control processes. These are: an adjustable decision criterion for response availability and the evaluation of trace strength; a process of selection of items for entry to the response buffer, and a facility for determining which locations in the response buffer are to be read out first. Between them, these provide the means by which information may be reorganised in active memory and give the key to how active memory may be used as a working memory in cognitive processing (see Sections 2.31, 2.32, 2.33). To try and specify these processes in any further detail would be to go beyond present needs and available data. I offer

them here as a heuristic. Our main concern will be to decide whether, given the possibility of such control processes, the main structural components of Morton's model are adequate for the job.

Two distinct forms of post-categorical active memory are thus proposed: decaying item-traces at logogens and items encoded as a string of articulatory commands in the response buffer. Of course, in immediate memory experiments there may also be contributions from inactive memory (which Morton calls the 'cognitive system') and from precategorical sensory stores, which are supposed to be located prior to the logogen system as part of the hierarchies of modality specific feature-analysers (see Section 3.131). Morton (1970) particularly stresses precategorical acoustic storage (PAS) (see Crowder & Morton, 1969; Morton, Crowder & Prussin, 1971). However, since presentation in my own experiments is visual, PAS need not greatly concern us here.

In the next two chapters I review evidence relevant to this model: in the first, experiments on short-term forgetting, and in the second data from studies employing 'memory-scanning' paradigms.

VERBAL STIMULI

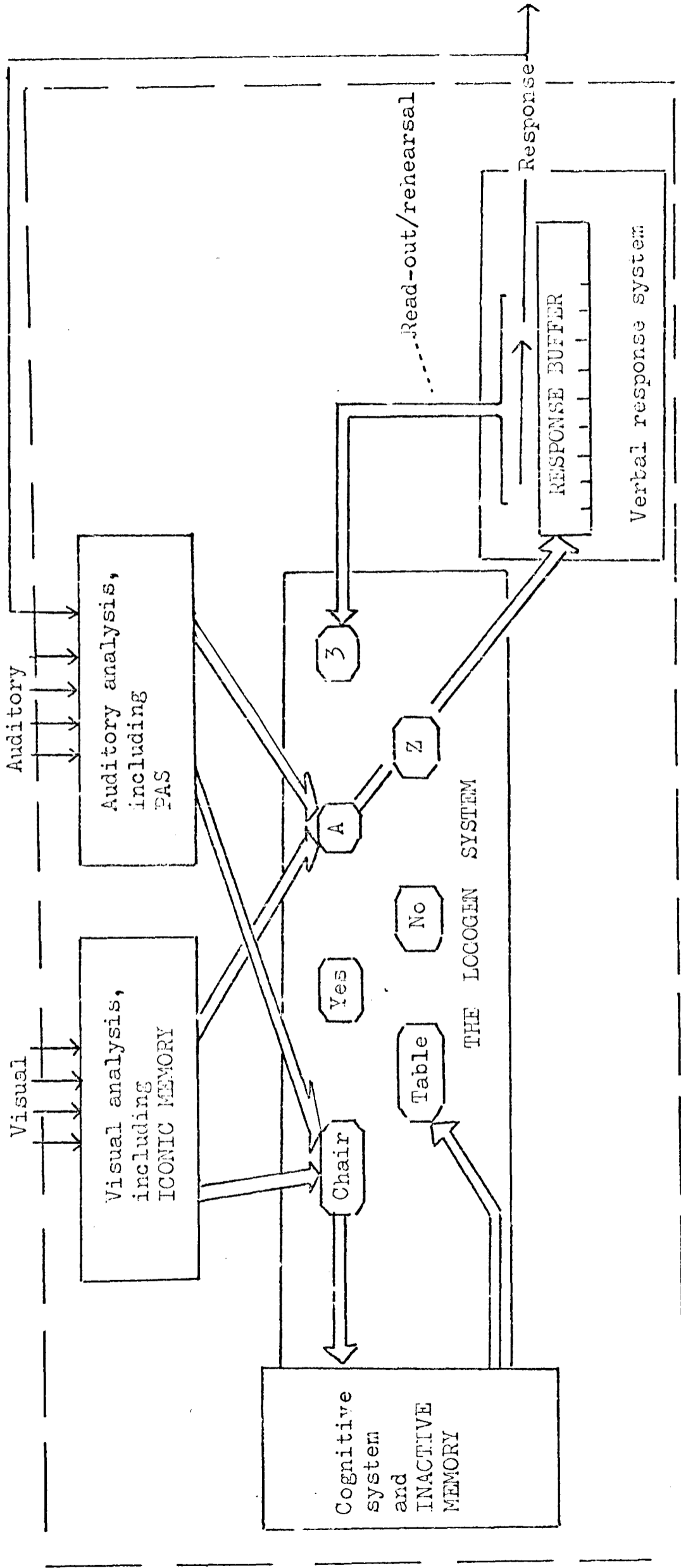


Figure 1.1 General model of active memory (after Morton, 1970), showing the structural components and pathways.

CHAPTER II. Short-term forgetting.

Introduction

- 2.1 The representation of order in immediate memory.
 - 2.11 The recall of serial order.
 - 2.12 Direct comparisons of order and item retention.
 - 2.13 Probed recall: position and order.

- 2.2 Short-term recognition and the concept of trace strength.
 - 2.21 Trace strength theory.
 - 2.22 Trace strength theory as an account of item recognition.
 - 2.23 Trace strength theory and order recognition.
 - 2.24 Is a strength measure possible for recall?

- 2.3 Active memory: capacity, function and coding.
 - 2.31 The capacity of active memory and speech comprehension.
 - 2.32 Broadbent's 'address register'.
 - 2.33 Active memory as a working memory.
 - 2.34 Sperling's recognition buffer.
 - 2.35 Articulatory versus acoustic coding in active memory.

Summary and conclusions.

CHAPTER II. Short-term forgetting.

In this chapter, I shall review a selection of studies on short-term forgetting of verbal material. We begin in Section 2.11 with the problem of the representation of order in immediate memory. Theories emphasising associative links will be contrasted with theories according to which temporal order is retained as an intrinsic property of the memory structure, as is the case in Morton's 'response buffer'. A consideration of the relation between order and item errors, and of the nature of order errors, will lead to a conclusion first suggested by Conrad (1965): that the second class of theory is adequate provided we assume an additional form of active memory with the property that it makes items available as responses as a direct function of their recency of occurrence. It will be claimed that 'item-traces' in the logogen system may serve this role. In Section 2.12 some experimental comparisons of item recognition and tasks requiring order retention will be described. In Section 2.13 data from experiments comparing the effectiveness of different kinds of probe for recall will be discussed in relation to the hypothesis that probed recall is mediated by retrieval from the response buffer.

In pursuit of enlightenment on the nature of 'item-traces' we shall turn, in Section 2.2 to a survey of experiments on short-term recognition conducted within the context of 'trace-strength theory'. The main assumptions of this theory will be described and evaluated briefly. I shall argue that they appear to be valid for item recognition and for judgements of recency. We shall also consider

comparable experiments on order recognition; here the picture is more confused and this, together with previous evidence, will be held to imply that short-term associations play no major role in active memory.

At this point the theoretical position will have hardened into a firm hypothesis positing two forms of post-categorical active memory, as introduced in Section 1.4 : items represented in active memory as temporary traces, subject to decay along a strength continuum, located at permanent content-addressable representations which I have identified with Morton's 'logogens'; items represented also in a limited-capacity, temporally-ordered store in terms of their articulatory features - the 'response buffer'. In Section 2.3 I shall attempt to show that this hypothesis sheds some light on problems of the nature of active memory in relation to its general functions, I shall consider some arguments of Broadbent (1970, 1971) for two forms of storage in short-term memory, and conclude by mentioning some evidence on encoding from sensory memory, and on articulatory coding in immediate memory, which appears to be compatible with the hypothesis.

2.1. The representation of order in immediate memory.

2.11. The recall of serial order.

Many experiments within the tradition of research initiated with the work of Brown (1958) and Broadbent (1958) have required immediate serial recall of short lists of verbal items chosen from a small vocabulary. Correct reproduction of order has been emphasised;

typically the total stimulus ensemble has been made available to the subject either overtly or through the use of well-known ensembles such as digits or letters. Retention of order is also required when recall of single items is probed with a contextual or a position cue. How is the order of a sequence of items represented in immediate memory? The literature on the topic is labyrinthine. Ariadne gave Theseus a ball of thread to guide him through the Cretan maze; in spite of her subsequent fate, I shall try to do the same for the reader here.

Two extreme theoretical positions may be discerned. According to the first, the order of items is preserved by the same (and only) store that records their occurrence: a pure 'slots' model. According to the second, information about the order of a sequence of items is encoded separately from information about which items occurred. I shall mention two versions of this viewpoint: the information-theoretic (which I shall ignore thereafter) and the associationist. Having introduced these theories, I shall go on to review evidence on the relationship between errors of order and errors of content; this yields difficulties for a pure slots model, as several variables influence transposition errors but not item errors. Evidence for the associationist viewpoint will then be presented: it will be argued that evidence for the short-term retention of inter-item associations is not strong, but that there is good evidence for item-position links. However, the latter could also be interpreted in terms of a slots model, provided that the first objection can be answered. At this point the Minotaur will make his appearance, in the shape of Conrad (1965) who both

obtained evidence indicating that order and item errors are not independent, and proposed a modified 'slots' theory, incorporating an additional items-only store. His model will be discussed in some detail, and then by a magical transformation shown to be a precursor of the model outlined in Section 1.4.

The pure slots model: Perhaps the simplest system one can imagine is one in which the order of the items is preserved by the store in the same way that a railway track preserves the order of the trucks upon it, or a pack of cards the order of the constituent cards; that is, a system in which a sequence of items to be memorised enters, remains in and leaves the memory store in the temporal order in which the items are perceived. This hypothesis does not distinguish the storage of the items themselves from the storage of their order. Neisser (1967) refers to it as a 'slots' hypothesis and it appears to be implicit in Atkinson & Shiffrin's (1968) 'rehearsal buffer' and Reitman's (1970) 'waiting room'.

The view from information theory: Brown (1959) and Crossman (1960) approach the problem from an information-theoretic standpoint. Order information is, they claim, encoded with less redundancy than item information in a limited capacity storage space; a particular amount of decay will on average produce a relatively greater loss of order than of item information, and the two kinds of information may be encoded and forgotten independently. However, the popularity of the information theory approach has waned, partly because the metric of 'bits' has not proved strikingly helpful in the discovery of underlying mechanisms and partly in the light of Miller's (1956) classic demonstration that memory span is limited not by

the 'bits' of information theory but by the number of 'chunks' that can be retained. Crossman (1960) attempts a reply, arguing that the equation normally used to calculate the information content of a list lacks a term for the order information, and that if one is introduced memory span data are better accounted for. But his data are unimpressive and his argument is simply false: the suggested equation includes order information twice over, though in fairness it may be argued that his 'random-address' model implies that order is in effect represented twice over (Hitch, 1972).

The associationist viewpoint: The two classical answers to the problem of the internal representation of serial order are those framed in terms either of inter-item associations ('associative chaining') or of item-position associations. The former is illustrated by Ebbinghaus' (1885) theory of serial learning which relies on forward, backward and remote associations between list items, and the latter by its more recent opponents (Ebenholtz, 1963; Jensen & Rohwer, 1965) a current view being that both forms of association are operative in long-term serial learning (Young, 1968). The same dialectic may be seen in physiologically-oriented discussions of serial organization. Lashley's (1951) classic attack on associative chaining has been influential, though his concept of an independent 'schema' for the control of order is not very clear. Milner (1961) has developed the idea of a schema in terms of item-position links, proposing a 'neural scanning chain' to which responses become linked. On the other hand, Wickelgren (1969b) has attempted (unsuccessfully, according to Macneillage, 1970) to answer Lashley's arguments on associative links between phonemes with the concept of 'context-sensitive associations' or 'allophones'. Bryden (1967)

moreover, considers that the influence on behaviour of long-term inter-item associations can be incorporated into a version of Lashley's schema.

The chief advocate for an associationist theory of short-term memory has been Wickelgren, whose approach is elaborated as part of a major theoretical system - 'multi-trace strength theory' (Wickelgren, 1970a, 1972). Among his theoretical distinctions is that between two basic types of memory structure, associative and non-associative. We may find his definition helpful:

"In associative memories, each event or concept has a unique internal representative, and the internal representatives have different degrees of association to each other depending on how frequently they have been contiguously activated. In non-associative memories, there is an ordered set of locations (cells, registers, boxes, etc.) into which the internal representative of any event or concept can be coded, and sequences of events or concepts are stored in order in this ordered set of locations."

(Wickelgren, 1972, p.20)

Wickelgren considers only very short-term sensory memory - e.g. iconic memory - to be non-associative. (Wickelgren & Whitman, 1970). In short-term verbal memory, the recent occurrence of an item is said to be represented by the temporary activation of an associative link between the item's 'unique internal representative' and a 'familiarity representative'. The order in which recent items have occurred is represented by the temporary activation both of associations between the representations of items contiguous in the sequence and of associations between item representations and position representations - which may be labels like 'beginning', 'middle' or 'end'. All three kinds of associations decay, but their strengths are independent. If a subject is given both a position cue and a context cue, the strength of the item-position and inter-item associations should add (see Section 2.13 for evidence).

Evidence on the differential behaviour of order and content errors:

Conrad (1959) showed that a large proportion of the errors made in serial recall are transposition errors; i.e. mistakes in the ordering of responses. In experiments in which the presented lists do not exhaust the stimulus ensemble, it is possible to score separately the incidence of errors of order (transpositions and semi-transpositions) and of content (omissions and extra-list intrusions). Now if the storage of both items and order in active memory is entirely by a 'slots' mechanism, then the incidence of these two classes of error should be similarly influenced by any experimental variable. This appears not to be the case for several variables:

1) Serial position. Marked differences in the serial position curves for order and item errors have been found both with serial recall (Aronson, 1968) and probed recall (Fuchs, 1969; McNicol, 1971b; Hitch, 1972, Exps. 1 & 2). The typical bowed serial position curve appears to be typical only of order errors; item errors show little variation with serial position, and such as there is in these studies is monotonic. (The lack of a bowed serial curve for item errors is contingent on the use of lists not greatly in excess of span.)

2) Probe type. Hitch (1972, Exps. 1 & 2) compared contextual with position with double probes (see Section 2.13). Such differences as he found were restricted entirely to order errors.

3) Acoustic similarity. Wickelgren (1965c) showed that acoustic similarity between the items to be remembered depressed serial recall, but that if the data were scored purely in terms of item errors, the effect was, if anything, reversed.

4) Grouping. Ryan (1969) presented lists interrupted by pauses

splitting the list into groups. Grouping improved performance compared to ungrouped lists, but the difference showed only in the order errors. Wickelgren (1967a) instructed subjects to rehearse in non-overlapping groups of various sizes, and found that group size significantly influenced order errors, three being the optimal size, while item errors were unaffected.

5) Presentation rate. The use of presentation rate as a variable is complicated by the fact that while fast presentation may allow inadequate time for initial acquisition, slow presentation may simultaneously produce a decrement in recall resulting from time-based decay and an increment as the result of increased opportunities for organisation and rehearsal. There is some evidence that item and order errors are differentially susceptible to presentation rate. Aaronson & Sternberg (1964) and Matthews & Henderson (1970) found that order errors increased more than item errors as presentation rate increased. But Aaronson (1967) argues that some order errors result from failure to discriminate successiveness and Scarborough & Sternberg (1967) showed that perception of the order of items may be impaired at rates which permit the identification of items. So this evidence is equivocal with respect to memory.

6) The time course of forgetting. The fact that transpositions frequently occur in the absence of item errors (Murdock & Vom Saal, 1967) suggests the possibility that order information may be lost before item information (though there is a problem of guessing probabilities here). The data of Fuchs (1969) and McNicol (1970) suggest that as retention interval is increased, order errors may become less frequent while item errors continue to increase.

On the face of it, such evidence is difficult for a pure 'slots' model to account for. It seems to suggest that items in adjacent or even more remarkably, distant slots can be exchanged, so that confusions of order can occur without loss of items. Whether this is to be accounted evidence for the separation of order and item information in active memory clearly depends on the account we can give of the nature of transpositions, and why they are affected by phonemic similarity. We shall return to this problem below.

Evidence for inter-item and item-position associations: Wickelgren

(1972) adduces two main pieces of evidence for the importance of inter-item associations in short-term verbal memory. The most striking is the occurrence, during serial recall of lists containing repeated items, of 'associative intrusions', that is: erroneous recall of items following a repeated item, at a level greater than chance (Wickelgren, 1965c, 1966d). For instance, in the list ABACD, the subject will tend to recall C after the first A, and B after the second. Associative intrusions are claimed to result from the representations of C and B both being associated to that of A; at A the subject is faced with a choice between two pathways; a wrong turn results in an associative intrusion. It should be noted however that other effects of rather greater magnitude occur with repeated items. Subjects tend to suppress recall of the second occurrence (the 'Ranschburg effect' - Jahnke, 1969) and Wickelgren's data (Wickelgren, 1965e) show that lists containing adjacent repeated items are better recalled than lists with no repeats, nor is this just a scoring artefact (McNicol, 1971c). Wickelgren's second argument is that there is both proactive and retroactive interference from phonemically similar material on order retention (Wickelgren,

1965a, 1966a), while within-list similarity may facilitate item retention (Wickelgren, 1965c). The relevance of this evidence depends on his claim that a list of letters is represented by associations between their constituent phonemes. Thus, the order of a pair of letters sharing a vowel phoneme is represented by associations between only their consonant representations, while the order of a pair of dissimilar letters is carried by all four associations from and to adjacent consonant and vowel phonemes. As for the retention of items, phonemic similarity has a facilitatory effect because the common phoneme acts as a retrieval cue which is itself certain to be recalled. A third line of argument, which Hitch (1972) considers supports Wickelgren's position, is that there exist several demonstrations of the influence of well-learned relationships between successive items on short-term retention of their order. His idea is presumably that short-term associations of greater strength should be set up where there already existed long-term associations than where there were none. Baddeley (1964) and Baddeley, Conrad & Hull (1965) showed significant correlations between letter-sequence predictability and immediate recall. Mayzner & Schoenberg (1964) and Kantowitz, Ornstein & Schwartz (1972) showed that immediate serial recall of high frequency consonants was better when they were arranged in high-frequency digrams than when the same letters were arranged in low-frequency digrams. However, this argument is hardly critical since these data may equally well be interpreted as reflecting the greater availability of high frequency and probable items as guesses during output.

Whatever the validity of Wickelgren's arguments for the existence

of short-term inter-item associations, they are certainly not sufficient to account for serial recall. Firstly, associative intrusions occur much less frequently than would be predicted on the basis of a memory structure consisting purely of inter-item associations (unless it be proposed that subjects can distinguish the strengths of associations early and late in the list). Secondly, an associative chaining theory should predict that recall would improve with the size of the rehearsal group, since the longer the group, the more pairs of adjacent items there will be that are rehearsed in contiguity. But Wickelgren (1967a), who varied the group size from one to five items, found three to be the optimal size. Thirdly, a pure associative chaining theory predicts serial position curves exhibiting pure primacy, rather than the bowed serial position curves typically obtained. (Even with visual presentation, a small recency effect is obtained - Morton, 1970.) Fourthly, Baddeley (1968, Exps. 5,6) showed that for serial recall of lists in which alternate items were phonemically similar, performance was worse on the similar items than on the alternating dissimilar items; an associative-chaining hypothesis predicts more errors on the items following similar items - since the role of similar items in producing intrusions is formally similar to that of repeated items.

Such evidence does, on the other hand, emphasise the importance of item-position links. There is more direct evidence. Wickelgren (1967a) found that when subjects were required to break the list into groups, order errors tended to consist either of transpositions within a group, or transpositions between items in the same position in different groups. Wickelgren (1972) concludes from this result

that items can have short-term associations both to group labels and to positions within a group. Neisser (1967), emphasising the beneficial effects on recall of rhythmic structure and grouping, seems also to favour the idea of item-position links, the advantage of rhythmic structure being the increase in the number of reference points to which items may be linked. The most compelling evidence for some form of item-position links in active memory is the occurrence of 'serial order intrusions' (Conrad, 1960, 1965; Fuchs, 1969) - that is, intrusions in recall from a given position in the list, of the item which occupied the same position in the previous trial, at a level well above that predicted by chance.

So far the only direct evidence for inter-item associations in active memory has been the phenomenon of associative intrusions. However, there is considerable evidence that items are linked to positions in some way. This evidence could be accounted for either in terms of item-position associations, or in terms of a slots model. We have already seen that the occurrence of transpositions and their relation to acoustic similarity are problematic for a slots model. They are no less problematic for a theory postulating item-position associations. Why, for instance, should distant transpositions be rarer than adjacent transpositions (Murdock & Vom Saal, 1967)? Moreover it is vital for the associationist account that item errors and order errors should be independent. There is evidence that this is not the case.

The relationship between transpositions and phonemic similarity;

Conrad's model: Conrad (1964) showed that when only one letter was wrongly reported during serial recall of a visually presented list, it tended to be substituted for by an acoustically similar letter.

And Conrad (1965) also showed that items which were acoustically similar tended also mutually to transpose in serial recall. The implication is that the forgetting of order and of items is not independent. How can this be reconciled with the evidence already mentioned? At this point it is appropriate to turn to the model proposed by Conrad (1965). Basically, this hypothesises a slots mechanism, called the 'fixed-address store', (FAS), to which is added a source of item information, called the 'response availability store' (RAS). The FAS consists of a fixed number of boxes representing the serial positions of the list, and items are represented in the boxes in terms of their phonemic features.

"Items enter the boxes in the order of perception (which could be different from the order of presentation). During recall, Ss read the name of the item from each box in an order determined by the experimental instructions. They may start anywhere in the sequence (Kay & Poulton, 1951; Posner, 1964) but there is no way in which the contents of the boxes can interchange."

(Conrad, 1965, p.168)

Transposition errors are held to be due not to an exchange of items but to two independent substitution errors. The a priori probability of two single substitutions being reciprocal is of course far too low to account for the proportion actually obtained without extra assumptions. The first of these is that the subject edits his output in order to avoid repeating an item when he knows that the list contains no repeats. Thus, if a subject, in recalling the sequence ****TP****, erroneously substitutes a P for the T, when he moves on to the next box and finds P, will, provided that he can remember that he has already said 'P', say something else

instead. Conrad's data do indeed indicate a clear tendency to avoid immediate repetitions during output, and while distant repetitions do occur in recall even though there are none in the stimulus lists, it is also the case that distant transpositions are relatively rare. This is explained by the premise that subjects have difficulty in keeping track of their own responses. What governs which items subjects will report in place of the ones they suppress in this manner? According to Conrad, "all other things being equal, items are available for response as a function of their recency of occurrence" (Conrad, 1965, p.167), a premise supported by Murdock's (1961) demonstration that intrusions in memory for single words, proactively inhibited by previous words, are most likely to be the immediately preceding words. The consequence of this is that the suppressed item is likely to be substituted for by an item presented on the present trial. When we add to this the third assumption that it will be an item acoustically similar to the suppressed item, it will be clear that in our example the probability of selecting 'T' as a substitute will now be quite high. Now in Conrad's description, the contents of the store of available responses are also determined by the subject's knowledge of the total stimulus ensemble, as well as their recency of occurrence, but the recency assumption is the crucial one, and, as Broadbent (1970) points out, it is often overlooked in references to Conrad's model; aside from present purposes it also distinguishes Conrad's account of phonemic similarity effects from the otherwise similar model of Sperling & Speelman (1970), who assume that substitutions are chosen at random from the total ensemble (see Ch. 5 for further discussion of this point).

Conrad's model has not been entirely free from criticism. McNicol (1971a) purports to test it by determining whether the observed frequency of transpositions was compatible with the assumption of selection of substitutions from RAS, and finds that on his calculations transpositions were too frequent. But there are two serious objections to this study. Firstly, McNicol's experiment involved dichotic presentation with order of recall post-cued, and secondly the assumptions he uses to derive his predictions appear to take no account of the dependence of the contents of RAS on recency. Hitch (1972) considers it problematic for the model that though digits have a relatively low mutual acoustic confusability compared to letters, transposition errors are still frequent in their serial recall. But this criticism is mistaken; Conrad does not and need not say that acoustic similarity is necessary for transpositions to occur, only that if they do they are most likely between acoustically similar pairs of items. The one remaining serious problem for Conrad's account of order retention is the phenomenon of associative intrusions. One possible answer to this, given their relatively low level of occurrence, is to attribute them entirely to the secondary memory component in immediate recall. This proposal is testable in at least two ways; firstly, they should appear with equal frequency in delayed recall; secondly, if one asked subjects to recall sub-span lists immediately at a maximum rate, and measured latencies for the output of each item (the technique of Anders & Lillyquist, 1971), since recall would presumably be entirely from active memory, items following repeated items should on my account be recalled no slower than other items, while according to Wickelgren, response competition should produce delays for these items.

Conrad is not very specific as to the properties of the response availability store, save that its contents are determined by the stimulus ensemble, the task requirements, and the recency with which the items have been presented. There seems nothing, therefore, to prevent the identification of the response-availability store with the logogen system, as described in Section 1.4. It will be recalled that item-traces at logogens influence availability as a function of their recency. Moreover since Conrad (1972) now agrees that the contents of the fixed-address store are articulatorily coded, it seems very similar to the response buffer. There are, however two problems to this identification. Firstly, the capacity of the response buffer is a fixed number of items (as is also the case for Atkinson & Shiffrin's rehearsal buffer) while Conrad's FAS appears to have as many slots as there are positions; forgetting from the response buffer is thus primarily by displacement while items in the FAS are lost through decay. Given the available evidence on decay versus displacement (e.g. Waugh & Norman, 1965) the displacement story now seems preferable. Moreover, the limited capacity of the response buffer suggests an additional mechanism by which order and position information may be lost while item information is not. The second problem is that Morton (1970, p.233) insists that "if recall is serial, information from the response buffer cannot be used by the subject, and will be eliminated by the act of recalling items at the beginning of the list". This is a consequence of Morton's assumption that input to the response buffer is automatic (see Section 1.4). However, if the contents of the response buffer are assumed to be utilised in free recall but not serial recall, then memory span should be four or five items greater for free than for serial

recall. Yet Crowder & Morton (1969) showed that running memory span for digits under free recall instructions was not more than five or six items altogether. Moreover, Morton cites experiments by Posner (1964) and Murdock (1968a) comparing serial recall with 'semibackwards' recall, in which subjects recall the last few items first and may therefore, Morton says, use the contents of the response buffer; but the overall level of recall does not differ for the two conditions, save for slow presentation in Posner's study. It is for these reasons that in Section 1.4 I concluded that the subject may select items for entry to the response buffer. It follows that the act of recalling items which occur at the beginning of the list need not displace the contents of the response buffer. There is additional positive support for this proposal in the finding of Waugh & Norman (1968) that in their probed recall task, a regularly occurring predictable item did not appear to displace other items stored in primary memory, while an unpredictable new item did. It does not follow that no output interference need occur. It remains possible that during output of the first part of the list, items are lost from the response buffer by decay, and this may account, for instance, for Murdock's finding of a smaller recency effect for serial than for probed recall.

It may be concluded, then, that the hypothesis of two forms of postcategorical memory, based on Morton's model as amended in Section 1.4, may be considered a more detailed version of Conrad's model. Order retention in serial and probed recall is mediated by the response buffer, which has a limited capacity for items coded as a string of articulatory features. When information about an item

is lost from the response buffer, by decay or displacement, during rehearsal or during recall, the item recalled in its place will depend on which traces are most active in the logogen system, which as a form of active memory does not appear to be subject to the same capacity limitation as the response buffer (Section 2.31). The most active traces will be those of recent items - i.e. items from the list. When no information about an item remains in the response buffer and the subject guesses at random, suppressing responses already made, then he will tend to recall an item from the list, rather than from the vocabulary as a whole. In addition, if any articulatory features of the item remain in the response buffer, when these are fed back to the logogen system the logogen activated will tend to be that of an item phonemically similar and from the list. Transpositions are therefore inevitable, and the items will tend to be phonemically related. Distant transpositions may be rare for two reasons: a) the subject tends to lose track of his own responses and thus repeats items (Conrad, 1965), and b) by the time he tries to recall late items in the list, early ones will not be very recent.

Of the remaining variables which were shown to influence order but not item errors, grouping and type of probe would appear to be factors influencing retrieval from the response buffer, and would therefore not be expected to influence item errors. The differential effects of rate of presentation, if any, may be accounted for by assuming that the maximum rate of input to the response buffer is lower than that at which logogens may be activated (see Section 2.34) so that confusions of order may occur at input.

In summary, I have tried to make a case for the claim that data on the nature of order errors, and on their relationship with item errors, are at least compatible with a view of active memory which supposes the existence of a) an order-preserving, phonemically-coded, 'slots' mechanism and b) a form of storage containing traces of items which influence the availability of these items as responses during recall, as a direct function of their recency. An associationist hypothesis according to which order is preserved entirely by item-position links does not account for transpositions and their relationship to phonemic similarity. A pure associative chaining hypothesis is counter-indicated by evidence for item-position links. Evidence on double probes which is difficult for a hypothesis combining both kinds of association will be reviewed in Section 2.13. Any hypothesis that item and order loss are entirely independent must cope with Conrad's (1965) evidence that they are not.

The task of serial recall presents an immensely complex problem of analysis, which is not helped by the probable existence of a contaminating inactive memory component. It was suggested in Section 1.3 that questions of the separability of attributes in active memory are better tested with probe techniques. It is to these that we now turn.

2.12. Direct comparisons of order and item retention.

In the previous section I mentioned some evidence that the incidence of item and order errors in serial recall is differentially affected by such variables as serial position, acoustic similarity, grouping, rate of presentation and retention interval. If this is

due to at least partial separation between the storage of items and of information about their order, as our model implies, then similar differences should be apparent when we contrast the retention of order and content more directly. This may be achieved by the use of suitable probe techniques: that is, we may compare item recognition either with order recognition or with probed recall. Ideally this comparison should be made within a single experiment, with all the other experimental factors, such as subjects, vocabulary, presentation rate etc. identical; then, any differences that do appear may definitely be attributed to the fact that we are asking the subject to retrieve different list attributes. Unfortunately, relatively few experiments have been done which fit this description. There is one interesting attempt to test the independence of position and item retention directly, by probing both on each trial and attempting to determine whether success on one class of probe was contingent on success on the other. This is a study by Cumming & Coltheart (1969), who conclude from their results that item and position information are stored as a unit and forgotten together. However, it must be doubted whether their experiment permits this conclusion. The positions were spatial rather than temporal, presentation was tachistoscopic (80 msec.), followed by a masking field, and the success rate was only twenty per cent. These conditions seem to implicate interdependence in terms of acquisition rather than storage or retrieval; that is, the economical explanation is that some positions on the display were simply not adequately encoded. Nevertheless, the principle is ingenious and would bear following up with an experiment involving serial presentation and a probe to which the subject is

obliged to respond both with a recognition judgement and with the position it occupied in the list; I would predict that the performance on the latter would be contingent on the former, but not vice versa.

Another negative result, in relation to the search for variables to dissociate our two stores, is found in studies of PI 'build-up' effects (Keppel & Underwood, 1962; Loess, 1964) in the Brown-Peterson paradigm. PI appears to build up over trials in the same way whether order or item information is required (Gorfein & Jacobsen, 1972; Petrusic & Dillon, 1972). However, this is not a very critical finding. Regardless of whether order and item information are stored independently or not, it may well be the case that both are stored regardless of the nature of the task. Moreover, there is some doubt whether the PI build-up effect is really an active memory phenomenon at all (see Baddeley, 1972).

Comparisons of order and item retention which do yield some differences may be summarised under the following headings:

1) Retention interval. Donaldson (1971) presented twelve-word lists, either visually or auditorily, followed by either an item or an order forced-choice probe, using the technique of Donaldson & Glathe (1969) (see Section 2.23). There were no modality differences, but marked serial position effects were obtained. While recency effects were apparent with both types of probe, they were much more marked for order probes, such that performance on these was little different from chance for the first half of the list. Moreover, a marked 'saw-tooth' effect was apparent in the serial position curve for the order probe condition, which Donaldson discusses in terms

of Norman & Wickelgren's (1965) postulate of rehearsal in non-overlapping pairs of items.

2) Acoustic similarity. It will be recalled that Wickelgren (1965c) demonstrated that the deleterious effect of acoustic similarity upon serial recall was restricted to its effect on order errors. We might therefore expect to find no effect of acoustic similarity upon item recognition. Wickelgren (1966c) performed an experiment on recognition of a single letter followed by a string of twelve distracting letters, whose acoustic similarity to the remembered item and the recognition probe was varied. The amount of forgetting did appear to be a function of the degree of similarity. But, Ingleby (1969) has pointed out that negative probes in this experiment were likely to have served also as interpolated items, and that Wickelgren's result would be obtained if subjects were inclined falsely to accept, regardless of acoustic similarity, any item presented on the current trial. In his own experiments, Ingleby avoided this problem by using digit shadowing as the interpolated task. He presented two letters (or trigrams), acoustically similar or dissimilar, to be memorised; after the digit-shadowing task, a probe was presented. When the subject had to say whether the probe had been present or not, there were no effects of acoustic similarity. But when he had to indicate whether the probe was the first or second item, there was a marked effect. However, effects of acoustic similarity on item recognition have been demonstrated when rather larger numbers of items are held in memory. Reicher, Ligon & Conrad (1969) had subjects shadow strings of up to twenty words followed by a two-choice recognition probe; acoustic similarity did produce a decrement in d' (see Section 2.2 for a discussion of this

measure). Moreover Hitch (1972), in a series of experiments on the effects of acoustic similarity on immediate recognition and recall, did find effects on recognition with eight-item lists. However, his experiments also suggest that these effects differ in a complex fashion from those obtained in recall. Further discussion will be postponed to Ch. 5, where my own experiments on acoustic similarity are presented.

3) Presentation rate. Henderson & Matthews (1971) compared order recognition and item recognition with eight letter lists at rates of presentation varying between a half and twelve items/sec. Alive to the possibility of confounding memory loss and perceptual loss when fast presentation rates are used, they also included a comparable monitoring task, with a probe presented prior to the list. The results showed that as presentation rate increased, $p(\text{correct})$ fell off more sharply with order probes for both the monitoring and the recognition tasks. The authors also calculate a 'pure' retention function, by subtracting monitoring from recognition $p(\text{correct})$, for each presentation rate (which assumes additivity). The result of this manipulation is not particularly clear; there is some suggestion that order retention is influenced more than item retention by presentation rate. However, the assumption that the monitoring task involved no memory load seems dubious: the two items of the probe had to be retained in addition to any demands made upon memory by the process of comparison.

None of these experiments, save perhaps for Ingleby's, leads to strikingly clear contrasts between the retention of order and item information. It will be suggested (in Section 4.1) that part of the

reason for this is the use of error rather than RT measures. Nevertheless, differences related to those discussed in Section 2.11 are apparent; acoustic similarity shows particular promise as a dissociating factor. Some further points of comparison between item and order recognition will be mentioned in the context of trace strength theory, in Section 2.23.

2.13. Probed recall: position and order.

I have suggested that the order of recently presented items may be preserved for some seconds in a temporally-ordered store which may tentatively be identified with Morton's response buffer. This hypothesis has been contrasted with Wickelgren's (1972) approach, which emphasises the importance of both inter-item and item-position associations. Memory for the order of a list of words may directly be probed:

- by presenting an item from the list and asking for its successor (a 'contextual' probe),
- by presenting an item from the list and asking for its position (a 'location' probe), or
- by presenting a position (indicated by display position or a number) and asking for the item presented at it (a 'position' probe).

It is of interest to compare the effectiveness of these different forms of probe, both singly and combined ('double' probe). On Wickelgren's account, we might expect the pattern of performance to differ with contextual and position probes, since these should selectively tap the two kinds of association. Moreover, a double probe consisting of both should be more effective than either singly, since the association strengths should add. According to the response

buffer hypothesis, however, the storage of order and position information are one and the same. A double probe would therefore not be expected to be markedly superior to a single probe, since a contextual probe and a position probe should represent largely redundant information to the subject. However, small differences in accuracy might appear between two kinds of probe as a result of different modes of access within the same store; I have suggested (in Section 1.4) that locations in the response buffer may to some degree be accessed directly given a position, but not a content, cue. Any differences found would therefore be expected to favour position probes, though differences due to mode of retrieval are more profitably investigated using latency measures (see Section 4.1 and Ch.6).

Murdock (1968a) used all three kinds of probe in different experiments and found broadly similar bowed serial position curves, with marked recency and a slighter primacy effect, for all, though there were slight differences, particularly for shorter lists. Woodward & Murdock (1968) directly compared the effectiveness of contextual, position (numerically defined) and double probes, with ten-item lists presented visually at a brisk rate, and immediate test. While a position probe was slightly more effective than a contextual probe, a double probe did not differ significantly in effectiveness from either. Nor could the double probe's lack of advantage be explained by subjects storing only order or position information, since it made no difference whether subjects had foreknowledge of the nature of the probe or not. Hitch (1972, Exp. 1) performed a similar experiment with eight-consonant lists presented sequentially across a spatial array. Recall of an item was probed with a light at its spatial position, the preceding item shown on a

separate display, or a double probe consisting of a light at its spatial position plus the preceding item displayed at the preceding position. He found that the double probe was slightly superior to the position probe, which in turn was superior to the contextual probe, though only the double versus contextual contrast reached significance. However, he argues that his result is not incompatible with that of Woodward & Murdock (1968) in that the slight superiority of the double over the position probe could in his experiment be an artefact: in the double probe condition, the presentation of the preceding item in the appropriate position would be likely to suppress the subject's tendency (seen in the error data for the position probe condition) to produce as an erroneous response the preceding item; allowing for this possibility on the basis of the error data virtually eliminates the difference. In a second experiment, Hitch (1972, Exp. 2) again showed that a double probe was only slightly superior to a contextual probe, and also that spatial position, when un-correlated with temporal position, was a very poor retrieval cue. The latter result supports the premise that order is temporally coded, and indicates that in the first experiment, spatial position functioned as an efficient probe only via a translation into temporal/ordinal position.

The evidence reviewed above is entirely compatible with the response buffer hypothesis, according to which the storage of order and position is not to be distinguished, though access via position may indeed be more direct; moreover, temporal rather than spatial representation of order is implicated. The general failure to find any additivity of the effectiveness of position and contextual cues is not congenial to Wickelgren's account in terms of inter-item and

item-position associations, and again implies that associative chaining has no important role in immediate memory. Of course, it remains possible for Wickelgren to argue that some attentional limitation prevents the subject from using both parts of a double probe; but Mcleod, Williams & Broadbent (1971) succeeded in showing an advantage for a double probe in a long-term memory situation, where presumably the same attentional demands were imposed.

2.2. Short term recognition and the concept of trace strength.

In the previous section, the discussion turned on how short-term serial recall and probed recall could best be accounted for: a theory according to which the temporal order of items was preserved in active memory as an intrinsic property of the memory structure was contrasted with an associationist viewpoint* emphasising associative chaining, item-position associations or both. I suggested that the former theory could account for the data on order and item errors provided that we assumed the additional and independent existence of a form of active memory influencing the availability of items as responses, in a manner determined by their recency. In the evidence that the factors influencing order errors did not have similar effects on item recognition, there was implicit the suggestion that the latter task was also mediated by this form of storage. We now consider this suggestion explicitly.

What is the nature of this 'availability' which is a function of recency? Tulving & Pearlstone (1966) point out that 'availability' in memory, which is a function of storage factors, must be distinguished from 'accessibility', which depends jointly on the condition of the stored information and the difficulty of the retrieval process. It has been widely assumed that problems of accessibility may be bypassed by the use of recognition tasks: for instance, Wickelgren & Berian (1971) say: "recognition appears to provide a pure test of loss in storage.....independent of retrieval interference (competition)...(This is) so plausible an assumption that most workers in verbal learning assume it to be true without proof". Similar views are expressed by Murdock (1968a) Broadbent

(1971) and Hitch (1972) with respect to short-term recognition, and Kintsch (1970) with respect to long-term recognition. Such an assumption implies a form of memory to which direct access is possible, one in which the representations are 'content-addressable' (Norman, 1968, 1973). We should note that the adequacy of this assumption for long-term recognition has recently been questioned by Tulving & Thompson (1971) and Mandler (1972), and that there is one major dissenter in the field of short-term recognition, namely Sternberg, whose work we shall consider in detail in Chapter 3. Theories of recognition memory have been dominated by the concept of 'trace strength' (and related concepts, such as 'familiarity' - Kintsch, 1970, and 'amount of information' - Freund, Loftus & Atkinson, 1969, to name but two). The major exponents of a trace strength theory for short-term recognition have been Wickelgren and Norman (Wickelgren & Norman, 1966, 1971; Norman & Wickelgren, 1965, 1969; Norman, 1966; Wickelgren, 1966c, 1967b, 1970a, 1970b, 1972; Wickelgren & Berian, 1971). I shall describe briefly the major assumptions of their theory, and then consider the validity of the assumptions separately for item and order recognition.

2.21. Trace strength theory.

The major assumptions are the following:

1) For every stimulus item there exists a unique and specific representation in memory (cf. Wickelgren, 1972, quotation in Section 2.11).

2) Upon presentation of a stimulus, the subject has direct access to its representation, without a search process.

3) Stored with or at the representation there may be considered to be a unidimensional continuous variable; the 'trace strength' which the subject is able to assess with respect to a decision

criterion located on that dimension.

4) Strength is incremented on presentation or rehearsal of an item. Subsequently it suffers exponential decay as a function of time and/or the number of intervening items processed. Now an item's strength is in fact held to be the sum of several components with different time courses, each of which show exponential decay: a short-term trace which is consolidated instantly and decays within 1-20 seconds or so, and longer-term traces, of which Wickelgren (1970a) now distinguishes two: intermediate-term and long-term memory. However, the latter have, if any, only a minor effect in most short-term recognition studies owing to their long consolidation period (Wickelgren & Berian, 1971) and may for present purposes safely be ignored.

Thus the subject's decision in recognition is claimed to be based on the value of a single unidimensional variable which is a direct function of the item's recency, though it may also be affected by other variables such as the duration and clarity of presentation. To this decision may be applied the machinery of statistical decision theory, given certain assumptions about strength distributions (see Wickelgren & Norman, 1966, for details). This enables the empirical separation of the memory process and the decision process. By plotting the way in which the probability of making a false positive judgement varies with that of making a correct positive judgement (the 'memory operating characteristic' or MOC), it is possible to derive a pure measure of memory strength (d') uncontaminated by the subject's bias towards responding positively, which is reflected in the criterion parameter (β). The relative trace strength, d' , is in these terms conceived as the distance between the means of the strength distributions of 'old'

and of 'new' items. By obtaining a measure of d' at various values of delay, etc., with rehearsal controlled, it is possible to study the manner in which trace strength depends on recency and other variables.

2.22. Trace strength theory as an account of item recognition.

I shall consider the four assumptions listed above in turn.

1) The assumption of direct access has been made almost entirely without empirical backing. It is in fact difficult to test if one uses only accuracy as a measure; I know of only one attempt - an experiment performed by Hitch (1972, Exp.7). Hitch reasons that if some form of search process is necessary to locate the item in memory even in recognition, then performance should be facilitated if, together with the probe, a cue is presented which tells the subject where the item is located in memory. Accordingly he compared recognition performance when the probe was presented with and without a spatial position cue, indicating the position it had occupied if it was in the list, for eight item lists presented sequentially across an array. The MOCs he obtained, and the variation of d' with serial position (which showed marked recency) were essentially identical in the two conditions - only the criterion showed any effect. Note, however, that to conclude in favour of direct access is to make two important assumptions:

a) that if search were necessary for retrieval in recognition, it would be through a memory ordered in terms of position, so that a position cue would provide an effective short cut. But, in spite of the fact that position is an adequate cue for recall (Murdock, 1968a) it remains possible that the store normally mediating

recognition is organised in terms of some other attribute, so that position could be useless as a retrieval cue even if search were necessary.

b) that if a position cue did cut short a retrieval process and so speed access, a significant amount of decay would occur during the time saved - so that a difference in d' would actually show up.

Hence latency would appear to be a more powerful means of testing the direct-access assumption; we shall consider RT studies of short term recognition in Chapter 3, and the issue of search versus direct access will prove crucial.

2) Evidence for strength theory in general, and the assumption of unique representations in particular, comes from experiments in which subjects are asked to make judgements of the recency with which an item has been presented. Yntema & Trask (1963) presented lists of words and asked subjects to say which of two words from the list was the later. Accuracy declined as both the spacing of the two words, and the delay between the second word and the test, increased. The authors interpret their result in terms of a 'time-tagging' hypothesis, but Morton (1968) suggests that subjects base their decision on a comparison of the two items' strength, which must decay non-linearly. (See also Hinrichs, 1970, for a detailed review of judgements of recency in these terms). Morton showed that if the first of the two words was repeated in this paradigm, performance declined to below chance, while repeating the second word improved performance; Fozard & Yntema (1966) obtained a similar result using picture stimuli. In like vein, Peterson, Johnson & Coatney (1969) who required absolute rather than relative judgements of recency, showed that repetition increased apparent recency, provided that

repetitions were not too widely spaced (Peterson, 1967). It is important that these results are not predicted by trace strength theory unless it is assumed that the two presentations of the repeated item increment the strength of the same trace: that is they imply the validity of assumption (2) - that the representations underlying judgements of recency are unique in the sense that they are specific to 'type' not to 'token'. Comparable effects of repetition have been found for short-term recognition (Wolf & Jahnke, 1968; Jahnke, 1972).

3) The assumption that recognition decisions are based on an underlying continuous dimension of strength seems to be borne out by the fact that if subjects are asked to rate the confidence with which they judge the probe item to be old or new, the resulting MCCs and a posteriori probability functions are of the form predicted by statistical decision theory rather than by two-state threshold models (Murdock, 1965; Norman & Wickelgren, 1965; Wickelgren & Norman, 1966). Also, performance in item recognition and multiple-choice recognition experiments appears to be mutually predictable on the basis of SDT (Norman & Wickelgren, 1969), though some discrepancies are apparent. However, as Anderson & Bower (1972) point out, these are successes for the technical machinery of SDT, not for the basic strength assumptions. For instance, it has been shown that the results from these rating scale experiments may also be predicted by a version of the low-threshold two-state model in which the non-overlapping rule relating states and rating categories is relaxed (Broadbent, 1966; Wickelgren, 1968). While this need not deter us from the use of d' as a bias-free measure (Broadbent, 1971) it does reduce the force of the argument that the applicability of

SDT justifies the assumption of an underlying continuous strength variable.

4) More impressive is the emergence of simple empirical relationships between strength as measured by d' and retention interval, when rehearsal is controlled - relationships which are not evident if the data are plotted simply in terms of error rate. For instance, Wickelgren & Norman (1966), who presented a probe immediately after lists of between two and seven three-digit numbers found that strength decayed as a simple exponential function of number of intervening items; Norman (1966) presented lists of digits, digit pairs and backward digits followed by a recognition probe and found the same decay function; Norman & Waugh (1968) found exponential decline in d' with both list and test position of probe; Wickelgren & Berian (1971) demonstrated exponential decay when the interval was filled with a counting backwards task. Moreover, the rate of decay appears to be unaffected by serial position (Wickelgren & Norman, 1966), type of item (Norman, 1966), number of repetitions (Wickelgren & Norman, 1971) and storage load (Wickelgre & Berian, 1971). The effect of these factors may be attributed entirely to differences in initial level of acquisition. Decay rate appears to be a function both of the number of intervening items and of time (Wickelgren, 1970b); also decay rate seems to depend on the nature of the intervening activity, being much smaller in studies employing the 'distractor' design than in single list probe studies (Wickelgren & Berian, 1971). However, the main point for present purposes is that in all these studies, in which rehearsal is controlled both by fast presentation rate and by instructions, the retention functions over the first ten seconds or so appear to be simple exponentials.

The simplicity of the decay functions argues both for the appropriateness of d' as a measure, and for the assumption that the decay of an effectively single representation for each item is involved.

The application of trace strength theory to item recognition is not without difficulties. Broadbent (1971) points out that the data of Wickelgren & Norman (1966) and the recognition data of Norman (1966) provide very little information on items separated from the probe by less than three or so items, since very few errors are made at so short a delay; moreover, for such data as are available for recall of very recent items, MOC curves (plotted on normal probability graphs) are not particularly well fitted by straight lines with a slope of unity. When errors are rare, the only other sensitive index is reaction time (see Ch.3).

A second problem concerns the nature of trace strength. The simplest way to conceive of strength is as a quantity intrinsic to an item's representation in the same way that degree of resonance is intrinsic to a bell recently sounded; the duration and intensity of its resonance depend on the force with which it was struck. Given the dependence of trace strength on recency, an item recognition task such as that employed by Norman (1966), in which a list of items is followed immediately by a probe, is performed essentially as a judgement of recency. If the strength exceeds a criterion, then the probe item must have been presented recently enough to have been in the list; if not, then it is new. (Of course, trace strength is a function not only of recency, but also of the duration and frequency of presentation.) Now, Wickelgren (1972) deviates somewhat from this metaphor in that he considers strength to be intrinsic not to the representation of the item itself, but to that of its association to a 'familiarity representation'. To

me this is both unclear and unparsimonious, but it really makes no substantive difference and may perhaps be accounted merely an associationist version of the same story. A more serious objection arises when we consider the views of Anderson & Bower (1972), who have argued forcefully that in long-term recognition subjects cannot be basing their decisions solely on the value of a unidimensional strength variable which is a function of recency, frequency and presentation duration. For one thing, subjects are able to differentiate sets of items, which on these grounds are presumably equal in strength, by means of other attributes, such as who said them and how and where they were presented. For another, it does not prove possible to trade frequency for recency in a manner predicted by their proposed functional equivalence (Winograd, 1968). They therefore invoke the idea that in recognition subjects test not the strength of the item, but the adequacy of the associative links between the probe item and contextual cues or 'list markers'. The claim that interrogation of contextual information is necessary for recognition is also made by Norman & Rumelhart (1970), who assert that we do not store items at all (they are permanently stored already) but rather store information about the context in which they have appeared. The subject bases his recognition decision on whether a sufficient number of attributes of the test item can be matched with a consistent and appropriate context. To be sure, it is possible to sum over these all-or-none 'vectors', and the resulting overall 'strength' measure will behave very similarly to that of classical trace strength theory (see Norman & Rumelhart, 1970, pp. 47-50). But the underlying conception of the memory structure mediating performance is very different: subjects judge the number

and adequacy of associative links to the item rather than its recency.

There are at least three reasons why we need not apply this approach to short-term recognition. Firstly, it would be surprising if the brisk presentation rates typically used permitted the formation of rich contextual associations. Secondly, contextual cues change very little from one trial to the next and barely at all during a trial; yet subjects are able to make quite good judgements of relative recency. Of course, it could be claimed that list position is a contextual cue, but Hitch (1972, Exp.7 - see above) has shown that presenting position context with the probe does not facilitate recognition. Thirdly, it is possible to influence apparent recency with frequency of presentation in both short-term recognition and judgements of recency (see -2- above). On the basis of present evidence, there is no need to bring contextual information into short-term recognition: most experiments have enabled subjects to base their decision on recency alone. In the rare studies where the material interpolated between list and probe does contain items which may also be in the list, (Wickelgren, 1966c; Norman & Wickelgren, 1965), the data presented do not permit one to determine whether most false positive responses were made on interpolated items which also served as list items; even if they were not, it remains possible that subjects could be basing their response on whether the strength of the probe item fell within a criterial 'band' of strength (Hinrichs, 1970).

A third problem arises from the question: what counts as an item? Wickelgren & Norman (1966, 1971) imply one possible answer to this question by treating three-digit numbers as items. It would be just as reasonable, a priori, to treat them as ordered

sets of three items; in this connection it would be interesting to know how the exact relationship between negative probes and list items in terms of shared digits or digit pairs affected performance in these experiments. At the other extreme, we find Norman & Rumelhart (1970) following Bower (1967) in treating even a single digit as a collection of independent component attributes. Another question is: do the items have to be familiar to the subject as items in order to have a 'unique internal representative' (Wickelgren 1972). Norman (1966) apparently obtained similar decay rates for digits played backwards, which were presumably not initially familiar, and for natural digits, though the acquisition parameters did differ.

In summary, there seem to be good empirical grounds for the claim that the recent presentation of an item is represented as an increment in a unidimensional variable located at a content-addressable representation specific to that item (qua 'type'), that this increment decays, when the item is no longer present, exponentially as a function of time and/or intervening items to an asymptote in roughly 10-30 seconds (or items), and that decisions based on this 'trace strength' may underly short-term recognition and judgements of recency. This conclusion is not yet well founded for very recently presented items, nor is it necessarily the case that this is the only way an item is represented. There are also certain conceptual problems outstanding concerning the degree to which the representation is unitary, and whether strength is a quantity possessed by an item's representation or by its associations to other representations.

2.23. Trace strength theory and order recognition.

We have so far encountered very little evidence for the proposition that inter-item associations play a major role in short-term memory. The comparison between single and double probes (Section 2.13) indicates the contrary, and of Wickelgren's arguments from the nature of order errors (Section 2.11) only that based on the phenomenon of associative intrusions bears any weight; however, as was previously remarked, it is possible that associative intrusions result from an inactive memory component in recall, rather than from associations represented as such in active memory. The use of recognition procedures and strength measures should yield more evidence on whether associations between items are represented as such in active memory.

Wickelgren (1967b) tested the assumption that d' measured by response probabilities in recognition of an inter-item association was independent of competing associations to the same 'stimulus' member. He presented twelve-digit lists in which only two items were repeated, and probed with an ordered pair. The subject had to decide whether the second item had followed the first in the list, and the first digit of the pair was sometimes a repeated item, sometimes not. The results provided no evidence that d' for the pair A-B was reduced by strengthening the pair A-C in the same list. (In fact, there was some evidence that when pairs at the beginning of the list shared a common first term, their strength was actually increased.) However, it should be noted that this experiment does not imply that there is direct access to the representation of an association, only that, given both members of a pair, a subject is not much influenced in his assessment of the probability that they

were adjacent in the list by the participation of the first item in an ordered pair elsewhere in the list. When we come to RT experiments on probed recall of lists without repeats (Section 3.3 et seq.) evidence will be presented which suggest that when a subject is required to recall the second member of a pair from the list given the first, he does not appear to have direct access to the stimulus member, as according to Wickelgren (1972) he should.

Are the MOC curves and a posteriori probability functions obtained in order recognition experiments compatible with the assumption of an underlying strength continuum for the representation of associations? An experiment by Norman & Wickelgren (1965) seemed to suggest that they were not. They required subjects to judge whether both members of a probe pair were in a list of digits, after some interpolated digits, and, if they were, whether their order was correct. They obtained discontinuous MOC functions for order recognition which appeared to be composed of two straight line components. But Donaldson & Glathe (1969) argue that since Norman & Wickelgren plotted the conditions with two old digits against that involving two new digits, they confounded order and item information. Donaldson & Glathe repeated the experiment so that order probes were always adjacent pairs from the memory list and subjects had merely to decide whether their order was correct: they obtained smooth, continuous and symmetrical MOC curves for both order and item probes (d' was markedly lower for order probes and d' plotted against serial position yielded serial position curves much more bowed for order than for item probes). Norman & Wickelgren (1969) performed an experiment using lists of four paired-associates (P-As) always composed of the same four stimulus and response terms,

probed with a P-A composed from the same stimulus and response terms (so that subject had to judge whether the association was correct or not), after three to six seconds of distracting activity. They also found smooth symmetrical curves and monotonic a posteriori probability functions. Murdock (1965), who followed six-item lists of P-As with a P-A recognition probe, also found that d' was independent of false alarm rate for the first four pairs of the list. But, for the fifth item there was some evidence that d' changed with false alarm rate, and the last item was, of course, almost always recognised correctly. So it is necessary to extend the caveat first mentioned in the context of item recognition: firstly, these experiments provide no evidence on memory for very recent associations, since performance on these is almost error-free; secondly, for associations separated from test by up to five to six items (if in the case of Murdock (1965) we include the test pair) evidence for the continuous strength assumption is not particularly good. For older items, however, the data is compatible with the assumption of a strength continuum for associations.

Does d' for associations suffer exponential decay? The evidence is meagre. Murdock (1965) found no change in d' over the first four items, though the presentation rate may have permitted some rehearsal, which would have obscured the retention function. Wickelgren (1967b) claims that his data show exponential decay, but, as he admits, the scatter introduced by the several conditions makes this conclusion highly tentative; only four delay values were used, and d' did not change between the longer two (six and eight intervening items).

In conclusion, the evidence for the applicability of trace

strength theory to inter-item associations in short-term memory is equivocal. The possibility of all-or-none representation is not ruled out for very recent items, and for items retained for longer periods, for whose associations the continuous strength assumptions do seem to hold, there is not yet good evidence for the exponential decay of those associations that one could expect if they were represented in active memory.

2.24. Is a strength measure possible for recall?

In view of the utility of SDT's strength measure d' , it would be useful to have a means of computing it from recall probability. Both Norman (1966) and Murdock (1968a) claim that this is possible. The problem may best be appreciated in the context of Murdock's data. He probed recall from five-, eight- or eleven-word lists with either a position or a contextual probe, and encouraged subjects to guess; plotting $p(\text{correct})$ against serial position probed he found marked recency and mild primacy effects for both kinds of probe. But when the data were plotted in terms of the serial position from which the response had come, it could be seen that though subjects responded less often with early than with late items, they were more likely to be correct when they did. Murdock (1966, 1968a) interprets this as evidence for a shift in decision criterion in the direction of greater risk for later serial positions. He therefore advocates the use of a d' measure which he calculates separately for each serial position, using as false alarm rate the probability of intrusions from that position given other probes. Norman (1966) uses the probability of intrusions at that position, on the assumption that each of the possible incorrect responses is equally likely.

These procedures seem to me entirely dubious. Firstly omissions

are ignored. Secondly and more seriously, in probed recall the subject needs to retain information about the order/position of the items. If, as has been repeatedly suggested in previous pages, the occurrence of items is retained in a form independent of the order-preserving store, Murdock's finding is predicted without reference to changing decision criteria. For, if the subject guesses when the correct response is not available, the likelihood of his guessing a particular response will reflect its availability as an 'item-trace', which, it has been argued, is a direct function of its recency of occurrence (Sections 1.4, 2.11, 2.22). Thus incorrect guesses will tend to be recent items, the tendency noted by Murdock. In any case, there seems no good reason why subjects should change their decision criteria for different serial positions.

In the light of the possibility that the occurrence of items may be represented in active memory independently of the form of storage in which their order is preserved, attempts to obtain measures of d' from the number and/or source of intrusions at different retention intervals in probed recall would appear to be fruitless.

2.3. Active memory: capacity, function and coding.

The studies of short-term forgetting reviewed in the last two sections have suggested two rather different conceptions of the nature of representation in active memory. The evidence on serial recall and on comparisons between the different methods of probing recall has given rise to the idea of active memory as a limited capacity, phonemically-coded store, which retains a short string of items in their order of input and thus retains temporal order information as an intrinsic property of its structure. On the other hand, data from studies of item recognition and judgements of recency, together with the fact that the response buffer hypothesis cannot by itself account for the nature of transposition errors, have seemed to require that the occurrence of an item is represented as an 'item-trace' - the temporary activation of a permanent content-addressable representation (a logogen) the degree of activation being a simple, probably exponential, function of recency. The former conception is implicit in Conrad's (1965) 'fixed-address store', Atkinson & Shiffrin's (1968) 'rehearsal buffer', and Reitman's (1970) 'waiting room'. The latter is represented in Norman's (1968, 1969b) description of 'primary memory'. I have suggested that the two may be combined in a hypothesis of two forms of post-categorical active memory, and that we may tentatively identify the first kind with Morton's 'response buffer' and the latter with Morton's 'logogens', with the amendments described in Section 1.4. Armed with this hypothesis, we now consider some rather general questions of the functions and format of active memory.

2.31. The capacity of active memory and speech comprehension.

Among the functions most commonly suggested for active memory (aside from the temporary retention of telephone numbers, a comparatively recent human need in terms of the evolutionary time scale) are speech comprehension (Neisser, 1967; Craik, 1971) and speech production (Morton, 1969a, 1970). Without any very profound analysis of comprehension, it is obvious that in listening to or reading continuous prose, a man must retain in memory at least the semantic identity of the words that have recently occurred, in order to understand the word currently perceived. He must do this in order, for example, to understand the reference of a pronoun, or of a verb divorced from its subject, to disambiguate a present word by using its past context, and to take advantage of the redundancy of language when the speech stream is masked by coughs or low-flying aeroplanes. Moreover, it is occasionally necessary for words that have already been seen or heard to be held in a relatively unanalysed form, because their disambiguation waits upon future context. It is very striking how large is the span of intervening material over which this retention faculty appears to operate, compared to conventional estimates of the capacity of 'primary' memory. Also, if one asks subjects to repeat continuous prose verbatim, their span (12-20 words) is considerably in excess of their span for unrelated material. Murdock (1972) finds these discrepancies so hard to swallow that he makes them the basis of his main argument against the 'consensus' view of short-term memory, represented by Waugh & Norman (1965) and Atkinson & Shiffrin (1968), of whom he says:

"Such models posit perfect memory for a very small number of items, while the redundancy of language demands an STM system capable of holding imperfectly a large number of items. To comprehend written or spoken language, one must keep in memory fairly large

segments of unprocessed or partly processed information before making linguistic decisions....." (Murdock, 1972, p.91)

Where do the estimates of capacity come from? Morton's (1970) estimate of the capacity of the response buffer, four or five words, is based on the observation that this is the upper bound of the ear-voice span in shadowing and the eye-voice in reading (Morton, 1964c). Also, order errors in reading may involve the repositioning of a word by up to five places (Morton, 1964c). Also, in running memory experiments, subjects typically recall no more than five or six items (Crowder & Morton, 1969). Atkinson & Shiffrin (1968) estimate a 'lower limit on rehearsal capacity' from digit span experiments, by identifying the series length at which the subject never errs, which they say is in the range five to eight items. Sperling & Speelman (1970) derive from the results of their partial report experiment an estimate of capacity which is, however, somewhat dependent on the assumptions of their model. They claim that 'auditory short-term memory' can contain 7.5 acoustically different letters, or 5 acoustically similar letters, plus the two words of the partial report cue which are presented after the memory list.

Estimates of capacity derived from free recall experiments tend to be somewhat lower than those we have already reviewed. Craik (1971) reviews several methods of estimating 'primary memory' capacity from the recency effect in free recall: Waugh & Norman's (1965) correction for the probability of recalling the item from secondary memory, Baddeley's (1970) estimate from the difference between immediate and delayed recall, and a method due to Tulving & Colotla (1970). He concludes that "most workers now accept that primary

memory can hold between 2.5 and 3.5 words", and that therefore the immediate memory span must include a considerable secondary memory component. My own feeling is that these methods underestimate what they attempt to measure, partly because they do not make adequate allowance for the effects of output interference, but more crucially because it is probable that in free recall, subjects employ part of their response buffer capacity to retain information from earlier in the list (see Section 2.32 for further discussion of this point). So these estimates may not in fact be incompatible with those we have so far mentioned.

How are these estimates of capacity, all derived from recall, reflected in recognition performance? The data of Donaldson (1971) (Section 2.12) show recognition performance to be good after up to twelve intervening items, while a comparable measure of order recognition declines rapidly to a level near chance after five or six items. Norman & Waugh's (1968) results imply that decay of item strength is still occurring after up to ten or twelve intervening items. According to Wickelgren & Berian (1971), decay of the short-term trace continues after more than ten seconds of counting backwards in threes. Moreover, as Tulving & Madigan (1970) point out, judgements of recency are possible for items located on the flat portion of the free recall serial position curve, and therefore presumably outside the compass of primary memory capacity as estimated above (Hinrichs & Buschke, 1968; Yntema & Trask, 1963). Thus it appears that information upon which item recognition and judgements of recency may be based remains available in active memory long after the limit on capacity derived from recall measures would suggest.

The classical answer to this dichotomy would be in terms of different 'thresholds of strength' for recall and recognition (Kintsch, 1970); but the data of Donaldson¹ (1971) suggest that the distinction

¹see Section 2.12

between order and item information is the critical one. So the dichotomy may more successfully be resolved by our hypothesis of two forms of active memory (Section 1.4), if we assume that the number of intervening items required (at standard presentation rates) for 'item-traces' at logogens to decay to an effective asymptote is considerably larger than the item capacity of the response buffer. On these grounds, we should not expect item recognition and judgments of recency to be limited by response buffer capacity, since they are mediated by judgments of the strength of 'item-traces' (Section 2.22). But this suggested resolution raises another problem: why is free recall, which by definition does not require the retention of order, apparently limited by response buffer capacity? Why can the subject not use item information stored in the form of item-traces? We must consider more closely the nature of access to 'item-traces' at logogens.

I have insisted that logogens are 'content-addressable' in Norman's (1968) sense, that is, they are activated directly by the physical and/or semantic features of the item they represent. By 'physical' features I intend to refer to necessary features of the item, like its visual shape, or auditory distinctive features or articulatory description, not to contingent features such as its location in time and space, its colour, pitch, etc. Semantic features are intended to include context and expectancy. For the item-trace at a logogen to influence behaviour, it is therefore necessary that a 'query' (Norman, 1968) be put to the logogen system formulated in terms of physical/semantic features, not in terms of contingent temporal or spatial properties. What this boils

down to is the following. The query, 'Give me recent items.' is not a legitimate one for the logogen system, save perhaps for the most recent one or two logogens to have been activated. Thus in free recall, the subject cannot make direct use of item-traces. There are, however, two forms of query which are legitimate:

1) 'How recent is (=what is the trace strength of) item X (i.e. an item with features a,b,c,d,e.)?'

2) 'Give me any item with features a,b,c.....' In this case, the first response to become available will be the item represented by the most active logogen addressed by those features.

Queries of the first form are involved in item recognition and judgements of recency. They are probably also involved in a task employed by Shulman (1970). Shulman presented a ten-word list and asked subjects to say whether a homophone or synonym of the probe word had been in the list. His data showed no recency-related dissociation between the two kinds of probe: the relative efficiency of the two kinds of probe was the same at all serial positions and both showed a recency effect. Shulman feels that these results imply both semantic and acoustic coding at all retention intervals and are therefore incompatible with conclusions drawn by Kintsch & Buschke (1969) from their experiment. They presented 16-word lists, followed by a probe; subjects attempted to recall the item following the probe. The list contained either synonyms or homophones of the probe. The serial position curve showed a flat early portion and a recency effect over the last few positions. When the latter positions were probed acoustic similarity had an adverse effect, while semantic similarity had none. For earlier positions, the reverse was true. The authors interpret this dissociation, congruent with the recency effect, as evidence for acoustic coding in primary memory, semantic

coding in secondary memory. However, the results of the two studies are not in conflict. The test item in Shulman's study constituted a recognition probe for a set of physical or semantic features, and would therefore provide access to item-traces in the logogen system whose strength could be compared against a criterion. In contrast, Kintsch & Buschke's subjects had to recall an item, and for the last five or six items this would chiefly be via retrieval from the response buffer, which is, of course, phonemically coded.

Queries of the second form are involved in the account given of the occurrence of transpositions in serial recall in Section 2.11, and explain how the logogen system can function as Conrad's 'store of available responses', from which responses are available as a function of recency. If no phonemic information about an item remains in the response buffer, the query may be simply: 'Give me a letter', if some information remains, the query will be: 'Give me a letter with phoneme Y'. Essentially this provides a mechanism for ensuring that guesses are recent items and often phonemically related to the items they replace.

Finally, I should like to suggest that the problem of reconciling immediate memory span and sentence comprehension which initiated this section is resolved, since connected prose will allow subjects at least some access to information retained in the form of item-traces, in that the redundancy of language will provide effective queries of the second form. Thus, sentence span will be related to the rate at which item-traces cease to be available, and need not be limited by response buffer capacity.

An example may help. A test case for any account of the role

of active memory in comprehension is the phenomenon of 'retro-active disambiguation'. A classic example is the sentence, 'Rapid righting with his uninjured hand saved from loss the contents of the capsized canoe'. The problem posed by this sentence is that most people hearing it initially interpret the second word as its homophone 'writing'. The fact that they are nevertheless able (on the whole) to interpret and remember the sentence correctly may seem to imply that some non-semantic representation of the word 'righting' must be retained over a span of some twelve words. However, the following informal account can be given on the basis of the present model. When the word 'righting' is heard, both the logogens [writing] and [righting] receive an equal increment in trace strength. Since the word 'rapid' will already have incremented the strength at [writing] more than that at [righting], the interpretation 'writing' will be dominant. The traces at both logogens now suffer decay, but both receive a roughly equal increment from 'with his uninjured hand', which is appropriate context for both words. The words 'capsized canoe' will serve as semantic input only to the logogen [righting], as a result of which its trace strength will preferentially receive an increment causing it now to exceed that of [writing]. Hence the interpretation 'righting' will finally become dominant. Essentially, the argument is that choice of the initial interpretation 'writing' does not destroy the activation of the logogen [righting]. (There may be additional processes at work - e.g. an inhibitory process preventing any but the dominant logogen of a homophone pair or triple from influencing the cognitive system.) This account is obviously very simple-minded; I dip an ignorant toe into these treacherous swamps only in order to suggest that given the account

of active memory favoured here, Murdock's problem is not intractable.

2.32. Broadbent's 'address register'.

Broadbent (1970, '71) has also argued recently that a single form of post-categorical short-term memory store is inadequate. He has accordingly proposed two. They are by no means homologous with the two forms of active memory which I have been advocating. He calls them:

- 1) 'primary sensory storage' (the 'in-tray') which, though distinct from iconic or echoic storage and post-categorical, is otherwise sensory in nature and subject to decay, and
- 2) the 'address register' (the 'desk-top'), which contains a fixed number of items, subject to displacement, which function as the addresses of locations either in primary sensory storage or in long-term memory. All categorised items also pass into the long-term store, but they tend to become inaccessible unless an appropriate address is held in the address register. The address register is thus a store which functions primarily as a control or retrieval system.

Two of the arguments put forward by Broadbent are critical for this hypothesis. The first is that, if both the recency effect in free recall and the effects of acoustic similarity are attributed to the same short-term memory component in recall, it is strange that "experiments such as those of Baddeley (1968) show acoustic similarity effects to be just as great at the beginning of a serial list as they are at the end" (Broadbent, 1970, p.345). Nevertheless acoustic similarity effects do die away after distracting activity. This lack of congruity between recency effects and acoustic similar-

ity effect implies, he argues, two forms of active memory; since as we have already seen, acoustic similarity appears primarily to affect recall rather than item recognition, Broadbent proposes that the store affected by acoustic similarity is the address register. Broadbent's second argument depends on some evidence of Tulving & Patkau (1962). They presented word lists considerably in excess of span for free recall and scored recall in an unusual way. Any sequence of responses contiguously reproduced was counted as a 'chunk' if the words had also been adjacent in the stimulus list. The number of 'chunks' reproduced appeared to be constant, at around seven. Broadbent interprets this as indicating the existence of a mechanism containing a fixed number of addresses, from each of which access can be had to the locations in long-term memory or in the short-term store at which the items constituting a chunk are stored.

Unfortunately, recent evidence implies that the first of these arguments is mistaken. The experiment of Kintsch & Buschke (1969) demonstrated an exact congruity between acoustic similarity effects and the recency effect (see Section 2.31). Then how are the data of Baddeley (1968) which Broadbent quotes to be accounted for? The simple answer would appear to be that when six-letter lists are presented at one item/sec. for immediate self-paced recall, rehearsal is not adequately controlled for Broadbent to draw the conclusions he wishes to draw; if the subjects rehearse, serial position is likely not to represent accurately the number of items or time since presentation. The differential effects of acoustic similarity on recall and item recognition are therefore better accounted for by

our own two-store hypothesis, according to which phonemic similarity affects response buffer retrieval but not decisions based on 'item-traces'.

The argument from Tulving & Patkau's (1962) finding is more convincing. But it does not seem necessary to me to postulate an additional store to serve as an address register. Inactive memory appears to be largely content-addressable. If this is the case it implies that items are addresses. Thus items stored in active memory, in whatever form, may function either as items in themselves or as retrieval cues or 'addresses' for access to inactive memory representations. In the experiment of Tulving & Patkau (1962), and probably in most free recall experiments, it is probable that subjects retain items in active memory from early parts of the list in order that they may serve in this fashion as retrieval cues for related items stored in inactive memory. (Note that this implies an ability to select items for retention in active memory: that this was possible for the response buffer was suggested in Section 1.4 and argued in Section 2.11.) The more they do this, the less space there will be in the response buffer for the retention of recent items. It was suggested in Section 2.31 that this explained why estimates of primary memory capacity from the recency effect in free recall (Craik, 1971) were so small.

In summary, Broadbent's arguments do not in critical respects necessitate his two-store model, and provide no difficulties for our own.

2.33. Active memory as a working memory.

I have mentioned the role of active memory in speech comprehension and in keeping track of information recently stored in inactive

memory. Another function sometimes suggested is the organisation and reorganisation of material currently being operated upon. For instance, Hunter (1964) emphasises the need for a 'working memory' in problem-solving. Posner (1967) speaks of an 'operational memory' for organising material reactivated from long-term storage. Some form of temporary storage is also implied by the need to keep track of one's progress during the process of retrieval.

Two schematic examples will serve to indicate the need for an active (as opposed to passive) organisational role for active memory. Firstly, consider all operations in problem-solving which require the combination, or inversion of the order of, two verbal elements A and B of a sequence stored in active memory. The evidence reviewed in this and the next chapter, and experiments reported subsequently, make it clear that the elements of an ordered sequence in active memory are not simultaneously and instantly available, but that access requires a serial retrieval process whose rate is roughly that of sub-vocal speech. This has the implication that when the item A has been retrieved, it must be held in some other immediately available form of representation while the item B is retrieved. It is suggested that this is as the most recent 'item-trace' at a logogen. Consider also the problem of keeping track of one's own recall from active memory. I shall in Section 3.3⁴⁴ mention evidence for Conrad's (1965) claim that in order to recall a list backwards, the subject is obliged to run through the list forwards from the beginning for each item (or pair of items) retrieved. The question arises; how does he keep track of where he has got to? Again, there is a necessity for some form of active memory independent of the store actually being searched.

It is an important feature of the two-store model of active memory introduced in Section 1.4 that it permits this organizational role. For this to be the case, the amendments made to Morton's account are crucial. It will be recalled that I proposed that the subjects could select which items made available by the logogen system were transferred to the response buffer, and could also control at which position in the response buffer he started the process of feeding articulatory descriptions back to the logogen system (sub-vocal rehearsal or read-out). These control processes provide a means whereby information may be 'shunted' back and forth between the two forms of storage, reordered and combined. Moreover, if one postulates only one store, and if this store must be serially searched, then the necessity to retain a retrieval cue while one searches for the required item generates a paradox: crudely, one has to let the cue go, in order to find the item; but one does not know whether one has found the right item unless one has the retrieval cue available; but retrieving the cue requires one to let go of the item again. The postulate of two forms of storage avoids the paradox (though other solutions are no doubt possible). Whether the elements contained in the model provide enough flexibility for a working memory (for verbal processing only, of course) must remain a topic for future work.

We have seen that the model of active memory proposed in Section 1.4 permits it to fulfil three functions: a temporary store in linguistic processing, an 'address register' for keeping track of information recently stored in inactive memory, and a working memory with a capacity for the reordering and combination of verbal items and the control of retrieval. In reality, of course, these functions are

highly interrelated. We turn now to consider, briefly, some evidence that representation in the logogen system is at the level of 'words', and that items in the response buffer are articulatorily coded.

2.34. Sperling's recognition buffer.

Sperling (1963) presented visual displays of letters for various durations, terminated by a visual noise field. When the number of letters reported was plotted against exposure duration, the resulting function consisted of two roughly linear components: as exposure duration is increased, the number reported increased very rapidly (about one item per ten msec.) up to about four or five items, after which the rate of increase was very slow, with no more than one item being added for every extra second. On the assumption that the mask did indeed terminate the icon (for which there is now good evidence - Merickle, Coltheart & Lowe, 1971) Sperling concluded that display items could be scanned serially from the display at a rate of ten msec. per item. However, he has more recently (Sperling, 1967) indicated that the process may not be entirely serial, since when report accuracy is analysed for different locations on the display individually, location n is seen to be reported well above chance long before maximum accuracy is achieved on location $n-1$; this implies either that scanning pattern varies from trial to trial, or that the extraction of items is at least partly simultaneous.

Recent models assuming simultaneous extraction of display items (e.g. Rumelhart, 1970) have proved quite successful. Nevertheless the important point for present purposes is that however items are extracted, they become available at a rate considerably in excess of the highest estimates of rehearsal rate (Landauer, 1962, obtained a maximum rate of around eight letters a second).

Impressed by the order of magnitude of this discrepancy, Sperling (1963, 1967) introduced into his model a component called the 'recognition buffer' between the scanning process, and the auditory information store (AIS). Inputs to this component are conceived as rapidly setting up programmes for the pronunciation of items (in spite of the acoustic nature of AIS), which are then executed at the rate of sub-vocal speech to transfer items to AIS.

I suggest that we may identify Sperling's recognition buffer with the logogen system, and AIS with the response buffer. Allport (1973) has recently reported some evidence in support of the first part of this claim. Using the Sperling paradigm described above, he showed that the rate of increase in number of items reported with increasing exposure duration was similar for three letter CVC words and for single letters, but much slower for CCC non-words. (Unfortunately, he did not try CVC pronounceable non-words.) Moreover, the same slope was obtained for two-syllable as for one-syllable words of roughly the same length; nor did the slope depend upon word length. Since the rate of acquisition is limited not by the number of visual characters to be encoded, but by the number of names, Allport concludes that subsequent in processing to the level of representation destroyed by the mask, there exists a level at which the unit of representation appears to be the word, rather than display location, letter or syllable. He identifies these lexical representations with Morton's logogens.

In summary, processing up to the level of logogens, which are lexical representations, is probably essentially parallel¹; thus if several items are presented at once, outputs from several logogens

¹This is not intended to imply that there is no selection of input prior to logogens.

may become available very close together in time; however, the rate at which they may be read into the response buffer is slow. Span of apprehension is, according to our model, limited by a combination of response buffer read-in rate, response buffer capacity, and the decay rate of item-traces, in addition to any limits there may be on focal attention.

2.35. Articulatory versus acoustic coding in active memory.

There is a wealth of evidence for the proposition that coding in active memory, at least for ordered recall is phonemic. Within-list phonemic similarity results in a recall decrement (e.g. Baddeley, 1966a; Conrad & Hull, 1964; Laughery & Pinkus, 1966; Wickelgren, 1965c); interpolated items similar to memorised items cause more forgetting than dissimilar items (Conrad, 1967; Wickelgren, 1965a); intrusion errors tend to share phonemes (Conrad, 1964; Wickelgren, 1965d) or, perhaps, distinctive features (Wickelgren, 1965b, 1966b) with the correct responses they replace. However, none of this evidence favours acoustic over articulatory coding or vice versa (Morton, 1970; Wickelgren, 1969a). According to Morton (1970) representation in the response buffer is articulatory in nature. Since theorists such as Sperling & Speelman (1970), who favour acoustic representation in a form of storage particularly compatible with auditory presentation, may still be found, we must consider the issue briefly. It should first be mentioned that experiments on this topic which use auditory presentation should include some form of control for recall from PAS (Crowder & Morton, 1969), which being a sensory store is necessarily coded in terms of acoustic features. Evidence on the issue may be summarised under the following headings:

- 1) Error analysis. On the basis of the data of Conrad (1964)

and Miller & Nicely (1955), Hintzman (1967) argues that confusion matrices compiled from recall data show the influence of place-of-articulation errors more than do perceptual confusion matrices. But the validity of his argument depends on the dangerous assumption that memory 'noise' is 'white'.

2) The effect of variation in compatible articulation. Murray (1966) had subjects mouth, whisper or read aloud the presented material, and found that increasing articulatory activity in this manner improved recall. But since the availability of acoustic information was also increased, the effect could be due to recall from PAS.

3) Minimisation of sub-vocal muscular activity. Glassman (1972) used an EMG recording technique to monitor sub-vocal activity (SVA) and trained subjects to suppress their SVA with feedback from their own EMG, in the Brown-Peterson paradigm, with visually presented phonemically similar material. While a higher proportion of errors were 'acoustic' when SVA was not suppressed, the overall level of performance was not affected. However, a fifteen second retention interval was employed, and the involvement of active memory in recall at this delay is dubious. Moreover, it may not be sufficient merely to suppress articulatory activity; Gumenik (1969) found no recall decrement when subjects were required to hold their tongues between their teeth. What is needed is:

4) Competing articulatory activity. Murray (1967) had subjects say 'the, the,....' during presentation of the list and did find a recall decrement. Levy (1971) gave visually presented lists accompanied by compatible or incompatible silent articulation. Probed recall performance was reduced almost to chance level by incompatible articulation, suggesting that articulation is necessary for immediate recall of visually presented material. But when she accompanied the

lists with simultaneous auditory presentation, the decrement due to incompatible articulation almost disappeared. Levy therefore favours a dual coding hypothesis. However, her experiment contains a small flaw: since articulation was suppressed only during list presentation, it remains possible that the subject may, in the case of auditory presentation, articulate the contents of PAS subsequent to presentation.

5) Visual presentation with auditory suffix. If visual information is, as Sperling (1967) and Sperling & Speelman (1970) suggest, encoded into an acoustic store, then an auditory suffix should interfere with retention in the same way when the presentation is visual as it does with auditory presentation. Morton & Holloway (1970) showed that it did not, even when the subject was forced to process the suffix.

6) Deaf people. Perhaps the most striking evidence for the articulatory coding hypothesis is to be found in Conrad's (1970, 1972) findings that among profoundly deaf children, those who develop 'acceptable' speech also tend to find phonemically similar sequences of words and letters difficult to remember, and to show typical 'acoustic' confusions in recall.

In conclusion, the effects of competing articulation, the dependence of acoustic errors on sub-vocalisation, and the finding of similar errors in normals and the profoundly deaf all favour the hypothesis of articulatory coding in active memory. However, the possibility, when presentation is auditory, of some additional acoustic coding, distinct from PAS, cannot on the basis of present evidence be ruled out.

Summary and conclusions.

In this chapter, I have surveyed a number of experiments on short-term forgetting, conducted in the context of a variety of theoretical issues. The first of these was the problem of the representation of order. That the storage of order and content in active memory is not mediated by a single mechanism was implied by:

- evidence of the differential effects on order and item errors in serial recall of serial position, probe type, acoustic similarity, grouping, and perhaps presentation rate and retention interval;
- evidence for contrasting effects of serial position, acoustic similarity and presentation rate on short-term item recognition on the one hand, and tasks requiring the retention of order on the other;
- Conrad's argument from the nature of transpositions for a 'store of available responses'.

That inter-item associations alone cannot mediate order retention in active memory was demonstrated by various pieces of evidence which did on the other hand favour the existence of some kind of links between items and positions. Positive evidence for the existence of any associative chaining in active memory was restricted to two points; the occurrence of associative intrusions, which, however, may result from an inactive memory component in immediate recall; and the evidence for the applicability of trace strength theory to order recognition, which was equivocal with respect to whether there is an active, decaying component to the strength with which an association is represented.

That item-position associations alone cannot mediate order retention was implied by the difficulty of accounting for trans-

positions and their relationship to phonemic similarity.

It was therefore proposed that order and position information are maintained in active memory primarily by the retention of items in a limited-capacity, articulatorily-coded, temporally-ordered store, the response buffer. This (and the inadequacy of Wickelgren's hypothesis of inter-item plus item-position associations) was implied by:

- evidence that position and contextual probes do not differ much in effectiveness, except inasmuch as is compatible with access via position but not via content;
- evidence that adding a contextual probe to a position probe does not facilitate recall;
- evidence that a (purely) spatial position probe is relatively ineffective;
- evidence for articulatory coding in active memory.

Experiments on short-term item recognition and judgements of recency conducted within the context of trace strength theory suggested that items are also represented as temporary traces located at permanent specific content-addressable representations or logogens, and that in the absence of rehearsal these traces are subject to decay along a single strength dimension as a simple, probably exponential, function of recency.

The hypothesis of two forms of active memory was further discussed in the context of estimates of the capacity of active memory, and was held to resolve the dichotomy between estimates deriving from recall measures on the one hand, and item recognition and judgements of recency on the other. This raised the question of why item-traces at logogens are not accessible in recall. It was concluded that access to all but the most recent item-traces is only direct via a 'query'

framed in terms of content. This mode of access was claimed to be adequate to answer the needs of speech comprehension. Other functions of active memory were mentioned; use was made of Broadbent's 'address register' concept, but his two-store dichotomy was rejected. It was also argued that active memory could serve in an organisational role, as a working memory. Finally, it was pointed out that the logogen system would fulfil the same theoretical need as the 'recognition buffer' proposed by Sperling.

The hypothesis of two forms of active memory (Section 1.4), based on Morton's (1970) logogens and response buffer, is robust, general and useful. The factors by which the two forms of representation may empirically be dissociated may include: acoustic similarity, position cues as compared to content cues, rate of presentation and capacity/serial position/retention interval. In the next chapter, we shall be concerned with reaction time experiments on immediate memory for sub-span lists. This line of research, initiated by Sternberg, has as a matter of history developed largely independently of the work reviewed in this chapter. So, in a sense we start again at the beginning, but I hope to show that the same questions arise, in spite of differences of emphasis, and that there may be grounds for favouring the same answers.

CHAPTER III. Immediate memory and reaction time.

Introduction.

3.1 Immediate recognition: the varied-set paradigm.

3.11 Five models.

3.12 Predictions from the five models and data.

3.121 The relationship between set size and mean RT.

3.122 Set size and the RT distributions.

3.123 Limiting set size.

3.124 The relationship between positive RT and serial position probed.

3.125 The effect of direct manipulation of recency.

3.126 Practice.

3.127 The effects of probability and isolation of the to-be-probed item.

3.128 The effects of response probability/bias.

3.129 Subsets.

3.13 The nature of the representation.

3.131 The idea of a sensory component of trace strength.

3.132 Evidence on the nature of the representation.

3.14 The varied-set immediate recognition paradigm: summary.

3.2 The constant-set recognition paradigm.

3.21 Sternberg's argument.

3.22 Differences between the c-s and v-s versions of the IRn paradigm.

3.23 Strategies available in the c-s version.

3.231 Rehearsal of the set in active memory.

3.232 Access to permanent representations (facilitated by recency).

3.233 Feature-testing.

3.24 The constant-set recognition paradigm: conclusions.

- 3.3 The other 'memory-scanning' paradigms.
- 3.31 Sternberg's contextual recall experiment.
- 3.32 The hypothesis of two forms of post-categorical active memory.
- 3.33 Evidence from the CR paradigm.
- 3.331 Serial position effects.
- 3.332 Format of the representation.
- 3.333 The effects of information on where and what.
- 3.334 Rates of presentation.
- 3.335 Limiting set size.
- 3.34 Evidence from other paradigms.
- 3.341 Order recognition.
- 3.342 Location (Ln).
- 3.343 'Metered memory search'.
- 3.344 Retrieval time in forwards and backwards recall.
- 3.35 Summary of Section 3.3.

- 3.4 Is this active memory?

- 3.5 Conclusions.

CHAPTER III. Immediate memory and reaction time.

The aim of this chapter is to review the results of some experiments on immediate memory for short sequences of items in which reaction time is measured and to consider their implications for theories of the nature of immediate memory. The experiments have these characteristics in common: lists of items ('memory sets'), chosen from a larger set (the 'stimulus ensemble'), are presented to subjects for memorization; subsequently a further item is presented (the 'probe') to which the subject must make a speeded¹ response by reference to the memorized list; reaction time is measured from the probe onset, and is the dependent variable of chief interest; typically performance is relatively error-free, in that although performance is speeded, error rates are usually kept below 10%. In general, by 'items' I mean constellations of perceptual features which are familiar, which if they lack names are at least readily identifiable and which may be treated as units; in practice this will involve the use mostly of letters, digits and words, though in some studies line figures, geometrical forms, faces (schematic or photographed) and well-learned nonsense forms have been employed. The 'question' which the probe item puts to the subject varies with the paradigm. Among those I shall discuss will be 'item recognition'² (IRn) in which the subject must make one response if the probe is contained in the memory set and another if it is not, 'contextual recall'² (CR) in which the subject must name the item following the probe item in the memory set, and 'location' (Ln) which requires the subject to indicate the ordinal position of the probe in the list. Other possibilities exist and have on occasion been used. As will become clear the bulk of published

¹i.e. 'as fast as you can while avoiding errors'. ²Sternberg (1969a)

studies have used the IRn paradigm. It is apparent that these different tasks may require different attributes of the memory set to be represented in and retrieved from memory, and may impose different modes of access to the stored representations. A distinction which I shall particularly stress, in line with the position argued in the previous chapter, is that between tasks requiring the use of stored order information and those that do not.

The chapter falls into three sections: In the first I shall consider in detail that version of the IRn paradigm in which the memory set is varied from trial to trial (Sternberg's 'varied-set' procedure). Five models will be outlined, and their predictions compared to data available in the literature. In the second section, the implications for immediate memory of the 'constant-set' version of this paradigm, in which the memory set remains the same over large numbers of trials, will be examined. In separating the two I deviate from other reviewers (Sternberg, 1969a; Rabbitt, 1971; Nickerson, 1972) on the grounds that the constant-set paradigm is unlikely to be a pure immediate memory paradigm. Now Sternberg (1969a) has argued that the subject performs the constant-set task as a (covert) version of the varied-set task in that he rehearses the constant set and so maintains it in active memory. My objection will be that while Sternberg's argument may hold for relatively unpractised subjects, there now exists evidence of marked changes in the pattern of performance with practice on a particular memory set, and since one cannot know a priori at what stage a subject is in whatever progression underlies such changes, a sensible precaution is to make use only of that paradigm which one is certain involves

active memory. (The grounds for this certainty will be discussed briefly at the end of the chapter.) In the third section I shall consider the relationship between the results obtained with the IRn paradigm, and some of the other memory-scanning paradigms. These will be argued to be compatible with the hypothesis of two forms of post-categorical active memory introduced in Section 1.4. Some predictions from the model, together with others derived in the first section, are tested in experiments described in later chapters.

Excluded from this review, which is otherwise intended to be fairly comprehensive, are some experiments in which several probes are presented at once, (as in some conditions of Nickerson, 1966, and Briggs & Blaha, 1969), those in which the memory set consists of connected discourse, and those in which 'multidimensional' items are used in an attempt to establish whether individual memory items are 'processed' in series or in parallel (e.g. Checkosky, 1971; Dumas, Gross & Checkosky, 1972). These seem to me to introduce extra complications which have not so far enhanced our understanding of immediate memory.

3.1. Immediate recognition: the varied-set paradigm.

Much research employing this paradigm has been published since Sternberg's (1966) ingenious pioneering work. The exhaustive scanning model which Sternberg proposed has exercised a powerful hold on the imaginations of some authors¹, to the extent that in a few cases its validity is so taken for granted that data are presented purely in terms of the slopes of the linear regression lines that best fit the set size functions and as such are assumed to represent 'memory comparison times'. Relatively few experiments have been directed at testing the exhaustive-scanning model directly against well-articulated alternatives. I shall describe five models: the two variants of Sternberg's scanning hypothesis - 'exhaustive' and 'self-terminating'; two versions of a 'trace strength' theory, which I have tried to specify in sufficient detail so that clear-cut and distinctive predictions can be made; and a mixed model, at present rather loosely formulated, for reasons which will become apparent. Though I have tried to formulate the models as precisely as is both possible and reasonable, I have concentrated on predictions of a qualitative rather than an exactly quantitative nature, and have not tried to fit models to data in order to estimate parameters; nor is it likely that such attempts would prosper without data and grounds for estimating parameters better than those we now possess.

3.11. Five models.

The two scanning models: these are well described by Sternberg (1969a). He and others (e.g. Clifton & Birenbaum, 1970) have suggested that these two models may in some sense represent

¹probably because it is one of the rare cases in which Donders'

alternative strategies available to the subject.

MODEL I. Exhaustive scanning.

As the items of the list are presented or rehearsed, they are represented in 'active memory'.¹ When the probe item is presented, it is compared successively with each item in 'active memory'. When all comparisons have been made, a decision is made as to whether a match has occurred or not and the appropriate response is emitted. Reaction time may be partitioned into four components:

$$RT = t_e + mt_c + t_d + t_r \quad (1)$$

t_e , encoding time, is the time required for perceptual analysis of the probe to whatever level is necessary for comparison to the memory representations. t_c is the time required for a single comparison. m is the number of items in the list. (t_c is usually supposed to be constant with respect to m .) t_d is decision time, and t_r the time required to programme and execute the appropriate response.

Note some consequences. 'Active memory' must be a quasi-ordered store, in that the subject is able to scan each item once and only once, and knows when he has finished. Moreover, since t_c is independent of m , a relatively stable representation is assumed. Were subsequent items to cause an item's representation to decay, then presumably comparison time would increase with amount of degradation (cf. Sternberg, 1967b) and t_c would increase with m . It is not specified how forgetting takes place, save that a sufficient period of distracting activity will displace the list from 'active memory'. For a particular set of experimental parameters and responses, $(t_e + t_d + t_r)$ will be constant and the slope of the set size

¹ 'Active memory', when in quotes, means whatever Sternberg means by it. It is not necessarily to be confused with active memory as defined in Ch. 1 and discussed there and in Ch.2.

function¹ may be taken as a pure index of comparison time. Variables which affect comparison time rather than the other stages will affect the slope rather than the intercept. To determine which of the other stages is influenced by which variables, further multifactorial experiments must be done using the additive-factor method (Sternberg, 1969b). Broadly, we may expect that t_e will be influenced by stimulus intensity, quality, etc., t_d by speed/accuracy instructions, t_r by response relative frequency, number of response alternatives and S-R compatibility.

MODEL II. Self-terminating scanning.

This is essentially the same as Model I save that the subject stops scanning the list in 'active memory' as soon as he has found a match². In consequence, the set-size parameter (m) in (1) must be replaced by the number of items searched; for negative probes, this will be equal to m ; for positive probes, if a) the position of the probe in the list is randomised and, b) the subject has no knowledge of its position, then since the subject must on average search $(m+1)/2$ items the slope of the set size function will be half that for negative probes.

Trace strength models: these share the assumptions that

a) each member i of the stimulus ensemble has a unique and specific representation L_i in memory (in the sense that has already been indicated for logogens in Section 1.4 and the representations assumed by Wickelgren & Norman - see Section 2.2).

b) the time elapsing since a given recent item i has been presented is represented, at the time that the same item is presented as a

¹ i.e. mean RT plotted against m . ² see Section 3.31 for a discussion of the relative efficiencies of exhaustive and self-terminating search strategies.

probe P_i , at L_i , in the form of a variable S_i , continuous between certain limits, which we may call 'trace strength'. While S_i may indeed be influenced by variables other than pure recency, it is proposed that in most IRn experiments, the judgement of presence or absence is in effect a judgement of whether the item i has occurred more or less recently than some criterion which the subject sets on the basis of his knowledge of the underlying distributions of S_i (in a manner similar to that assumed by signal detection theory).

c) no search is necessary; L_i is in some sense directly accessed by P_i . Two versions may be distinguished:

MODEL III. Trace strength 'inspection'.

This model is derived directly from Baddeley & Ecob (1970) and the trace strength theory reviewed in Section 2.2.

a) Representation. When an item i is presented or rehearsed, the representation L_i has its trace strength S_i 1) raised to a ceiling, OR¹ 2) raised by a fixed increment, OR 3) raised by a quantity determined by its and other representations' prior strength.

b) Forgetting. When the item is removed, the quantity S_i decays as an inverse monotonic function of 1) time AND/OR 2) subsequent items.

$$S_i = f(1/T_i) - \text{the DECAY FUNCTION} \quad (2)$$

T_i is the time (items) since presentation of item i . The form of the decay function is not specified, but a reasonable guess is that the rate of decay is initially rapid and declines to an asymptote

c) Nature of decision. When the probe P_i is presented, the

¹These represent possible alternatives. The discussion will be couched in terms of the first mentioned, without the intention of ruling the others out.

subject inspects the trace strength S_i at L_i and responds positively if it exceeds a criterion C , negatively if it does not.

d) Decision time. This is assumed to be inversely and monotonically related to the difference between the trace strength and the criterion, $|S_i - C|$.

$$t_d = f(1/|S_i - C|) - \text{the } \underline{\text{Decision}} \text{ function.} \quad (3)$$

The form of this function is also unknown, save that it is assumed to have an upper and a lower limit.

e) Reaction time. This consists of the time taken to encode the stimulus to a point where the decision process can start - t_e , time to decide - t_d , and the time to organise and emit a response - t_r .

$$RT = t_e + t_d + t_r \quad (4)$$

I do not, it should be noted, intend to imply that stimulus encoding and decision are necessarily separate and independent processes. This will depend on precisely which decision model proves applicable. For instance, the dependence of decision time on difference between strength and criterion may be due to the fact that when this difference is small the subject must sample stimulus information longer before deciding than when the difference is large. However, I shall not consider the relative merits of continuous, fixed sampling, and other models of choice reaction time, since they do not affect the major predictions. (See Laming, 1968; Broadbent, 1971; and Audley, 1971, for reviews.)

Combining (2), (3) and (4) it will be obvious that reaction

time can be expressed as some function of T_i , the time since the item was last presented, and C .

$$RT = f(T_i, C) + X \quad - \text{the } \underline{\text{Recency}} \text{ function,} \quad (5)$$

where $X = (t_e + t_r)$. RT will be a monotonic function of the recency of the probe P_i ; increasing as T_i increases for positive probes, decreasing for negative probes. The exact form of the recency function will depend on that of its two component functions, the decay function and the decision function, and is therefore also unspecified.

f) Criterion. As in signal detection theory, the subject's optimal criterion placing will depend on the underlying distributions of S_i , and the probabilities and pay-offs. Fig. 3.1 shows a schematic version of the distributions subsequent to presentation of lists of various lengths. (Decelerating decay to a floor is assumed here, but with appropriate transformations of the horizontal scale the picture will serve for any decay function.) If positive and negative probes are equally frequent, then for maximum speed the subject should place his criterion midway between the means of the two compound (memory set and non-memory set, positive and negative) distributions, provided these are symmetrical. For accuracy he must place it somewhere near the intersection of the weakest positive item and the strongest negative item distributions, that is, so as to make the false positive/false negative ratio roughly unity. The shape of the distributions will determine which side of the speed criterion it¹ will lie, if indeed it is different at all. It is, of course, possible that for some reason the subject may be unable to optimise his criterion placing.

¹the accuracy criterion.

MODEL IV. The 'default' model.

This is derived from Nickerson's (1972) rather imprecise description. I shall describe only its points of difference from Model III, the 'inspection' model.

a) Representation. As in Model III, save that in this case it may be helpful to conceptualise the increment in S_i as reflecting an increase in sensitivity (or lowering of threshold) rather than an increment in some level of activity¹.

b) Forgetting. As in Model III.

c) Nature of decision. When the probe P_i is presented, the subject determines whether some activity or count in L_i passes a criterial level K within a criterial time T_c . The subject responds positively if it does and as soon as it does, and negatively if it does not (i.e. by default).

d) The time taken for L_i to pass some level K given P_i is assumed to be a monotonically decreasing function of the 'sensitivity' S_i . Hence for positive probes,

$$\text{decision time } t_{d+} = f(S_i, K) \quad (6)$$

while for negative probes it depends purely on the criterion T_c :

$$\text{decision time } t_{d-} = T_c \quad (7)$$

$$\text{e) Reaction time. } Rt = t_s + t_d + t_r \quad (8)$$

where t_s represents the time elapsing between probe onset and the starting point for timing T_c . This time is occupied by some form of initial sensory registration. In this model there is no distinction between an encoding stage and a decision stage, in that basically the decision is based on whether the encoding of the probe has

¹The inspection model could also be interpreted in these terms.

proceeded to a criterial level within a criterial time or not.

f) Criterion. A similar argument for the placing of the criterion applies here as did in Model III. The difference is now that the distributions to be considered are those of the times to reach the critical level K for items of different levels of recency. Fig. 3.1 is again appropriate, but the horizontal scale is now time not strength. The larger the value of K , the greater the spread of and the distance between these distributions. The level of the criterion time T_c does not affect positive RTs, but decrease in negative RT. can be traded for an increasing proportion of false negatives by reducing the level of T_c (i.e. moving the criterion to the right).

The essential difference between the two trace strength models may perhaps be captured by the following analogy¹: items are conceptualised as ringing specific 'bells'; in the 'inspection' model, the subject determines whether the probe item was in the memory set by checking to see if that item's 'bell' is still resonating loudly enough for it to have been struck in the recent list; in the 'default' model, the subject responds 'yes' as soon as any bell rings so long as it rings within a fixed time after the probe onset, 'no' when that time has elapsed silently.

An interesting point is raised by the question: how, according to Model III does the subject 'inspect' the trace strength at the logogen addressed by the probe? If the representations $L_1 \dots L_n$ are logogens, terminal units in the identification mechanism, then logogen L_i must be used to identify the probe P_i . In terms of the bells metaphor, how can the subject tell whether the resonance of the bell exceeds a criterion if the probe confuses matters by

¹ borrowed from Nickerson (1972).

striking the bell again? The answer may be that the bell is rung louder by the probe when it was already resonating than when it was not, and it is the loudness that is judged. Rather than assuming that the subject really 'inspects' the trace left by the item, it may be preferable to assume that the trace strength generated at L_i by the probe P_i adds to that already present, if any, and it is this compound strength that is inspected (compared to a criterion).

MODEL V. Mixed models.

The trace strength models are undoubtedly more complex than the scanning models, but their scope is greater in that they include a forgetting axiom, and in that they relate closely to general models of pattern-recognition and the models for accuracy (Section 2.2). The mixed model suffers from even greater complexity without any marked advantage in generality. Such a model was first described for a long-term recognition situation by Juola, Fischler, Wood & Atkinson (1971). I mention it here because of recent suggestions that such a model may apply to the IRn paradigm, (Sternberg, 1973; Darley & Arabin, 1972).

Briefly, the basic premises are these. The representation and forgetting axioms are those of trace strength theories: on presentation the item's representation receives some increment in strength or 'familiarity', which subsequently decays unless the item is rehearsed. When the probe item is presented, the subject inspects the 'familiarity' of that item. If it is higher than a high criterion C_H he responds 'Yes', if lower than a low criterion C_L he responds 'No'. Responses having this basis are held to be fast, and according to Juola et al. (1971) not dependent on the trace

strength - that is, effectively constant. If the strength falls between the criteria, the subject performs some form of search. This could be either exhaustive or self-terminating and presumably requires in addition to dynamic traces in logogen-style representations, an additional 'active memory' with the properties specified in Models I and II. The resulting decision time and hence reaction time will be a probabilistic mixture of fast 'familiarity judgements' and the slower outcomes of searches. The more recent a negative item, and the less recent a positive item, the greater the likelihood that its familiarity will fall between the two criteria and the longer the mean RT. (See Fig. 3.2)

I shall not attempt to work out the mixed model in any greater detail for two reasons. The first is its complexity relative to the 'one-process' models. There are both a number of alternative versions that could be specified, and a number of 'loose' parameters within those alternatives, which argues that the sheer effort should be postponed until all the one-process models have been shown to fail. The second is perhaps more substantial: without making very strong assumptions, concerning the nature, parameters and error-proneness¹ of the two forms of retrieval, it seems to me that it will make almost identical predictions to Model III, the inspection model. To see why this is so, consider what I have called the decision function in that model: decision time is held to be a continuous monotonically decreasing (for positive responses) or increasing (for negative responses) function of strength. The

¹For instance, Darley & Arable (1972) according to Sternberg (personal communication), base their parameter estimations on the assumption that all errors are pure errors in the sense that they derive from the tails of the strength distributions which fall the wrong side of the appropriate criteria. In spite of the very pretty fit of their model to their data, it is unlikely that their subjects are not subject on occasion to erroneous anticipations, response blocking, hesitations, distractions and other ills to which the C.N.S. is heir, and overall error rates are very low.

decision function for the mixed model turns out to be no more than the discrete-state version of this continuous function. On any given trial the subject will be either in the state in which a fast 'familiarity' decision is possible, or one in which a slow search is required (which could count as one or several states, depending on whether search is exhaustive or self-terminating), and the probability of being in the former rather than the latter state is a direct function of strength, which lies along the same underlying scale in both models. Hence we are plunged into the difficulties usually contingent upon attempting to decide between otherwise equally adequate continuous and discrete-state models. So I shall not, with an exception or two, spell out the predictions of a mixed model separately.

The well-read reader may wonder why I have not discussed parallel scanning models. Of course, Sternberg (1966) briskly rules out simple parallel models, but Murdock (1971) has described a parallel model which predicts linear increase in RT with set size, parallel positive and negative slopes, and (almost any) serial position effects. In fact the 'default' model, Model IV, may be considered a self-terminating parallel scanning model, if one so chooses, in which the comparison times are determined by recency in a manner rather less ad hoc than Murdock's suggestions of a linear primacy component and a logarithmic recency component. Corcoran (1971) suggests a 'limited-power' parallel scanning model, somewhat lacking psychological plausibility, in order to account for linear set size functions; but the function relating mean comparison time to set size, deriving from this suggestion must be very similar to that derived by Murdock,

who makes comparison time a function of distance from both ends of the list, and so Corcoran's model may be considered merely as a weaker version of Murdock's.

3.12. Predictions from the five models and data.

3.121. The relationship between set size and mean RT.

Predictions:

(i) Mean positive RTs. Models I and II predict that RT should be a linear increasing function of set size. The predictions from the trace strength models derive from the recency function.

$$\text{Since } RT_i = f(T_i, C^1) + X \quad (5)$$

then, if m is the set size, and the rate of presentation/rehearsal is constant over and within set size, and each serial position probed equally often, then mean positive RT

$$RT_+ = \frac{\sum_1^m (RT_i)}{m} + X = \frac{\sum_1^m f(T_i, C)}{m} + X \quad (9)$$

If RT is a monotonic increasing function of recency, and the initial increment in strength and rate of decay are assumed to be independent of the number of prior items, then the set-size function should be similar in form to the recency function, though less markedly non-linear if the recency function is non-linear. Two important consequences follow. If the recency function (whose form

¹ or K , in Model IV

may be inferred from the effects of serial position with constant list length, or by means of a continuous recognition paradigm - e.g. Okada, 1971, provided that rehearsal is controlled in both cases) is non-linear, the set-size function cannot be linear. Secondly and more crucially, the slope of the set-size function should be considerably less than that of the serial position function when rehearsal is controlled. Both consequences may however be avoided if it be assumed that not only is amount of decay determined by time and/or the number of subsequent items, but also the initial increment in strength is influenced by the strengths at representations activated by prior items. This would be the case if, for instance, the representations were linked by relations of mutual inhibition. One way in which a linear set-size function could be predicted on this sort of basis is, as Baddeley & Ecob (1973)¹ have recently suggested in an elaboration of their 1970 formulation, to assume that the system is so constructed that the total trace strength, summing over all representations, is roughly constant. Then, if the criterion is placed midway between the means of the two compound distributions, and decision time is inversely and linearly related to $|S_i - C|$, decision time = $a(2m/S_T) + b$, where S_T is the total strength and a and b are constants. It is no doubt possible to devise comparable solutions which do not assume a linear decision function and constant S_T . While the set-size function is in Baddeley and Ecob's "limited capacity" model determined by the total strength of the system, the serial position function will reflect how that strength is distributed over the

¹BADDELEY A.D. & ECOB J.R. (1973) Reaction time and short-term memory: implications of repetition effects for the high-speed exhaustive scan hypothesis. Quarterly Journal of Experimental Psychology 25 229-240.

representations of the memory set items. If each new item takes a relatively large share of the total strength, the serial position function will be relatively steep; if a small share, then shallow.

(ii) Mean negative RTs. Models I and II differ here in that while both predict a linear increase in mean RT with increasing set size, according to the exhaustive model the rate of increase should equal that for mean positive RT, while if the search is self-terminating the slope ratio¹ should be two. Any overall intercept differences between positive and negative RT are attributed to stages other than comparison. According to the inspection model, to maximise his speed for both responses, the subject should place the criterion C midway between the means of the positive and negative distributions (see Fig. 3.1). As set size is increased, this results in $S_i - C$ decreasing at the same rate for both response types, so that the positive and negative set size functions must necessarily be parallel. A concomitant feature of parallel slopes in Model III is that since the variance of the positive distribution of strength increases with set size while that of the negative distribution does not change, false negatives should increase at a greater rate than false positives with increasing set size. On the other hand different placings of the criterion, which may because of the particular distributions involved lead to greater accuracy, will produce slope ratios other than unity, so we might expect to find a range of slope ratios, on either side of unity. In the default model, only negative RT is affected by the criterion T_c . For greatest accuracy the subject should place T_c so as to cut off as little of the list and non-list distributions as possible. Assuming that the latter remains relatively stable with changing set size, and that the distributions

¹slope ratio = negative slope/positive slope, of the set size function.

are reasonably symmetrical, this may be achieved by placing the criterion roughly midway between the means of the compound distributions. It may be seen that this will result in negative decision time increasing with set-size at a slower rate (about half) than positive decision time - i.e. a slope ratio considerably less than unity. Higher slope ratios would result if the subject placed the criterion to the right (see Fig. 3.1) of centre for small set sizes and to the left of centre for large set sizes. But a) it is still unlikely that this would result in slope ratios as great as or greater than unity and b) if slope ratios of this order were found, a necessary consequence is that false positives should increase very much faster with set-size than false negatives. On the other hand a

range of slope ratios above and below unity is permitted by the inspection model, with the restriction that if the ratio is unity, false negatives should increase more rapidly with set size than false positives.

The data:

Linearity: With the exception of two studies, I know of no variable-set IRn studies where marked deviations from linearity are found. The exceptions are Smith (1967) and Nickerson (1966) both of whom obtain distinctly logarithmic set-size functions. However for present purposes both are flawed: Smith's is a between-subjects design and each memory set was used for several trials in a row; Nickerson's is an exploratory experiment and designed so that each subject only underwent some of the conditions. In general, though, it is the case that linear trends are found to account for an impressively large proportion of the variance (e.g. Sternberg, 1966; Burrows & Okada, 1971; Chase & Calfee, 1969; Wingfield & Branca, 1970; Baddeley & Ecob, 1970; Klatzky & Atkinson, 1971). The slopes obtained depend on the nature of the items, being about 33 msec./item for digits, 40 msec./item for letters and 50 msec./item for common words (Cavanagh, 1972). Nevertheless, considering the data as a whole,

certain small but consistent departures from linearity do stand out. In the first place, there are quite consistent deviations of the mean positive RTs from the best fitting regression line, for small set sizes ($m=1,2$): the value for $m=1$ being often too low while that for $m=2$ is sometimes too high (Sternberg, 1966; Clifton & Birenbaum, 1970). Secondly, there is very frequently a small negatively accelerated trend in the set size function, particularly for positive RTs; in some studies (Morin, Derosa & Ulm, 1967; Corballis et al., 1972) this results in a statistically significant quadratic trend, and in others in which no test is made it is certainly evident (Juola & Atkinson, 1971; Klatzky & Smith, 1972; Forrin & Morin, 1969; Sternberg 1969a, Exp. 4). Such departures from linearity, if they may be considered reliable, are difficult for either of the scanning theories to explain. Supposing one were to introduce into the scanning models an assumption that representations in active memory decayed, leading to slower comparisons: this would lead to positively not negatively accelerated set-size functions. Unduly fast positive RTs for $m=1$ might possibly be accommodated to the scanning models by merely supposing some special strategy for that case (Nickerson, 1972), but without further evidence this argument is both casuistical and unconvincing. Trace strength theories, on the other hand are quite compatible with non-linear set-size functions, and negative acceleration will be expected if the recency function from which the set-size function derives itself decelerates to an asymptote. Moreover, for small set sizes, or for cases where presentation is fast and probe delay short (e.g. Morin, Derosa & Ulm, 1967), there may be an appreciable sensory component to the decay curve (see Section 3.13 for full discussion) leading to more marked non-linearity.

There is other evidence for a non-linear recency function.

Okada (1971) using a continuous recognition paradigm found that if RT was plotted against number of items since previous presentation, the data were best fitted by an exponential decay function. Since in this case we would expect the criterion to remain constant, the resulting function should be a pure representation of the effect of recency on RT. The combination of roughly linear set-size functions and evidence for a non-linear recency function suggests that some form of 'limited capacity' (or mutual inhibition) version of the trace strength model may be necessary.

Slope ratios: In the bulk of studies (Sternberg, 1966; Sternberg, 1969a, Exp. 4; Chase & Calfee, 1969; Baddeley & Ecob, 1970; Juola & Atkinson, 1971; Klatzky & Smith, 1972; Morin, Derosa & Ulm, 1967; Swanson et al., 1972) the mean slope ratio is close to unity. However, there is some suggestion that averaging across subjects may conceal quite large individual variations. For instance, Clifton & Birenbaum, (1970) found that out of 12 subjects the data of 9 showed a slope ratio of unity, while for the rest the negative slope was twice the positive slope. Other studies also reveal slope ratios greater than unity: Corballis et al. (1972) report two subjects with a ratio of 2, Klatzky & Atkinson (1970) obtained a mean ratio of 1.75 and Klatzky, Juola & Atkinson (1971) found ratios ranging from 1 to 2.25 under different conditions. Large slope ratios are problematic for Model IV, since the upper limit on slope ratios, given reasonable accuracy, should be not much greater than one. Moreover such data can be satisfactorily accommodated by Models I and II only if it is supposed that these describe alternative strategies, as Clifton &

Birenbaum suggest. Unfortunately for this suggestion, several studies find slope ratios less than unity. Burrows & Okada (1971) obtained ratios of 0.8 and 0.85, Corballis et al. (1972) a mean ratio of 0.68 (with 16 out of 20 subjects less than unity) and Forrin & Morin (1969) 0.57. I have argued that only Model III permits such a wide range of slope ratios; the actual values obtained will depend on the nature and placing of the underlying strength distributions and on response frequencies, though it is also the case that by and large the optimal placing of the criterion for speeded performance will be such as to give parallel functions. Is there, then, any support for the concomitant prediction in Model III that when the slope ratio is unity, false negatives (FN) should increase faster than false positives (FP) with set size? Error rates are rarely presented in sufficient detail to answer this. Clifton & Birenbaum (1970), 9 of whose subjects showed parallel slopes, present error data pooled over subjects, and if one fits regression lines it transpires that FN increase by 0.87% FP by 0.59% for each item added to the memory set. Nickerson (1966, Exp. 4) with 4 well-practised subjects, but only 2 set sizes ($m=1,4$) found that the difference in FNs between the two set sizes was greater than the difference in FPs. Ideally one should look at the relationship between RT slope ratios and error-frequency slope ratios within individual subjects.

In summary then the slope ratio data appear to favour Model III, and the set size functions generally obtained are also compatible with this model. Nor is the success of this model so far due merely to lack of specificity in its predictions, as the point about error frequencies demonstrates.

3.122. Set size and the RT distributions.

(i) Variances. Quite strong predictions can be made:

Model I: The variance here will be the cumulative variance of the successive comparisons, and since the variance of the sums = the sum of the variances, RT variance should be a linear function of set size, with an intercept above the origin, for both positive and negative responses.

Model II: The same prediction is made for negative RT as in Model I, but since on average only $(m+1)/2$ comparisons need to be made for a positive response, beyond $m = 1$, the slope of the positive response function will be approximately half that for negative responses.

Model III: Though in general RT variance will be expected to increase with set size, negative RT variance should increase markedly less than positive RT variance, since the underlying distribution of strengths for negative items is not changed¹.

Model IV: Unless the criterion becomes more unstable with increasing set size, variance for negative RT should not increase at all .

Unfortunately I know of no data in the literature against which to evaluate these predictions. Sternberg (personal communication) found that in his data, RT variance increased linearly with set size, but the intercept was below the origin. My own data will be discussed in Ch.7.

(ii) RT distributions. There are two points of potential interest. Firstly, according to the strength models the fastest positive RTs should not increase with set size; indeed, according to the inspection

¹ However, in some circumstances, this prediction may have to be modified, depending on the nature of the decision function (see Ch.7, discussion).

model, they might be expected to decrease as the criterion is lowered. (A related point will be discussed in the context of serial position effects - see Section 3.124.) Secondly, detailed examination of the RT distributions could in theory be one way of distinguishing Models III and V. According to a mixed model, the RT distributions should consist of two components, a stationary distribution of fast familiarity judgements, and a slow distribution whose behaviour should be that predicted by one of the scanning models. The fastest and slowest RTs could therefore be assumed to come from different distributions. (A related argument is put forward by Lively (1972) and Lively & Sanford (1972) who present 'fastest RT' data from a constant-set experiment in an attempt to evaluate a strategy-mixture hypothesis.)

Unfortunately, unless clear evidence for bimodal distributions were found, the sophistication of such analyses would probably be excessive for the real world. The difficulties with arguments based on extremes of distributions are a) that there is a major sampling problem, and b) that the extremes of the distributions are almost certain to be contaminated on the one hand with fast anticipations (Yellott, 1967) and on the other with hesitations, failures to depress the key first time, response blocking, etc.

3.123. Limiting set size.

What upper limit on memory set size would we expect in this task? The scanning models appear to require an 'active memory' which is a quasi-ordered store with a definite capacity; if this store is what is conventionally described as short-term or primary memory, then we would expect some discontinuity in IRn performance as set size is increased beyond conventional memory span. Trace strength

models, on the other hand, predict no sharp cut-off, but presumably a gradual increase in RT and error rate to an asymptote dependent on the temporal characteristics of decay. Wingfield & Branca (1970) have shown that, provided that the possibility of the subject coding in terms of the complementary set is avoided, mean RT continues to increase smoothly up to set size = 12, with a concomitant rise in error rate to only about 20%. (See discussion of capacity measures in Section 2.31)

Also, Cavanagh (1972) has demonstrated a highly impressive correlation between the slopes of the set size function for different types of material, as gleaned from the literature, and estimates of the memory span for the same material. This is easily accounted for by a trace strength theory, since both memory span and slope will be functions of rate of decay, which may well depend on type of material. To accommodate the scanning models to this correlation requires the addition of some rather strong assumptions (see Cavanagh, 1972, p.528).

3.124. The relationship between positive RT and serial position probed.

(i) Serial position functions:

Predictions: Model I predicts no serial position effects. Model II can accommodate any serial position effects on a purely ad hoc basis, on the assumption that those positions for which RT is fastest are those which are located first. As for strength models, since RT_i is a direct function of recency T_i , provided that rehearsal is controlled (i.e. either there is no rehearsal or the whole memory set is rehearsed an integral number of times) pure recency serial position curves are predicted.

Data: Of the twenty-odd published experiments for which serial position data are available, I know of only three in which serial position effects are not obtained. These are those of Sternberg (1966), Baddeley & Ecob (1970, Exp.2) - both of which were characterised by slow presentation rates, 2 sec. probe delay and set size varying from trial to trial, and Wingfield & Branca (1970) who used simultaneous presentation and highly variable set size. These conditions are such that subjects would find it both necessary to rehearse the memory set and difficult to time a fixed number of full rehearsals before the probe item is presented. In general, however, strong recency effects are obtained. Nor is it the case that such a finding is restricted (as Sternberg, 1969a suggests) to those cases where presentation rate is fast and probe delay short. Baddeley & Ecob (1970, Exp.1) with presentation rate and probe delay similar to Sternberg's and list length constant at 4 items found significant recency effects. Kirsner & Craik (1971) with similar parameters and set size constant at 8 items found unequivocal monotonic recency effects (even when the serial position data were 'corrected' for any possible priming effect on identification of the probe item). Clifton & Birenbaum (1970), using slow presentation and variable list length found recency effects provided the probe delay was short. And Kennedy & Hamilton (1969) and Burrows & Okada (1971, slow condition) found bowed serial position curves even with slow presentation, long delay and variable list length. Nevertheless, it is certainly the case that recency effects are more pure, marked and reliable with relatively fast presentation and short probe delay (Corballis, 1967; et al., 1972; Derosa & Beckwith, 1971; Forrin & Morin, 1969; Krueger, 1971; Morin, Derosa & Stulz, 1967; Morin, Derosa & Ulm, 1967).

There are two consistent exceptions to the pattern of a decelerating increase in RT with backward serial position:

1) In all studies cited so far in which recency effects are found, with the exceptions of Forrin & Morin (1969), Kirsner & Craik (1971), Corballis (1967; fast condition) and Derosa & Beckwith (1971, short delays), there was a small primacy effect, in that RT to the first position was faster than RT to the second.

2) Workers at Stanford (Klatzky & Atkinson, 1970; Klatzky, Juola & Atkinson, 1971; Klatzky & Smith, 1972) have reliably found that RT increases with forward serial position. The procedure used in these studies is however unusual in that the items of the memory set were displayed simultaneously, for a short period of time or ad lib, and the subject initiated the probe himself when he was ready. But Juola & Atkinson (1971) used a procedure similar in all respects save that the list was presented auditorily and then repeated aloud by the subject before he triggered the probe, and they found recency effects, with primacy on the first item only. So display position appears to be the crucial determinant of the serial position effects found by Klatzky et al.: perhaps items read first off the display are rehearsed more frequently or intensively prior to the probe.

The available serial position data may perhaps be summarised thus: in experiments in which rehearsal is either minimised by fast presentation rate and immediate test, or at least partially controlled in the sense that subjects are able and likely to 'programme' a whole number of complete rehearsals of the list on most trials, RT is found to be a monotonically increasing and decelerating function of backward serial position, with the exception of 'first item

primacy; in other experiments in which rehearsal is likely to be random or idiosyncratic, serial position effects vanish or are controlled by other factors such as display position.

Such findings appear to constitute powerful refutation of an unmodified exhaustive scanning theory. Model II can cope by supposing a backward scan of the list, but this supposition is vitiated by two features of the data: the serial position effects are non-linear, and, more crucially, serial position effects of this form have been obtained in experiments in which set size has also been varied and in none of these studies has a slope ratio greater than unity been obtained (Burrows & Okada, 1971; Corcoran, 1971; Morin Derosa & Stulz, 1967; Forrin & Morin, 1969; Corballis et al., 1972). In contrast, the serial position findings are highly congenial to the strength models, with the exception of the small primacy effects found. One possible way of accounting for the latter is to suppose that the first item to be presented receives a greater initial increment in trace strength than the rest, as a result of the lack of activity, in the identification mechanism, deriving from prior items. A redundant but identified prefix should on this account eliminate the primacy effect: Kirsner & Craik (1971), who probed positions 2-8 only of their lists, found no primacy effect (and see also Exp. VI, Chapter 7).

Can Model I be modified to account for the effects of serial position? Burrows & Okada (1971) attempt to introduce an assumption that an item can be in a special state of 'high accessibility' or attention, with a probability that depends on its position in the list, but any such model must predict that any RT differences due to

serial position should be less than the slope constant of the set size function; this follows from the fact that the advantage gained by a position which is always highly accessible, over one which is never so, cannot be greater than the time for a single comparison; this prediction fails in most of the studies cited. Sternberg (1969a p.428) hints that recency effects may be related to sensory matching (cf. Posner, 1969) but as recency effects a) are not limited to short probe delays, and b) are obtained in cross-modal conditions (Kirsner & Craik, 1971; Juola & Atkinson, 1971), any sensory component must occupy a secondary role (see Section 3.13). There is one possible modification¹ which works: suppose that the 'flag' in the match register (the mechanism which retains information as to whether a match has been found until the end of the scan) decays rapidly to a floor, and that the list is always scanned forward - then recency effects are obtained, together with negative acceleration of the set size function for positive RTs. But, for the present, this hypothesis is rather too heavy with extra assumptions to merit lengthy consideration.

(ii) The relation between negative RTs and serial position curves. According to Model IV, mean negative RT should be quite distinctly longer than the longest mean positive RT characteristic of a single serial position assuming that the response stage is not faster for negative responses than for positive. Unfortunately, there are practically no studies which combine the conditions necessary for testing this (i.e. key-pressing responses, half positive half negative, with response allocation balanced across hands) which report serial position effects as well as negative RTs. Corballis et al. (1972) appear to have fulfilled these conditions, and,

¹I am indebted to Caro Williams for this suggestion.

at least for the longer list lengths, negative RTs appear from their data to be shorter than those for the slowest serial position.

(iii) Interaction between serial position and set size effects.

On the basis of Model III, we would expect¹ a slight decrease in RT and error rate at any given backward serial position as set size is increased, since the level of the criterion C should be correspondingly reduced. Models II and IV predict no change in RT, but the latter predicts a reduction in errors. There are two problems with testing this prediction: firstly, rehearsal must be controlled absolutely; secondly, it is usually the case that equal numbers of trials per set size are usually given rather than per combination of serial position and set size, so that a practice effect may militate against the effect predicted from Model III. There are two studies in which the frequency of trial types were determined by the latter procedure. The data of Morin, Derosa & Ulm (1967) indicate that RT for the later serial positions increased slightly with set size, though errors decreased markedly. But in the case of Kennedy & Hamilton (1969) where subjects were given rather less practice, RT for the last serial position decreased slightly with increasing set size. So that although this is a possible point of difficulty for Model III, the evidence is so far equivocal and dependent for significance on gaining absolute control of rehearsal, which is probably impossible.

3.125. The effect of direct manipulation of recency.

(i) Recency of the probe. According to the strength models, trace strength remaining from items presented on recent prior trials would be expected to be a factor influencing RT when those items are

¹ unless the initial increment in strength is a function of the strengths of the prior items in the list -- see p.103a.

presented in the memory set or as the probe on the current trial. While proactive effects might also be predicted by the search models, they would presumably be attributed to an effect upon the encoding stage, and should therefore be in the same direction for both positive and negative responses and not interact with set size.

Positive RT: Whether effects on these are predicted by the strength models depends on which version of the representation axiom (a)¹ is assumed. If the presentation of an item in the memory set raises the strength S_i at L_i to a ceiling, then there can be no effect of prior activation. If on the other hand, the presentation of item i adds an increment to whatever value of S_i the representation L_i already possesses, then the resultant trace strength will be greater when the item i has occurred in a recent trial than when it has not, and positive RT should be shorter. Sadly, though Baddeley & Ecob (1970) recommend the collection of appropriate data, none is available. Smith (1967) finds a facilitating effect of presenting a probe a second time for the same list, but his design is such as totally to confound recency with expectancy.

Negative RT: The predictions made here by the models are distinctive and crucial. No model in which the decision is based on a search of the memory set, be the search serial or parallel, self-terminating or exhaustive, can predict an effect of recency of negative probes - unless the recency of the negative probe is held to affect encoding time, in which case there should be no interaction with set size. Model III predicts that since the strength of more recent negative items will be closer to the criterion, more recent negative probes should be rejected both slower and less accurately. In Model IV, since negative RT does not depend directly on the

¹see p.96.

strength of negative items, but only indirectly via the criterion T_c placing, no RT difference is predicted but subjects should make more errors on recent negative probes. Zechmeister (1971) and Darley & Arabin (1972) in experiments in which the memory set on each trial was composed of novel words showed that it took longer to reject a word when it had occurred in a recent trial than when it was novel, which appears to favour Model III (or V, as Darley & Arabin wish to argue) and there was an interaction with set size. My own data on this point, together with further discussion, are presented in Ch.7.

(ii) Recency of the memory set. There are two ways in which this could be manipulated. Reducing rate of presentation, since it spreads the list out more in time, should, if decay is time-dependent, increase the slope of the set-size and serial position functions. However, as attempts to solve the problem of decay-vs-interference in STM demonstrate, there is a considerable problem in equating level of acquisition and amount of rehearsal for different rates.

A second possibility is to vary the delay between memory set and probe by interposing a brief rehearsal-preventing task. (The principle is that of the Peterson 'distractor' technique (Peterson & Peterson, 1959), though in this case the task will have to be both highly controlled and brief enough for IRn performance to remain 'error-free'. What do the trace-strength models predict? If, as I have suggested, the recency function is such that the rate of increase of RT decelerates as time since presentation increases, the consequence of a filled probe delay would be that the average slope of the portion of the recency curve tapped should be smaller than when test is immediate. Hence a filled probe delay should

reduce the slope of the set-size and serial position functions. Now such a finding would be very difficult for the search models to explain; if they predicted any change in slope of the set size function, it would surely be an increase - due to longer comparison times produced by deteriorating representations in active memory. An attempt is made to test this prediction in Experiment VI (Ch.7). I know of no other relevant data; Burrows & Okada (1971) and Connor (1972) report a smaller slope with longer probe delays, but neither study controlled rehearsal during presentation or delay.

(iii) The effect of repeating items in the memory set. Baddeley & Ecob (1970) in an elegant pair of experiments show that when items which are repeated in the memory set are probed, RT is shorter than for items from control lists without repeats whose backward serial position is the same as the most recent presentation of the repeated item. Such a finding cannot be explained by Model I. Model II is counter-indicated by the finding of parallel positive and negative set-size functions. Nor can the results be due to a bias against non-repeated items, or to a strategy of deleting the repetitions from memory (cf. Derosa, 1969) since RT to non-repeated items in lists with repeats is unaffected. Either of the strength models are sufficient to account for the results. (Note that were one certain that subjects always rehearsed the list an integral number of times, this finding would also argue against that version of the representation axiom in which trace strength is raised to a ceiling on presentation. However, since the serial position effects are rather small, there are no grounds for such certainty.) It is also worth mentioning that in a continuous recognition paradigm,

both Hintzman (1969) and Okada (1971) found facilitating effects of repetition on positive RT.

3.126. Practice.

A marked effect of practice on the slope of the set size function would be troublesome for the strength theories, since it would imply that either decay rate, or the nature of the decision function could be changed by practice. Scanning models make no strong predictions though it would be quite surprising if comparison time were not reduced by practice. But Nickerson (1966) practised 4 subjects for 17 days, with only two set sizes, and found no change in slope. And Burrows & Murdock (1969) who practised two subjects for 15 days in a multiple-probe experiment say that the effects of memory load were not attenuated by practice.

3.127. The effects of probability and isolation of the to-be-probed item.

Darley, Klatzky & Atkinson (1972) show that if the subject is certain which item is to be probed there is no effect of set size on RT. This is unsurprising according to any model; the subject is presumably able to rehearse the indicated item selectively. But Klatzky & Smith (1972) manipulated the probability that an indicated member of the memory set would be probed, and found that biasing the subject towards an item reduced the intercept of the set size function, but not the slope, compared to control lists. Morin, Derosa & Ulm (1967) showed that isolating an item spatially led to faster RTs when it was probed, but only for longer lists. Model I can presumably only account for such effects by supposing that isolation

or probability affects the encoding stage, but then the interaction with set size in the Morin et al. study is a problem. Model II is obliged to say that the 'biased' item tends to be scanned earlier in the sequence. The strength models can accommodate such findings by supposing that the biased item receives more rehearsal/attention/time and thus a stronger trace is established.

3.128. The effects of response probability/bias.

As far as I know, this has not been systematically studied for the varied-set paradigm. This is a pity, since some interesting predictions arise. The search theories predict only intercept effects. On the basis of strength theories, we would expect this variable to affect the point at which the subject places his criterion between the distributions. Without definite knowledge of their shape, precise predictions cannot be made, but it remains that an effect of response bias on the slope ratio would not be incompatible with strength theories, and according to Model IV any effect that is found should be on the negative slope only, since positive RT does not depend on the location of the time criterion T_c .

3.129. Subsets.

There exists some evidence which may seem to suggest that the subject can use information specified in the probe to restrict a search of the memory set to a subset within it. Kaminsky & Derosa (1972), using mixed lists of digits and letters, obtained evidence that with overall set size constant, RT increases with the size of the subset probed. (See also Exps. VII and VIII, Ch.8) Such evidence clearly argues against a pure exhaustive scanning strategy, and

may pose problems for trace strength theories. This will be further discussed in Ch.8 in the context of my own data, but there are points worth making here. Firstly, the result does not appear to be highly reliable, in that only about half of Kaminsky & Derosa's subjects showed an effect of the size of the subset probed; also, for every subject, RT increased at a slower rate with increases in subset size than with the same increases in set size in pure lists. Secondly, the available evidence appears to suggest that specification in the probe of the relevant subset is only effective if the subset is defined by semantic category. Burrows (1972, Exp.II) shows that specifying the subset by modality of presentation does not help, and I shall be presenting data in Ch. 6 which show little or no facilitating effect of specifying the subset by spatial or temporal position.

3.13. The nature of the representation.

The issue of the format of the memorial representations underlying performance in the immediate recognition paradigm is to some degree independent of the question of how they are accessed and what their dynamic properties are. It remains that the necessity in the trace strength models for a pre-existing, content-addressable representation L_i for each item i of the stimulus ensemble favours, on the grounds of economy, an abstract or amodal representation of the form of Morton's (1969a, 1970) logogens; the search models impose less theoretical constraint, while a mixed model must require two forms of storage.

Burrows (1972) suggest that four possible forms of memory coding are possible in this paradigm: a) auditory (or articulatory)

(Conrad, 1964; Sperling, 1963), b) visual (Sternberg, 1969, appears to favour this, but perhaps for visual presentation only), c) bimodal, that is, items represented in stores specific to their modality of input (Murdoch, 1972), d) amodal (Morton, 1969a, 1970). Before considering the evidence, it may be worth risking a somewhat speculative elaboration of a fifth, which we may call e) logogen + sensory component.

3.131. The idea of a sensory component of trace strength.

Logogens are conceived as essentially terminal units in a 'passive' (Morton & Broadbent, 1967) recognition system based on a hierarchy of feature analysers (cf. Selfridge, 1959; Norman, 1968; Treisman, 1969; Morton, 1969a, 1970). They receive input from both visual and auditory feature analysers. A logogen is activated therefore as the end result and apex of a particular 'pyramid' of activation spreading upwards through the hierarchy from the receptor surface. I suggest that there may be a sensory component in trace strength due to traces surviving for a brief period in feature analysers, as well as the longer lasting traces intrinsic to logogens themselves. Such sensory traces will facilitate the activation, and therefore contribute to the trace strength, of any logogen provided that and to the degree that the probe activates the logogen through the mediation of those feature analysers retaining traces. Thus if the probe P_i has the same physical form as the presentation in the memory set of that item i , the trace strength of that item will (for as long as the sensory trace lasts) be enhanced by the sensory component; if the physical form differs, as for example when the probe is presented in a different modality, there will be no enhance-

ment. The net effect may be to change the effective rate of decay in cases where the probe has the same physical form as the stimulus, compared to cases where it does not.

The issue is complicated by the work of Posner (Posner, 1969; Posner et al., 1969) which raises the possibility that, in the terms of the present discussion, subjects may generate, 'efferently', sensory traces from a name code, provided that they have half a second or so in which to do so. There may, therefore, be an effective sensory component even where item and probe are not physically identical. However, Posner (1969) presents data suggesting that the subject's ability to generate sensory codes, of letters at least, simultaneously, is limited to one or two items. Thus physical identity will in general be expected to make a difference only with larger memory sets. The net effect of physical identity of probe and list items would on these grounds be smaller slopes of the set size functions, than when probe and item were of different format.

These ideas correspond nicely with suggestions made by Eichelman (1970). In a serial naming task, using upper and lower case letters, he looked at the difference between the facilitation of RT due to repeating the same stimulus (e.g. aa) and that due to repeating the same response (e.g. Aa) as a function of R-S interval; he found that the difference decreased as R-S interval increased from 200 to 700 msec. in a manner closely analagous to the relationship found by Posner's group between the advantage of physical identity over name identity matches and ISI, (Posner & Keele, 1967; Posner & Mitchell, 1967; Posner, 1969). Eichelman proposes that the same principle underlies both phenomena:

"The advantage of PI (physical identity) over NI (name identity) may not be due to a terminal decision reached at a lower stage. Rather,

we may think of the process as a savings in the time required to 'read-in' the stimulus when it is physically identical to the immediately preceding one. It would be like regenerating a decaying trace or organization of neural elements that correspond to the 'elements' making up the physical event." Eichelman (1970, p.95)

Returning to the IR_n paradigm, in addition to the effects on the slope of the set size function mentioned above, a sensory component, decaying relatively rapidly compared to trace strength at a logogen, may account for part of the non-linearity of the set size function. If this idea has any substance, we would expect set size functions to be most markedly non-linear when the probe delay is short and the probe of the same physical format as the list; the data of Morin, Derosa & Ulm (1967) appear to support this, and similar data of my own will be presented in Ch.7.

What sort of decay time should we attribute to the sensory component? If we take Posner as our guide, decay back to baseline level would appear to occur in one to two seconds for visual letter features (but may be delayed by 'rehearsal' or 'attention'). If such traces are also the basis of iconic memory, Sperling's (1963, 1967) 'VIS', then Sperling's data would suggest a rather more rapid decay. But it should be noted that traces which are not strong enough to make responses available (in the sense of Section 1.4) for recall may nevertheless be adequate to facilitate the activation of logogens. There is some evidence from a recognition study that visual information survives for several seconds (Philips & Baddeley, 1971). As for the auditory system, the work of Crowder & Morton (1969) suggests that memory traces for speech features survive, in what they call PAS, in a form which is adequate to make verbal responses available for a period of two seconds or so, and the results of Norman (1969a) on memory for the rejected message while shadowing, and of Treisman

(1964) on the lags at which the identity of the shadowed and rejected messages are noticed, lead to a similar conclusion. We might therefore conclude that traces in PAS (below response threshold) may contribute to logogen activation for periods considerably in excess of two seconds. Thus the sensory component apparently has a slower decay rate for auditory (speech) features than for visual features.

One problem over which the argument above skates arises when we consider the question of whether the sensory component is disrupted by a masking stimulus. Posner (1969) fails to find an effect of masking, and concludes that his form of sensory storage is not the same as Sperling's. However, it does not seem to have been remarked that Posner used a checkerboard mask, while Sperling used a masking stimulus consisting of letter features. Moreover, Morton et al. (1971) found that the less similarity the suffix bore to the list items, in terms of features like pitch and timbre, the less it destroyed the recall advantage, to the last items, of auditory presentation. It therefore seems not unreasonable to conclude that in both modalities a masking stimulus will only be effective to the extent to which it activates the same set of feature analysers as the masked stimulus, and that Posner's and Sperling's sensory stores may well be the same in the sense discussed above.

This attempt to inter-relate iconic and echoic memory, Posner's work, stimulus repetition effects and the immediate recognition paradigm is perhaps dangerously speculative. However, bearing in mind the present interpretation of trace strength as being located primarily at amodal logogens, but possessing a relatively short-lived sensory-specific component, let us turn to the evidence relevant to memory format in the IRn paradigm.

3.132. Evidence on the nature of the representation.

(i) Cross-modal experiments, in which the memory set is presented in the same modality as or in a modality different from the probe. The argument usually runs: if the probe item has first to be transformed into a format compatible with the memory set items in storage, an increase in RT will be expected. In the context of the search models it is further argued that if the probe is translated into the format of the memory set, this will affect only the intercept, while if the memory set items are transformed, the slope of the set size function will be increased. Chase & Calfee (1969), who used letters and a rather long unfilled probe delay, found a small but significant interaction such that the unimodal slopes (AA, VV) were less than those in the cross-modal condition (AV, VA). The more general finding is however exemplified by Burrows (1972) who, using digits, found that VV, AV and VA slopes did not differ, while in the AA condition RT was shorter overall and the slope smaller. Kirsner & Craik (1971), who used words and a constant list length, again found that mean RTs were ordered: AA less than VV = AV = VA, and that the slope of the serial position function was slightly reduced for AA. Goldring (1969) replicated Chase & Calfee's design using words: unfortunately the data are reported only in terms of slopes - AA less than VV less than AV less than VA.

The data seem somewhat contradictory; how do the various models fare? The models postulating a single format predict no interaction between list modality and probe modality - only the latter should be effective provided adequate time is given for encoding the list. The bimodal model predicts that unimodal conditions should be superior to cross-modal conditions, and also that mixed-modality lists

should be harder than pure lists: Burrows (1972) included the appropriate condition to test the latter prediction and showed that it failed. Burrows (1972) concludes in favour of a logogen model, supplemented by PAS. I am inclined to agree: I have suggested above that the presence of an effective sensory component may result in smaller slopes. Why then the apparent asymmetry in the results, in that AA is superior to VV, which shows little advantage over the cross-modal conditions? It is noteworthy that both Kirsner & Craik and Goldring used visually complex material (words) and a long probe delay, while Burrows' probe delay was filled with a visual display saying 'Probe'. Chase & Calfee, with an unfilled probe delay and single letters found smaller slopes for both unimodal conditions. So that an advantage of AA over VV is only found when a visual sensory component is likely to be minimised.

(ii) The effects of visual and acoustic similarity. The argument here is obvious: if a particular memory format is hypothesised, then similarity of the material in terms of this format should have a disruptive effect. Moreover, according to the search models, we should expect an increase in slope. There are three problems in investigating similarity effects. First, as Conrad (1964) points out there is the problem of controlling for the effects of similarity on perceptual processes rather than on memory, so that ideally we should test for the effects of, say, acoustic similarity with visual presentation and test. Secondly, there is the problem of rehearsal. I have suggested (Section 1.4) that though the traces underlying IRn are themselves located at amodal logogens, they may be revived by re-recognising the items from a response buffer in which items are articulatorily coded - so that effects of phonemic similarity could

arise during this rehearsal process. Thus rehearsal should be avoided if one wishes to test the logogen model. Thirdly, if we select two vocabularies which differ in similarity, there is also the danger that they will differ in some variable confounded with similarity, such as encoding difficulty. This must be controlled for (see design of Exps. I, II, III, Ch.5)

Now although these conditions have not simultaneously been met in any published experiments, it remains the case that such effects of acoustic¹ and visual similarity as have been found are both small and equivocal. Chase & Posner (1965) found no effect of acoustic confusion in a recognition task with visual presentation and test. Chase & Calfee (1969) with both visual and auditory presentation found a small effect of acoustic similarity, unreliable in that it occurred in only one of two experiments, and then mainly for negative probes. They felt that their data justified the assertion that "the Sternberg (1966) recognition memory test procedure largely bypasses the acoustic/articulatory processes that result in acoustic interference". Wilcox & Wilding (1970) who used auditory presentation found a significant increase in slope with acoustically similar lists, but their data showed rather wide deviations from linearity in both directions, and the slope difference was due mainly to faster RTs for acoustically similar lists when $m = 1$ and 2. Connor (1972) who presents set size slopes for negative trials only, showed no effect of visual or acoustic similarity at short probe delays, and mixed and messy effects at longer delays.

So far, evidence on similarity effects is adequate neither to reject a hypothesis of amodal representations, nor to support one of a format specific to a particular modality.

¹by 'acoustic' similarity is meant no more than phonemic similarity

(iii) 'Semantic' factors. If the memory representation had the form of an acoustic or visual code, one would expect no effects of any semantic features of the items on RT. Some such effects would, in contrast, be congenial to a 'logogen' model, since there are grounds for supposing that logogens are in some sense associatively related: Morton (1969a), in order to explain the effects on word identification of context and redundancy proposes that the sensitivity of individual logogens is partially a function of the prior activation of semantically and syntactically related logogens, and also accounts for the effect of word frequency by supposing a lower threshold for high-frequency words (cf. Treisman, 1969). In line with these premises, we would expect that the trace strength of a given item in the memory set would be enhanced when it enjoyed a strong semantic relationship with other items in the memory set. Positive RT should therefore be increased by semantic relatedness, while a negative probe should be easier to reject the greater its semantic distance from the memory set items.

The evidence is as follows. Morin, Derosa & Stulz (1967) and Derosa & Beckwith (1971) showed that when sets of four digits were used as the memory set, if the digits formed a numerically consecutive set a) RT was shorter b) the purity and size of the serial position recency effect at short probe delays was somewhat moderated and c) negative RT was faster the more remote the probe was numerically, from the positive set. Also, at short probe delays, it did not make much difference whether the consecutive set was presented forwards, backwards or in a scrambled order. With the additional but reasonable assumption that numerically adjacent digits are the most strongly associated, these data are entirely compatible with the above suggestions. Krueger (1971) using sequentially presented memory sets of

letters which did or did not comprise a word showed a marked facilitatory effect of this form of redundancy on both RT and accuracy. Smith (1967) has shown that high frequency words are accepted more quickly than low frequency words and that this effect does not interact with set size. This is unlikely to be a pure encoding effect, in Sternberg's sense, as the effect does not reach significance for negative probes. No one has yet manipulated semantic similarity in this paradigm in a manner analogous to formal similarity. This would seem to be worth doing; does it, for instance, take longer to reject a probe word and less time to accept it, if the memory set contains a synonym? A final point on semantic attributes: Smith (1967) and Juola & Atkinson (1971) both find that the general pattern of results obtained using memory sets consisting of category names and probing with exemplars of the categories does not differ markedly from that obtained with memory sets composed of words in the usual manner. The latter authors discuss their results in terms of various scanning strategies, but a notion that may be preferable, if vague, is of categories functioning, like superordinate nodes in a Collins & Quillian (1972) - style semantic memory network, as 'super-logogens' with their own decay characteristics.

(iv) Translation experiments. By this I mean experiments in which the probe itself is not, in the experimental conditions, in the list, but the subject has to indicate whether an item associated to the probe by some rule or by previous learning is. That is, in the control condition the probe is physically identical to its occurrence in the list, while in the experimental condition some translation operation is required. Klatzky, Juola & Atkinson (1971) in an improved version of the experiment of Klatzky & Atkinson (1970)

presented memory sets of letters and probed with either a letter or with a picture of an object - the first letter of whose name was the effective probe. In both studies the slopes of set size functions were in general greater for the picture probes, which required translation, than for the physically identical letter probes. Swanson, Johnsen & Briggs (1972) had subjects learn 2-digit numerals as names for a set of random octagonal shapes, and then factorially crossed the memory set form and probe form in an IRn experiment with visual presentation. They found greater slopes for associational identity than for physical identity. Cruse & Clifton apparently (Sternberg, 1973) obtained equivalent results in an experiment in which subjects learned to associate letters to digits.

Note that these results are analagous to those obtained in cross-modal experiments; a similar theoretical argument may be applied, namely - with small memory sets, even when the physical form of the probe and memory sets differ, the subject is able usefully to prepare himself for the physical form of the probe, by the 'generation' of sensory traces; thus the facilitatory effect of the 'sensory component', in conditions where probe and memory set have the same physical format, is seen only at longer list lengths. Search models must presumably account for the difference in slopes with and without physical identity by the assumption that the subject translates each item in the memory set before comparing it with the probe, rather than vice versa.

(v) The range of possible stimuli. One possible difficulty for a 'logogen' theory is the range of stimuli which have been found to give 'typical' results in the IRn paradigm. While the notion of logogen-type representations seems reasonable not only for words, digits

or letters but also for any constellation of features which is familiar as an individual entity (e.g. the nonsense shapes used by Swanson et al., 1972), Krueger (1971) used line figures and Sternberg and Treisman (Sternberg, 1969a, Exp.4) used faces and nonsense forms, and in both cases 'typical' results were found, even though the stimuli were initially unfamiliar and relatively little practice was given. Of course, that logogens (if they exist) have associated sensory attribute sets criterial for their activation does not imply that a system cannot be built up for identifying (as a familiar entity) a (disjunctive) set of sensory attributes without having as its terminal unit a node representing a name. But it remains an open and interesting question how fast such a system can be built up. It may be a relevant fact that one only has to see a face once before identifying it, days later, as a face one has seen before.

3.14. The varied-set immediate recognition paradigm: summary.

This section has consisted of an attempt to combine a presentation of some models of performance in the varied-set version of the IRn paradigm and the predictions deriving from them, with a review of the existing literature and of some points at which more data is needed. A brief summary of the argument may at this juncture be useful.

I believe I have shown that while all the models mentioned are compatible with the typical set-size functions obtained, the trace strength models can cope better with the small deviations from linearity; moreover, the inspection model has the advantage of permitting the wide range of slope ratios reported in the literature, together with related error frequencies. The widespread finding of marked recency in the serial position data is extremely problematic

for either search model, and while limited ad hoc solutions may be found, strength models are again favoured. The effect of the recency of the negative probe is held to be crucial: existing data appear incompatible with both the default model and any pure search model; further data will be presented in Ch.7. Baddeley & Ecob's (1970) repetition experiment is further evidence against Model I. The effects of practice and stimulus probability, and the gradual nature of the deterioration of performance as set size is increased to large values, while not decisive, is at least as compatible with strength as with scanning models. Turning to the nature of the representations, I have suggested that a trace strength model is most easily combined with a hypothesis of amodal representations like logogens, and that there is evidence for this hypothesis in the semantic effects found and in the failure so far to show any very convincing effect of visual or acoustic similarity. To account for the findings in cross-modal and translation experiments of smaller slopes when probe and memory set are of the same physical form, I introduced the idea of a 'sensory component' of trace strength which may, for one or two items, be efferently generated. While this is the most speculative part of the argument, it has the advantage of relating the present paradigm to the stimulus repetition effect and the various forms of evidence for sensory memories. Nor is the performance of the other models in relation to this aspect of the data impressive.

On existing evidence I am inclined to favour the trace strength inspection/logogen (+ sensory component) model, although I have pointed out the difficulties of distinguishing the trace strength inspection model (Model III) empirically from a mixed model (Model V), to which it is however to be preferred on the grounds of pars-

imony. Among the possible difficulties for the favoured model are: the small primacy effect in the serial position data, the apparent facilitation resulting from information allowing the subject to restrict his 'search' to a semantic subset, and the results obtained with unfamiliar non-symbolic material. Aspects of the data which merit greater attention are: the relationship between RT slope ratios and error rates, the effect of set size on RT variance, the effects of directly manipulating the recency of the probe and the memory set, response frequency, and the effects of semantic and acoustic similarity with appropriate controls. Some of these points will receive attention in following chapters.

A final point: Sternberg (1973) was at pains to refute the argument from plausibility which he had heard levelled against his exhaustive scanning hypothesis, remarking that the engineers at Bell Telephone Laboratories could see many advantages in such a strategy for retrieval in an information-processing system. One takes the point, but it remains the case that the brain is made of neurons not ferrite disks and magnetic tape. In this context it is worth mentioning that the concept of trace strength has had some currency in the neurophysiological literature (e.g. Konorski, 1961), nor is the concept of a logogen without its physiological relatives (e.g. the 'cell-assembly' of Hebb, 1949, the 'gnostic' units of Konorski, 1967, and the 'grandmother cells' spoken of by Lettvin and co-workers at M.I.T. - Teuber in lectures).

3.2. The constant-set recognition paradigm.

3.21. Sternberg's argument.

In this version of the recognition paradigm, in which the memory set is presented only at the beginning of a block of trials or a session, and the sequence of trials consists merely of a sequence of probes to be evaluated with respect to this constant set, it seems reasonable that the memory set has entered inactive memory. Sternberg (1969a) remarks that subjects who had roughly ten minutes practice with particular positive sets were able to recall them days later. That the set is represented in inactive memory does not, of course, necessarily imply that the inactive memory representations mediate performance in this task. Sternberg (1969a, 1973) has argued that they do not. His argument has two horns: the first is that the results obtained using the two versions (c-s and v-s)¹ of the paradigm are similar in that he finds parallel and linear set size functions of approximately the same slope and intercept in both, leading to the claim that "even when a set is contained in long-term memory, it is transferred into active memory and maintained there by rehearsal in order to be used in the item recognition task". The second is a "small preliminary experiment" (Sternberg, 1969a, Exp.5), intended to test this idea by restricting the opportunities for rehearsal immediately prior to the presentation of the probe. This was achieved by presenting in the experimental condition a string of seven letters before the probe on each trial. On half the trials the letters had to be recalled, on the other half a digit probe was presented for classification with respect to a constant set of 1, 3 or 5 digits. Compared to a control condition without the letter recall task, performance was markedly slowed: the intercept increased by about 130

¹ abbreviations: c-s = constant-set, v-s = varied-set.

msec. while the slope increased to approx. 105 msec./item, though parallel and linear functions were still obtained. So it would appear that "the retrieval process is radically altered when the information to be retrieved is not being rehearsed and is therefore not in active memory".

This evidence is indeed persuasive argument for the contention that in some circumstances the subject covertly turns the c-s paradigm into the v-s paradigm by rehearsing before every trial. (This suggestion is independent of which model, exhaustive scanning or trace strength, actually holds for immediate recognition.) It is however the purpose of this section to question the generality with which such a strategy may be supposed to underly c-s recognition performance. To start with, there are several ways in which the data now available show important differences between results obtained in the two versions. I shall mention three examples.

3.22. Differences between the c-s and v-s versions of the IRn paradigm.

(i) Linearity of the set-size function and practice. In Sternberg's (1966, 1967b) c-s experiments, each digit was associated with both the positive and the negative response during the course of the experiment; also subjects were relatively unpractised. There is now considerable evidence that when the assignment of stimuli to responses is consistent, practice rapidly leads to the development of markedly non-linear set-size functions. Briggs & Blaha (1969) using random shapes as stimuli, consistently assigned, and set sizes 1, 2 and 4, found linear functions early in practice, but later, RTs were as fast for two-item positive sets as for one¹. Usually, however, RT is found to be an approximately logarithmic function of set size. Ross

¹Oddly, Swanson & Briggs (1969) who used, as far as one can tell, identical procedure and design, plot their data, for theoretical reasons, on log. coordinates and obtain impressively straight lines - i.e. RT varied as \log_2 (set size). They do not comment on the discrepancy.

(1970) gave subjects extended practice in a c-s recognition experiment on sets of 1, 2, 4 and 8 alphanumeric characters, which were both consistently assigned and nested and found that even on the first day, the set-size data were well fitted by a logarithmic function, whose slope was continuously reduced by practice. Simpson (1972) using non-intersecting but consistently assigned sets of 1, 2 and 4 letters over 3 sessions also obtained a log. relationship and in a second experiment with set sizes of 1, 4 and 8 found a change within the single session from a linear to a log. function. In both studies error rates were comparable to Sternberg's. Marcel (1970) using a between-subjects design for set size, so that response assignment was necessarily consistent, found a similar, clearly non-linear set-size function. Kristofferson (1972b), who practiced subjects for 36 sessions with positive sets of 1, 2 and 4 digits, both nested and consistently assigned to responses, found non-linear functions (of approx. log. form) from the first day; the slope decreased and the non-linearity increased with practice. The role of consistent response assignment is thrown sharply into relief by the fact that in another study (Kristofferson, 1972a) identical in all respects save that non-nested sets were assigned inconsistently to responses from day to day, she found no effect on slope of practice and linear functions. That consistent response assignment rather than the use of nested sets is crucial is suggested by Simpson's (1972) results (see above); also Burrows & Murdock (1969), who practised two subjects for 15 sessions with nested sets whose members were not consistently assigned to responses, found in a multiple-probe experiment that latency was a linear function of memory set size. The function's slope was not changed by practice. (This study is, however, unlike the standard c-s design in various other respects.)

It is clear that when responses are consistently assigned to stimulus items, that is, when the subject is able to learn a particular stimulus-response mapping, non-linear and usually logarithmic set-size functions are obtained (which is reminiscent of the logarithmic relationship between set size and number of alternatives in CRT experiments); also, such effects are progressive with practice. In the absence of this condition, and in the v-s paradigm (with a couple of dubious exceptions - see Section 3.121) marked deviations from linearity are not found, and slope is not changed by practice.

(ii) The effects of probability/frequency/recency. Stimulus probability is well known to be an important determinant of CRT (Welford, 1968). In Sternberg's c-s experiments stimulus probability is confounded with set size, as Theios et al. (1971)¹ point out. There is now impressive evidence that in the c-s paradigm, when stimulus probability is increased independently of set size, RT to both positive and negative probes is reduced. Hawkins & Hosking (1969) demonstrated the effects of stimulus probability within (presumptive) positive sets, and Krueger (1970) found RT to both positive and negative probes to be shorter the greater the relative frequency of the probes; however since the positive and negative sets were of equal size it is possible that subjects were coding in terms of the complementary set, though they claimed the contrary. Theios et al. (1971) in a study in which set size and stimulus probability were varied in a partial factorial design, but again with positive and negative sets equal in size, found that RT was monotonically decreasing function of stimulus probability for both responses. It is likely that in these studies probability was effective as a variable via recency - i.e. the more probable the stimulus the more recently

¹Theios' model will be outlined anon.

it must on average have occurred. The connection is well demonstrated in an impressive experiment by Walter reported by Theios (1972). Walter used Sternberg's (1966) design so that the negative set was always larger than the positive set but in addition varied stimulus probability, with two R-S intervals (0.5 and 2 sec.); also more subjects were used and more sessions were given than in Sternberg's (1966) experiment. Sternberg's parallel set size functions were replicated though the departures from linearity were significant. Again, RT was found to be inversely related to stimulus probability. In addition a stylish analysis of sequential effects revealed a stimulus repetition effect extra to the effect of response repetition (cf. Rabbitt, 1968; Bertelson, 1965) and this effect was only slightly reduced at the longer R-S interval.

These studies show clearly that in at least some experiments using the c-s paradigm, a negative probe is rejected faster the more recently it has occurred. This is in complete contrast to the results obtained in the v-s paradigm (Section 3.125 and Exp.VI, Ch.7), in which a more recent probe is rejected more slowly and with more errors.

(iii) Semantic factors. A third region of difference between the c-s and v-s data is to be found in the effects of semantic features of the stimuli. For instance, both Derosa & Morin (1970), and Marcel (1970), who used positive sets of consecutive digits in the c-s version, showed that RT decreased the more remote the negative probe was numerically from the positive set. This result is found with the v-s version (Section 3.132 iii). But effects of numerical position within the positive set appeared which are not a feature of the v-s data. Derosa & Morin found longer RTs for items at the boundary of the set, and Marcel, whose sets all began with the digit 1, found that

the nearer to the upper bound of the set the probe was, the longer the RT.

These pieces of evidence imply that a general claim for similarity between the results from c-s and v-s versions of the IRn paradigm may now no longer be made. In addition, the evidence includes quite contradictory features, such that no single model is likely to prove adequate to explain performance in all constant-set experiments. For instance, Sternberg (1966, 1973) showed no effect of the size of the negative set. Yet, the smaller the negative set, the greater the frequency of individual negative items will be, which should lead to results like those of Theios and his associates. On the other hand the 'logarithmic' set size functions found pose obvious problems for a serial scanning model as do the effects of frequency/recency. Attempts to account for the latter in terms of an effect of recency on encoding time in Sternberg's stage model are frustrated by an interesting feature of the data from Walter's experiment (Theios, 1972). Namely, the effects of stimulus probability on negative RT interacted with positive set size. According to Sternberg a slope change implies an effect on the comparison stage, not the encoding stage.

3.23. Strategies available in the c-s version.

Though the purpose of this discussion is primarily to exclude most of the c-s studies as being of at best ambiguous relevance to immediate memory, which is our main concern, it may be worth indicating the lines along which possible solutions to the above problems may be found. They lie, I suggest in the idea that the use of the

same memory set over many trials, and the consequent possibility of its representation in several forms, including inactive memory, offers a range of strategies to the subject, which will be differentially favoured under different circumstances. Let us consider some that may be possible.

3.231. Rehearsal of the set in active memory.

This has already been mentioned as Sternberg's preference. In what circumstances would such a strategy be ineffective? If the interval between successive probes were too short for complete rehearsal of the list then it would not be possible to base judgements on pure recency (or for that matter on an exhaustive scan of the contents of 'active memory' since 'active memory' would contain recent negative probes). In Theios' experiments the R-S intervals were typically short (1 sec. in Theios et al., 1971, 0.5 and 2 sec. in Walter's experiment) and it is in these experiments that data not predictable from this strategy were obtained (effect on negative RT of frequency, non-linear set size function). If the subject is given some task which prevents him rehearsing the list at all, this strategy should again be disrupted or prohibited. Sternberg's (1969a) Exp.5 shows that performance is markedly changed by a concurrent irrelevant memory load, though one should note that since the subjects could still perform the task, other strategies must be available.

3.232. Access to permanent representations (facilitated by recency).

Suppose that the subject accesses some permanent representation of the probe item in inactive memory and is thus directed to a representation of the appropriate response. What might the nature of this access be?

I shall describe first a model developed by Theios (Theios & Smith, 1970; Theios et al., 1971) and recently elaborated further (Theios, 1972). As well as performance in the c-s paradigm it is intended to account for stimulus probability effects in RT (Hyman, 1953; Falmagne, 1965) and stimulus sequence effects (Bertelson, 1965, Remington, 1969). Theios et al. (1971) say:

"We assume that encoded representations of all the elements of the set of possible stimuli are stored in memory with codes of their respective responses, as paired associates. Further, memory is conceived of as a dynamic stack or hierarchy in which representations of the set of possible stimuli are changing position from trial to trial, but are ordered roughly on the basis of the recency and frequency of their occurrence." (Theios et al, 1971, p.2)

In order to determine the response, the probe item is searched for in this 'stack' by means of a self-terminating scan, starting from the top. The more recent and frequent stimuli will be higher in the stack and so found first. The crucial feature of the model is that negative items are included in the memory stack, thus accounting for the effects of frequency/recency on negative RT.

Now, Theios (1972) has elaborated this model considerably.

Briefly, the following additions are important:

1) The serially searched memory stack is now held to be a short-term buffer¹ with a limited capacity, in which only a subset of the S-R pairs (the most frequent/recent) are stored. All pairs are also stored in and retrieved from LTM (long-term memory), and may be considered as occupying a terminal slot in the buffer. Retrieval from LTM, which is slow, occurs in parallel with search of the buffer, and its speed determines the effective capacity of the buffer since once the LTM representation is found, no further search of the buffer is needed.

The point of this complication is to account for the fact that choice RT is a non-linear function of number of alternatives and

¹for 'buffer' do not read 'response buffer' (as used in Section 1.4 and Ch.2)

tends towards an asymptote (as Theios, 1972, demonstrated in a task involving key-pressing to digits). The asymptotic limit is determined by the time required by retrieval from LTM.

2) Very familiar responses such as names are held to be in some sense 'directly' available, and do not require a buffer search. Theios (1972) in a serial naming task with parameters identical to the key-pressing task just mentioned, found no effect of number of alternatives. Also he found no effect of stimulus probability in a serial naming task (though a slight trend is discernible in his data).

3) There are now held to be two short-term buffers - one containing S-R pairs as already mentioned, the other containing the response output programmes appropriate to each response. Though the latter buffer can safely be ignored in the c-s recognition task, its function is to account for the separate effects of stimulus and response frequency, repetition or number of alternatives.

There is no doubt that this model accounts for Theios' data better than Sternberg's exhaustive scanning model; it also has advantages of generality, in that a wide range of choice, RT phenomena are encompassed. Aside from quibbles such as the fact that the estimated capacity of the buffer varies rather widely when the model is fitted to data from different experiments, the main disadvantage of this model is its postulation of complex mechanisms for which we have no independent evidence and which are not strong in terms of psychological plausibility. I should like to suggest that a simpler version can be articulated which retains the formal properties of this model, but relates closely to the concept of trace strength developed in Section 3.1, and is compatible with the trace strength inspection model derived for performance in the v-s paradigm.

A trace strength model for the c-s paradigm: We make an assumption identical to that made in Section 3.1: namely that every stimulus item has a permanent representation in memory which has the property that subsequent to activation it retains a decaying 'trace' which may be conceived of as an increment in sensitivity or increase in level of activity. We assume also that when a subject has worked with the constant set for a few trials, response representations, with similar properties, become associated to the stimulus representations (logogens) in that the former may be activated by the latter. (The response representation might in fact be a logogen - e.g. that for 'yes'.) That is, activation of an associated response representation is an output from a stimulus item representation when the latter is itself activated to a criterial level by appropriate input. Suppose that given a particular input state, the time taken for either kind of representation to reach a criterial level is a function of its prior trace strength, and that the prior trace strength depends upon the recency with which it has been activated (as well perhaps as other factors like expectancy), as in the immediate memory model; then RT will be a function of the recency of both the stimulus and the response, independently, as in Theios' model. Now since Theios does not specify the precise nature of the recency function, our model will have identical formal properties to his, provided we assume an asymptotic recency function, as we do for the v-s paradigm. To summarise, in Theios' model, the subject scans 'downward' through a stack in which 'height' is a function of recency; in the trace strength model stimulus representations are content-addressable without search from the probe, but the time taken for a particular level of activation to be reached is the same function of recency.

The assumption of similar properties for response as for

stimulus representations in our account answers the same needs as moved Theios (1972) to propose two separate short term buffers. But why on this account should naming stimuli be different from classifying them (as Theios, 1972, showed). One possible answer is that, as in Morton's account, the articulatory features of the name are a direct output from the logogen, and do not require the activation of mediating response representations. Nevertheless we would still expect some effect of stimulus recency: it is interesting that though Theios' results from a serial naming task showed virtually no effect of stimulus recency, Kirsner & Craik (1971) and Kirsner (1972) in experiments which involved discrete trials did find a marked 'priming' effect on naming of earlier presentation of the item, which was dependent on the recency of the previous presentation (which the subject did not have to name), for both letters and words. Moreover, the effect was greater for words than for letters, which Kirsner (1972) attributes to the fact that letters come from a small set and are over-used. Perhaps the conditions of Theios' (1972) naming experiment were such (digits, short ISI) that priming effects were effectively at a ceiling.

How does this use of the trace strength concept differ from that proposed for the v-s paradigm? The trace strength models for the v-s paradigm assume that the subject's decision is actually a judgement about the magnitude of the active trace at the representation addressed by the probe. In the c-s paradigm, it is suggested, the subject retrieves the response from inactive memory via the same representation, and this retrieval is facilitated by any active trace remaining from prior presentations of the probe. The nature of the active trace and its dependence on recency are the same in both cases.

3.233. Feature-testing.

A third strategy which may be possible in the constant-set task is for the subject to attempt to discover a small set of features that differentiates the positive items from the negative items, or a disjunctive set of such sets. Rabbitt (1967) raised this possibility in the context of speeded visual search and stimulus classification tasks, and spoke of the subject testing for a 'cue-set' of critical features, rather than making any sort of comparison between the item to be classified and items as such in memory.

There is now substantial evidence that the development with practice of a strategy of this sort is indeed predominant in visual search and underlies what has often been spoken of as a transition from serial to parallel processing in that task. Neisser's (1963, 1964) original data seemed to indicate that subjects could compare display items to ten items in memory as fast as to one, after practice. But it is now clear that the result was dependent on the very high error rate permitted by Neisser (Wattenbarger, 1968), though even with low error rates a decrease in slope and a marked non-linearity appear with practice (Kristofferson, 1972c). Moreover Rabbitt (1967) showed in a card-sorting version of the task the importance of the nature of the negative set items, and Neisser's own data demonstrate that the degree of commonality of features between the target items and the background items is crucial, while at the same time subjects have very little awareness of the background items as identified items. The idea of a 'cue-set' is supported also by Brand (1971) who showed that without prior practice, subjects could search for any of the set of digits in a letter background as fast as for a particular digit - a case in which subjects have already had long experience in distinguishing the two classes of items.

However, we cannot conclude from these results that even after practice the subject is basing his response purely on the outcome of a test for the presence of certain combinations of physical features; Graboi (1971) using a visual search task with a relatively high proportion of targets and accurate performance gave subjects 11 sessions practice on positive sets of 1, 3, 5, and 7 words and found the customary marked decrease in slope of the (negatively accelerated) set size function. The introduction at this point of new target sets returned performance to the initial level, which confirms that specific rather than general practice is necessary; but changing the type face from upper to lower case did not slow performance nearly so much, in spite of the fact that this should presumably disrupt any strategy based entirely on physical features.

Now, to what extent do these arguments apply to the constant-set recognition paradigm? Kristofferson (1972b,c) showed that the pattern of results obtained in the recognition paradigm is closely similar to that obtained in a visual search task, provided that level of practice, error level, positive set structure and response consistency are identical, and argued for a hierarchical feature-testing model, as did Simpson (1972). Foss & Dowell (1971) carried out a version of the recognition paradigm in which the subject had to say whether or not a word began with one of a particular set of phonemes. They reported the usual linear increase in RT with set size for positive sets consisting of dissimilar items. But when the items of the positive set were similar in that they shared several distinctive features, markedly non-linear functions appeared. Foss & Dowell therefore suggested that subjects test for features common to members of the set, though their results did not show the development with practice of the pattern that one would expect. Ross

(1970) after practising subjects for 20 days in a c-s recognition paradigm task found (as did Graboi, using visual search) that changing the stimuli to lower case increased RT, but only by a relatively small amount compared to the overall effect of practice.

To conclude, it is probable that some form of feature testing strategy is characteristic of performance in constant-set recognition, given practice and appropriate material. But the relatively small effect of changing the type-case of the probe (notwithstanding the fact that upper and lower case letters do share some features) seems to imply that such a strategy cannot be the whole story.

3.24. The constant-set recognition paradigm-: conclusions.

The three strategies I have discussed, rehearsal in active memory retrieval from inactive memory and feature-testing, by no means exhaust the possibilities. Consider for instance a situation in which the subject has no opportunity to rehearse the memory set and does not have an unambiguous stimulus-response mapping established but is able to retrieve the memory set items from inactive memory (e.g. Sternberg's 1969a, Exp.5). The subject could read the items into active memory after the arrival of the probe, or read the items from inactive memory checking them one at a time against the probe as represented in active memory.

As a rule, we should perhaps expect a mixture of strategies, across subjects, perhaps within subjects at different stages of practice, conceivably even within a trial - that is, it is possible that two forms of retrieval may occur simultaneously. That retrieval from inactive memory and judgement with respect to a set in active memory may occur in parallel is suggested by the results of an

important experiment by Forrin & Morin (1969) in which subjects were required to evaluate probes by reference both to a varied set presented on each trial and to a constant set. (The two sets never overlapped.) The most striking features of their data were these: although adding the constant set produced an increase in intercept for negative and varied-set positive probes, there was no effect of the size of the constant set. Nor was RT to constant-set positive probes much affected by the size of the varied set. The authors argue that the two sets must therefore have been consulted independently and presumably simultaneously. Can a feature-testing strategy be combined with the other two? It does indeed seem to be the case that subjects may reject probes on the basis of a feature test while other strategies are being employed. Lively & Sanford (1972) in a constant-set study with all-letter or all-digit positive sets and probes from either category showed that while the set-size functions were standard for trials in which the memory set and probe were of the same category, negative probes of a different category were rejected faster, for the larger set sizes. Ellis & Chase (1971) performed a varied-set experiment in which it was possible to reject any probe of a different size or colour without reference to the memory set. It appeared that this feature-based rejection occurred in parallel with consultation of active memory, in that performance was as good as on pure versions of each task (i.e. a) IRn without probes of different form b) judgments of size or colour).

This section has had a threefold purpose. The first was to present some evidence (on non-linear set-size functions, on the effects of recency/frequency on negative RT, on the effect of semantic features) to oppose Sternberg's contention that the data from

the v-s and c-s versions of the recognition paradigm were so similar as to argue mediation by the same process. The second was to argue that the c-s recognition data contain elements irreconcilable within the constraints of any single model, and that subjects can and do use one or several of a range of strategies, depending on the circumstances. The third was to discuss three strategies for which there is evidence, and in particular to point out that Theios' impressive but complicated model can be reinterpreted in a way that is more parsimonious than his model in terms of underlying mechanisms, relates closely to the trace strength models developed in the previous section, but preserves the formal properties and scope of the original. As for the implications of data from the c-s paradigm for immediate memory, it should by now be clear that although active memory traces may be implicated in several of the possible strategies the manner of their implication differs, and until we can establish with certainty what strategy a subject is pursuing, no clear conclusions can be drawn with respect to active memory. I hope, however to have shown that none of the results obtained with the constant-set task need contradict the trace strength models developed for the v-s task in Section 3.1, as is sometimes claimed (Sternberg, 1973), and that the trace strength concept may be of use in interpreting some of the constant-set, as well as the varied-set, data.

3.3. The other 'memory-scanning' paradigms.

We have concentrated so far on the task of recognition. We turn now to consider experiments whose parameters are similar to the IRn studies discussed in Section 3.1, but in which the subject, rather than indicating the presence or absence of the probe in the memory set, must base his response on some other aspect of the relationship between the probe and the list. The main question I shall want to consider is this: given that similar small memory sets are presented to the subject in the same way in all these paradigms, and therefore might naïvely be expected to be represented initially in active memory in a similar way, are we in fact able to claim that performance in these tasks is based on a single form of representation accessed in a set manner, or must we suppose different strategies of retrieval, or even different forms of storage to explain the results obtained? Experiments on short-term forgetting (see Ch.2) have indicated that there is some difficulty in accounting for order and item retention in terms of a single form of storage. It will be recalled that I attempted to resolve these difficulties by the hypothesis of two forms of postcategorical active memory, item-traces in the logogen system and articulatory coding in the response buffer (see Section 1.4). In reviewing the remaining RT data I shall follow a parallel strategy and concentrate on the differences between the v-s IRn data discussed in Section 3.1 and data obtained from those paradigms which require the subject to use or report information about the order or position of items in comparable memory sets. In one respect the reviewer is fortunate: in contrast to the work on recognition, the burden of relevant literature is relatively light and a number of the available studies are somewhat thin in theoretical content.

The major exception is, of course to be found in the work of Sternberg, and the discussion may conveniently be focussed around an experiment by Sternberg (1967a, 1969) which employed the contextual recall paradigm.

3.31. Sternberg's contextual recall experiment.

On each trial the subject was presented a new random list of 3 to 7 digits, with presentation conditions and rate, and probe delay, similar to those used in his IRn experiments. The probe was always randomly chosen from the digits in the list, excepting the last, and the subject's task was to report the digit that had followed the probe in the list; vocal RT was measured. The salient features of the data were these:

1) Performance was slow compared to IRn. This is perhaps unsurprising in that the nature of the response and the amount of response entropy differ across tasks. More importantly,

2) RT was a more or less linear function of set size. But the slope of the function was very much greater (124 msec. per item) than that reported for IRn (38 msec./item).

3) The pooled serial position functions showed monotonic increases in RT with serial position. This contrasts both with the absence of serial position effects obtained by Sternberg with IRn and more importantly with the marked recency effects found by most researchers (see Section 3.124) using that paradigm. Moreover, the subjects divided into two groups: one showing no serial position effects and the other showing steep, reasonably linear and superimposed pure primacy effects.

4) In contrast to IRn, errors increased very markedly with set

size, such that when $m = 7$, subjects made errors on 25% of trials.

How are these results to be accounted for? Sternberg's ingenious explanation incorporates an account of why scanning should be exhaustive in IRn. The serial position data, it is argued, indicate a self-terminating scan of the list to locate the probe item: some subjects' data show no serial position effects because they start their search at a random point in the list, depending on where they are in their rehearsal cycle when the probe is presented; others start at the beginning. The slope of the set-size function on this account gives us a scanning rate of about 4 items/sec., which is approximately seven times as slow as the rate indicated for the search process proposed for IRn. Why this enormous difference? According to Sternberg, in both tasks the subject scans and compares the probe to the same representations, but the difference in scanning rate derives from the fact that in order to determine whether a comparison has yielded a match, the subject (or an internal 'homunculus') must perform the time-consuming operation of switching to examine the contents of the match register. This operation, or the process of changing to or from it, is very slow compared to comparison. In the IRn task it can be postponed to the end of the list - hence the search is exhaustive; in CR, the requirement to locate the probe in the lists necessitates the operation after every comparison. (See Fig. 3.3 for flow charts.)

Unfortunately for this account, the evidence reviewed in Section 3.1 indicated that the exhaustive scanning model is not an adequate account of IRn performance. If this is indeed the case, Sternberg's explanation of the differences between IRn and CR must also fail. What alternative account can be given? In a footnote in his 1969a paper, Sternberg acknowledges another possibility.

"One interesting alternative (which existing data cannot reject) is that memory representations which can carry order information are different from those that need only carry item information and that

the observed differences in retrieval result from the fact that different kinds of memory representation are being scanned."

(Sternberg, 1969a, p.447)

At this point, it is appropriate to recall that the essential dichotomy that emerged from the short-term forgetting studies reviewed in Ch.2 was also one between tasks requiring the recall of order and tasks involving recognition (and judgements of recency). Apart from the model introduced in Section 1.4, only one other model exists which makes a serious attempt to encompass performance in both kinds of task - Wickelgren's (1972) multitrace associative theory. But the latter cannot account for Sternberg's CR data. According to Wickelgren, access to the probe item's representation should be direct, and the time taken to retrieve the response, once the probe item is accessed, should depend on the strength of the association linking the two. Since the strength of the association should be greater the more recently the items were presented, the model cannot predict the pure primacy serial position effects obtained, unless very bizarre assumptions concerning rehearsal are made. (Any such assumptions will prove inadequate when we consider data from experiments in which rehearsal is prevented - see Exp. I, Ch.5) It therefore seems that we must turn again to the logogen system/response buffer model introduced in Section 1.4.

3.32. The hypothesis of two forms of post-categorical active memory.

The reader is referred to Section 1.4 for a detailed discussion of the model, and to Fig. 1.1, which depicts the relevant structural components. In the present context, the salient features may briefly be summarised as follows. When a sub-span sequence is presented to the senses, each item is represented in active memory, subsequent to

identification, in both of two labile forms. The first is as a decaying trace at a permanent amodal representation or logogen specific to that item. This trace is assumed to have the dynamic properties enshrined in the trace strength inspection model for IRn described in Section 3.11. It is assumed that when a probe item is presented, the trace strength at the logogen addressed by it may be compared against an adjustable criterion; this process is assumed to mediate IRn. Secondly, each item is represented, in order, as one of a string of articulatorily coded items in the response buffer. In order to locate an item given as a probe, in the response buffer, its contents must be recirculated, in the order in which they were stored, back to the logogen system until the logogen of the searched-for item is reactivated. It is proposed that this process mediates retrieval of the probed item in CR, and that the response is then available as the next item read from the response buffer. Three further properties of the response buffer are relevant. Firstly, the search process, which is the same as sub-vocal rehearsal, is slow. Landauer (1962) asked 'highly motivated and reliable subjects' to say to themselves subvocally various well-known sets of items as fast as they could, which turned out to be approx. 4-5 digits a second, 8-9 letters a second. Under normal circumstances, with unpredictable and unlearned sequences we would expect a slower rate. Secondly, the subject may be able to control the point in the sequence at which he starts the read-out sequence. Thirdly, it will be recalled that there was evidence to suggest that rate of input to the response buffer was limited, compared to that at which logogens could be activated (Sections 2.11, 2.12, 2.34).

Now, it is important that this model makes the claim that if a

short list of items is presented at an appropriate rate and attended to, it will be represented simultaneously in both forms of storage. However, because of the differences in their nature and manner of access, the two stores will differ with respect to their utility in different tasks. For empirical purposes the two-store hypothesis really has two parts:

a) There exist two forms of immediate memory representation.

b) Performance in Task A is mediated by one form of storage while the other is the basis of performance in Task B.

Even if a) is true, it can only find empirical support if there are two tasks A and B such that b) for the most part holds. It is my contention that the v-s IRn paradigm and the CR paradigm will adequately occupy these roles. At the same time it is probably impossible to find two tasks that are in this sense 'pure'. While a strategy of trace strength inspection is, according to the model, the quickest and most efficient way of performing the IRn task, since the list is also represented in the response buffer there is nothing save time pressure to prevent the subject from searching the response buffer as well, in order to check his response - indeed this may pay dividends in accuracy. (Such a strategy would constitute a form of the mixed model - see Section 3.11 - Model V.) Similarly while traces at logogens are inappropriate as a primary source of information in CR because they do not code order directly, the subject may be able to make some use of them. For instance, the strength (i.e. recency) of the probe item may tell the subject what part of the list to search; or, the subject may be able to set a criterion such that if the probe exceeds it in strength he responds with the last item made available by the logogen system, which would remove the need to store the final item in the response buffer. Evidence for such mixed

strategies will be mentioned in later chapters.

In the next section I shall indicate some of the points at which comparison of IRn and CR performance may be fruitful, and discuss existing evidence.

3.33. Evidence from the CR paradigm.

3.331. Serial position effects.

Sternberg's (1967a) finding of more or less linear increases in RT with serial position probed, which provided the starting point for this discussion, has in general found confirmation. Chi & Chase (1972) using a fast presentation rate but self-paced probe delay found RT to be a linear though messy function of set-size, and obtained serial position effects; RT increased linearly with serial position. Their data imply a scanning rate of about 6 items/sec.. Derosa & Baumgarte (1971) in a CR experiment in which lists of 5, 6 or 7 digits were presented at a rate of 2 items/sec. with a 2 sec. probe delay found strong primacy effects; some equivocal recency effects were also obtained. Ideally, however, one would prefer rehearsal to be more strictly controlled. Data collected under conditions likely to prevent rehearsal will be presented in due course (EXP I). It may be of interest that serial position curves exhibiting marked primacy are also obtained when similar small memory sets are in inactive memory (Kennedy, 1968), and Sternberg (1969a, Exp.7) has shown that degree of learning does not affect the set-size function, but does cut down errors. These results could imply either that well-learned lists were, in these experiments, maintained in the response buffer by rehearsal, or that the process of reading an ordered sequence from inactive memory involves processes similar to those

involved in reading the contents of the response buffer.

3.332. Format of the representation.

I discussed in Section 3.13 some evidence for the claim that the representations underlying IRn performance are amodal and abstract in character; I pointed out that there were certain shortcomings in the design of experiments on the effects of acoustic similarity on IRn and that in any event the evidence for such effects was equivocal. Chi & Chase (1972) have recently presented data from a CR experiment comparable in some respects to the IRn experiments of Chase & Calfee (1969): their data show a marked decrease in inferred rate of search when the members of the memory set were acoustically similar to one another. However, this study is also susceptible to all three of the possible objections raised in Section 3.132 (ii). Nevertheless, acoustic similarity appears to be implicated as a possible factor for dissociating our two stores. In Ch.5 I shall present a set of experiments on the effects of acoustic similarity on performance in the tasks IRn, CR and Ln, with experimental parameters otherwise matched, controlled rehearsal, visual presentation, and a design intended to control for other differences between vocabularies. There is one feature of Chi & Chase's data which is puzzling: they found faster inferred search rates (smaller slopes) when the list and probe were presented in the same modality than when they were not. They suggest that some sensory matching occurs, but this is difficult to accept in view of the lack of any interaction between the effects of presentation modality and of acoustic similarity, since acoustic sensory matching should be adversely affected by acoustic similarity.

3.333. The effects of information on where and what.

According to the model outlined above, in IRn logogens are accessed directly from the probe, whereas in the CR task the response buffer must be scanned to locate the probe. It was proposed that the subject can vary the starting point of his scan of the response buffer's contents; this was prompted by the simple idea that if one is searching a store ordered in terms of temporal sequence rather than content, it might help to be told where in the sequence the probe item had occurred, whereas one would expect no such facilitation in the IRn task. The validity of this idea is explored in two experiments (Exps. IV & V) described in Ch.6, which compare the utility of giving position information along with the probe in the two paradigms. There already exists a little evidence that in CR the subject can vary the starting point of his search. Some of Sternberg's (1967a) subjects appeared to vary their starting point, while others always started at the beginning. Derosa & Baumgarte (1971) introduced pauses into half the memory sets in their CR experiment, splitting them into two groups. Though grouping was not beneficial overall, there was evidence for serial position effects within groups in that RT was shortest for the first member of a subgroup, longest for the last. But both these findings could depend on how the probe onset interrupted the rehearsal cycle, and do not require the assumption that the subject need not start the search with the first item in the response buffer.

If telling the subject where the item was in the sequence does not facilitate IRn, then since logogens are held to be lexical representations, semantically associated, perhaps it may help to tell the subject what the item is: for instance, what semantic category it belongs to. Exps. VII, VIII are relevant to this possibility.

3.334. Rates of presentation.

It was suggested that the rate of readout from logogens into the response buffer is limited. This raises the possibility of separating the two forms of storage with the variable of presentation rate. In Ch.9 some preliminary experiments will be mentioned which suggest that the limiting rate of presentation for adequate performance on the CR task is lower than that for IRn. A traditional line of argument on this point would run: the subject must in the CR task transmit more information than is required by IRn; thus he must take longer to encode it to reach the same level of performance. This is not good enough. It remains a real empirical question as to whether the limiting rates do differ, and if they do, then it remains an important theoretical question why it should take longer to encode the order information; at least some single store models should predict no difference, whereas here we consider the strong answer that the subject must employ an additional storage system whose input rate is limited.

3.335. Limiting set size.

The available evidence on IRn showed no sharp limit to capacity (Section 3.123), which fact suited a trace strength model. The two-store model outlined would imply that CR performance should, in contrast, reflect the limited capacity (5-6 items?) of the response buffer. Now Sternberg (1967a) found an accelerating increase in errors up to 25% and Chi & Chase (1972) up to 13%, for 7 item lists. Unfortunately longer lists have not yet been explored, though we may note that Waugh & Norman (1965) in an unspedded probed recall task found a very sharp fall-off in performance as lag increased such that the success rate had reached a low asymptote by the time

the lag was 10 or 11 items (cf. Wingfield & Branca's ,1970 ,IRn data). It is of course the case that since the CR task requires the retention of an ordered pair where IRn requires a single item to be remembered, it can be argued that IRn will naturally be more accurate at the same overall level of retention. So the absolute difference in accuracy is indeed not very exciting; what is important is whether the error rates change in the same manner with increasing set size, and though the available data lend some support to the suggested difference, appropriate experiments have not yet been done.

3.34. Evidence from other paradigms.

So far I have concentrated on the ways in which performance in the CR paradigm differs from that in IRn. The requirement for order retention is of course not the only feature distinguishing the two tasks (see Sternberg, 1969a, p.452). One would like to be able to show that such features as vocal response, the increase in number of alternative responses with list length, etc., are not critical. This requires consideration of other paradigms.

3.341. Order recognition.

Sternberg (1969a, Exp.8) performed an experiment whose parameters and subjects were identical to his CR experiment, but in which subjects were given as probes pairs of digits that had been adjacent in the memory set and required to indicate whether their left-right order matched their temporal order in the list. In spite of the possible availability of a range of strategies, the set-size function was linear and identical in slope to that obtained for CR. Moreover, the serial position curves show an overall increase in RT with serial position.

3.342. Location (Ln).

In this paradigm, the subject is required to respond with the position of the probe item in the list. This task makes demands slightly different to CR, in that while the latter requires the subject to retain the relative position of two items, Ln requires the retention of the absolute position of a single item. Were position retained by item-position associations while order was coded as inter-item associations (cf. Wickelgren, 1972) we might expect differences between CR & Ln, whereas if the above model holds, both tasks may be accomplished by searching the response buffer. Note, however, that the subject may also be able to make some use of a trace strength strategy in this task: this depends on the extent to which he is able accurately to classify the strength of the unidimensional trace addressed by the probe with respect to a number of criterion values, (i.e. he might be able to infer position from recency). On the assumption that this faculty is fairly limited, so that only the last few items of the list, whose strengths will be widely separated, may be classified in this way, we might expect to find evidence for a mixture of strategies: earlier items being responded to on the basis of a response buffer search, later ones being categorised on the basis of trace strength. Fuchs (1971) obtained bowed serial position effects peaking at position three for 5-item lists. Moss & Sharac (1970), using 7-item lists, show an increase in RT with serial position up to positions four or five and then a sharp recency effect. These authors do in fact argue for a dual retrieval process (but not the duality suggested here). An experiment of my own (Exp.3, Ch.5) investigates the extent to which the acoustic similarity effects found for CR (Exp.1) are also obtained in the Ln paradigm, and provides serial position data.

3.343. 'Metered memory search.'

Weber and associates (Weber, Cross & Carlton, 1968; Weber & Castleman, 1969; Weber & Blagowsky, 1970) have worked with a paradigm in which the subject retains in memory a 'circular' sequence of digits or letters. On presentation of a probe from the sequence, his task is to respond with the item a specified number of steps forward in the list. The sequence is said to be circular because the first item is to be treated as following the last. RT increases with the size of the 'transformation' and the inferred scanning rate is found to be between 3 and 5 items/sec., which is entirely compatible with Landauer's (1962) figures. The scanning rate is unaffected by whether the vocalization is covert or overt, but is affected by the order of the items: i.e. arbitrarily-ordered digit lists are scanned slower than lists in reverse numerical order while scanning is fastest for naturally-ordered lists. These results are entirely compatible with the proposition that the limit on scanning rate is the maximum rate of covert articulation.

3.344. Retrieval time in forwards and backwards recall.

I have suggested that retrieval from the response buffer is unidirectional. An assumption of this sort has been fairly general with respect to ordered recall from STM (Broadbent, 1958; Conrad, 1965; Wickelgren, 1966a); Conrad in particular argues that backward recall requires the subject to retrieve each item by scanning through the list again in a forward direction. Anders & Lillyquist (1971) gave subjects lists of digits that approximated the memory span in length and timed their responses while they recalled the lists either forwards or backwards as fast as they could. For forward recall the rate of report was about 3 items/sec., which is well within the range

with which we have become familiar in this section; backward recall (average list length = 5-6 digits) was about half as fast. Given that subjects reported retrieving items in groups of two, this is well in line with Conrad's suggestion.

3.35 Summary of Section 3.3.

In this section I have considered some 'memory-scanning' paradigms other than recognition. These share the feature that they all require subjects to retrieve information about the absolute or relative position of items in the memory set. Such information may not, it appears, be retrieved from the representations which in Section 3.1 were suggested to mediate immediate recognition. Accordingly, an additional form of representation in active memory was proposed: a limited capacity 'response buffer' which contains the names of items, articulatorily coded, in order of entry, and must be serially searched for the probe item unless a position cue is given. The two-store model accounts for the markedly different serial position effects obtained in the IRn and CR paradigms, and unlike Sternberg's two scanning strategies model, predicts that other variables will differentially influence performance in the two tasks: it is suggested that these should include acoustic similarity, the utility of position or semantic cues and rate of presentation. Some existing evidence from CR and other paradigms was mentioned; my own will be presented in later chapters.

3.4. Is this active memory?

Throughout this chapter, the assumption has been made that, at least in circumstances in which a new memory set is presented on every trial, performance is based on active memory representations. The differences postulated between different forms of storage have been differences within active memory. Inactive memory is only involved insofar as item-traces are held to be temporary states of activation at the permanent representations, logogens, which mediate identification of the items. It is worth mentioning briefly some of the evidence upon which this assumption is based. There are really two questions to consider.

1) Does the memory set as such get into inactive memory? For subjects to show perfect or even good retention over the long term of verbal material after a single presentation generally requires at least one of the following conditions: the material must be highly meaningful or salient, the list must be the first or only one presented, or the subject must have time to elaborate imagery or other mnemonics (Bower, 1970; Paivio, 1971). To take these in order, in many of the studies I have reviewed, and in all the experiments I shall present, alphanumeric characters have been used; secondly, many lists are presented in brisk succession, usually chosen from small and well-known stimulus ensembles, a procedure that has generally been held to minimise the contribution of inactive memory to immediate memory performance, owing to the rapid build-up of proactive inhibition (Keppel & Underwood, 1962); thirdly the studies I shall describe all used presentation rates in no case slower than 1 item/sec., which should greatly restrict the time available for the discovery of mnemonically useful patterns or associations. In

addition to these three points, there are data which indicate that if subjects know that they are going to be able to retrieve information from active memory, they do not bother to code it into inactive memory. A 'negative recency effect' has been demonstrated both for free recall (Craik, 1970) and for paired-associates learning (Madigan & McCabe, 1971): when subjects are given an unexpected delayed test on material learned for immediate test, they show very poor long-term retention of the last few items presented, which is of course the converse of the result found for immediate recall.

2) Even if the list, or part of it, is securely represented in inactive memory, does this representation mediate performance? I have already mentioned evidence (Sternberg, 1969a, Exp.5) which appears to demonstrate that even when it is certain that the memory set is in inactive memory, subjects prefer, if they are able, to employ an active memory strategy. Wattenbarger & Pachella (1972) used a v-s IRn paradigm, with set-size varying between 1 and 5 items, and on half the trials, instead of a probe they presented a stimulus for binary classification. Classification RT was unaffected by memory set size, and they argue that this implies that IRn performance is mediated by 'STM' since in those conditions in comparable studies where the memory span is exceeded (Shulman & Greeneberg, 1971) or is expected to be exceeded (Sanders & Van Borselen, 1966) a concurrent memory load does affect CRT. Waugh (1969, 1970) showed that in both a paired-associate task and a task using a form of probed recall, latencies for items retrieved from primary memory (as operationally defined on the basis of Waugh & Norman, 1965) were much shorter than for secondary memory items, and the data implied that PM was searched before SM.

In conclusion, it seems safe for the present to conclude that

these experiments do tap active memory in a relatively pure way.

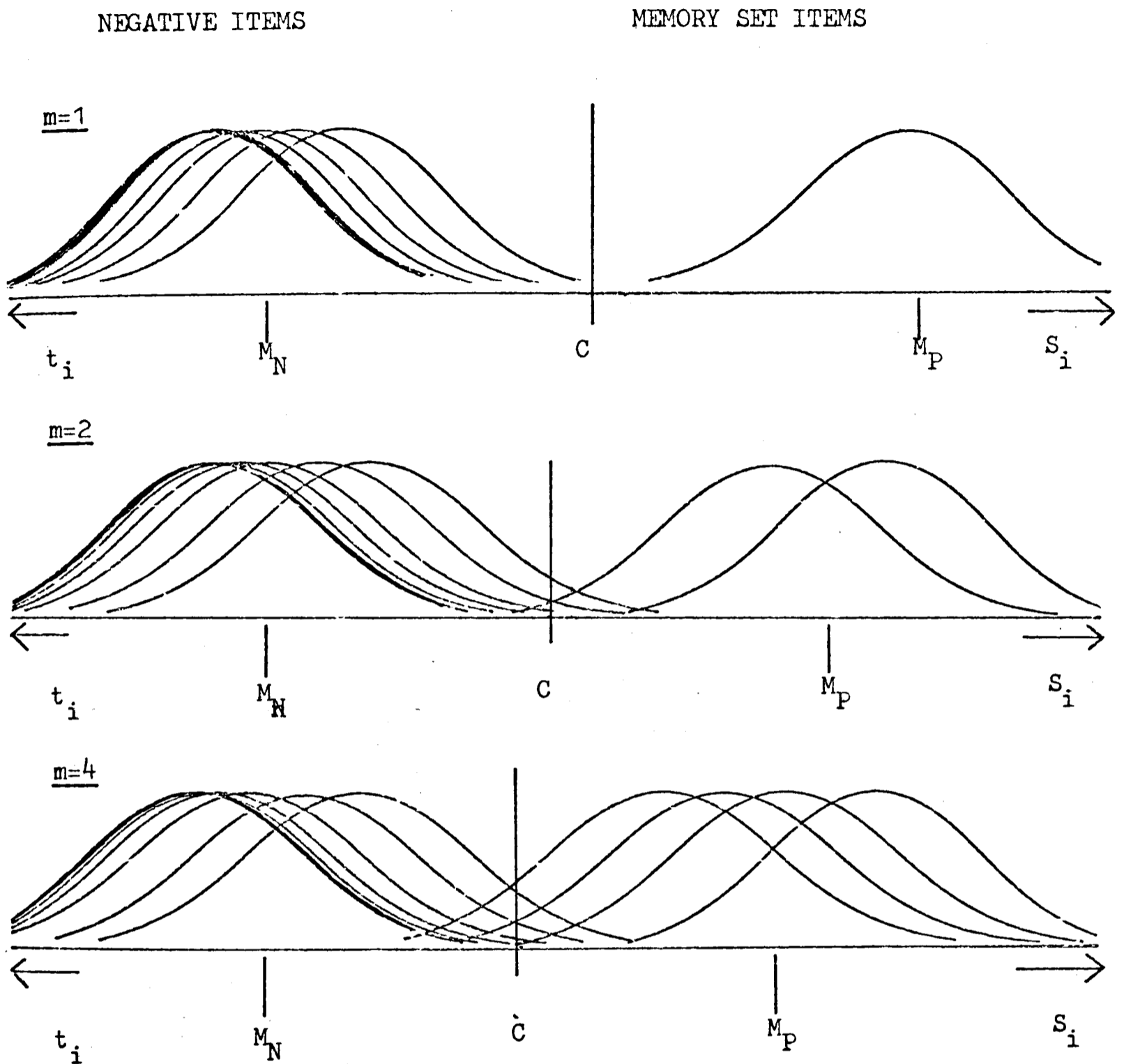
Ideally experiments should be done to show that with extended periods of distracting activity between list and probe, performance in the 'memory-scanning' paradigms does break down in the same way as in those short-term forgetting paradigms for which the concepts of short-term or primary memory were originally introduced.

3.5. Conclusions.

Two themes have dominated this chapter: the possible advantages of the concepts of trace strength and of amodal content-addressable representations in accounting for immediate speeded recognition, and in some cases, perhaps, for constant-set recognition; and the introduction of the hypothesis of two forms of active memory to account for the empirical discrepancies emerging between IRn and comparable 'memory-scanning' paradigms requiring the retention of order information. Two lines of empirical attack are thus implied. Much evidence on IRn is already available but experiments are still needed to test precise and crucial predictions distinguishing the various models. In contrast, we have little evidence on how performances in the different paradigms compare, and experiments must be carried out to identify factors which will produce a reliable and gross dissociation between patterns of performance in the two classes of paradigm when all other relevant parameters are identical.

As I have discussed them, the two themes are clearly related. But they are logically separable. For, we may find differences between the paradigms irreconcilable within a single store model without any implications for the validity of a trace strength model of recognition. Equally, it might conceivably prove possible, though I doubt it, to account for, say, contextual recall data wholly in terms of some strategy of retrieval from a memory having the form of labile traces in logogens. Nevertheless, it is obvious that some experiments will bear dual implications: an experiment which compares recognition and recall may provide evidence which relates to the comparison between specific models of recognition. In Chs. 5 and 6, experiments will be described which were performed to compare the

effects of, respectively, acoustic similarity and a position cue in some memory scanning paradigms. Of these experiments, three (II, IV and V) employ the IRn paradigm, and yield some data relevant to the models outlined in Section 3.1. More direct tests of those models are provided by Exp.VI in Ch.7. In Ch.8, I describe some experiments (VII and VIII) which were carried out to examine the utility of specifying the semantic class of the probe in IRn. But some attention must first be given to methodological issues involved in the use of reaction times. This, together with a general description of the apparatus and procedure used in the experiments, is undertaken in the next chapter.



Schematic illustration of the trace strength models. showing the states of the representations $L_1 \dots L_n$ of the stimulus ensemble items subsequent to presentation of memory sets containing 1, 2 or 4 items respectively.

For Model III, the distributions are those of the trace strength (S_i) at each representation L_i . C shows the optimal placing of the criterion C midway between the mean (M_P) of the compound memory set distribution and (M_N) of the negative set distribution.

For Model IV, the distributions are those of the time taken (t_i) for the representation L_i to reach a critical level of activation K on presentation of the probe P_i .

(Given the distributions shown above, the optimal placing of the criterion T_c would be very similar to that of C.)

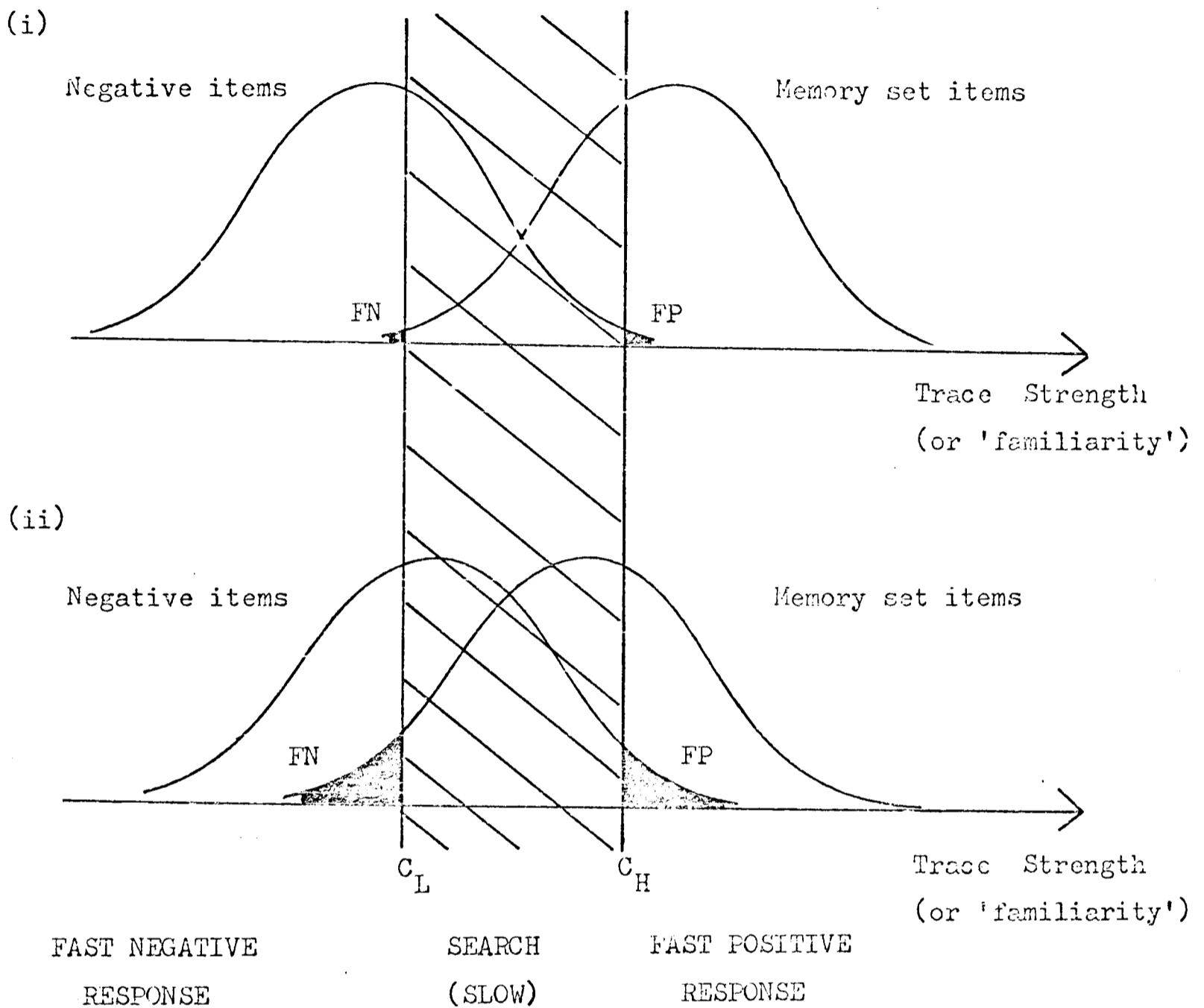
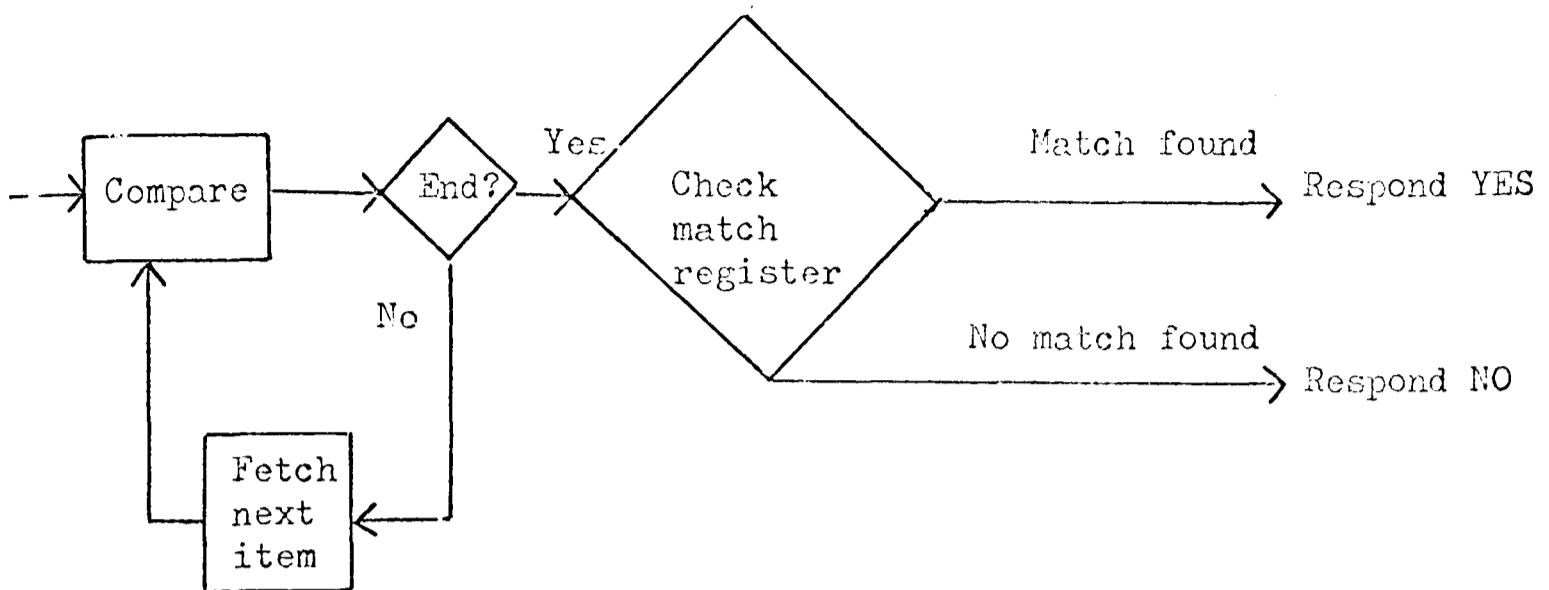


Figure 3.2. A mixed model for IRn (after the long-term recognition model of Juola et al., 1971). The two distributions shown are the compound strength distributions of the memory set items and the items which constitute the negative set on that trial. Comparison between (i) and (ii) shows that the less recent a positive probe (or the more recent a negative probe) the more often a search is necessary and the more false negative - FN (or false positive - FP) responses are made.

EXHAUSTIVE SCANNING: (in item recognition)



SELF-TERMINATING SCANNING: (in contextual recall)

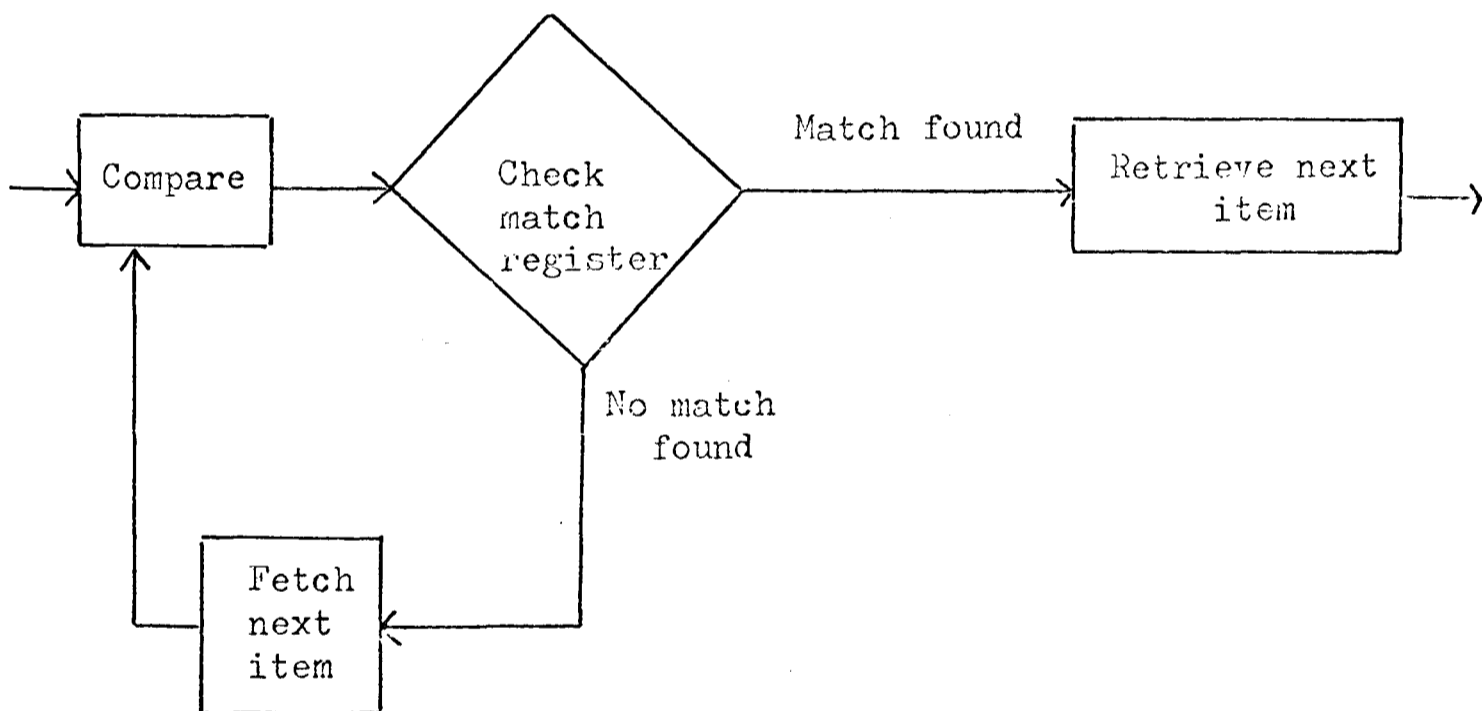


Figure 3.3 Flow charts to illustrate Sternberg's account of the differences between immediate recognition and contextual recall. The stage 'check the match register', or the process of switching from or to it is held to be extremely slow in comparison to the time taken by the 'compare' stage.

PART II. The experiments.

CHAPTER IV. Methodology and Method.

Introduction.

4.1 The use of reaction times.

4.11 RT distributions.

4.12 Speed/accuracy trade-off.

4.13 Reaction time and strength.

4.14 Individual variability.

4.2 The Experiments: Apparatus and General Procedure.

4.21 Subjects.

4.22 Apparatus.

4.23 Display.

4.24 Procedure.

4.25 Trials.

4.26 Timing.

CHAPTER IV. Methodology and Method.

In this chapter, as a prelude to the experimental work to follow, I shall consider briefly some of the advantages and difficulties involved in the use of reaction times as a measure, and shall then give an account of those features of the apparatus and procedure which are general to all the experiments to be described.

4.1. The use of reaction times.

There are three principal reasons for my choice of reaction times as a measure:

1) RT is the only measure we can use when performance is error-free, or nearly so. We have already seen (Section 2.2) that the assumptions of trace strength theory may not adequately be tested for very recent items using accuracy as a measure, since the ratio of correct to false positive responses is too great to obtain a reliable measure of d' .

2) RT is a much more appropriate measure than accuracy if one is concerned to test hypotheses about the nature of access to a memory - the nature of the retrieval process. Consider a hypothetical store containing a number of locations from which information is lost, in the absence of rehearsal, over a period of several seconds. Let us assume that the locations represent positions in a list, and that the subject has direct access to the item at a location given the position as a retrieval cue. Assume also that the store is not content-addressable, so that if the subject must say what position a given item occupied in the list, he is obliged to search serially through the contents of the store. Suppose that search rate is of the order

of 10 items/sec. and that six-item lists are given. Then the mean difference in RT between the two types of probe, other things being equal, will be about 300 msec.. The point of this example is that in 300 msec., relatively little forgetting is likely to occur, so that any difference in accuracy due to probe type will be small; in contrast a difference in RT of 300 msec. would constitute powerful evidence for a difference in mode of access between the two kinds of probe. It was for this reason that I mentioned in discussing some of Hitch's (1972) experiments (Sections 2.13, 2.22) that in the context of the questions they were designed to investigate, RT might be a more appropriate measure than accuracy.

3) RT also has marked advantages over accuracy if one's aim is to disentangle the properties of two memory stores with overlapping retention periods. This is best shown by another hypothetical example. Suppose that an item is initially represented in both of two stores A and B. Let access to A be rapid and access to B slow. If we either present lists longer than the capacity of either store, or present shorter lists and induce forgetting with an interpolated distracting task, then the accuracy of unspeeded recall will depend on whether the subject can retrieve the item from either store. Hence p (correct recall) will be a compound function of the factors influencing both forms of storage. If, on the other hand, we present lists within the capacity of both A and B, ask the subject to respond as fast as he can and measure RT, then since access to A is fast and all the items are represented in A, performance will be mediated purely or at least primarily by retrieval from A and RT will be a function only of the factors influencing store A.

There are at least two real-life applications of this argument. Firstly, it is widely agreed that in immediate unspeeded recall, there

is usually a 'secondary memory component' (Waugh & Norman, 1965). However, in Section 3.4, I presented some arguments favouring the idea that the memory-scanning paradigms were likely to tap only active memory. Secondly, we saw in Section 2.12 that when the accuracy of, say, item recognition was compared with that of probed recall in respect of the effects of factors like acoustic similarity, the evidence for our two-store hypothesis was not very clear. I shall hope to demonstrate, by the experiments reported in the next chapter, that when RT is used, a much clearer dissociation between the two forms of storage emerges.

The use of RT to test a hypothesis that information is represented simultaneously in two stores, or that access to a given store may be via two processes, must be distinguished clearly from the use of RTs to separate functional 'processing stages' - as embodied in the 'additive-factor' method advocated by Sternberg (1969b). He assumes, as do I, that the time elapsing between stimulus and response can be divided into a number of successive 'processing stages' whose durations are additive components of the total RT. Sternberg proposes that the stages involved in performance of a particular task may be discovered by searching for experimental factors whose effects on mean RT do not interact; such factors must influence independent stages. It should be noted that the 'additive-factor' method is for the discovery of successive stages in a single task. In contrast, our concern will be to compare effects on RT, due to a given factor, between tasks, or interactions which may be interpreted as being due to alternative processes involved in a single task. As Sternberg (1969b, p.309) points out, "the additive-factor method cannot distinguish processes, only processing stages". If, for instance, we were

to treat 'tasks' as a factor, then an interaction between tasks and some other factor would lead only to the conclusion that they influenced the same processing stage. We shall try to interpret such interactions, so as to discover the processes occurring within what the additive-factor method is obliged to term a single processing stage. The general strategy will be to try and isolate the feature distinguishing the two tasks that is critical to the effect of the 'tasks' factor.

The use of reaction times may have considerable advantages over the use of accuracy, but there are problems consequent upon it, part theoretical, part practical, to which we must seek at least ad hoc solutions in order to know what kind of experiments to design. We will consider these problems under four headings.

4.11. RT distributions.

The acceptance of a general additive model of reaction time entails that a difference in RT induced by varying the level of some experimental factor is a difference in the amount of real-time processing required. A difference in RT of 50 msec. involves the same amount of extra processing after 1000 msec. as after 400 msec. of previous processing. Since the underlying scale is thus already specified by the nature of the measure, a logarithmic or a reciprocal transformation will not be appropriate. As a measure of central tendency, the arithmetic (rather than the harmonic or geometric) mean should be used; the median should not, since medians are not additive (Sternberg, 1969b).

However, empirical RT distributions are typically positively skewed; that is, they exhibit a tail of long RTs. This raises two

problems:

(i) Since we want to use analysis of variance as a statistical tool, we must be sure that the data conform to its assumptions. Among these is the assumption of normality of distributions. However, this need not be a problem if instead of raw RTs we use as observations the means of small samples of RTs; this has the additional advantage that errors need not give rise to empty cells in the data matrix. These observations will be distributed as the sampling distribution of means from a non-normal population distribution: the Central Limit Theorem assures us that the sampling distribution very rapidly approaches normality as the size of the sample increases (Hays, 1963). Another assumption of the ANOVA is that of homogeneity of error variance: unfortunately, there is some tendency for RT means and variances to be positively correlated. However, it appears to be current statistical opinion (e.g. Hays, 1963) that provided one is testing differences between means, and provided that each cell contains an equal number of observations, both the normality and the homogeneity-of-variance assumptions may be violated without serious risk.

(ii) The second problem arises out of the question: what do these long RTs constituting the tail of the distribution represent? This is a theoretical nettle which no one has yet adequately grasped, but has a practical consequence that in the small sample of RTs typically obtained for a given treatment combination and subject, one or two long RTs may exert a disproportionate influence on the mean, thus making it a less unbiased and efficient estimator. Two contrasting attitudes to this problem might be adopted:

a) It could be assumed that the occasional long RTs do not derive from the same underlying distribution as the rest, but result from

subjects hesitating, or rechecking their decisions, or 'response blocking', etc.. Thus the skewed distribution may be compounded from a symmetrical distribution representing the variability in 'pure' RTs, and a second very broad distribution resulting from the insertion of extra processes like rechecking. On these grounds we should try to exclude long RTs from the main RT analysis, but could treat them as a special class of error; the difficulty is: what exclusion principle should one use?

b) It may, on the other hand be thought that (a) involves an unacceptably strong assumption on which to base the discarding of data; also there is no obvious criterion to choose. Furthermore, even if long RTs do result from rechecking or other 'extra' processes, it may still be argued that they should be represented in the mean RT. The adverse effect of an experimental factor on, say, retrieval may be due either to a change in the mean of the 'pure' RT distribution, or to an increase in the number of times a subject must perform some rechecking operation. To attempt to discard RTs stemming from the latter source is to prejudge this issue.

In analysing the data from the experiments to follow there are usually insufficient data from any given treatment-combination x subject cell to specify the distribution adequately. In general, RTs above a high ceiling (usually 2 sec. for recognition, 3 sec. for recall) are excluded. In a few experiments RTs more than three standard deviations from the mean are excluded. In practice, I have found that using such an exclusion principle has little effect upon the overall pattern of results, save that it reduces overall mean RT by about 10 msec..

4.12. Speed/accuracy trade-off.

It is well known that, other things being equal, subjects may trade accuracy against speed in choice reaction time tasks. This may be due to the subject varying either the amount of 'evidence' he accumulates before making a response, or the proportion of trials on which he makes fast guesses (Yellott, 1967; Swensson & Edwards, 1971). It is general practice to guide subjects toward an optimal balance between speed and accuracy with admonishments like: 'Please try and respond as fast as possible while avoiding errors'. But one must remain alive to the danger that comparisons of RT between different conditions may be artefactually exaggerated or diminished by differences in speed/accuracy trade-off. In general, in making comparisons for which such a possibility exists, we should trust differences in RT only if there are also error frequency differences in the same direction; in all analyses, I shall therefore report the percentage of trials of a given type on which errors were made, together with the RTs.

One conclusion that does not follow from the existence of the speed/error trade-off is that: for a true comparison of RTs between two conditions, error rates should be equalised. Consider, for instance, the set-size function for IRn in relation to Model III, the trace strength inspection model (Section 3.11). As set size is increased, the differences between the means of the list items' and non-list items' strength distributions and the criterion decrease, so that RT is increased; also, the two distributions overlap more, so that error frequency should also increase, even though the subject remains just as cautious. The point is, for many comparisons, error rates should increase along with mean RTs.

4.13. Reaction time and strength.

Murdock (1963b, 1968b) has called into question the use of RT as a measure of trace strength in memory. Using Murdock's (1963a) probe technique with paired associates, he had found that with an unlimited recall interval, response latency covaried very nicely with d' - i.e. the most recent items had the shortest latencies. However, he found that curtailing the response interval permitted had a greater effect on recall of the last pair than earlier pairs. This suggested to him the possibility that latency reflected not how much time the subject must take to respond, but how much he chooses to take; response latency therefore reflected the placing of some sort of decision criterion rather than strength. In a later, somewhat inconclusive experiment (Murdock, 1968b) he found no support for the idea that latency reflected d' .

However, Murdock's arguments need not trouble us greatly in relation to the present experiments. Firstly, his (1963b) subjects were not asked to respond as fast as they could, and the response latencies were correspondingly of a rather different order to ours. Secondly, and more importantly, the trace strength theories outlined in Section 3.1 posit that RT reflects not absolute strength, but the difference between the strength and the criterion; RT changes reflect either changes in criterion or changes in strength; by and large, criterion changes will only be expected to occur when the subject knows what kind of trial he is going to have in advance of the probe onset.

4.14. Individual variability.

That there is considerable variation in absolute RT among

individuals implies that, whenever possible, comparisons should be made within subjects. The use of multi-factorial within-subject experimental designs and the need to collect a number of RTs per subject for each treatment-combination demand, even with on-line control of experiments, a number of experimental sessions per subject for the collection of sufficient data. Moreover, in the case at least of naïve subjects, there is usually a sharp decrease in RT in the initial stages of practice; it is also desirable to allow performance to stabilise, in the sense of eliminating inefficient or ineffective strategies, before collecting data. Hence it is usually necessary to provide several practice blocks (or even sessions).

These considerations lead inevitably to the use of small numbers of subjects. The role of idiosyncratic variations in performance thus becomes rather important. This is a coin with two sides. On the one hand, we should not be prepared to accept as conclusive a difference or trend in the data unless it is shown by all or most of the subjects; this is largely taken care of statistically if we use analysis of variance procedures in which subjects are treated as levels of a random variable in the mixed model, so that effects are tested against their interaction with subjects. On the other hand, individual differences in performance may be a positive advantage. The more sessions served by each subject, the less we should be prepared to interpret differences between one subject's performance and another's as the result purely of chance factors. This is particularly important in the context of hypotheses emphasising the possibility of alternative strategies underlying performance. For instance predictions may often be made on the basis of the premise that in a particular situation, though two strategies may be available (e.g. retrieval from stores A or B), one will be more efficient

and faster and thus may be expected to mediate speeded performance. However, occasional subjects may nevertheless opt for the inefficient strategy. If the hypothesis under consideration is adequate, it should be capable of accounting not only for the way that most of the subjects behave, but also for the manner of any idiosyncratic departures from the norm. I shall therefore report and discuss individual data wherever apparently reliable variations appear.

Another inevitable consequence of the use of small groups of subjects is the sacrifice of generality. It is already commonplace that the typical subject in most human performance experiments is not randomly selected from the population at large, but is a young adult of above-average intelligence. Moreover, constraints on the availability of shared apparatus mean that one tends to use subjects with previous experience of RT experiments because one knows they will not waste time, and one rejects subjects who, through quirk of personality, motivation or dissipation, fail to meet a reasonably strict error criterion in the first session. Thus the demands of methodology and technology combine to impel a trend, retrogressive in some respects, towards the 'elite' subject. It is true that we may set criteria of adequacy that are more operational than, say, Wundt's and assess a subject's consistency by the variability of his RTs rather than by assessing the quality of his introspections; but we must nevertheless be on our guard lest, by training, we have made him in some sense an abnormal subject. On the other hand, we should also take advantage of his skill, and question him on his strategy. It may even be possible to solve the problem of spurious long reaction times by getting subjects to indicate those trials on which they feel that they hesitate or that their performance

is disturbed in some other way. (This was not attempted in the present experiments.)

These methodological issues are in themselves a fit topic for further research. For the present, one is obliged to accept that one's criteria for preferring particular ad hoc solutions will have a sociological as much as a logical basis.

4.2. The Experiments: Apparatus and General Procedure.

All the experiments reported in this thesis were performed on-line; a Digital Equipment Corporation LINC-8 computer was used to present visual stimuli on a cathode-ray tube (CRT) and to time and record both vocal and key-press responses. The sequence of trials and of events within trials was under the control of one of a set of inter-related programmes written by me for the purpose called MNEMON. I shall describe here features general to all the experiments.

4.21. Subjects.

These were young adults, 17 to 30 years old, who had volunteered their services in response to public advertisements or personal pleas. They included secretaries, teachers, graduate and undergraduate students, and were selected without regard to sex. They either had good vision, or wore the corrective spectacles prescribed for them. They were paid 25p (later, 30p) a session. In most experiments, a mixture of naïve and experienced subjects was used. Subjects who exceeded a reasonably strict error criterion were replaced.

4.22. Apparatus.

The subject sat in a 'quiet' testing room adjacent to that containing the computer and its associated teletypewriter. The room was dimly and indirectly illuminated with a heavily shaded 40-watt bulb. The subject rested his face in a rubber mask at eye level, 40 cm. in front of which was the CRT screen, whose dimensions were 11 cm. by 9 cm.. The screen was surrounded by a black cardboard surface which extended to fill most of the subject's visual field, and was so arranged that no distracting reflections were visible in the screen.

On a table beneath the screen a keyboard was arranged so that the subject could rest his hands comfortably upon it. Positioned under each finger was a piano-style key operating a microswitch contact-closure connected to the computer.

For the timing of vocal responses, a microphone was suspended some 5 cm. in front of the subject's mouth and connected to a voice key (Electronics Developments), which was adapted so that it could be reset by the computer, and which operated a contact-closure connected to the computer. For experiments involving vocal responses, the experimenter sat at a table in the testing room in order to record erroneous responses and to indicate via a second keyboard one of four possible outcomes to the computer: a) Correct response, b) Error, c) No response (e.g. 'don't know'), d) Rejected trial. The last category was used when the voice key was either accidentally triggered by the subject, extraneous noise or inherent instability, or failed to be triggered at all.

4.23. Display.

The CRT (TEKTRONIX, 561 A/RM, modified) contains a very fast phosphor (P15) whose rise time is 5 microseconds and whose decay is rated as to 0.1% of its total brightness within 50 microseconds. Points displayed on the screen appear as bright and greenish on a dim, dark green, background. The computer possesses a hardware character display system which permits the efficient display of any subset of a 4 x 6 matrix of dots. The display patterns chosen for the alphanumeric characters are shown in Fig. 4.1. The individual characters are 4 mm. high and 3 mm. wide. They therefore subtended an angle at the eye of 27 minutes of arc horizontally, 36 minutes vertically. The average luminance of the display points is impossible

to control and measure precisely, and varied somewhat throughout the series of experiments, but never within a session; in general, the characters were bright enough to be easy to read without strain.

4.24. Procedure.

An effort was made to arrange that the several sessions required of a subject should be on consecutive days, at a more or less constant time of day. This ideal was rarely achieved, owing to the heavy pressure of other users on the apparatus. At the start of the first session, salient features of the apparatus were demonstrated to the subject, he was seated comfortably and at a proper height, a display of the whole character set was shown to him and the task explained. A few practice trials were usually given to familiarise him with the task and apparatus before the formal practice blocks. Instructions were read to him and questions answered. At the start of subsequent sessions, subjects were reminded of crucial details of the instructions.

Sessions were divided into blocks containing 30-120 trials. At the beginning of a block, the word 'Ready' was displayed until the subject pressed a button to initiate the sequence of trials. During a block, trials followed the responses to previous trials automatically after an interval of 2 or 3 secs. (see individual experiments for precise values). However, in most cases, the subject was able, if he so chose, to interrupt the sequence, between trials, by pressing a 'panic' button. The computer then displayed the word 'Pause' until the subject pressed the same button again, when the sequence of trials was resumed. The subject was instructed to use this facility sparingly, but whenever he was overwhelmed by fatigue, dizziness, fits of sneezing or other temporary disabilities. In practice he

made use of it rarely more than once a block, if that. At the end of a block the word 'End' was displayed, and the subject was free to emerge from the testing room while data were saved on tape and a summary of the mean RT and errors for that block typed out. Sessions usually lasted between 30 and 60 minutes. Subjects were encouraged to report their introspections.

4.25. Trials.

Details of MNEMON, the set of programmes used to compile blocks of trials, run them, record responses and analyse the data are not necessary here. For the present, only the following features are relevant. A 'display' consisted of a character (or group of characters) from a vocabulary of up to 30 characters, displayed in one (or several) of up to 16 predefined locations on the screen, for one of up to 8 predefined durations. A display might or might not expect a response; if it did, it might or might not be terminated by that response. A trial consisted of a sequence of such displays, separated by a preset inter-display interval; a background display of up to 60 points could be programmed to remain on throughout the course of a trial in order to serve as a reference frame for the display locations. In later experiments, a trial might be followed with a feedback display indicating whether the response(s) was(were) correct.

Thus, a typical trial might have the following form:

1) Onset of background display, coupled with a warning signal - an arrow pointing to the location at which the first character is to be displayed.

2) Four or five letters, one after the other (perhaps at different spatial locations) comprising a 'memory set'.

3) A delay - the 'probe delay', consisting of no display.

4) A single probe item, displayed either until 4 sec. had passed or a response was made, after which the display and background vanished.

5) Feedback display: a tick if correct, a cross if wrong, a query if no response.

6) A dark interval of a couple of seconds, followed by resumption of the cycle at 1).

The selection of the items of the memory set, and the sequence of trial types was under the control of a stimulus-generating programme specifically written for each individual experiment. This referred to a powerful set of sub-routines incorporating a pseudo-random generator which enabled the randomisation, control and balancing of several sets of relative frequencies simultaneously. Among the properties of this programme was that the relative frequencies of different trial types were not only perfectly controlled over the whole block, but were also kept roughly constant over sub-sections of the block. This had the consequence of limiting maximum run lengths for a given trial type. The degree of this constraint could be adjusted, but in general it was kept so that a run of more than six trials with the same response (in the case of two-response experiments) was prevented.

4.26. Timing.

Responses were timed by means of a programmable electronic clock, whose resolution was of the order of 1 microsecond. However, owing to storage limitations, MNEMON records RTs to an accuracy of 2 msec. only. RTs of more than 4 sec. were not recorded.

Further details will be given in descriptions of the relevant individual experiments, to which the following four chapters are devoted. The experiments are not presented in the chronological order in which they were performed. The actual sequence was:

VII, VIII, IV, I, II, III, V, VI.

A B C D E F G H I J
K L M N O P Q R S T
U V W X Y Z

0 1 2 3 4 5 6 7 8 9

Figure 4.1. The characters as they appeared on the screen of the CRT.

CHAPTER V. Experiments on the effects of acoustic similarity.

Introduction.

5.1 EXPERIMENT I. The Effects of Acoustic Similarity on Contextual

Recall.

5.11 Method.

5.12 Results.

5.13 Discussion.

5.2 EXPERIMENT II. The Effects of Acoustic Similarity on Item

Recognition.

5.21 Method.

5.22 Results.

5.23 Discussion.

5.3 EXPERIMENT III. The Effects of Acoustic Similarity on Probed

Location.

5.31 Method.

5.32 Results.

5.33 Discussion.

5.4 Conclusions and general discussion.

CHAPTER V. Experiments on the effects of acoustic similarity.

Sternberg (1967a) showed, and later studies have confirmed, that performance in the IRn and CR paradigms is markedly different (Section 3.31). The slope of the set size function is typically very much greater for CR than for IRn; moreover, for the CR task, RT increases more or less monotonically with serial position - a primacy effect (Section 3.331). Sternberg (1967a, 1969) proposed that these differences were due to different strategies of scanning 'active memory': a self-terminating scan to locate the item in CR, an exhaustive (and therefore faster) scan to determine its presence in IRn (Section 3.31). However, evidence against a scanning model for IRn has steadily accumulated (Section 3.1). Also, the serial position functions typically obtained with IRn exhibit marked non-linear recency effects (Section 3.124). I therefore proposed (in Section 3.3) that the differences between IRn and CR were better accounted for by the hypothesis of two forms of active memory outlined in Section 1.4. According to this hypothesis, IRn is mediated by decisions based on the trace-strength at the logogen addressed by the probe item, while CR requires a relatively slow serial search through the temporally-ordered, articulatorily-coded contents of the response buffer. Since items in the response buffer are coded in terms of their articulatory features, one would expect that phonemic similarity between items stored in the response buffer would retard retrieval from it; hence phonemic similarity would be expected to have an adverse effect on CR. No such effect would be expected on IRn, since logogens are abstract, conceptual or amodal representations. In contrast, on Sternberg's hypothesis, since the same store is scanned in both tasks, phonemic similarity should affect both tasks

equally, if at all. I therefore performed three experiments on the effects of acoustic similarity, which I report in this chapter. (I shall refer to acoustic, rather than phonemic or articulatory similarity; in terms of the material used they are effectively synonymous).

Various data already exist suggesting a dissociation between two forms of storage in active memory in terms of the effects of acoustic similarity. It will be recalled that Wickelgren (1965c) showed that acoustic similarity increased order but not item errors (Section 2.11) and that Ingleby (1969) showed that, for small sets of items held in memory, acoustic similarity affected probed location but not item recognition (Section 2.12); however, with longer lists, effects of acoustic similarity on the accuracy of item recognition have been reported (Hitch, 1972; Reicher, Ligon & Conrad, 1969). Turning to reaction time studies; such effects of acoustic similarity on IRn as have been shown are both small and equivocal (Section 3.132 ii), while Chi & Chase (1972) have recently found, in a CR experiment comparable in design to the IRn study of Chase & Calfee (1969) a marked increase in the slope of the set size function when the items were from an acoustically similar vocabulary (see Section 3.332). However, in discussing these experiments I pointed out (in Section 3.132 ii) that their interpretation is rendered difficult by certain problems involved in the use of acoustic similarity as a variable:

1) Perceptual errors. If acoustic presentation is used, then acoustic similarity may increase the occurrence of perceptual errors or the difficulty of auditory encoding, and thus have effects which may be confounded with possible effects on storage or retrieval. Hence visual presentation was used in the present experiments.

2) Rehearsal. According to the model outlined in Section 1.4, rehearsal is the recirculation of items from the response buffer back

to the logogen system. Acoustic similarity of the items could result in more errors during readout from the response buffer, or in a slower rate of rehearsal. Thus if rehearsal is not prevented, acoustic similarity could influence item recognition even though the traces directly mediating recognition judgements are, according to our hypothesis, not acoustic in nature. In the experiments reported here, I hope to have prevented rehearsal by the following means:

(i) No-rehearsal instructions (cf. Waugh & Norman, 1965)

(ii) The use of sub-span lists

(iii) Immediate test.

(iv) The use of fast presentation rates. This requires selection of an appropriate rate - i.e. one that is slow enough for comfortable identification of the items, but fast enough to allow no time for rehearsal. The rate chosen, on the basis of a preliminary experiment on CR, was 2 items/sec..

3) The manipulation of acoustic similarity. Typically, researchers have compared performance on memory lists chosen from a vocabulary of acoustically similar or confusing (C) items with that on lists chosen from acoustically dissimilar or non-confusing (N-C) items. Suppose that we found a difference in RT in such a CR experiment. How would we know that it was due to a difference in acoustic similarity, not to, say, a difference in voicing time, or some unknown feature distinguishing C and N-C items? This may be controlled for, at least as regards within-list similarity, by including both C and N-C items in the same list. Accordingly, in the experiments to be described, acoustic similarity was manipulated in the following manner. The total stimulus ensemble consisted of twelve consonants, half of which shared a vowel phoneme and constituted the C group, half of which were acoustically dissimilar and formed the N-C group.

Each memory list contained five items, four chosen randomly without replacement from one group, one from the other. This resulted in two types of list:

- Acoustically similar (AC SIM), containing four C items and one N-C item,
- .. Acoustically dissimilar (AC DIS), with one C item and four N-C items.

The 'odd man out' in each kind of list I shall call the 'isolated' item or I-item, though only in the AC SIM lists was it acoustically isolated.

This design has various advantages; in addition to looking for any overall effect of acoustic similarity we can look at the interaction between list type and probe type. For instance, does it help to locate the probe item if it is acoustically isolated (the I-item) in an otherwise similar list? (Note that the equivalent I-item in an AC DIS list provides a control for any differences resulting from probabilistic rather than acoustic isolation.)

At the time this series of experiments was begun, no experiment had been published on the effects of acoustic similarity in Sternberg's CR paradigm (a deficit since remedied by Chi & Chase, 1972). In the first experiment, I therefore investigated the effects of acoustic similarity, manipulated as I have described, on contextual recall.

5.1 EXPERIMENT I. The Effects of Acoustic Similarity on Contextual Recall.

This experiment concerned the effects of acoustic similarity on the time taken to recall an item from a five-consonant list in active memory, presented visually and sequentially, given, as a probe, the preceding item. Rehearsal was prevented by instructions, fast presentation rate and immediate test.

On the hypothesis that CR requires the subject to search serially the contents of a phonemically-coded, temporally-ordered store (the response buffer) in order to locate the probe item and thus retrieve the subsequent item, we should expect:

- 1) Longer RTs and more errors with AC SIM than with AC DIS lists.
- 2) An increase in RT with serial position probed.
- 3) An interaction between serial position and list type such that the effect of acoustic similarity increases with serial position, since the decremental effects should be cumulative as successive comparisons are made.
- 4) An interaction of list type and probe type such that:
 - a) access in memory to the probed item in an AC SIM list should be facilitated if either it or an item preceding it in the list is the I-item.
 - b) an I-item should be easier to recall than a non-I item, in an AC SIM list.

Wickelgren's (1972) associationist theory of the representation of order-and the effects of acoustic similarity does not predict 2), 3) and 4a). In an associative memory access to the representation of the probe item should be direct, acoustic similarity should only hinder retrieval of the response.

5.11. Method.

Six subjects each served for six sessions, each lasting about fifty minutes and consisting of two blocks of 84 trials; between the blocks there was a five-minute break. The first two sessions were practice, and data from only the last four sessions are presented here.

A trial: was initiated by the onset of the background display (which was visible throughout the trial). This consisted of six dots arranged as the corners of two horizontally adjacent squares, and served to define two locations about 1° of visual angle apart. The items of the memory set were displayed one after another at the left-hand location, followed by the presentation at the right-hand location of the probe item, which remained visible until a response was made (or 4 sec. had elapsed). The subject received immediate verbal feedback as to whether he was right or wrong, and another trial followed after $2\frac{1}{2}$ secs.. The memory set items were displayed for 400 msec. each, separated by 100 msec. of darkness. The background display preceded them by 600 msec. and served as a warning, and the probe followed 250 msec. after the offset of the last item.

The vocabulary: consisted of twelve consonants. The C-group was B D G P T V, and the N-C group was F H K L R Y. These were chosen primarily on the basis of the common vowel sound of the C-items. An idea of their relative confusability may be obtained as follows. I define the 'confusion ratio' of the vocabulary as: the ratio of the average probability that a C item rather than a N-C item will be substituted for a given C item, to the average probability that a N-C item rather than a C item will be substituted

for a given N-C item. If we abstract for our vocabulary (as Clarke, 1957, says we may) from Conrad's (1964) confusion matrix of listening errors in white noise, the acoustic confusion ratio of our vocabulary is 5.3:1. We can derive a comparable measure of visual similarity. However, visual confusion matrices are dependent to some extent on typeface (Fisher, Monty & Glucksberg, 1969) and unfortunately no confusion matrix is yet available for LINC-8 display characters. The best we can do, given the inadequacy of an overlap measure, is to take the average of the confusion ratios for several different type faces. Using those of Hodge (1962), Pew & Gardner (1965) and Fisher, Monty & Glucksberg (1969), for ordinary upper-case typefaces, we obtain a (geometric) mean visual confusion ratio for our vocabulary of 0.99: 1. Thus on this measure the two groups of items do not differ in visual similarity.

The design: is a within-subject design, and all the possible trial types were presented in a random order in every block. On each trial a memory set was presented containing four items chosen from among the C-items (or N-C items), and one N-C (or C) item. There were two list types (AC SIM, AC DIS), five positions of the isolated item (I-item) and four positions probed (subjects never had to recall the first item). Each block contained two examples of each of the 40 possible trial types, and the block was preceded by four warm-up trials. Selection of items was further constrained only by the requirement that each possible letter occur equally often as a response for each trial type. 24 blocks were generated, each used once in each two sessions of the experiment. All the blocks done by a given subject were different.

The instructions: were, "Do not attempt to rehearse or go back over the items in your head either during or after presentation. At any moment attend only to the item most recently presented; try to treat all the items the same and not to group them. Please try to respond as fast as possible while avoiding errors."

5.12. Results.

Raw data were pooled over blocks 5, 8, 9 and 12 and blocks 6, 7, 10 and 11 to give two observations per trial type per subject, with the contribution of any practice effects to the error variance minimal: each observation was based on a maximum of 8 trials. Overall error rate was 8.5%. and 0.9% of trials were rejected owing to voice-key failures.

Correct reaction time data were submitted to a four-way analysis of variance (mixed model, subjects random); the factors were: List Type, Serial Position (of probe), Position of I-item and Subjects. In all figures, RTs are graphed and error frequencies given as bar histograms.

1) List type and serial position. Fig. 5.1 shows mean RT plotted against Serial Position, separately for AC SIM and AC DIS lists. Overall RT is 69 msec. slower when the list items are acoustically similar than when they are dissimilar. This effect is significant at the $p < 0.025$ level ($F(1,5) = 12.50$), but interacts with Serial Position probed ($p < 0.05$, $F(3,15) = 3.39$). A post-hoc comparison shows the interaction to be due to significant ($p < 0.05$) differences between AC SIM and AC DIS lists at positions 2 and 3, but not at positions 4 and 1. It may be seen that the serial position

probed had a marked effect; RT is increased by about 180 msec. per item over the first three positions. However, the effect of serial position is significant only at the $p < 0.05$ level ($F(3,15) = 3.37$): this is the result of a highly significant interaction with Subjects ($p < 0.001$, $F(15,240) = 64.7$). The subjects in fact fall tidily into two groups: Fig. 5.2 shows the same data as Fig. 5.1, but separately for the two groups. For four subjects RT was fastest when the last position was probed; for the other two, RT at this position was the slowest. Moreover, for the latter pair of subjects the increase in the effect of List Type continued to the last serial position.

2) Position of the isolated item. For the AC SIM lists, it does appear to matter what position was occupied in the list by the isolated item. This may be seen in the List Type \times Position of I-item interaction ($p < 0.01$, $F(4,20) = 7.1$) which is depicted in Fig. 5.3. The effect of acoustic similarity was greatly diminished when the isolated item was at positions 2 and 3. However, this interaction does itself interact at a lower level of significance with serial position probed ($p < 0.025$, $F(12,60) = 2.2$). The three-way interaction is plotted in Fig. 5.5. It may be seen that:

- At serial position (SP) 1, such difference as there was between the AC SIM and AC DIS lists was due entirely to a probably spurious low RT for AC DIS lists when the I item was at position 4.
- At SPs 2, 3, and 4, the effect of acoustic similarity was greatly reduced when either the probe or the response was the I-item.
- At SP 3, the effect was also reduced when the item preceding the probe was the I-item.

Part of the sense of this complex interaction may be captured if we compare the cases where: the probe was the isolated item; the response

was the isolated item; and neither was the isolated item: this is shown in Fig. 5.4. (None of the comparisons plotted here reach significance on a Sheffé' post-hoc multiple comparison, but since some of the points summarise only a small number of observations, significance levels here are very stringent.) It may be seen that the effect of acoustic similarity was due largely to trials in which neither the probe nor the response item were acoustically isolated. It is interesting that it took longer to recall a C item from a dissimilar list than an N-C item. Could this have been due simply to a difference in voicing time between the two classes of item? That there was little difference between list types at serial position 1 and (for four subjects) serial position 4 implies that it could not. Certainly the whole effect of acoustic similarity cannot have been due to a voicing difference, or the interaction in Fig. 5.4 would be entirely symmetrical.

Errors. Error percentages were subjected to an arc-sine transformation and submitted to the same four-way analysis of variance. The main effect of List Type is significant at the $p < 0.001$ level ($F(1,5) = 47.8$). The error rates were 12.6% for AC SIM lists, 4.3% for AC DIS lists. In addition there is a main effect of Position of I-item ($p < 0.05$, $F(4,20) = 3.24$). As may be seen in Fig. 5.3, subjects made considerably fewer errors in acoustically similar lists when the list was interrupted by a dissimilar item than when the dissimilar item was at either end.

Of all errors, only 11% were extra-list intrusions or omissions. For the remaining errors, the source of the substituted item is shown in TABLE 5.1, as a percentage of the total transposition errors for a given serial position of probe. It will be seen that substitution

errors came largely from the positions immediately adjacent to the positions of the correct response and the probe. Also, all other things being equal, an erroneous response was twice as likely to be from the same class as the correct response, as from the other class, regardless of list type.

5.13. Discussion.

In relation to predictions 1, 2 and 3, the data show a clear effect of acoustic similarity on contextual recall RT and errors; the RT difference increased with the serial position probed, save for the terminal position in the case of four subjects; also, for the first three positions at least, RT increased with serial position by about 152 msec. per item for dissimilar lists and 195 msec. per item for similar lists. Thus with the exception of the fast RTs for terminal items produced by some subjects, the data are compatible with the hypothesis that retrieval involves a serial search through a temporally-ordered store, at a rate of between 5 and 7 items/sec. Acoustic similarity appears to increase the time per item by about 30 msec..

When the final item was required as the response, some subjects appeared to retrieve it in the manner just described, while others appeared to have rapid access to it in a manner unaffected by acoustic similarity. How is this rapid access to be explained? I suggest that subjects may have made some use of the 'item-traces' at logogens postulated in the two-stores model: that is, the subject may have judged the strength of the trace addressed by the probe - if it exceeded a criterion he would have known that he was required to recall the last item, which may still have been available from the logogen system; hence no search for this item was necessary.

In connection with prediction 4, retrieval seems to have been facilitated more or less equally when the probe item, and when the response item, were acoustically distinct from the rest of the list, though less errors were made when the response item was isolated. Moreover there is some sign that, at least at serial position 3, acoustically distinct items prior to the position probed facilitated retrieval. So in all major respects the data appear compatible with our hypothesis of a serial search through a phonemically-coded store, and incompatible with an associationist account.

Two features of the data remain unexplained. The first is that there appears to have been an effect of the position of the isolated-item on both RT and errors which was partly independent of the serial position probed, and therefore, presumably, of retrieval factors. Subjects did remark that lists containing four rhyming letters in a row were very difficult: perhaps this evaluation affected their motivation on such trials, or perhaps the effect is due to processes occurring during time in store. The second question is, why did it take longer to recall an item from a dissimilar list when it was one of the confusing group of items? The implication is that the effects of acoustic similarity were not restricted to within a trial, but that there was a more general effect of the similarity of the response (but not the probe item) to the whole vocabulary, over trials.

Having seen how acoustic similarity influences contextual recall, we now turn to consider its effects on item recognition.

5.2. EXPERIMENT II. The Effects of Acoustic Similarity on Item Recognition.

If Sternberg (1967a, 1969) is right, and subjects perform the IRn task by searching the same store as they search in CR, then since search rate is apparently slowed by acoustic similarity in CR, we should be able to demonstrate a comparable effect in IRn.. Indeed, if subjects perform an exhaustive scan for IRn, but a self-terminating scan for CR, then we might expect a greater effect on IRn, since on average more items must be scanned. The slower speed of a self-terminating scan is due, Sternberg suggests, to the necessity for the subject to insert an 'inspect the match register' stage in between each comparison; it is difficult to see how acoustic similarity could influence inspection of the match register rather than the comparison process. According to the two-store hypothesis, however, item recognition is mediated primarily by item-traces at logogens. Provided that presentation is visual and rehearsal prevented, we should expect little or no effect of acoustic similarity. We should also expect that, in contrast to the primacy effect found for CR, the serial position data would exhibit a non-linear recency effect (Section 3.124).

I therefore performed an experiment comparable to the last in all major respects save the following:

- 1) The subject was required to indicate by means of a key-pressing response whether the probe item had or had not been present in the memory set. This required that negative probes be presented. What proportion of these should be of the same class as the majority of the list items? If negative probes were chosen from the two

classes in the same proportion as list items i.e. 4:1, then the composition of the list could lead the subject to expect a particular class of probe; such expectations could influence encoding so as to mask any effect due to acoustic similarity. Negative probes were therefore chosen with a frequency close to that which would have resulted if they had been selected randomly from the remainder of the vocabulary. In fact they were of the same class as the I-item twice as often as they were not. It remains possible that the nature of the probe item could then bias the subject towards a particular response, but at least the effects would be opposite for positive and negative responses, and an acoustic similarity effect could not be masked.

2) The experiment was replicated with two different vocabularies: in Exp. IIa, four subjects were tested with the vocabulary used in Exp. I; in Exp. IIb, four subjects were tested with a different vocabulary.

3) Since key-pressing responses were used in this experiment, we have no guarantee that subjects were identifying the initially unfamiliar computer-style letters correctly. In the two practice sessions, subjects were therefore given practice at overt naming of the letters.

5.21. Method.

Eight subjects each served in six sessions, each lasting about 40 mins.. In each of the first two sessions, which were considered practice, in addition to a block of letter-naming practice subjects received two blocks of 80 trials on the main task. In each of the four experimental sessions, from which the data presented here derive, subjects received three blocks of 80 trials.

A trial: was exactly as in Exp.I, save that the probe terminated after 3 sec. if no response was made, and feedback was visual. Subjects pressed the right index finger key if the probe was a member of the memory set (a positive response), the left index key if it was not (a negative response).

Vocabulary: as mentioned above, two vocabularies were used:

	<u>C-items</u>	<u>N-C items</u>	
In Exp. IIa:	B D G P T V	F H K L R Y	(Vocab 'E')
In Exp. IIb:	F L N S X Z	D G H K Q R	(Vocab 'M')

The acoustic confusion ratio (as defined above) of Vocab 'M', on the basis of listening errors in white noise (Conrad, 1964) is 8.3:1. However, this is probably an overestimate of 'true' phonemic similarity since white noise emphasises sibilant errors. Also the majority of confusion errors with Vocab M is due to fewer pairs than is the case for Vocab E, making it a less suitable vocabulary for present purposes. The visual confusion ratio of Vocab M calculated from the same source as before is 1.08: 1.

Design: Again, all possible trial types were presented in a random order in every block, and the memory lists were composed in the same way as before. Positive trials: there were two list types (AC SIM, AC DIS), five serial positions of the probe and five positions of the I-item - making 50 trial types in all. One example of each occurred in every block.

Negative trials: For each combination of list type and isolate position (of which there are ten), the probe could be either the same or different in class to the I-item. Two examples of the former and one of the latter were presented in every block.

Thus, there were 80 trials in each block, and the ratio of positive to negative trials was 5:3. New blocks were generated for every session.

Instructions: were as in Exp. I, save that it was emphasised to subjects that while 'no rehearsal' did not preclude (silent) naming at presentation, they were to stick to the names on which they had been trained.

5.22. Results.

Raw data were pooled over blocks 5, 7, 9, 12, 14 and 16, and blocks 6, 8, 10, 11, 13 and 15, to yield two observations per trial type per subject; each observation was based on six trials (or, in the case of negative trials where the probe was of the same class as the I-item, twelve trials). Overall error rates were 4.5% in Exp. IIa and 6.6% in Exp. IIb. Data were analysed separately for the two replications, lest the two levels of acoustic similarity were not comparable. In Exp. IIb I excluded from the analysis all trials in which the memory set contained a G and the probe was a D or vice versa; the reason for this will be encountered in the description of Exp. III. For present purposes, it should be noted that since D and G are in fact acoustically confusable but were here part of the N-C class, this exclusion should increase any effects due to acoustic similarity.

Correct reaction time data were submitted to a five-way analysis of variance (mixed model, subjects random): List Type x Probe Type (Same class as I-item or different) x Response Type (Yes or No) x Position of I-item x Subjects.

1) There was no significant main effect of List Type with either vocabulary. In Exp. IIa, RTs were on average 6.8 msec. slower with AC SIM than with AC DIS lists ($F < 1$). In Exp. IIb, RTs were on average 13 msec. faster with AC SIM than with AC DIS lists ($F(1,3) = 1.25$). The 6.8 msec. difference in Exp. IIa was due to one subject, S_4 - and there is an appropriate Subjects x List Type interaction ($p < 0.001$, $F(3,160) = 5.5$).

2) List type does not interact significantly with any fixed variable, in either replication. It is of particular interest to inspect the List Type x Probe Type interaction, insignificant though it is (Exp. IIa: $F < 1$; Exp. IIb: $F(1,3) = 2.99$). This is shown in Fig. 5.6a and 5.6b, both separately for individual subjects and pooled. If recognition were facilitated when the probe was from the same class as the I-item, the interaction should have the form indicated as A in Fig. 5.6a. On the other hand there could be an interaction due merely to one class of item being more difficult to encode than the other, in which case it should have the form indicated as B in Fig. 5.6a. In Exp. IIa only S_4 shows an interaction of the type A. In Exp. IIb, though there are signs of an interaction, for subjects 5 and 8 it looks like B, S_7 fits neither pattern, and only subject 6 shows an A-type pattern. In short, evidence for a List Type x Probe Type interaction compatible with mediation by an acoustic representation is restricted entirely to two subjects, S_4 and S_6 .

3) The analysis also shows a main effect of Response Type (IIa: $p < 0.025$, $F(1,3) = 20.6$; Exp. IIb: $p < 0.01$, $F(1,3) = 66.1$), and significant interactions between Position of I-item, Probe Type and

Response Type, but these are due to a confounding of Position of I-item and serial position probed, which is best analysed more directly.

A four-way analysis of variance was performed with the factors: List Type x Serial Position Probed x Position of I-item x Subjects (random).

4) In both replications, the main effect of serial position is highly significant at the $p < 0.001$ level (IIa: $F(4,12) = 15.1$; IIb: $F(4,12) = 12.6$). Pooled serial position curves are shown in Fig. 5.7 and individual data are shown in Fig. 5.8. A Tukey post-hoc comparison shows the following comparisons to be significant at the $p < 0.01$ level. Exp. IIa - positions 3 vs 4, and 4 vs 5; Exp. IIb - positions 2 vs 3, and 4 vs 5. Save for a small primacy effect on the first item, there was a monotonic accelerating decrease in RT with increasing serial position: a non-linear recency effect.

5) There is no interaction between Serial Position and List Type. The small difference in the pooled data from Exp. IIa at positions 2 and 3 may be seen to be due entirely to S_4 .

6) In Exp. IIa, but not in Exp. IIb, there is a significant interaction between List Type and Position of I-item ($p < 0.01$, $F(4,12) = 5.6$). This reflects longer RTs, for AC SIM lists only, when the I-item was one of the last two items. I have no explanation for this interaction, but it is not reliable since it did not appear in Exp. IIb.

Errors: There were too few errors for statistical analysis. In Figs. 5.6, 5.7 and 5.8 they may be seen to follow a pattern roughly

similar to RTs. Overall error rates for the two list types were:

	<u>AC SIM</u>	<u>AC DIS</u>
Exp. IIa:	4.4%	4.5%
Exp. IIb:	5.1%	8.1%

5.23. Discussion.

There was no evidence of any overall effect of acoustic similarity between list items on recognition reaction time or errors.

While there was some sign of a small interaction between list type and probe type in the data of some subjects, for most of them it was not particularly compatible with the idea that recognition is facilitated by acoustic dissimilarity between probe and list items. A marked and highly reliable recency effect was obtained in the serial position data. Since the experimental parameters, for Exp. IIa, at least, were in all relevant respects those of Exp. I, these findings are incompatible with Sternberg's hypothesis that in IRn and CR subjects are searching the same store (albeit in different ways). On the other hand the lack of systematic effects of acoustic similarity, and the marked recency effect in the serial position curve are completely in line with our two-store hypothesis, according to which IRn is mediated by 'directly-accessed' item-traces at amodal logogens. Of course, this puts us in the position of trying to prove the null hypothesis: all we can say is that if there is an effect of acoustic similarity on IRn, it must be very small or very unreliable; and comparable neither in magnitude nor in nature to the effect on CR.

Why did one subject (S_4) differ so markedly from the rest?

Subjects were asked to report on how they performed the task. Most said that, at least by the end of the practice sessions, their performance had become 'automatic'; as one subject remarked, 'You get

the idea you have seen it (the probe) or you haven't, it's familiar or unfamiliar.' Some subjects also said that, occasionally, they searched the list. S_4 was one of these, and was also one of the only two subjects who said that lists with acoustically similar letters were difficult. I suggest that S_4 was being hyper-cautious, and that whenever she was uncertain about a 'trace-strength decision' (e.g. for items early in the list) she checked it by searching the contents of the response buffer. This suggestion is supported by her extreme accuracy on AC DIS lists, her very long negative RTs for AC SIM lists and her unusually marked primacy effect on the first item (Figs. 5.6a, 5.8). This suggestion may smell of sophistry, but it raises an important general point. According to the two-store hypothesis, items are in general represented in both forms of storage; while the claim is made that it is both possible and faster and more efficient to use a logogen strategy in this task, it is only the emphasis on speed that prevents a subject from checking his decision by searching the response buffer as a 'back-up' store. Some subjects' tendency towards caution may be strong enough to defeat the intention of the speed instructions. So the two-store hypothesis must not be interpreted as implying that it is impossible to obtain acoustic similarity effects in item recognition, only that if subjects are encouraged to respond fast and rehearsal is prevented, most will not show such effects. That is what the data from this experiment demonstrate.

5.3. EXPERIMENT III. The Effects of Acoustic Similarity on Probed Location.

In Exp. I effects of acoustic similarity on CR were demonstrated. Exp. II showed no such effects on IPn. The difference was accounted for by the hypothesis that the two tasks were mediated by the consultation of different forms of representation in active memory: representation in the response buffer and as item-traces in the logogen system. The first task requires access to order information; the latter does not. However, it is conceivable that the contrasting results might be accounted for on the basis of other differences between the tasks. For instance, they might depend on the use of a verbal response in CR, or (evidence to the contrary) on some property specific to inter-item associations. Again, it might be argued that both tasks are in fact mediated by the same store, but that in recognition, the use of the item as probe bypasses the problem of retrieval that exists in recall, and that acoustic similarity influences only retrieval of the item (cf. Hitch, 1972).

I therefore investigated the effects of acoustic similarity on performance in a third task, probed location (Ln), in which the subject was required to indicate with a key-pressing response which position in the list was occupied by the probe item. Hence no verbal response was required, position rather than order information was requested, and the presentation of the relevant item as the probe should, according to the third view, have bypassed any retrieval process. On any or all of these views, no effect of acoustic similarity is expected. What does the two-store hypothesis predict? The order and position of items is represented in the response buffer, to which

access is via a serial search if only a content cue is given. Hence the subject should be obliged in this, as in the CR task, to search the response buffer, and we should expect effects of acoustic similarity comparable to those obtained with CR. It will be recalled that in Exp. I there was some evidence for subjects using a 'trace strength' strategy for the recall of the terminal item in the list. Is such a strategy possible in Ln? It may indeed be possible to allocate items late in the list to their positions on the basis of strength. However, it is unlikely that the subject's capacity accurately to judge the absolute position of items on the basis of their traces' strengths extends to the earlier items of the list, whose strengths will be closer together. Hence we might expect a mixture of strategies: a search of the response buffer, adversely affected by acoustic similarity, for items early in the list, direct judgements of recency, unaffected by acoustic similarity, for items exceeding a certain strength criterion - i.e. those late in the list. Thus we might expect bowed serial position curves and maximal acoustic similarity effects in the middle serial positions.

5.31. Method.

Exp. III was identical to Exp. II, including replication with the two different vocabularies, Vocab E and Vocab M, in all respects save the following:

Display: the background display is depicted in Fig. 5.9. The items of the memory set appeared in sequence, successively, in the five positions of the spatial array. (Note: spatial position was always correlated with temporal position.) The list was preceded by a warning signal - an arrow pointing to the first display position

for 400 msec. followed by 100 msec. darkness. The probe item was presented 300 msec. after offset of the fifth item, in the box in the lower right hand corner of the display. Presentation rate was as before, but was slowed to one item per 600 msec. for the first practice session.

Response. The probe was always a member of the memory set, and subjects indicated the position it had occupied by pressing one of five keys: the middle and index fingers of the left hand, and the index, middle and ring fingers of the right hand were used to indicate positions 1-5, in that order. No subject had any difficulty with this mapping. Note that there was a complete confounding of serial position and finger used.

Sessions. Each block contained 50 trials, consisting of one example of each of the combinations of List type (AC SIM, AC DIS), serial position of probe and position of I-item, in random order. Each experimental session contained three blocks and lasted half an hour.

5.32. Results.

Data from the experimental sessions were pooled in the same way as in Exp. II. Overall error rate was 8.8% in Exp. IIIa and 4.7% in Exp. IIIb. In Exp. IIIb, subjects mentioned that they tended to find the letters D and G confusing. Since these were members of what was intended to be the non-confusing group of letters in Vocab M, it was decided to exclude trials containing both these items from the analysis. It was for this reason that a similar exclusion principle was applied in Exp. IIb (q.v.).

The correct RT data were subjected to a four-way analysis of variance (mixed model, subjects random): List Type x Serial Position of probe x Position of I-item x Subjects. The mean RT data are completely summarised in Figs. 5.10a and 5.10b.

1) For both replications there was a significant main effect of List Type (IIIa: $p < 0.025$, $F(1,3) = 20.8$; IIIb: $p < 0.05$, $F(1,3) = 11.9$). RT was in both cases longer for acoustically similar lists, by an average of 62 msec. in Exp. IIIa and 34 msec. in Exp. IIIb.

2) Serial position. The effect of serial position is highly significant in Exp. IIIb ($p < 0.001$, $F(4,12) = 13.3$), but fails to reach significance in the case of Exp. IIIa ($F(4,12) = 2.5$); this is because of a highly significant interaction with subjects ($p < 0.001$, $F(12,200) = 73.0$). Inspection of Fig. 5.11, in which individual serial position functions are shown, reveals that the latter interaction was due to S_3 . Otherwise a bow-shaped serial position curve was generally obtained, with the slowest RT at position 4 for four subjects, and position 3 for the remaining three. Of course, serial position was confounded with response finger, and serial position data must therefore be interpreted with caution, but the magnitude of the serial position effects is greatly in excess of differences normally found between different fingers for simple RT (Rabbitt, personal communication).

3) The interaction between serial position and list type in neither replication approaches significance; however, Fig. 5.12, in which the difference between RT for the two list types is plotted against serial position, shows the effects of acoustic similarity to

have been somewhat greater at middle positions.

4) Position of the isolated item. The only other significant effects in this analysis are a three-way interaction between the fixed effects in Exp. IIIa ($p < 0.01$, $F(16, 48) = 3.0$) and an interaction between List Type and Position of I-item in Exp. IIIb ($p < 0.001$, $F(4, 12) = 16.2$). Inspection of Fig. 5.10 shows the interactions to be due to a reduction or disappearance of an acoustic similarity effect when the isolated item is probed or precedes the probed item in the list. (In Exp. IIIa this effect interacts somewhat with serial position.) Most of the sense of these interactions is retained if we abstract from the data the comparison between the cases where the I-item was probed and where it was not, for both list types. This is shown in Fig. 5.13. When the probed item was acoustically dissimilar to the rest of the list items, the effect of acoustic similarity was greatly reduced. That this is the result of acoustic isolation is implied by the lack of a comparable effect for AC DIS lists.

Errors. In Exp. IIIa, the error rates were 13.9% for AC SIM lists, 3.8% for AC DIS lists. In Exp. IIIb, the comparable frequencies were 6.9% and 2.5%. Analysis of erroneous responses from Exp. IIIb, whose RT was generally longer than that of correct responses, shows that erroneous responses tended to indicate positions adjacent to the correct position. (Unfortunately, the data of Exp. IIIa was accidentally wiped from tape before a detailed error analysis had been performed.)

5.33. Discussion.

Experiment III clearly demonstrated an adverse effect of acoustic similarity between list items on both RT and accuracy in probed location, an effect comparable in magnitude (at least for Exp. IIIa, where the same vocabulary was used) to the effect on contextual recall found in Exp. I. Thus the effect on CR cannot have been due to the use of a verbal response, to the testing of order rather than position retention, or to the non-presentation of the item to be located in memory at the time of test. It seems to be due, rather, to the requirement to base a response on the order or position of items in the sequence. This is entirely compatible with our two-store hypothesis.

Moreover, there was evidence of the mixture of strategies suggested in the introduction to this experiment. Bow-shaped serial position curves were obtained, acoustic similarity effects were greater in the middle of the list, and several subjects remarked that while they were obliged to search the list most of the time, their decision was 'automatic' when the probe was the last item in the list. The serial position data of S_3 (Fig. 5.11) is instructive. He appears to have relied more on a trace strength strategy than other subjects: hence the recency effect and the ridiculously high error rate on the first positions. It would appear that after two or three intervening items, the rate of decay of 'item-traces' slows to a point at which absolute list positions may not reliably be discriminated on the basis of strength.

A final point needs mentioning before we summarise the general implications of these experiments. Subjects were closely questioned after each experiment as to whether they had successfully avoided

rehearsal (they had) and whether they thought the composition of the memory sets was entirely random or had discerned the structure (i.e. 4 from one class of items, 1 from the other). While many subjects noticed that some lists were more acoustically similar than others, none of the subjects thought that the sets were not entirely random. It is therefore not possible to explain away the effects relating to the acoustically isolated item by the claim that subjects strove to spot it on each trial and give it special attention.

5.4. Conclusions and general discussion.

The results of Exp.I and Exp.III appear to demonstrate a marked effect of within-list acoustic similarity on both reaction times and errors in contextual recall and probed location; Exp.II indicates that Sternberg's item recognition task can, under comparable experimental conditions, be performed without adverse effects from acoustic similarity of the items. I have interpreted these results as rejecting Sternberg's hypothesis that CR and IRn are mediated by different strategies of search through the same store, and as favouring a two-store model. The finding of marked and reasonably linear primacy effects in CR and Ln compared to the non-linear recency effect in IRn, the interaction of serial position and list type in CR and Ln, the effects due to the acoustic isolation of probe, response and preceding items, and the apparently rapid access for terminal positions in CR and Ln - have all proved readily interpretable within the framework of the two-store model outlined in Section 1.4. According to this, IRn is mediated primarily by judgements of the strength of 'item-traces' at amodal logogens, while CR and Ln are mediated primarily by a serial search of a phonemically-coded response buffer, though it seems that recall at location of the last one or two items of a list may also be based on item-traces.

It remains only to consider what implications these experiments have for theories of the way in which acoustic similarity influences recall.

1) Within-list versus within-vocabulary similarity. Sperling & Speelman (1970) propose that items are stored phonemically in short-term memory, that phonemes are retained or lost independently, and that if at recall only one of the two phonemes of the letter is available, the subject chooses at random from among those letters of the whole stimulus ensemble that contain the retained phoneme in the same position. The last of these basic assumptions carries the strong implication that recall of a given item should not be influenced by how many other acoustically similar items there are in the list, but only by how many similar items are contained in the vocabulary. The results of Baddeley (1968) and Hitch (1972) implied the contrary: they showed that, with a constant stimulus ensemble, as the number of similar items in a list increased, so the probability of correct recall of a similar item declined, the one for serial, the other for probed, recall. The data of Exp. I are slightly ambiguous in this respect: if we compare recall of a C item, given as probe a N-C item, when the list is AC SIM and when it is AC DIS (the appropriate comparison is between AC SIM + Probe is I-item and AC DIS + Response is I-item in Fig. 5.5), more than twice as many errors were made when list similarity was high, which agrees with the evidence of Baddeley and Hitch. However the force of this evidence is somewhat weakened by the fact that there is a slight RT difference in the opposite direction.

Other features of the data suggest that if within-list similarity is important, it is not the only determinant of the effects, at least as regards retrieval of the response: a C item takes longer to recall, but not to find, than a N-C item when embedded in a dissimilar list.

2) Storage or retrieval? At what stage do acoustic similarity effects occur in recall? Posner & Konick (1966) proposed an 'acid-bath' theory, according to which forgetting occurs during storage as a function of the number and similarity of other items stored. Other theorists (e.g, Broadbent, 1971; Wickelgren, 1972) have emphasised the role of retrieval factors. The views of Conrad (1965) and Sperling & Speelman (1970) also implicate retrieval as the locus of the acoustic similarity effect. They hold the rate of forgetting to be the same for similar and dissimilar items; as a result of the smaller phonemic redundancy of similar items, a given amount of forgetting will produce more errors during their retrieval. Baddeley (1968), using the distractor technique, was unable to show any difference in rate of forgetting for similar and dissimilar material. Hitch (1972) attempted to test the storage hypothesis by using an item recognition task, on the grounds that retrieval is bypassed thereby. But the experiments reported in this chapter have carried the implication that different forms of storage mediate recall and item recognition, so this test may be inappropriate. How else can the storage hypothesis be tested? I have suggested (Sections 1.4 and 2.13), and data presented in the next chapter will support, the idea that given a position cue, the subject may have access to a location in the response buffer without, or with an abbreviated, search. So that if recall is probed with a position probe, then, if the storage hypothesis is wrong and acoustic similarity effects are purely a function of retrieval, then the effect of acoustic similarity should be small or nonexistent, as it is for the first item in contextual recall (see Exp. I). Hitch (1972) in his first experiment on acoustic similarity effects, did in fact use a position probe. His eight-item lists contained 6 or 2 similar items. He found no

effect of acoustic similarity on accuracy. In a comparable experiment using a location probe (cf. Exp. III) he did find an effect. Thus retrieval rather than storage is clearly implicated as the locus of acoustic similarity effects. However, in later experiments, in which a higher level of acoustic similarity was used (i.e. all 8 items were similar or dissimilar) Hitch did find an effect of acoustic similarity on recall probed with a position. He also found an effect of acoustic similarity on item recognition, but again only at high levels of similarity. He therefore concludes that there is also an effect of acoustic similarity on loss from storage, but only when a relatively large number of similar items is stored. Unfortunately, as regards this conclusion Hitch's experiments have a major flaw: subjects were entirely free to rehearse. I have suggested that acoustic similarity should result in less efficient and accurate rehearsal. Given the evidence that item recognition and probed recall are mediated by different forms of active memory it seems at least as plausible to conclude from Hitch's data that it is the effects of acoustic similarity on rehearsal (rather than on rate of forgetting) that are apparent only when large numbers of similar items are held in memory. So, though there is evidence that acoustic similarity is disruptive at retrieval, there is not yet good evidence for the view advanced by Posner & Konick (1966).

3) The associationist view. According to Wickelgren's (1972) theory the process of accessing the stimulus representation (the probe), given that the probe is presented visually, should not be influenced by its acoustic similarity to other items in the list; only response similarity should have an effect: Exp. III quite clearly demonstrates the contrary and in Exp. I it appears that

acoustic isolation of the probed item had at least as strong a facilitatory effect on RT as acoustic isolation of the response. The present data are thus in line with the results of Baddeley (1968) who showed that more errors were made on similar items than on the dissimilar items following them in serial recall.

TABLE 5.1

Source of error relative to correct response

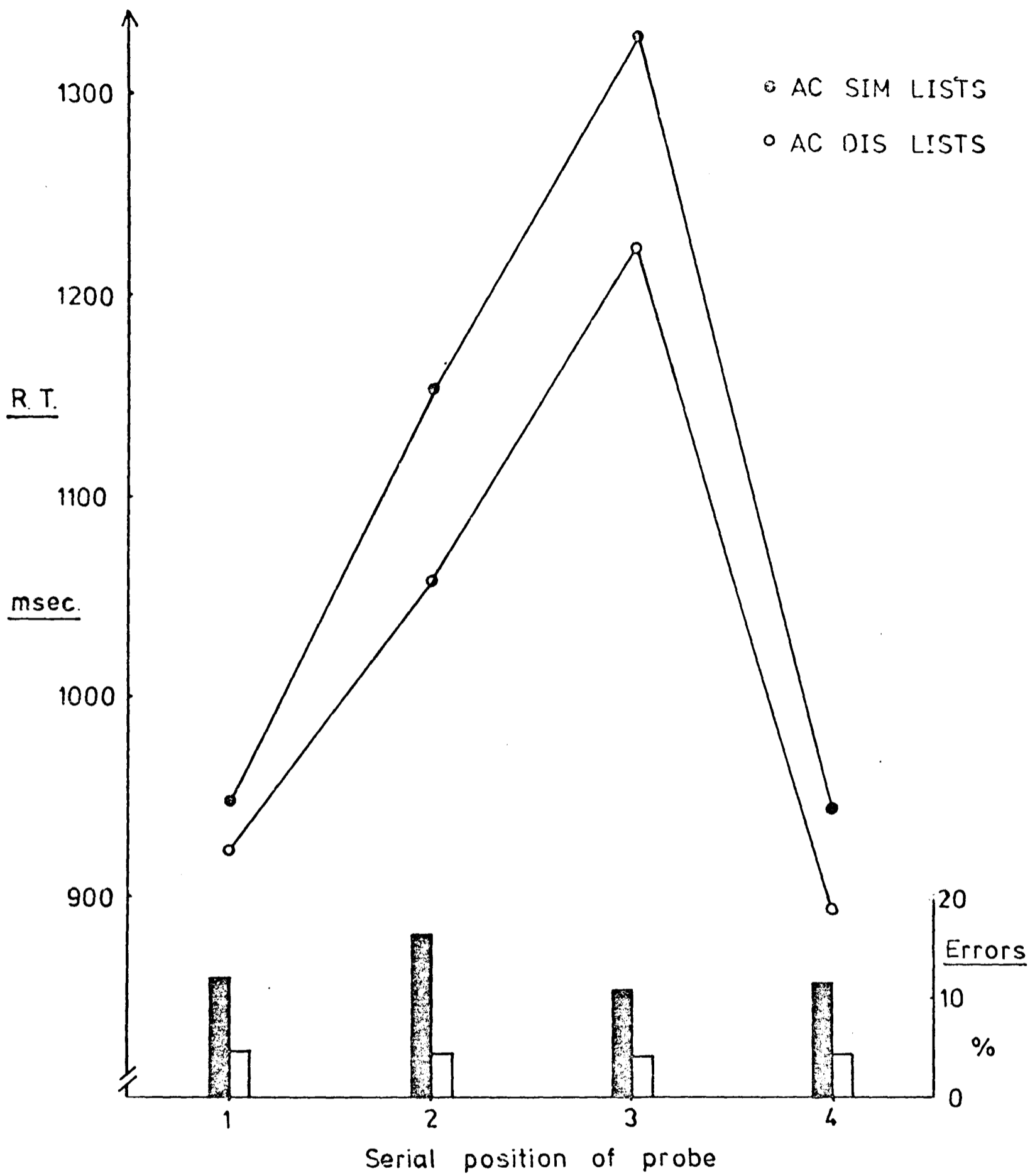
-4 . -3 . -2 . -1 . 0 . +1 . +2 . +3

Serial position
of probe:

1						62	24	2
2			9			79	12	
3		4	48			48		
4	1	14	85					

TABLE 5.1. Experiment I - CR.

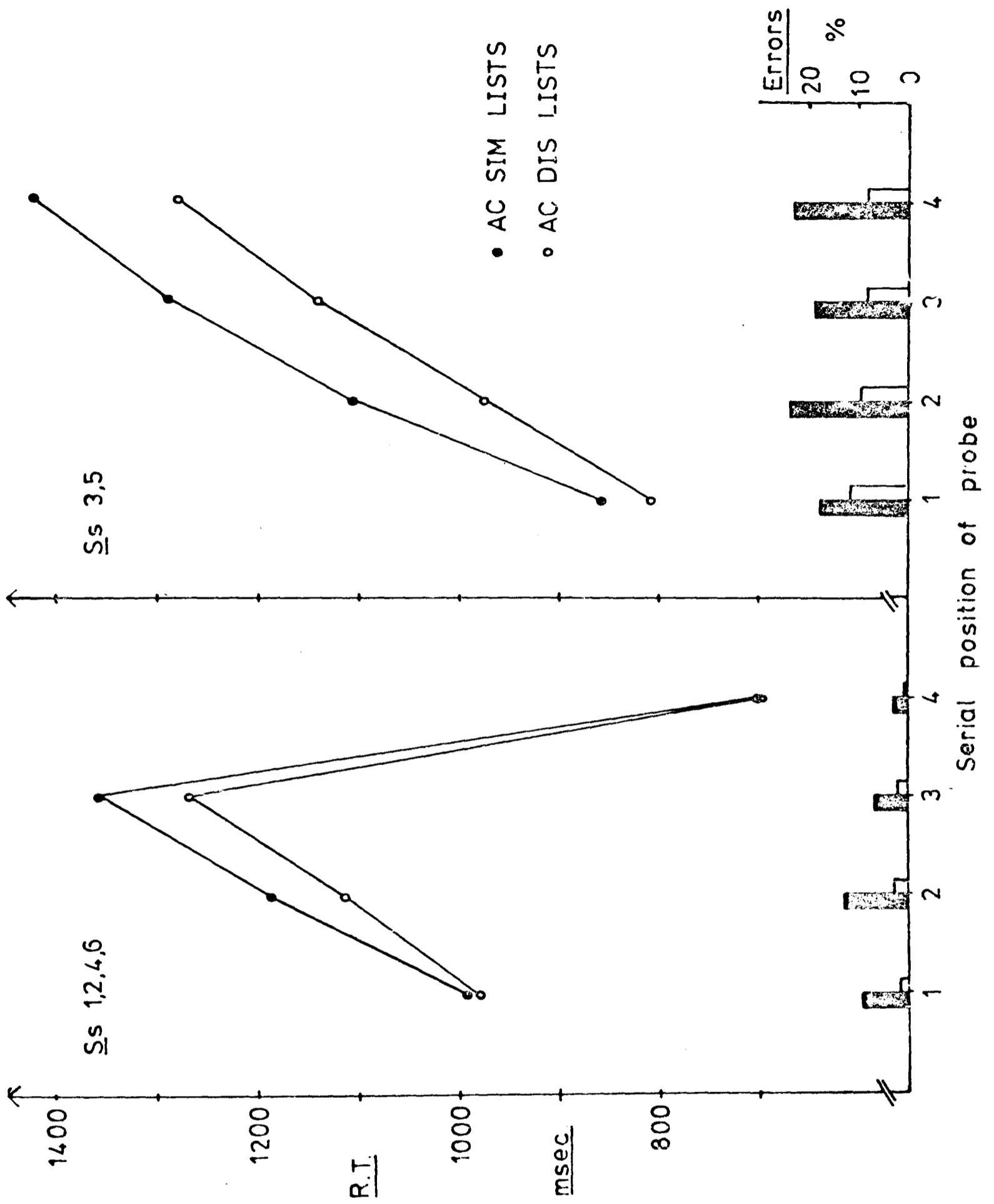
Each cell shows the percentage of erroneous intrusions from within the list that came from that position in the list when the indicated serial position was probed.



EXP. I - CR.

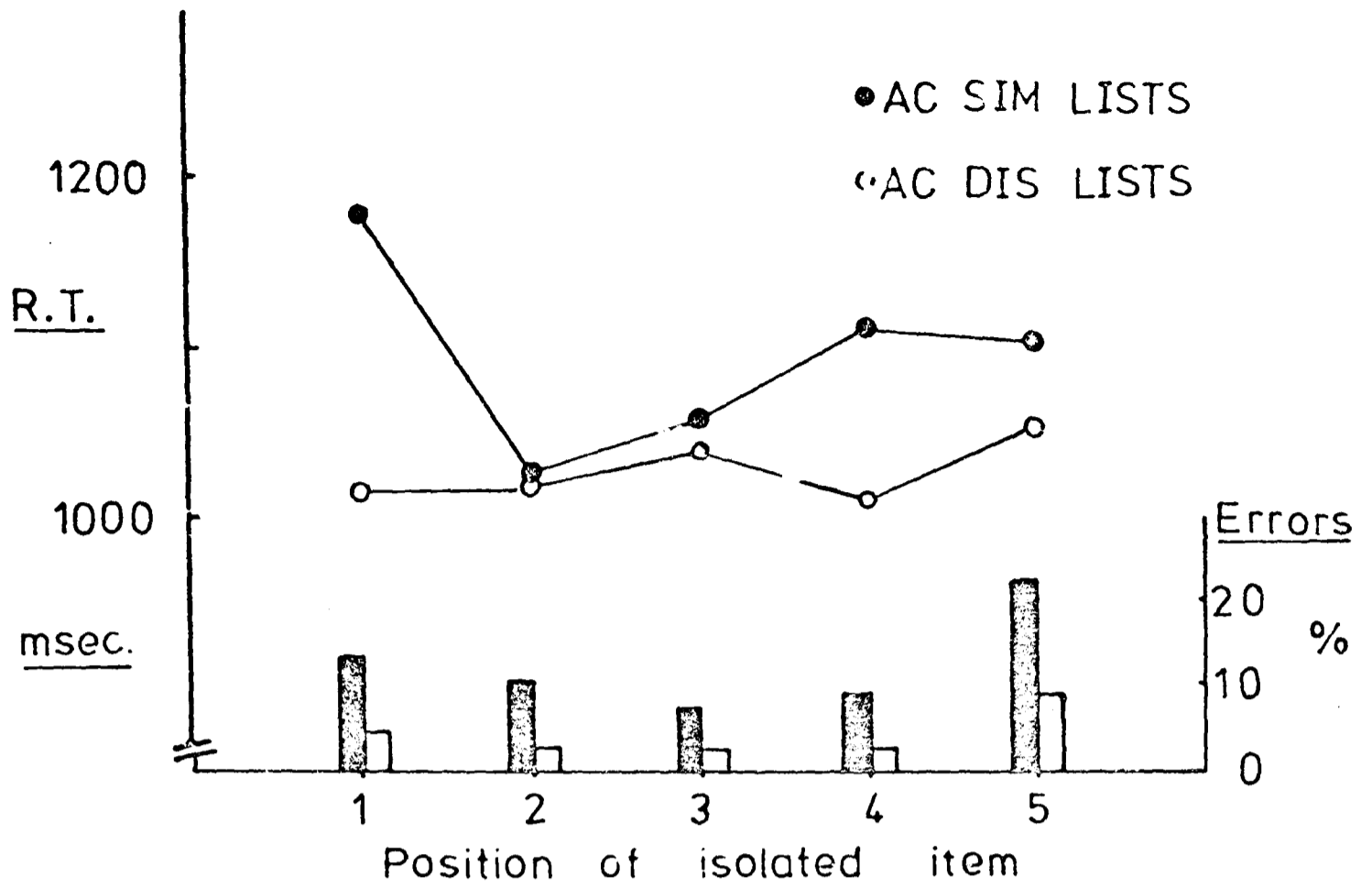
Serial position: group data.

FIG 5.1



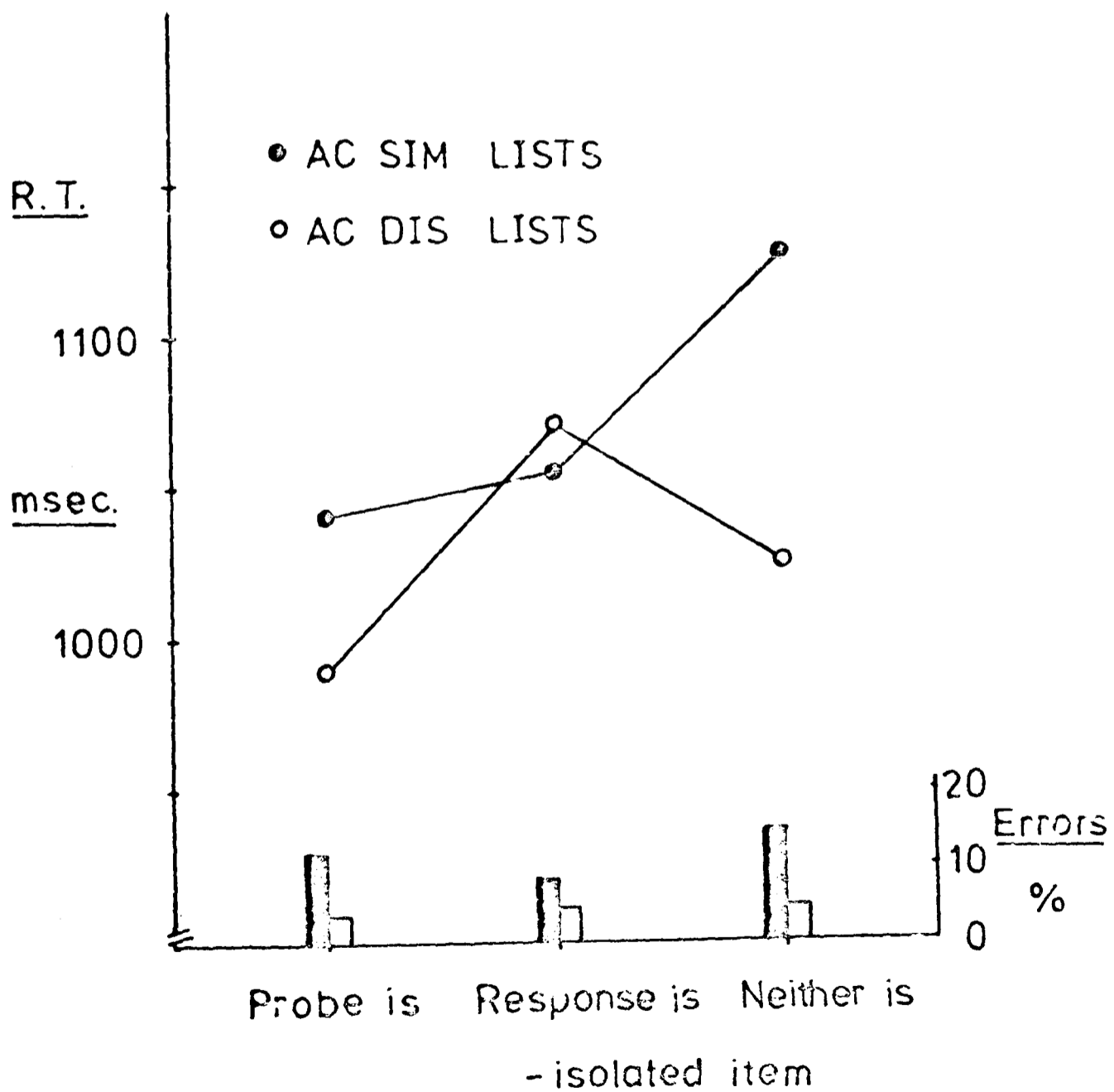
EXP. I - CR Serial position data plotted separately for the two groups of subjects.

FIG. 5.2



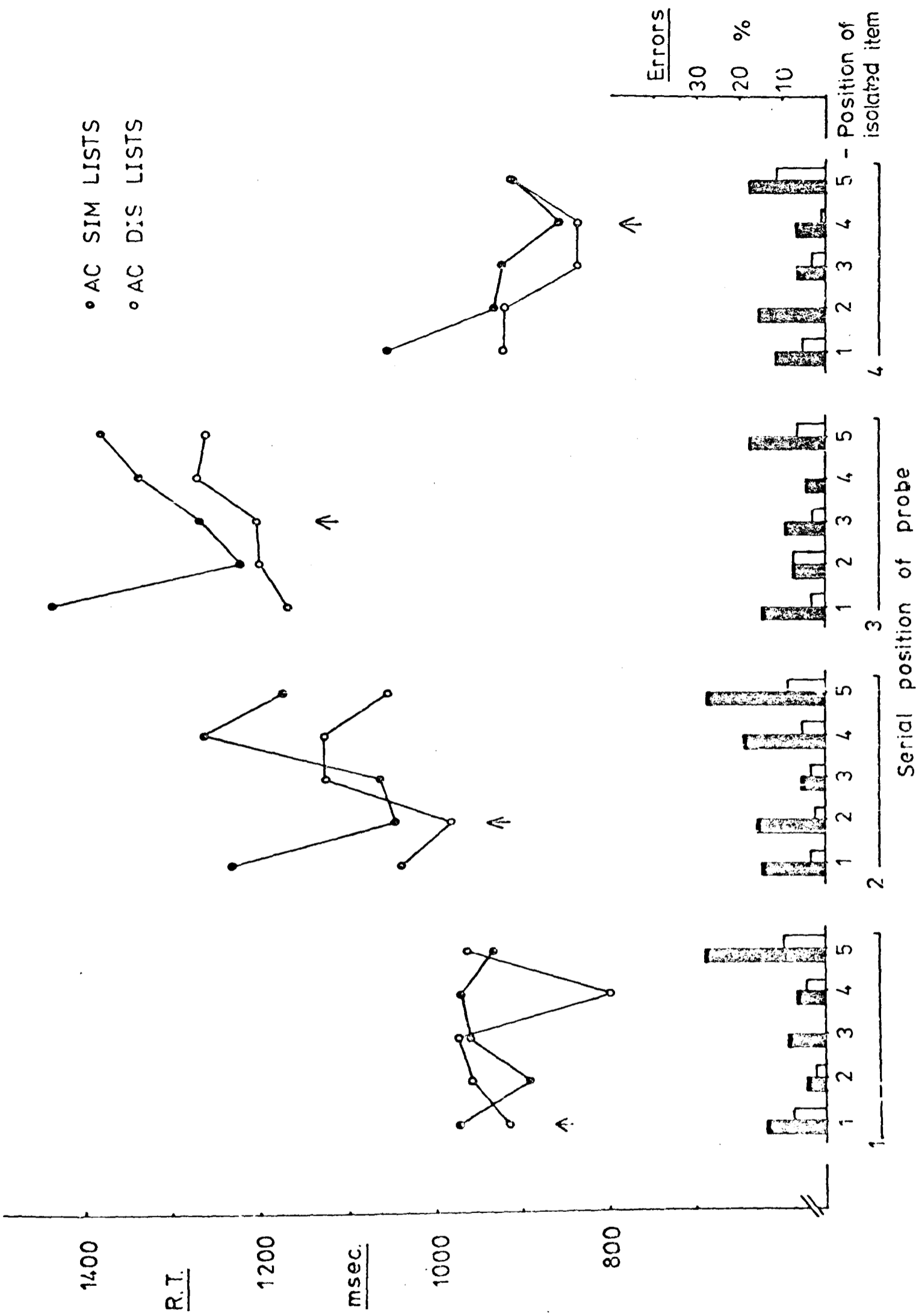
EXP. I - CR. Interaction between list type and position of i-item.

FIG 5.3

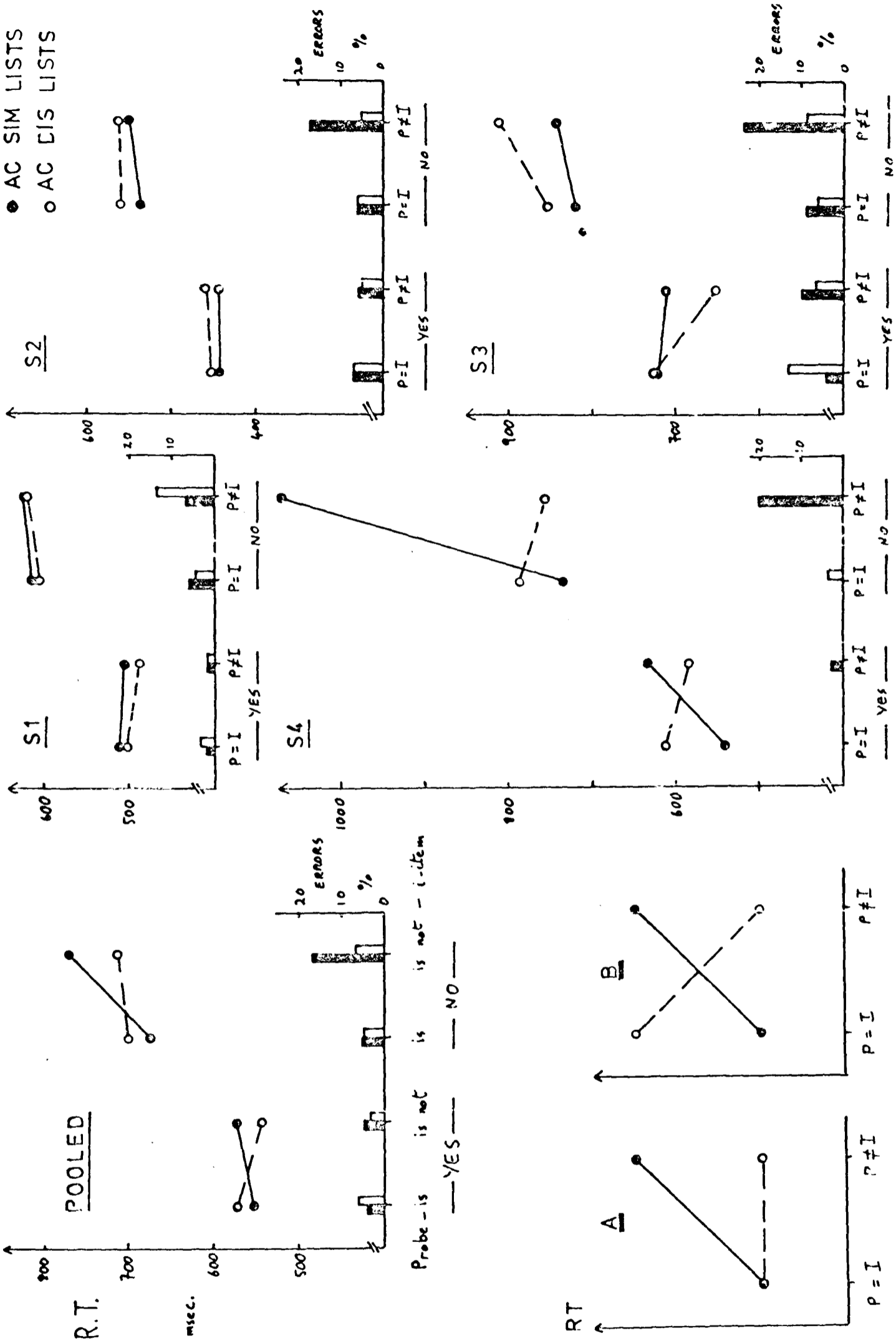


EXP. I - CR Effect of isolation of probed item and response item.

FIG 5.4



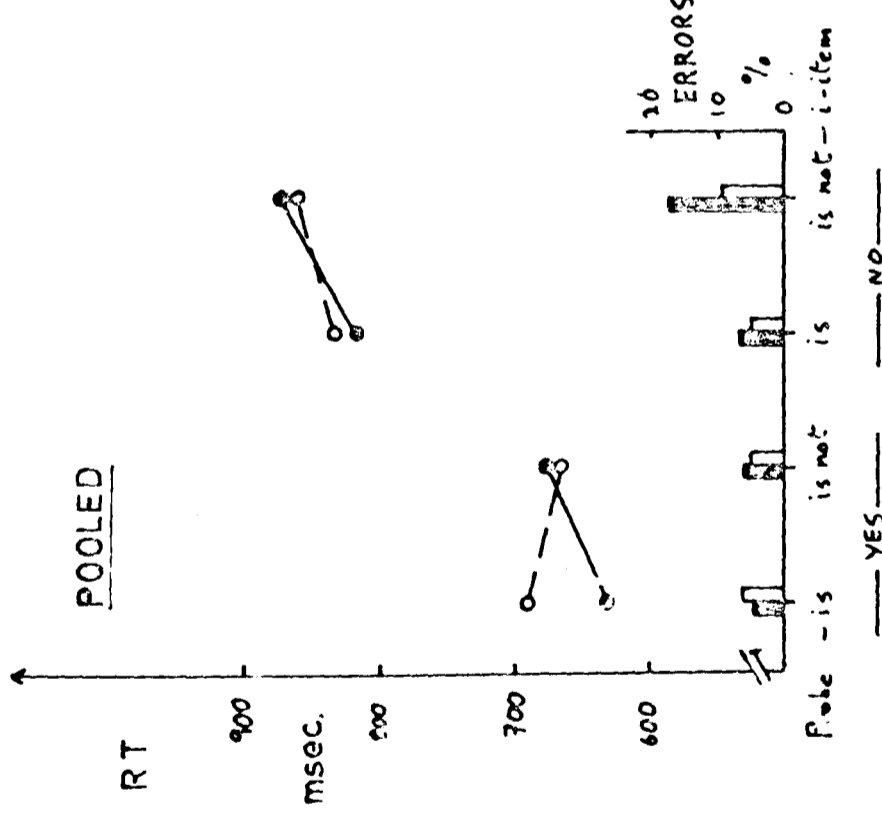
EXP. I - CR Interaction of list type, serial position probed and position of i-item.



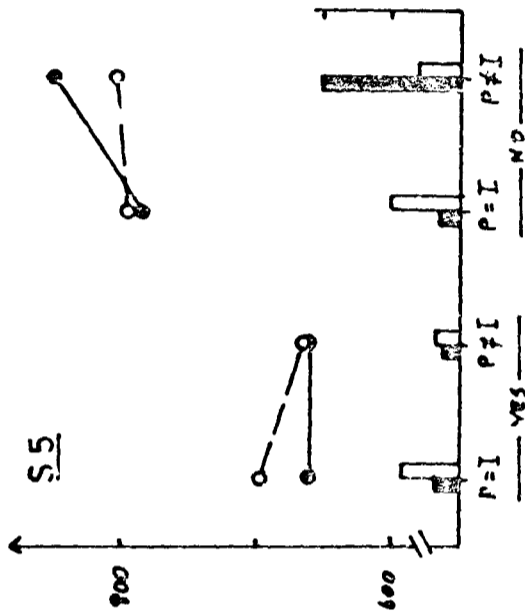
EXP. IIa - IRn - VOCAB E Interaction of list type and probe type

- AC SIM LISTS
- AC DIS LISTS

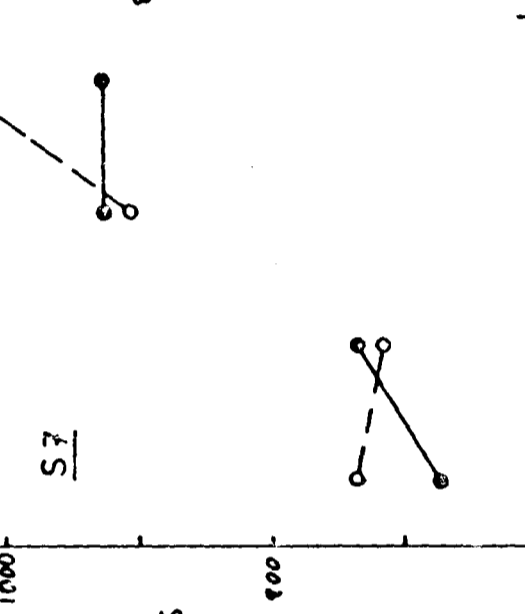
POOLED



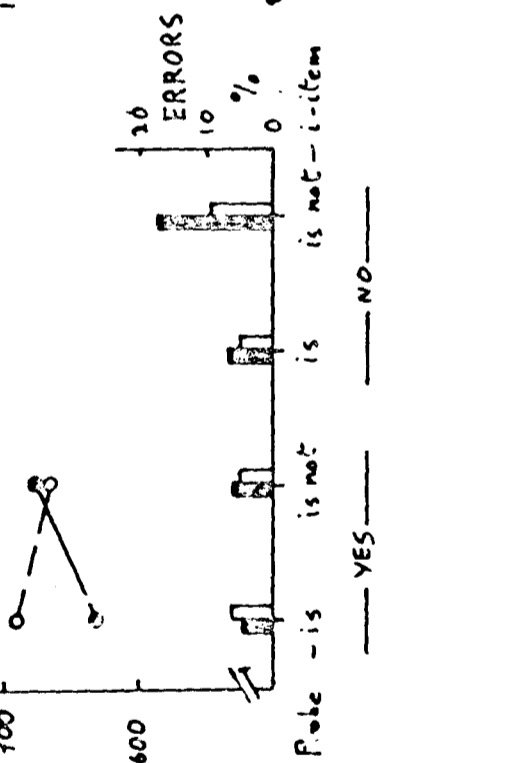
S5



S6



S7



S8

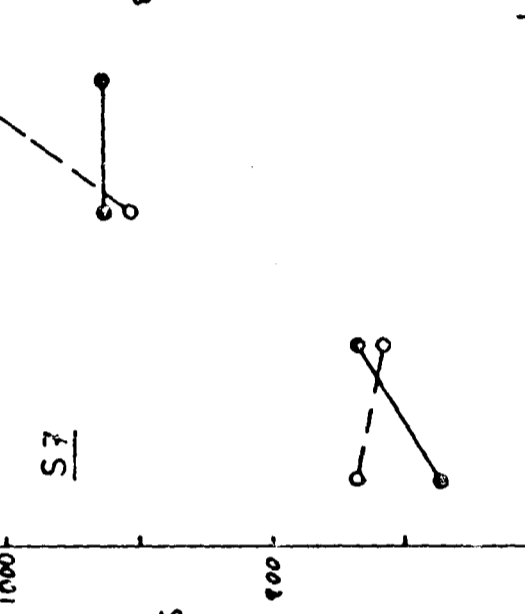
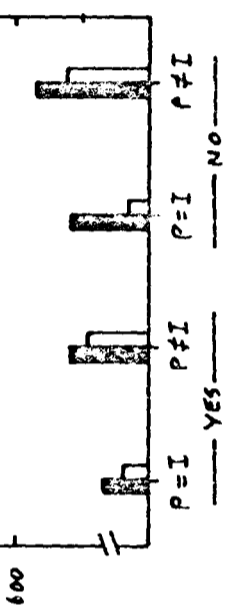
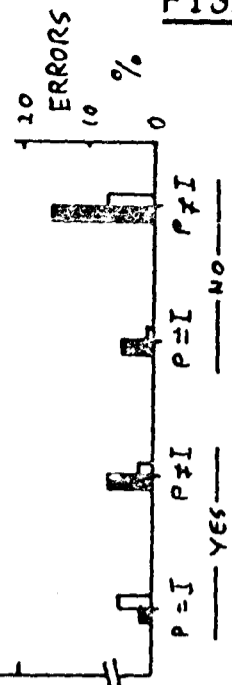


FIG. 5.6b



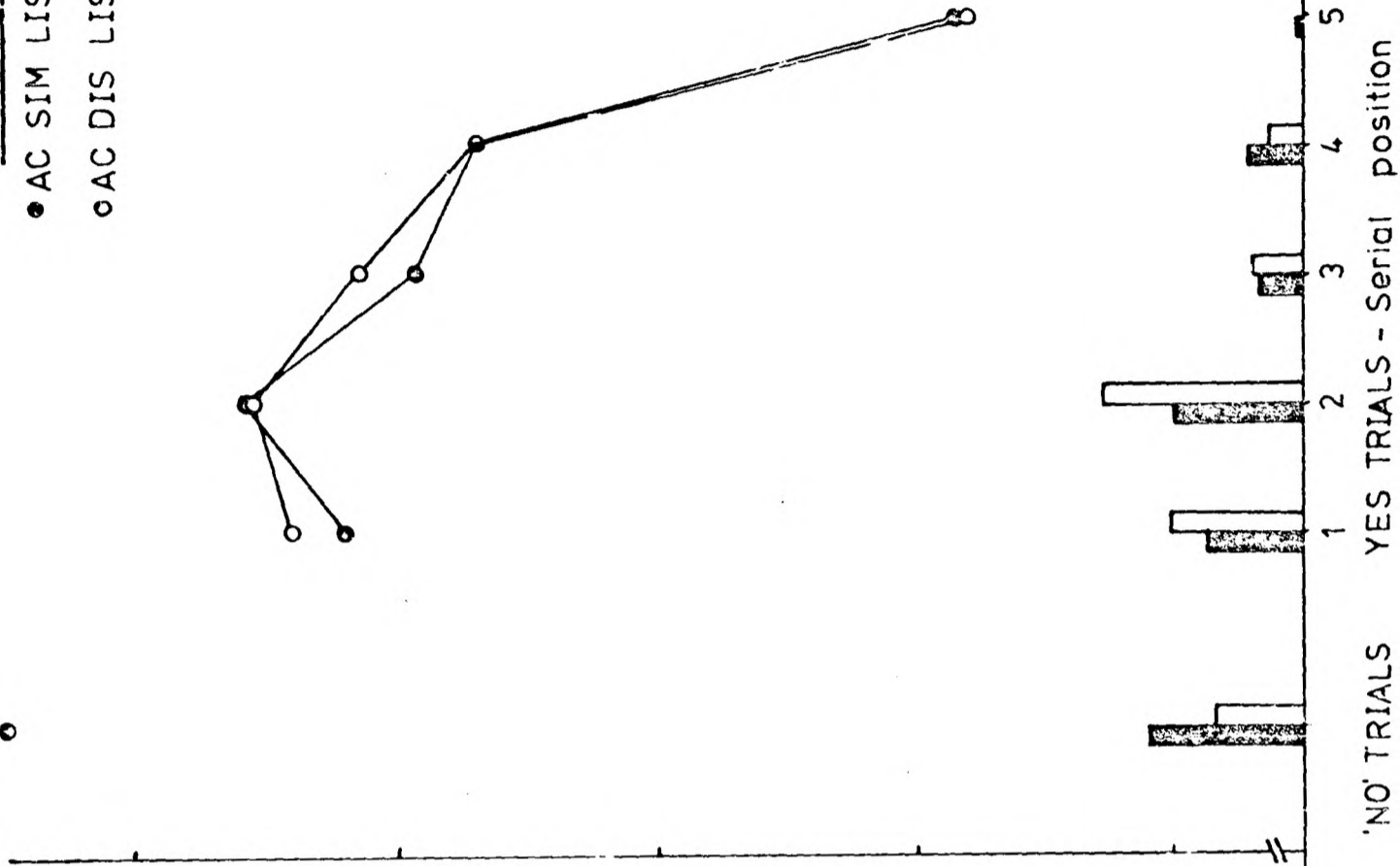
EXP Ib - IRn - VOCAB M Interaction of list type and probe type.

EXP Ib - IRn - VOCAB M

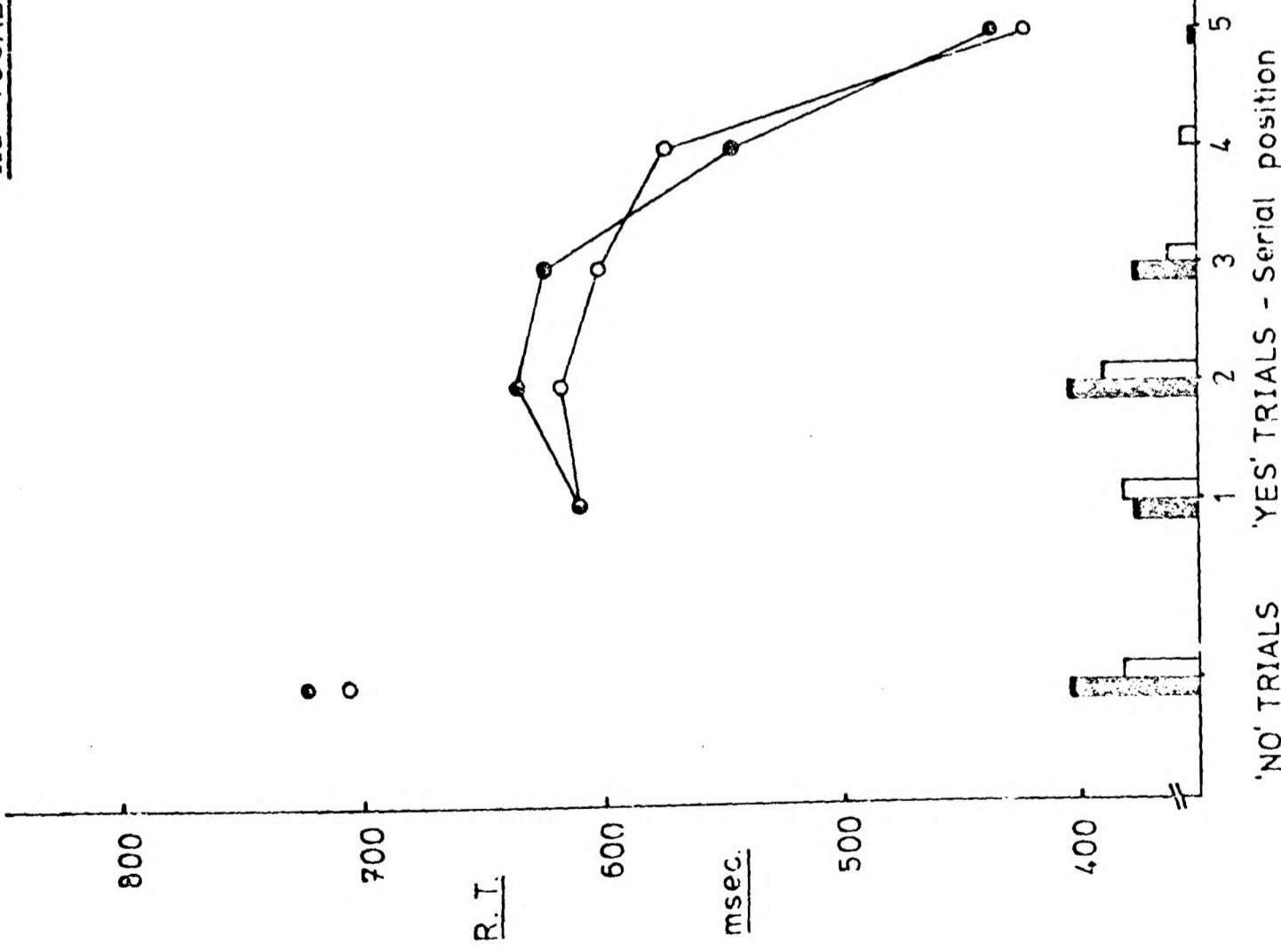
FIG. 5.7

Iib - VOCAB 'M'

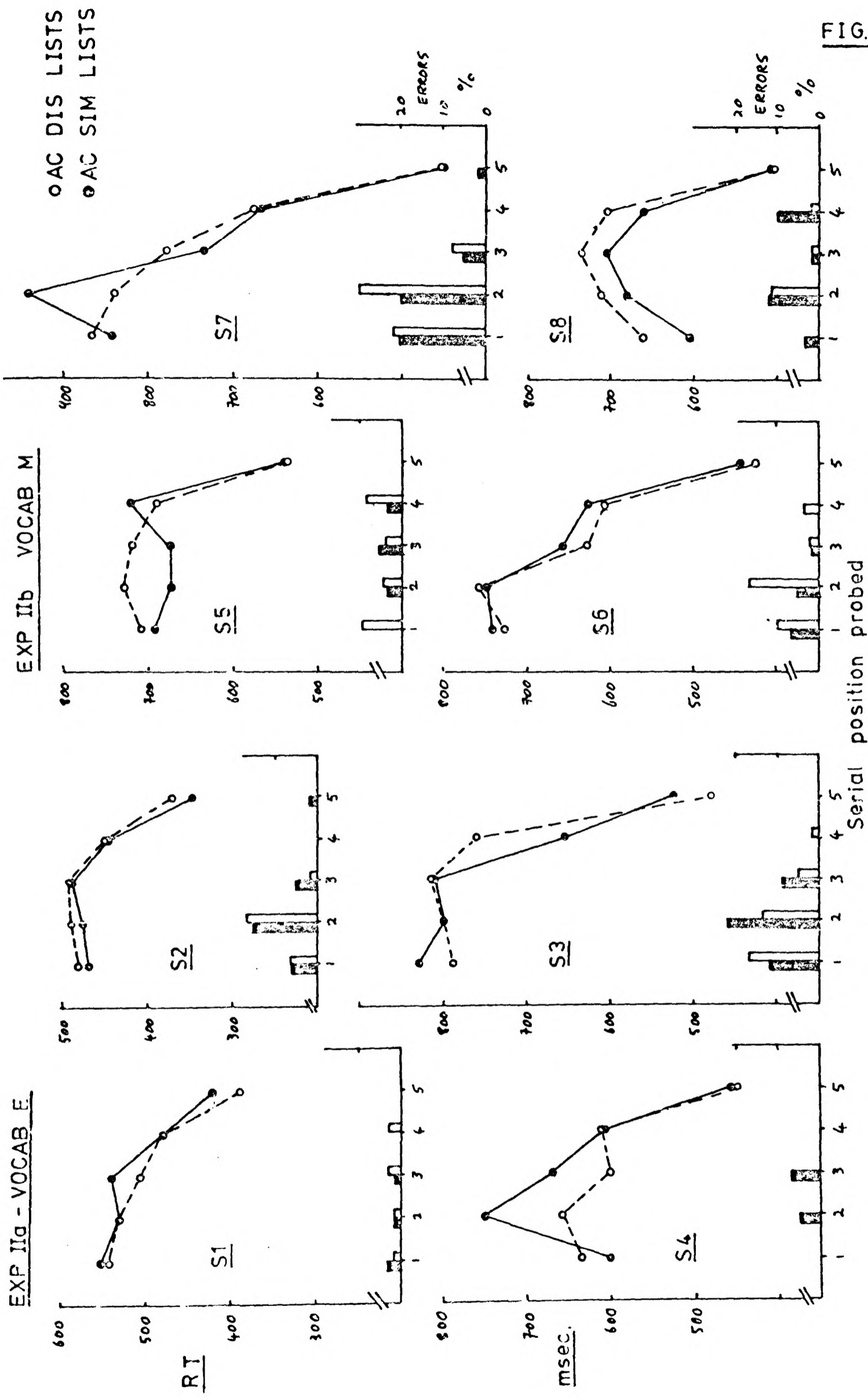
- AC SIM LISTS
- AC DIS LISTS



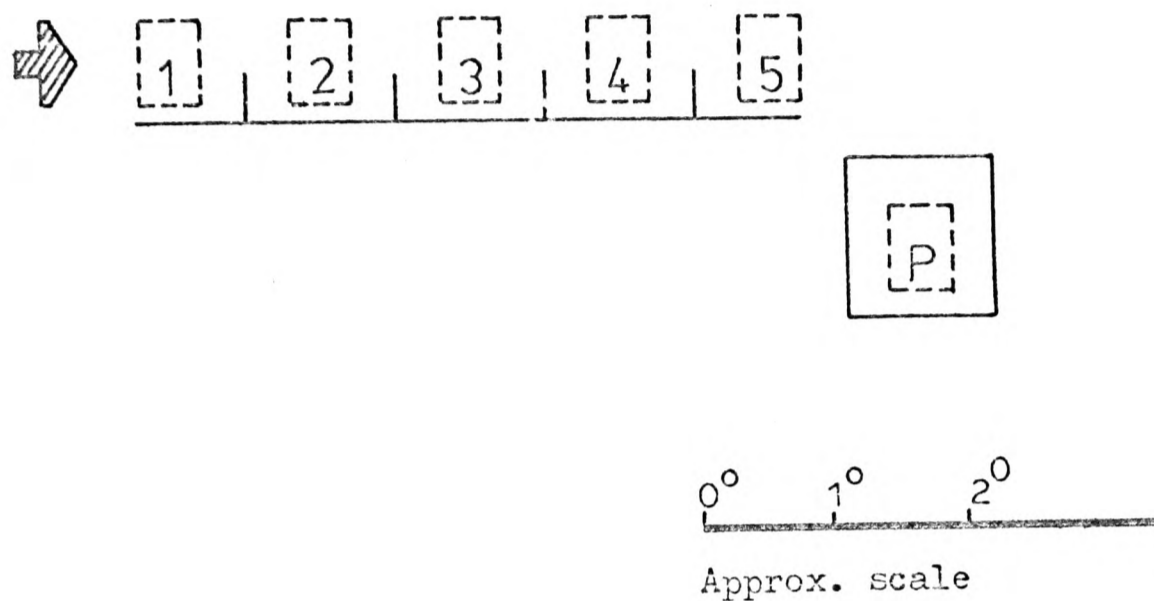
Iia - VOCAB 'E'



EXP II - IRn Serial position: Group data



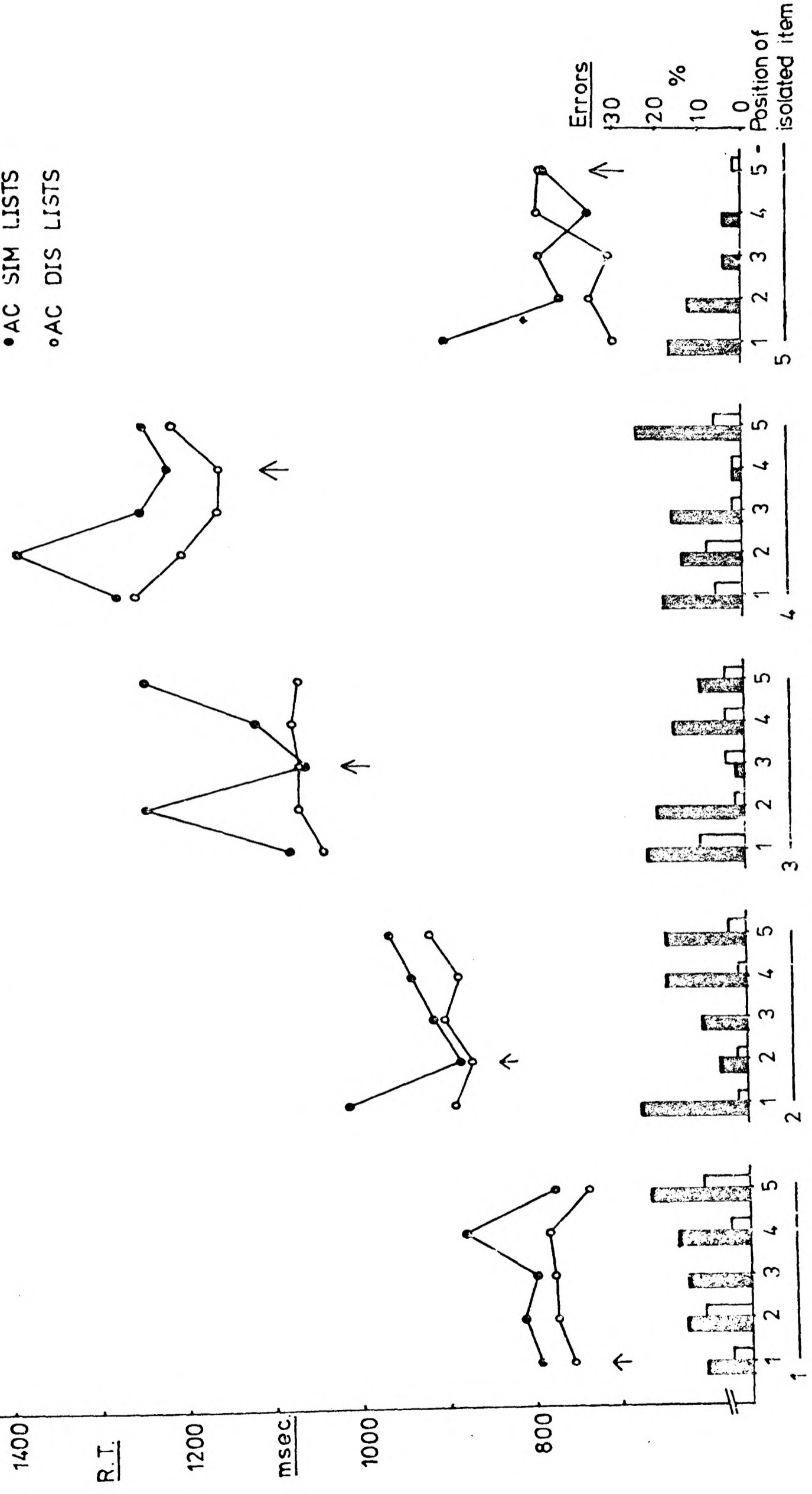
EXP II - IRn. Individual serial position data.



The dotted rectangles show the display locations of the memory set items (order as shown) and the probe item (P). The thin continuous line shows the background display. The warning arrow is shown (shaded) on the left.

EXPERIMENT IIa - Probed location - VOCAB 'E'

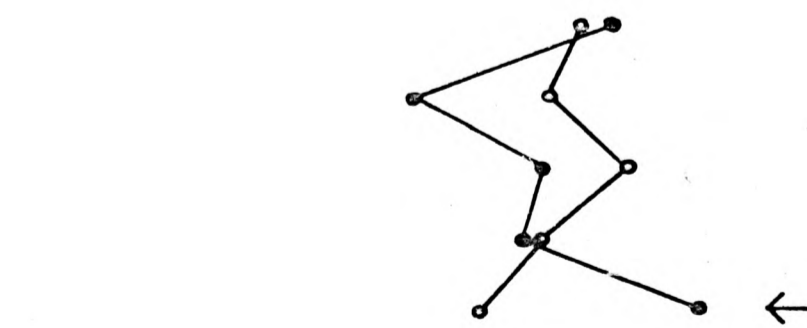
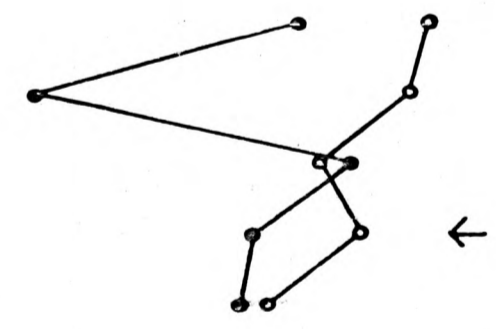
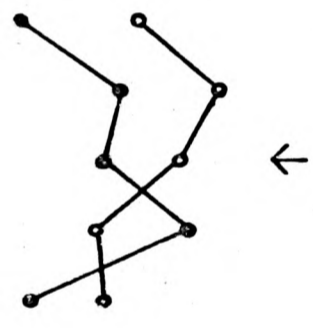
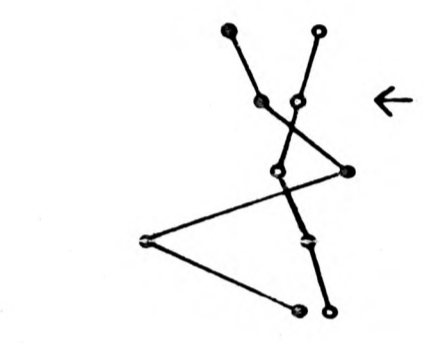
● AC SIM LISTS
○ AC DIS LISTS



Serial position of probe

EXPERIMENT IIIb- Probed location - VOCAB 'M'

- AC SIM LISTS
- AC DIS LISTS



1200

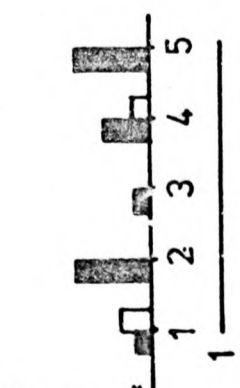
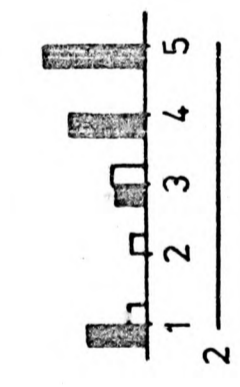
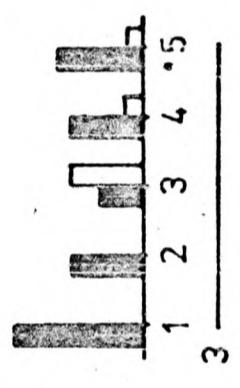
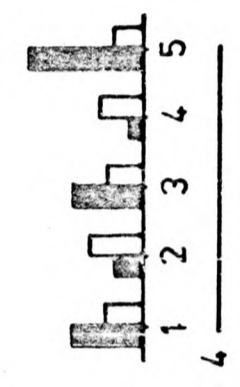
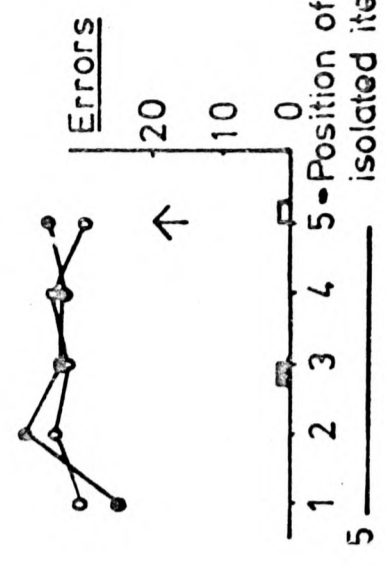
R.T.

1000

msec.

800

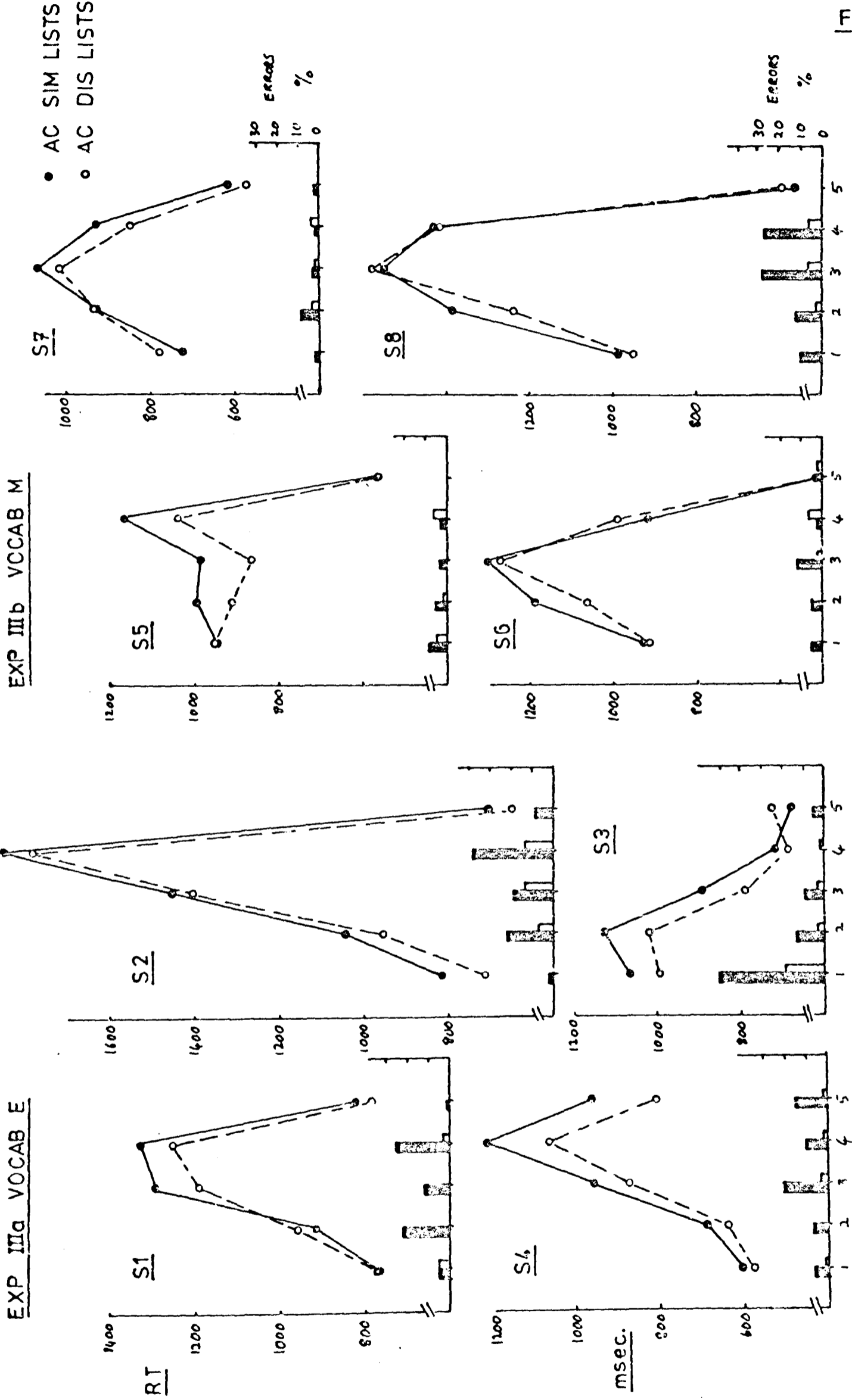
600



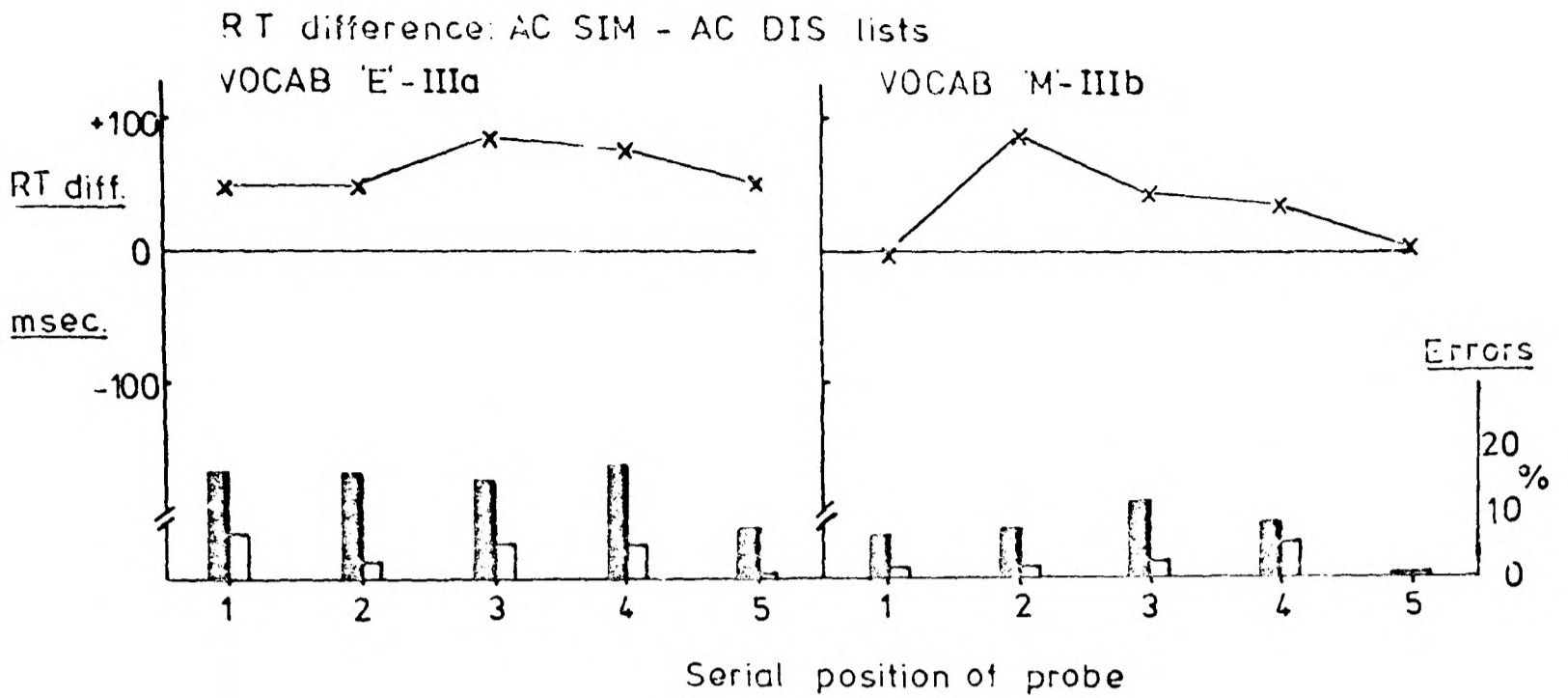
Serial position of probe

Position of isolated item

FIG. 5.10b

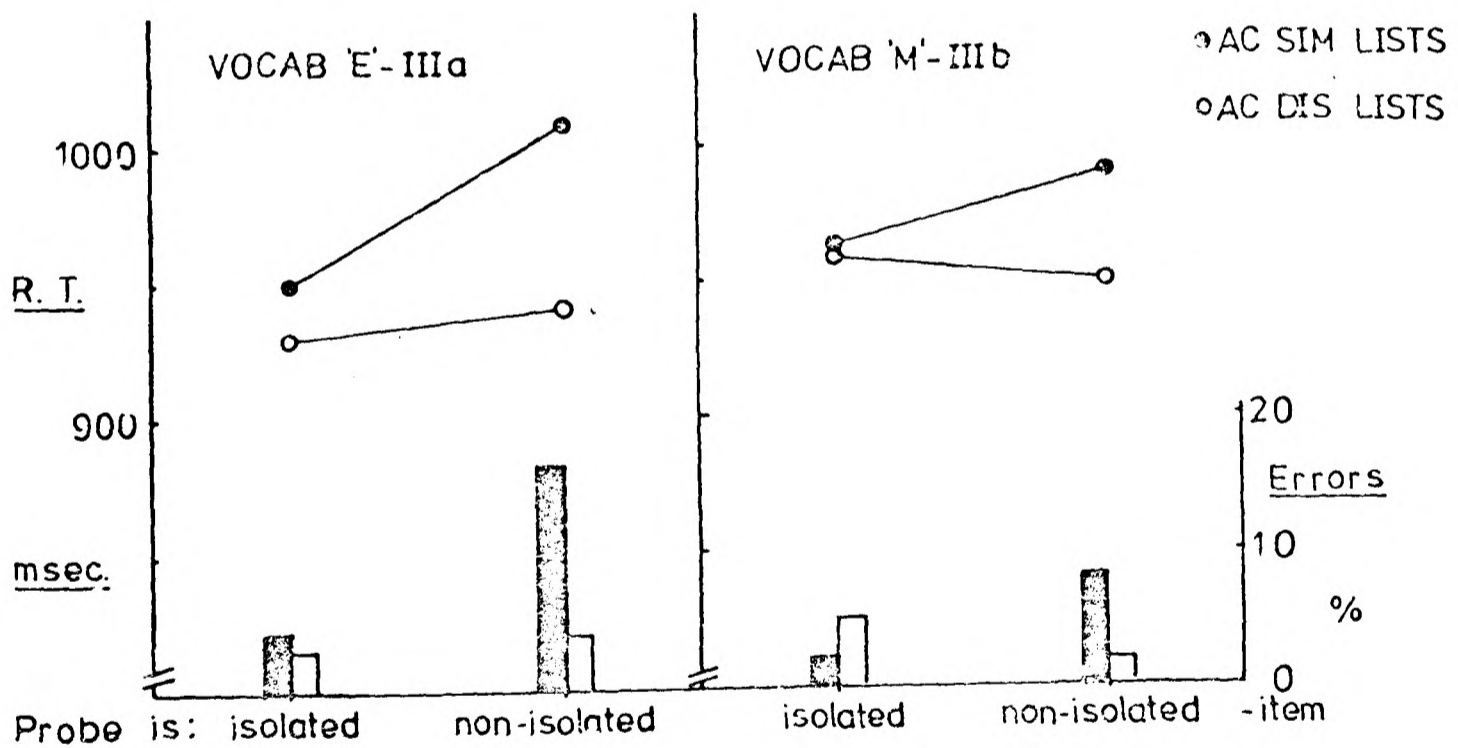


EXP III - Ln. Individual serial position data.



EXP III - Ln. Difference due to list type, plotted against serial position.

FIG. 5.12



EXP III - Ln. Interaction of list type and probe type.

FIG. 5.13

CHAPTER VI. The utility of a position cue.

Introduction.

6.1 EXPERIMENT IV. The effects of a position cue in IRn.

6.11 Method.

6.12 Results.

6.13 Discussion.

6.2 EXPERIMENT V. A comparison of the effects of a position cue
in IRn and CR.

6.21 Method.

6.22 Results.

6.23 Discussion.

Summary.

CHAPTER VI. The utility of a position cue.

The experiments to be reported in this chapter are prompted by the question: does information given with the probe about its ordinal position in the memory set facilitate either IRn or CR? The rationale is simple: if the items of the sequence are represented in a store whose contents are arranged in their temporal order of input, and if a search is normally required to locate a given item, then access to an item might be expected to be more efficient if its temporal position is specified. If either premise is lacking, then there is no reason to expect any facilitation from a position cue.

Our interest in this question is aroused by the hope that it may supply an additional basis for the distinction between two forms of storage in post-categorical active memory hypothesised in Section 1.4 and supported by the evidence on the effects of acoustic similarity. The basic predictions made by the two-store model are clear. Logogens are content-addressable. Hence access by the probe to the item-traces mediating IRn would not be expected to be facilitated by a position cue. That there is no facilitatory effect of position information on short-term recognition accuracy is implied by an experiment by Hitch (1972, Exp.7), who compared recognition with and without a position cue and found no d' difference (see discussion in Section 2.22,ii). I know of no comparable study employing RT as a measure. It may also be recalled (Section 3.129) that the finding of an effect on IRn RT of information given with the probe which enables the subject (at least logically) to restrict his 'search' to a subset of the memory set, has been restricted to studies where the subset is specified by semantic class (see Ch.8). The response buffer, on the other hand is a temporarily-ordered store.

Moreover in Section 1.4, I added to Morton's description the premise that there exists some capability for direct access to locations in the response buffer, but not, of course, to items, in that the subject may select a starting point for reading a sequence of items from the response buffer. Hence, since the process of serial scanning of the contents of the response buffer is held to be a slow process proceeding at not faster than about eight items a second, position information should enable considerable saving of time in retrieval, and we would expect a marked reduction in RT to result from the giving of a position cue along with the probe in CR. A certain amount of evidence already exists for this proposition in the finding of Hitch (1972, Exps. I & II) that recall was more accurate with a double (contextual + position) cue than with a contextual probe alone (see Section 2.13). However, as explained in Section 4.1, RT is a more efficient measure for tackling this question; I know of no clearly relevant RT study to date (Section 3.333). Thus, the two-store model predicts no facilitatory effect of a position cue on IRn, a marked effect on CR.

What do other theories predict? According to Wickelgren's (1972) account (Section 2.11), there is no difference in the nature of access in the two tasks, IRn and CR; it is held to be direct in both cases. However, position information would be expected to have a facilitatory effect on both RT and accuracy in contextual recall, on the grounds that the strength of the position-item association should add to that of the item-item association. The present experiments will therefore not be critical in testing Wickelgren's account against our own. But it will be recalled that the studies described in Section 2.13 failed to show an appropriate advantage in accuracy for a double over a single probe, and that the findings of

Exp. I were not congenial to Wickelgren's theory.

Any theory which postulates that item recognition and contextual recall both necessitate searching the same store in the same way must predict that the effects, if any, of position information on the two tasks should be identical. According to Sternberg (1969a), however, IRn and CR are mediated by different strategies of search in the same store (Section 3.31). A self-terminating, and hence slow, scan is held to be necessary for CR because the subject must continually monitor the outcome of his comparisons in order to stop searching at the appropriate point. Whether or not a position cue will be helpful depends on the separate question of whether the subject is able and/or willing to vary the starting point of his search. Sternberg (1969a) clearly considers active memory to be an ordered store - it must be for CR, and even for IRn it must be at least quasi-ordered in that the subject must 'know' which items he has scanned and which he has not. Thus, if the subject is able to start his search part-way through the sequence, even if in IRn he then searches exhaustively to the end of the list while his CR search is self-terminating, the position cue should have a facilitatory effect in both tasks, though the time saved would be smaller in the case of IRn.

6.1. EXPERIMENT IV. The effects of a position cue in IRn.

In this experiment, the memory set on each trial consisted of four digits, presented sequentially. The list was split into two pairs, one presented at one display location, the other at another. The subject was required to indicate by pressing one of two keys whether the probe item was or was not a member of the memory set. There were two conditions. In Condition P, the probe item was always presented at one of the same two display locations used for the memory set items, and the subject knew that if it had been in the list, the probe was one of the pair presented at that location. In Condition N, the probe was presented in one of two different locations, and the display location of the probe carried no information as to which portion of the memory set was relevant. If the subject is obliged to search active memory and is able to use a position cue to determine the starting point of his search, then he should respond faster in Condition P than in Condition N when the second pair in memory is indicated. If in addition he is able to use the position cue to stop searching when the relevant subset is exhausted then he should also be faster when the first pair is indicated. However, which pair in memory is 'first' will depend on whether the subject rehearses the list an integral number of times; as it happens rehearsal was not controlled in the present experiment. In view of the results obtained, this does not matter. Another consequence of the fact that this experiment was conducted before I had started to use fast presentation rates and instructions to control rehearsal is that serial position effects will not be very meaningful.

There is one other feature of the experiment that requires

explanation. At the time it was performed, I favoured a scanning model for IRn, and was inclined to attribute recency effects (see Section 3.124) to some (rather unspecified) form of sensory matching for recent items. In order to ensure that no sensory matching took place in this experiment, I interposed a half-second mask of dynamic visual noise between memory set and probe. (It may be worth mentioning that a later informal experiment suggested that the mask had little effect on the serial position curve when the probe delay was $\frac{3}{4}$ sec. It did however reduce RT at all serial positions; this appeared to be due to its offset serving as a warning signal which allowed $\frac{1}{4}$ sec. preparation for the probe onset.)

6.11. Method.

Eight subjects each served for two sessions, one practice, one experimental. Each session lasted about fifty minutes and was divided into two halves. In one half, subjects were given three blocks of 36 trials under Condition P, in the other, three blocks under Condition N. Between the two halves, a five-minute break was given.

Procedure: A trial was initiated by the onset of the background display, which remained on throughout the trial. It consisted of a large X, subtending a visual angle of some 5° horizontally and vertically. This served as a reference frame for four display locations each located centrally in one of the four quadrants - left and right, top and bottom. One second after its onset, two of the memory set items were displayed in one of the quadrants one after the other, followed by the remaining two in the opposite quadrant. Each item was displayed for 750 msec. and followed by a dark interval of 250 msec.. 250 msec. after the offset of the last item, a mask of

quasi-random dense visual noise, covering the area defined by the background X, changing every 50 msec., was displayed for 500 msec.. (Subjects were encouraged to fixate a blob at the centre.) 250 msec. later, a probe was presented. In Condition P it was displayed in one of the two quadrants in which the items of the memory set had appeared. In Condition N, one of the other two quadrants was used, randomly. Subjects responded with their right index finger for a positive response, left index finger for a negative response. The response terminated the display. No feedback was given until the end of a block. The next trial followed after an interval of $3\frac{1}{2}$ secs.

Design. The order in which the subjects underwent the two conditions in practice was reversed in the experimental session. Half began with P followed by N. For half the subjects in each order of conditions the memory set items were presented in the top and bottom quadrants, for the other half, in the left and right quadrants.

Each block contained 32 experimental trials and was preceded by 4 warm-up trials: 16 positive trials, including one of each of the combinations of serial position of the probe item, presentation order - i.e. left-right (up-down) or vice-versa, and, in Condition N, probe display location; 16 negative trials, eight for each probe display location.

The four items of the memory set were sampled without replacement from the digits 1-9.

New blocks were generated for each session.

Instructions. It was carefully explained to subjects that, while in condition N the display location of the probe meant nothing,

in condition P, if the probe was a member of the memory set it would always be displayed in the same quadrant as in the memory set, so that there was in principle no need to check the probe against the two memory set items displayed in the other quadrant. The instructions were thus clearly biased toward a 'directed search' strategy. Subjects always knew which condition held for the block they were doing. They were asked to respond as fast as possible while avoiding errors.

6.12. Results.

Correct RTs from the experimental session were pooled over the three blocks given under each condition. Overall error rate was 4.8%. A further 1.8% of trials were excluded from the analysis because the RTs were more than 3 standard deviations from the mean of the relevant (subject x response type) cell.

The correct RT data were submitted to a four-way analysis of variance: Conditions x Response type x Presentation order x Subjects (random); each of the two observations in a cell was based on 12 trials. Only the main effect of response type reaches significance: positive responses were 63 msec. faster than negative responses ($p < 0.001$, $F(1,7) = 57.2$).

1) There was effectively no overall difference in either RT or errors between the two conditions. This may be seen in TABLE 6.1, where mean RTs are shown for each response type and condition. It may be seen that individual scores show a difference in the direction: P slower than N, at least as often as in the other direction. Position information appears to have had no facilitatory effect at all

upon IRn in this experiment.

2) The only other effect remotely to approach significance in this analysis is the Conditions x Presentation Order x Response Type interaction ($F(1,7) = 4.7$). This interaction is summarised in TABLE 6.2.

3) Correct positive RTs were submitted to a four-way analysis of variance: Conditions x Serial Position x Presentation order x Subjects (random); each observation was based on only three trials. The interaction between Conditions and Presentation order is significant at the $p < 0.05$ level ($F(1,7) = 5.87$) - see TABLE 6.2, positive trials; no other effects approach significance. Serial position data are summarised separately for each condition in TABLE 6.3. There is clearly very little effect of serial position, save perhaps a hint of the effects of pair-wise organisation in the slightly longer RTs at positions 1 and 3.

6.13. Discussion.

The outcome of this experiment was quite straightforward. A cue which, when presented with the probe, potentially restricted the set to which reference was necessary to a subset defined by spatio-temporal position, neither benefited nor hindered IRn performance. Of course, faced with a negative result, we lack statistical certainty in accepting the null hypothesis, but some comfort may be derived from the exactness of the result.

The virtual absence of serial position effects implies, according to a trace strength hypothesis, that subjects did indeed rehearse (Section 3.124). That there is a small significant interaction between

Conditions and Presentation order for positive trials but not for negative trials suggests no more than that patterns of rehearsal and their dependence on presentation order may have differed somewhat between the two conditions. Nor do the serial position data provide any further information beyond a hint that the memory set pairs were indeed rehearsed as pairs.

May the main result be interpreted in any way other than as demonstrating that subjects are unable to make use of position information in IRn? There is at least one possibility. Suppose that a search process is involved in immediate recognition, but that there is a certain delay involved in selecting a subset (or 'finding' the appropriate starting point). It is conceivable that the experimental parameters were such that this delay either exactly cancelled out the saving of time achieved by restricting the search to a subset, or, so exceeded it that subjects deliberately avoided the use of a 'directed search' strategy. However, that the first possibility is unlikely is implied by this consideration: it is very improbable that the delay involved in selecting a subset should both be exactly cancelled out and be exactly equal for both of the subsets - yet there is no sign of an interaction between conditions and serial position. Another possibility is that the cue used may have been an inadequate means of specifying the subset. Temporal position was cued by spatial position in a manner that varied from trial to trial. Also, it might be rather more helpful to restrict the relevant subset to a single item rather than two. Both these points were taken account of in the next experiment, in which the utility of a position cue in IRn and CR was directly compared.

6.2. EXPERIMENT V. A comparison of the effects of a position cue
in IRn and CR.

In this experiment, four consonants were presented one after the other in a spatial array, with a given display location always corresponding to the ordinal position of the item in the list. Again there were two conditions: in Condition P, a positive probe item always occurred in the same display location it had occupied as a member of the memory set; in Condition N, the display location of the probe was entirely random. There were two parts to the experiment: in part (a) the task was recognition, with subjects pressing one key if the probe was a member of the memory set, another if not; in part (b) the task was contextual recall, and subjects had to respond with the item following the probe in the list. The same subjects served in both parts, half doing the two tasks in one order, half the other.

Before describing the experiment, it is worth mentioning briefly the results of an earlier version of it on which four subjects were tested. In that version, five-item lists were used, displayed sequentially across five positions in a horizontal row. In the CR task, only probes from positions 1-4 were presented; in the IRn task all five positions were probed. Presentation rate was 600 msec./item. The results were very striking, showing a vast difference between the effects of a position cue in the two tasks. In CR, mean RT was 1096 msec. in condition N and 815 msec. in condition P, a difference of 281 msec., significant at the $p < 0.025$ level. The corresponding error rates were 14.1% and 4.7%. In IRn, mean RT was 678 msec. in condition N and 648 msec. in condition P, a difference of only 28 msec., which did not approach statistical significance. Error rates

were 5.8% and 4% respectively. However, this experiment was flawed in certain important respects. In the CR task, the first item was never recalled. One subject pointed out that in condition P, it was not necessary to remember the first item at all, since the display location of the probe was adequate merely as a position probe. In condition N this is not the case, as the display location of the probe is random. Hence, the large difference in RT and errors between the two conditions might be due not to the position cue being superior to a contextual cue in terms of access (as the response buffer hypothesis requires) but merely to effective memory load differing between the two conditions. Secondly, that there was a difference (albeit small and non-significant) between the two conditions of the IRn task was puzzling. However, two of the subjects had unusually long RTs for IRn. Moreover, there was some suggestion of a facilitatory effect of presenting the probe in the same location as it had occupied in the list even in condition N, when it conveyed no position information: but the design of the experiment was inadequate to allow the evaluation of this effect, and its relationship with recency could not be established as rehearsal was uncontrolled.

Thus, in designing the present experiment, the following changes were made. Recall of the first item was required in CR whenever the last item was presented as a probe, so that in both conditions the subject had to remember all four items. Four-item rather than five-item lists were used to reduce the error rate. No-rehearsal instructions were given. In an attempt to reduce the variability in RT due to display location, a circular display array was used.

6.21. Method.

Eight subjects each served for four sessions, each lasting about 45 minutes. Every subject performed one task for two sessions,

followed by two sessions on the other task. Half the subjects did the IRn task first, half the CR task. In each half of each session there was one block of trials under condition N, one under condition P. For half the subjects the order of conditions in all sessions was PNNP, for the other half NPPN, balanced across order of tasks. The first two blocks of the first session in each part of the experiment served as practice, and data from them were discarded. A five-minute break was given half-way through each session.

A trial. The display locations and background display were as depicted in Fig. 6.1. A trial was initiated by the onset of the background display, and by a warning arrow, presented for 1 sec., pointing to the first display location. The four items of the memory set were then displayed in the sequence shown, each for 500 msec.. The shaded square at the centre of the display was then displayed for 100 msec. (This was intended to help guide the subject's eyes back to the centre of the display.) The probe item then followed in one of the four display locations: note that the probe display locations were slightly displaced with respect to those of the corresponding list items. All displays were separated by 100 msec., so that presentation rate was 600 msec./item and probe delay 300 msec..

In the IRn task, the subject pressed the right index key if the probe was in the list, the left index key if it was not.

In the CR task, the subject spoke the name of the letter following the probe item in the memory set, or the first item if the last was presented as the probe.

The response terminated the display; visual feedback was given for 1 sec. followed by a 2 sec. intertrial interval.

The four items comprising the memory set for each trial were

chosen randomly (without replacement) from the set: BDFGHJKLNRSZ.

A block:

Exp. Va, IRn. In each block there were 52 trials, of which the first 4 were warm-up trials, subsequently discarded. Of the remaining 48, 32 were positive trials. Of the 8 trials for each serial position probed, 2 were, in condition N, allocated to each display location of the probe. There were 4 negative trials for each display location. New blocks were generated for every session.

Exp. Vb, CR. In each block, there were 51 trials, of which the first 3 were warm-up trials, subsequently discarded. Each serial position was probed 12 times; in condition N these 12 trials were distributed equally among the 4 display locations. 8 blocks were generated for each condition, each used 4 times, once in each half session, always with different subjects.

Instructions. Subjects were instructed not to rehearse and to respond as fast as possible while avoiding errors. They were told:

"In condition P, if the probe item was in the list, it will always appear in the same quadrant of the display as it occupied in the list. You may find this helps you.....

(in CR) - to find the item.

(in IRn) - since you have only to decide whether the probe is the same as the one that occurred in that quadrant or not.

In condition R, the display quadrant of the probe tells you nothing, and is entirely random."

The subjects were shown examples, questioned to be sure they understood, and were reminded of the nature of the condition before each block.

6.22. Results.

Mean correct RT for each trial type was extracted for each block of trials. These values were then pooled in the analyses that follow. The results of each of the two parts of the experiment will be discussed separately. Since the analysis is somewhat complex, the salient features of the results are summarised at the end of the analysis for each task.

Exp. Va, IRn.

In this part of the experiment, the overall error rate was 2.3%. Mean RTs for each combination of condition and response type are shown separately for individual subjects and for blocks in TABLE 6.4. An overall comparison between the two conditions shows that RT was on average 13 msec. shorter in condition P than in condition N.

1) The overall effect of the position cue. A four-way analysis of variance was performed: Conditions x Blocks x Response Type x Subjects (random). There are main effects of both response type ($p < 0.001$, $F(1,7) = 31.1$) and blocks ($p < 0.001$, $F(2,14) = 28.0$). There is no main effect of Conditions ($F < 1$) and though TABLE 6.4 shows a slight increase in the effects of conditions with practice, the interaction of Blocks and Conditions does not approach significance ($F(2,14) = 1.6$). No other effects are noticeable save an interaction between Conditions and Response type, which is significant at the $p < 0.05$ level ($F(1,7) = 6.2$). The interaction is shown in Fig. 6.2. It may be seen that it is due largely to an advantage of some $\frac{3}{4}$ msec. for condition P on positive trials only.

2) The effects of display location. In order to understand the

Conditions \times Response Type interaction, we turn to a more detailed analysis of the positive trials of condition N. In a sense, the true controls against which to compare the positive trials of condition P are those N trials on which the probe was (by chance) displayed at the same location as in the memory set. The two classes of trial in condition N were therefore separated into two sub-conditions, mean RTs from which are also shown in Fig. 6.2. When a positive probe item was displayed in the same location as in the memory set (sub-condition N_S), mean RT was 584 msec. When it was not (sub-condition N_D) mean RT was 621 msec.. Mean positive RT in condition P was 579 msec.. Thus the small difference in positive RT between conditions P and N does not appear to have been due to the subject making use of the display location as a cue in condition P. It appears rather to result from some facilitatory effect when the probe was presented in the same quadrant as it occupied during presentation of the list, regardless of condition. Is this facilitatory effect significant? The data from condition N (pooled across blocks to obtain an observation based on at least six trials) were submitted to a three-way analysis of variance: Sub-conditions (N_S , N_D) \times Serial position probed \times Subjects (random). The effect of Sub-conditions is indeed significant at the $p < 0.025$ level ($F(1,7) = 10.9$). Neither Serial Position nor its interaction with Sub-conditions reach significance.

3) Serial position. Serial position data for condition P and sub-conditions N_D and N_S are shown in Fig. 6.3a. They are, however, virtually meaningless: serial position is confounded with display location, and, moreover, differently confounded for N_S and N_D . The mean RTs for each display location, collapsing across serial position probed, were computed for the positive trials of condition N and

used to correct the serial position data from the sub-conditions N_S and N_D for the effects of display location. The corrected serial position data are shown in Fig. 6.3b. It is clear that, if the correction procedure is acceptable, the facilitatory effect of presenting the probe in the same location as in the list (which I shall call the effect of 'location congruence') was most marked for the terminal item.

Unfortunately, the serial position data from condition P cannot be corrected in this way. An analysis of the negative trials from both conditions showed a significant interaction between Display Location of the probe and Conditions ($p < 0.05$, $F(3,21) = 3.07$), implying that attentional strategies differed in the two conditions. Nor may the data from negative trials of condition P be used to provide a correction. If, in condition N one compares the effects of display location for positive and negative probes, the pattern differs. (Conceivably, the decision criteria or bias are different for individual display locations.)

The significance of serial position effects in condition N was evaluated by collapsing the data across display locations and performing a three-way analysis of variance: Serial Position x Blocks x Subjects (random). The effect of Serial Position just fails to reach significance ($F(3,21) = 2.7$). The effect of Blocks is significant ($p < 0.01$, $F(2,14) = 9.0$), but there was no interaction between the two ($F < 1$).

The main findings from Exp. Va may be summarised thus:

a) IRn performance was neither improved nor hindered by cueing the subject to the relevant ordinal position of the memory set with a temporal and spatial position cue.

b) There was, however, a small but significant facilitatory

effect of 'location congruence' - presenting a positive probe in the same display location as it had occupied during presentation of the memory set.

c) The facilitatory effect of location congruence was considerably more marked for the last item in the list.

d) There was an effect of serial position which, though not quite significant, showed the usual recency effect and primacy on the first item.

Exp. Vb, CR.

Overall error rate in this part of the experiment was 3.9%. 2.9% of trials were lost through voice-key failures. Mean correct RT for each condition is shown separately for subjects and for blocks in TABLE 6.5.

1) The overall effect of the position cue. Mean RT was 306 msec. longer in condition N than in condition P. The correct RT data were subjected to a three-way analysis of variance: Conditions x Blocks x Subjects (random). Each observation was based on 48 trials. The difference between the two conditions is significant at the $p < 0.001$ level ($F(1,7) = 53.7$). The effect of Blocks is also significant ($p < 0.01$, $F(2,14) = 10.2$) but does not interact with Conditions ($F(2,14) = 1.7$).

2) The effect of location congruence. The correct RT data from condition N were subjected to a four-way analysis of variance: Blocks x Serial position x Display location of probe x Subjects (random). (Each observation is based on only 3 trials.) There are significant main effects of Blocks ($p < 0.01$, $F(2,14) = 8.5$) and Serial position

($p < 0.001$, $F(3,21) = 9.5$). No other F-ratio exceeds unity save that of the interaction between Serial position and Display Location which is highly significant ($p < 0.001$, $F(9,63) = 5.3$). The interaction is plotted in Fig. 6.4. It is quite evident that RTs in Condition N were shortest when the probe was displayed in the location it had occupied during presentation of the memory set. As in Exp. Va, one can separate the data from condition N into two sub-conditions. In condition N_S , when the location congruence holds, mean RT was 1013 msec.. In condition N_D , when it does not, mean RT was 1198 msec., a difference of 185 msec.. Is the effect of location congruence sufficient to explain away the difference between conditions P and N? The initial three-way analysis of variance was repeated, with the data from sub-condition N_S replacing those of condition N. The difference (157 msec.) in mean RT between condition P and sub-condition N_S is still highly significant ($p < 0.01$, $F(1,7) = 22.9$). Thus the effect of conditions may not be accounted for merely by the effects of location congruence.

3) Serial position. In condition P, serial position is ineluctably confounded with display location. In condition N it is not, and the four-way analysis above showed serial position to have a highly significant effect. Serial position data are plotted for sub-conditions N_S and N_D in Fig. 6.5, corrected for display location effects in the same way as in Exp. Va. The un-corrected data of condition P are also shown. There is a hint of a slight increase in the effects of location congruence with serial position, in that the difference between N_S and N_D increases slightly.

The main findings from the CR task may be summarised as follows:

a) CR was greatly facilitated by cueing the relevant position with the display location of the probe. The difference in RT was of the order of 300 msec..

b) There was also a large effect, within condition N, of display 'location congruence', such that RT was 185 msec. faster, when the display location of the probe was that at which it had occurred as a list member, than when it was not.

c) The effect of location congruence was not nearly sufficient to account for the overall effect of position cueing (as the much smaller effect in Exp. Va was). Hence, we can conclude that subjects were genuinely using and being aided by position information in condition P, in this task.

d) That the effect of location congruence was so much greater in Exp. Vb than in Va implies that it was not necessarily the same effect. However, there were some signs of an interaction with serial position of the same order in both tasks.

e) There was a highly significant effect of serial position.

6.23. Discussion.

In Exp. IV there was no effect on IRn of a position cue restricting the relevant subset of the memory set to two out of four items. In Exp. Va this finding was confirmed and extended. There was no overall effect on IRn of a cue telling the subject which one out of the four items of the memory set was relevant to the recognition decision. In this experiment, unlike Exp. IV, display locations were consistently associated with particular serial positions. Of course, it is still possible to argue that the time taken to 'select' a subset for directed search, or 'find' the starting point, may, in this experiment also, have cancelled out the RT advantage gained by a

directed search - but that the cancellation should be so exact in two rather different experiments is less than likely.

In contrast to the absence of any effect of a position cue on IRn, there was a very marked effect on CR in Exp. Vb. The size of the effect was moreover of an order compatible with the idea that the position cue enabled the subject to short-cut an otherwise slow search through the items of the memory set (see Section 4.1 for a hypothetical example). Hence the differential effects of providing a position cue on IRn and CR, and the direction and nature of the effects, were in complete accord with the predictions made by our two-store model. They were also, it seems, antithetical to Sternberg's position as interpreted in the introduction to this chapter. It does perhaps remain possible to argue that it is more efficient to start scanning the list from the beginning, ignoring any position cue, than to use the position cue, when the search is exhaustive but not when the search is self-terminating, but there seems to be no a priori reason why this should be the case.

Three features of the data remain to be accounted for.

1) The small difference between conditions P and N, on positive trials only, in Exp. Va turned out to be attributable entirely to a curious effect due to what I have called 'location congruence'. The effect was most marked for the final serial position. This suggests a facilitation of sensory encoding. I outlined in Section 3.13¹ the idea of a 'sensory component' of trace strength: the idea that feature-analysers intermediate between receptors and logogens might remain active for a brief period after the removal of the stimulus, and facilitate recognition in a sensory-specific manner. The present data suggest that it may be necessary to add the premise that the

sensory component is partially location-specific. Since subjects certainly moved their eyes in this experiment, I do not mean specific to retinal location, but location defined with respect to background contours. Sensory after-effects specific to objective location are not unknown - the phenomenon of visual persistence is one. The present finding would bear following up.

2) The effect of location congruence in Exp. Vb was, however, of a different order. The explanation seems quite simple. It is probable that at least some of the subjects, some of the time, carried over into condition N the strategy that they employed in condition P - that is, starting their search (of the contents of the response buffer) at the position indicated by the probe. In condition N, this would not have benefited them overall, but would have the result that their RT would have been shorter the nearer the display location was to the original position in the memory set - as Fig. 6.4 shows. This explanation does not deny the possibility that a small component of the location congruence effect in CR may not also have been due to the same (putatively sensory) facilitation postulated for IRn: there is a little support for this in that the location congruence effect in Exp. Vb was some 30 msec. greater for the last two than for the first two serial positions.

3) The overall shape of the serial position data in condition N of the CR task is in rough accord with previous evidence. Ignoring the case where the last item was presented as a probe, the shape is reminiscent of the serial position effect found for four subjects in Exp. I - with evidence for rapid retrieval when the last item was required as a response. That RT was relatively slow when the first item was required was presumably due to the fact that the probe and response were not adjacent in the response buffer: the subject had

to discover that the probe was the last item and then go back to the beginning. But there is one puzzle - if in condition P, RT was faster because the subject had some form of direct access to locations in the response buffer, why was the serial position effect found in condition N not entirely eliminated in condition P? The serial position effect was in fact attenuated quite considerably, and it must not be forgotten that serial position was confounded with display location. However, since the latter is not adequate to explain the remaining effect away, there are two possibilities. One is that access to locations in the response buffer is either not very precise or difficult, so that subjects were in fact obliged to search some of the time even when a position cue was given. The other is that some part of the serial position effect reflects storage rather than retrieval factors. The present data cannot distinguish these possibilities.

Summary.

Two experiments were performed to test whether access to representations in active memory is facilitated by a position cue. In Exp. IV, four-digit lists were presented sequentially in two pairs, each pair at a different spatial location. A recognition probe was presented which in one condition indicated by its display location the relevant pair or subset of the memory set, and in the other did not. No difference in RT or error rate between the two conditions was found. In Exp. V, the effectiveness of a position cue on IRn and CR was compared, the subjects and all relevant experimental parameters being identical for the two tasks. Four-consonant lists were presented sequentially in a spatial array of four locations, each

location being consistently associated with a list position. The probe in one condition always occupied the same location as during presentation of the list (if it was in the list), and in the other occurred randomly in one of the four locations. Again there was no effect of a position cue on IRn. There was evidence for a small facilitatory effect when the probe was presented in the same display location as in the list irrespective of whether it conveyed position information - which was interpreted in terms of the possible location- or background-specificity of a hypothesised sensory component. In contrast to the absence of effect on IRn, a position cue markedly reduced RT and errors in CR. This difference was argued to generate difficulties for a model assuming that the same active memory representations are consulted in both tasks, but to be completely in accord with the hypothesis of two forms of active memory. Access to logogens, which are content-addressable, is not expected to be facilitated by a position cue. CR however requires access to the contents of the response buffer, which is held to be more direct given a position cue than when only the item is given as a probe.

TABLE 6.1

Response type:	<u>YES</u>		<u>NO</u>		
Condition:	<u>P</u>	<u>N</u>	<u>P</u>	<u>N</u>	
Subject: S1	428	530	511	559	
S2	559	537	637	656	
S3	542	537	574	561	
S4	628	565	736	643	
S5	429	465	482	508	
S6	573	528	608	583	
S7	489	489	541	583	
S8	364	364	436	414	
Group mean:	501.5	501.1	565.5	563.4	(msec.)
Errors:	5.5	4.8	2.5	2.5	(%)

TABLE 6.1. Experiment IV - IRn.

Individual subjects' mean RT with (P) and without (N) the position cue for both positive and negative responses; group mean RT and error frequency.

TABLE 6.2

Response type:	<u>YES</u>		<u>NO</u>		
Presentation order:	R-L/T-B	L-R/B-T	R-L/T-B	L-R/B-T	
Condition P:	489 (6.0)	513 (5.0)	560 (1.5)	571 (3.5)	msec. (%)
Condition N:	506 (5.0)	496 (4.5)	557 (3.0)	569 (2.0)	msec. (%)

TABLE 6.2. Experiment IV - IRn. Group mean RT and error frequency to show the interaction of Conditions, Response type and Presentation order (i.e. Right-Left or Top-Bottom versus Left-Right or Bottom-Top)

TABLE 6.3

Serial position probed:	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	
Condition P:	510 (4.0)	497 (8.0)	509 (7.0)	489 (3.0)	msec. (%)
Condition N:	514 (6.0)	495 (3.0)	498 (6.0)	497 (4.0)	msec. (%)

TABLE 6.3 Experiment IV - IRn. Group mean RT and error frequency for each serial position of probe, positive trials.

TABLE 6.4

Conditions:		<u>P</u>		<u>N</u>		
Response type:		<u>YES</u>	<u>NO</u>	<u>YES</u>	<u>NO</u>	
Subject:	S1	593	732	635	699	
	S2	492	642	583	747	
	S3	563	678	572	720	
	S4	624	745	572	631	
	S5	565	694	594	694	
	S6	635	688	665	649	
	S7	569	793	641	755	
	S8	593	653	642	667	
Group mean:		579	703	612	695	msec.
Errors:		(2.0)	(1.8)	(1.7)	(3.6)	(%)
Block:	1	644	770	653	738	
		(3.1)	(1.6)	(0.8)	(3.1)	
	2	557	673	606	671	
		(1.6)	(1.6)	(3.5)	(3.9)	
	3	537	667	580	678	
		(1.2)	(2.3)	(0.8)	(3.9)	

TABLE 6.4 Experiment Va - IRn. Individual subjects mean RT with (P) and without (N) the position cue for both positive and negative responses; group mean RT and error frequency overall and separately for each of the three blocks.

TABLE 6.5

Conditions:	<u>P</u>	<u>N</u>
Subject: S1	917	1243
S2	769	1015
S3	802	1048
S4	804	1231
S5	788	896
S6	926	1254
S7	788	1062
S8	972	1466

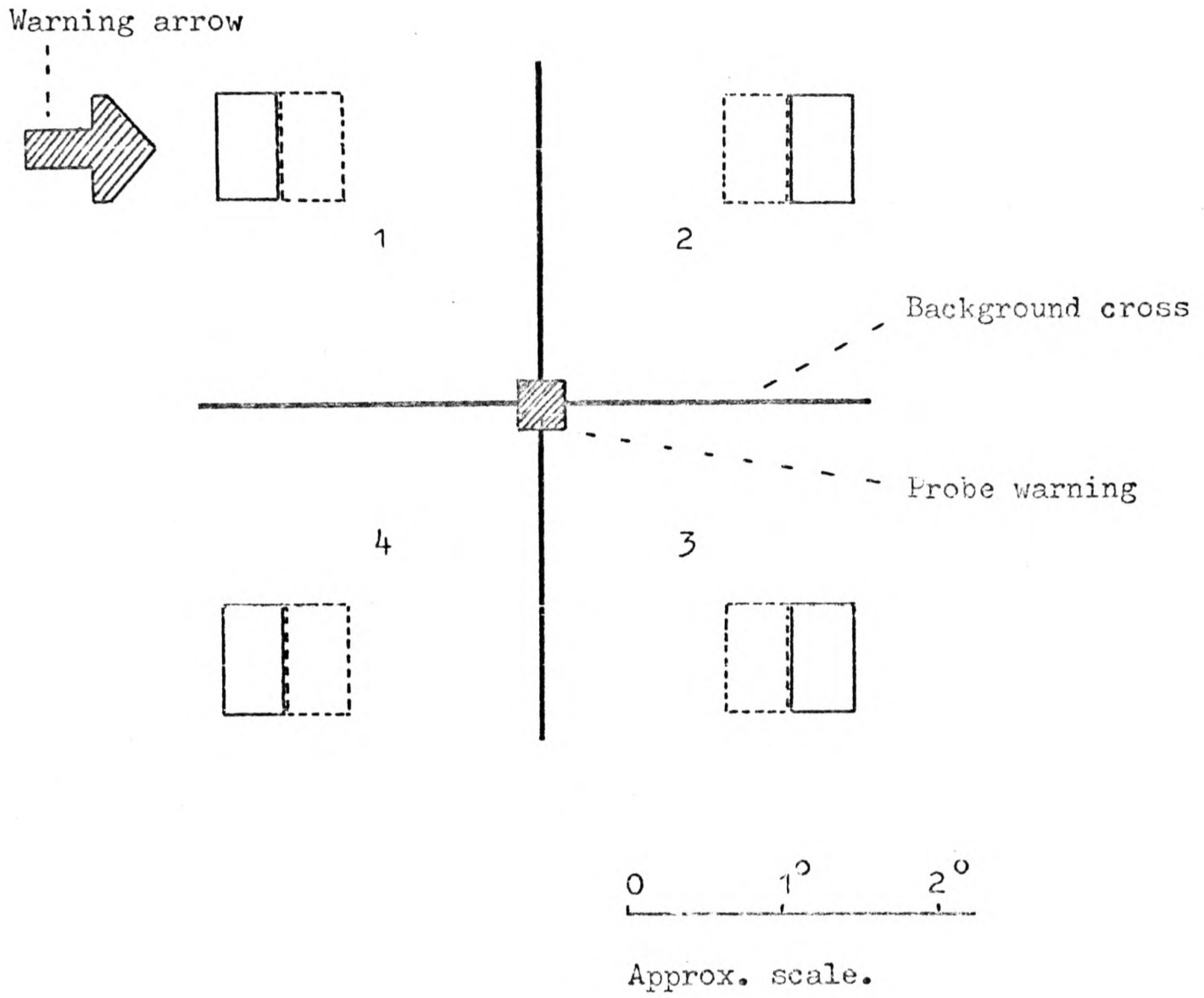
Group mean:	846	1152
Errors:	(2.4)	(5.3)

msec.

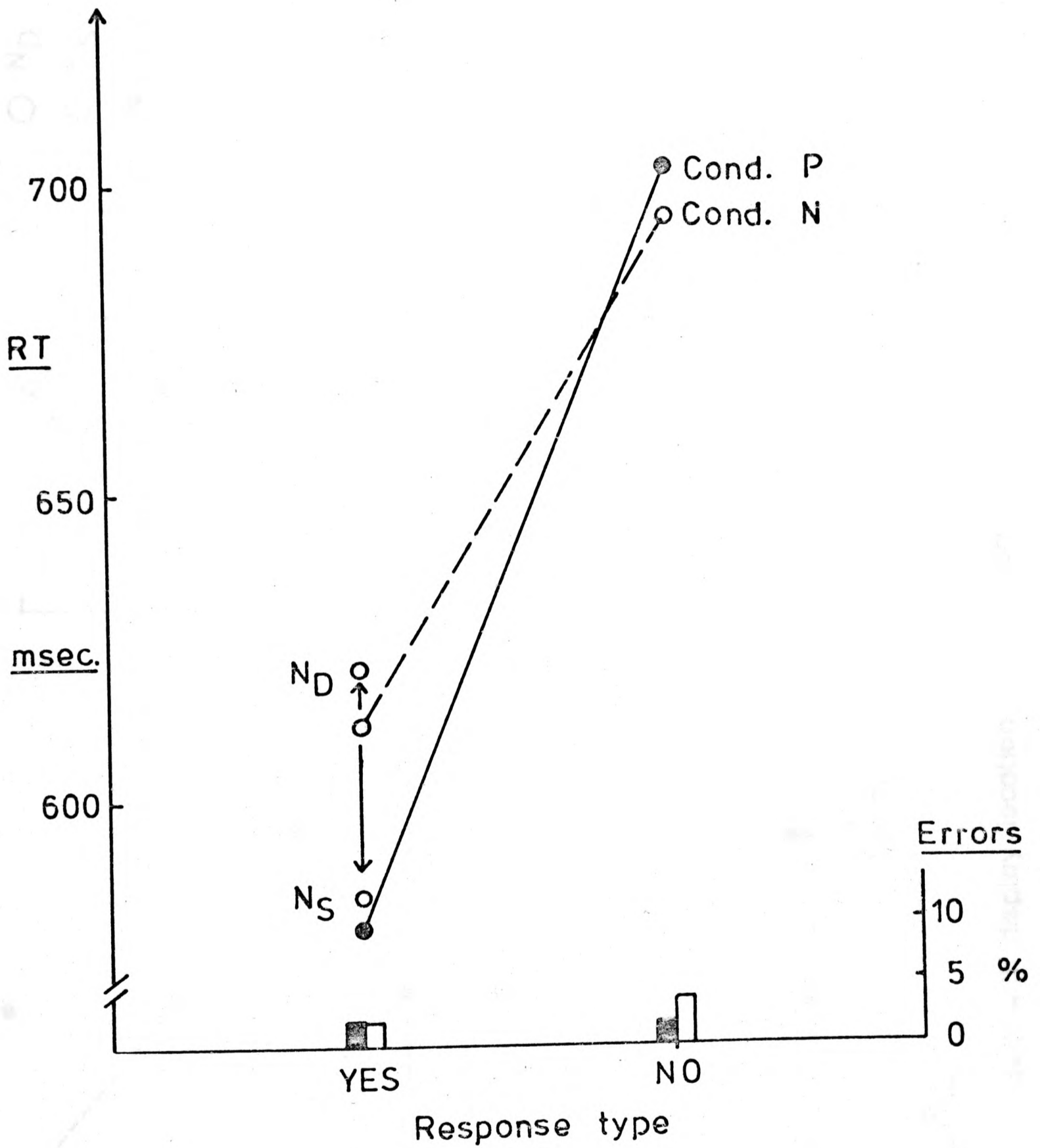
(%)

Block: 1	896	1227
	(2.8)	(7.0)
2	824	1101
	(2.0)	(5.3)
3	818	1128
	(2.3)	(3.8)

TABLE 6.5. Experiment Vb - CR. Individual subjects' mean RT with (P) and without (N) the position cue; group mean RT and error frequency overall and separately for each of the three blocks.



The background and display locations used in Experiment V. The hard-outline rectangles show the display location of the memory set items, the dotted-outline rectangles those of the corresponding probe items. Order of presentation was as indicated by the numbers in the quadrants.



EXP Va - IRn. Interaction of Conditions and Response type.

FIG. 6.2

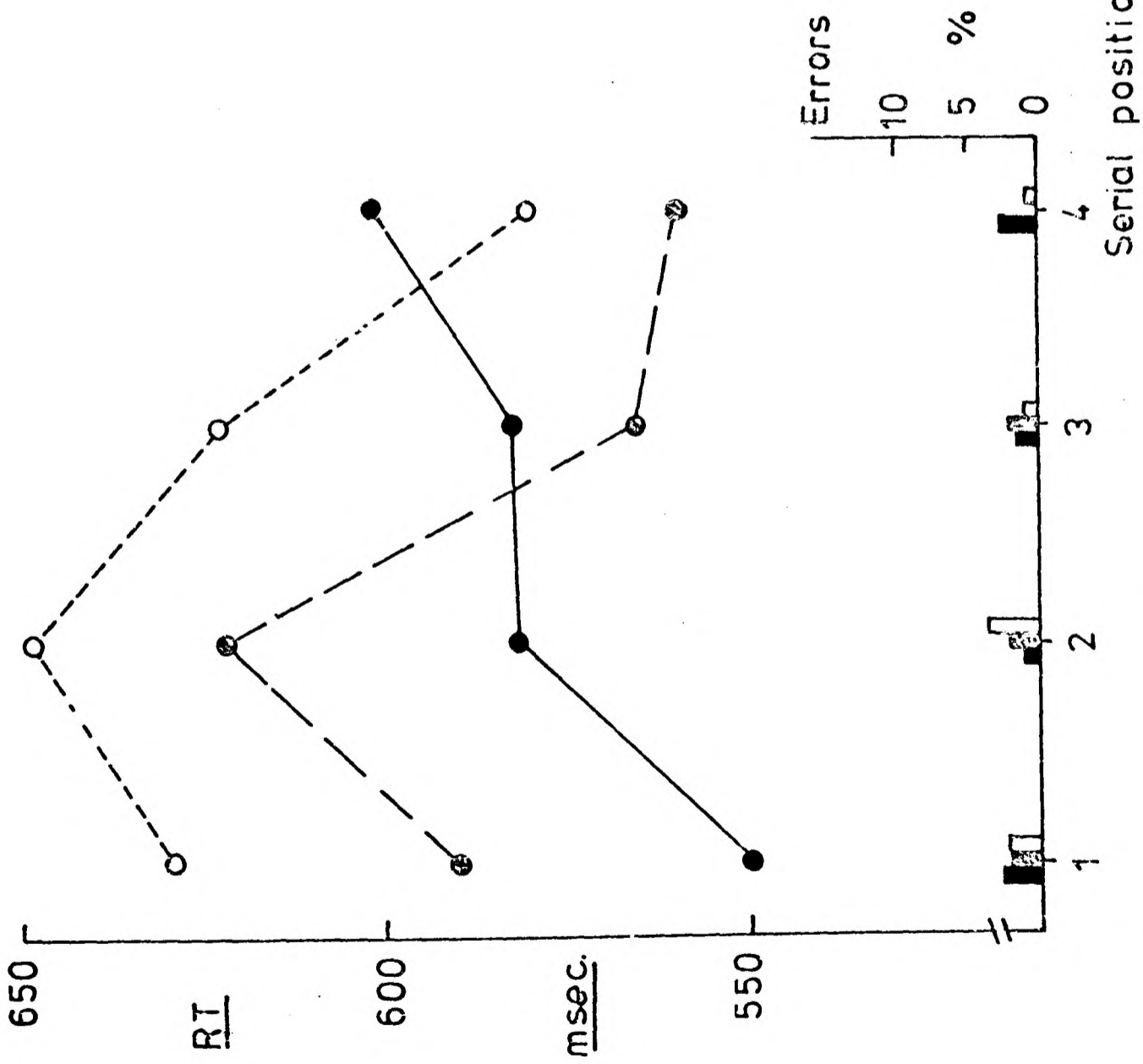


FIG. 6.3a. EXP. Va - IRn.

Serial position [confounded with display location].

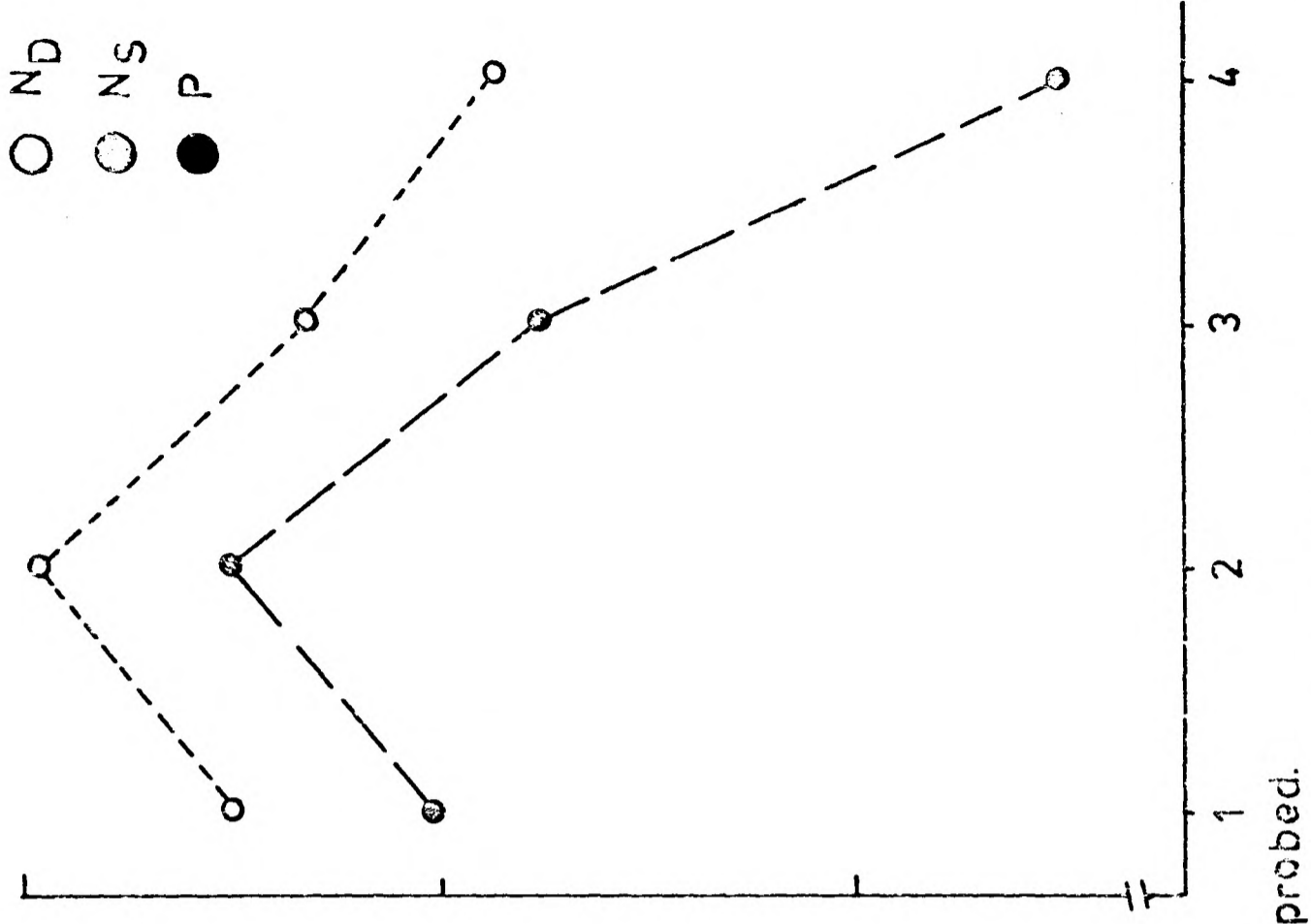
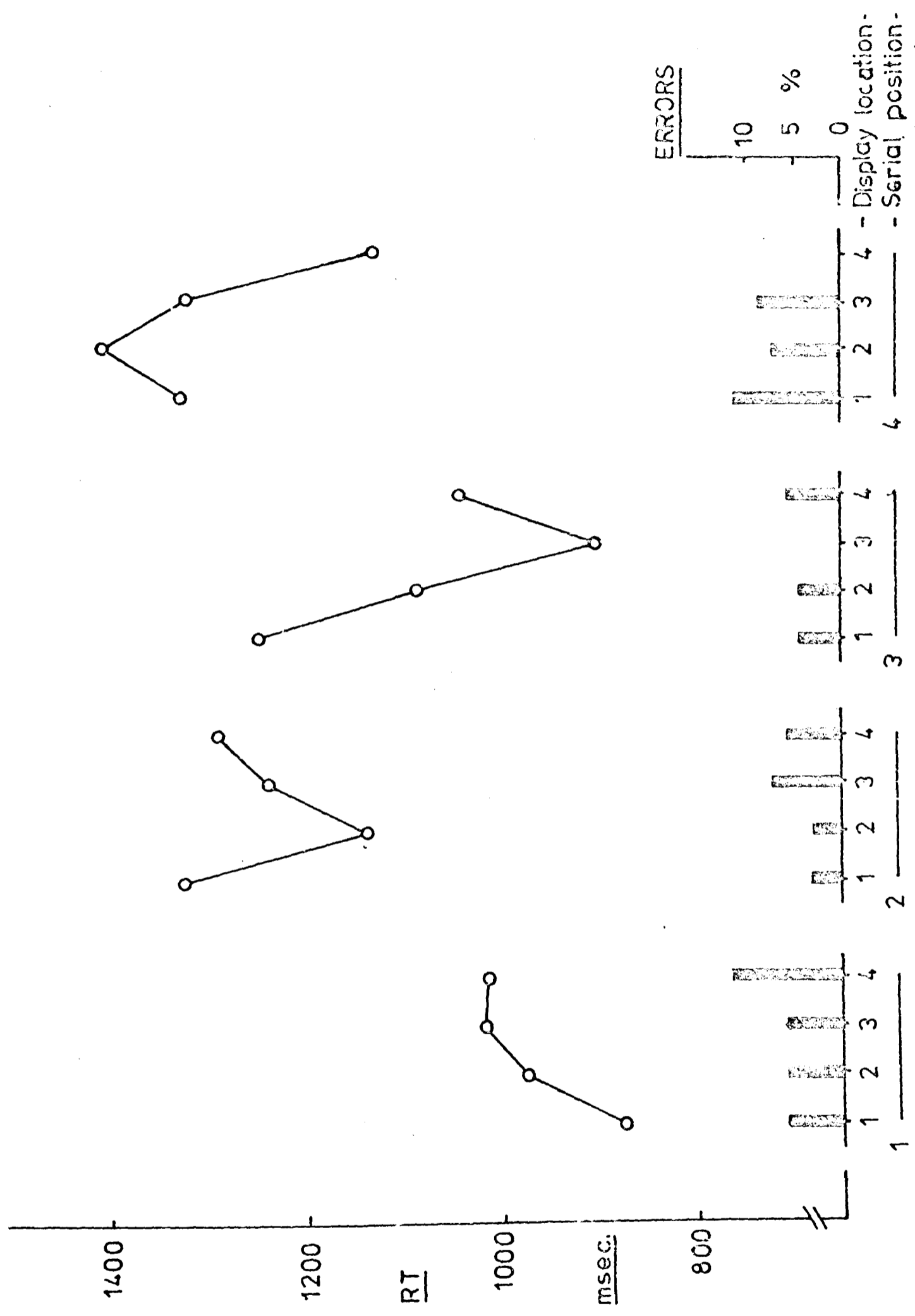


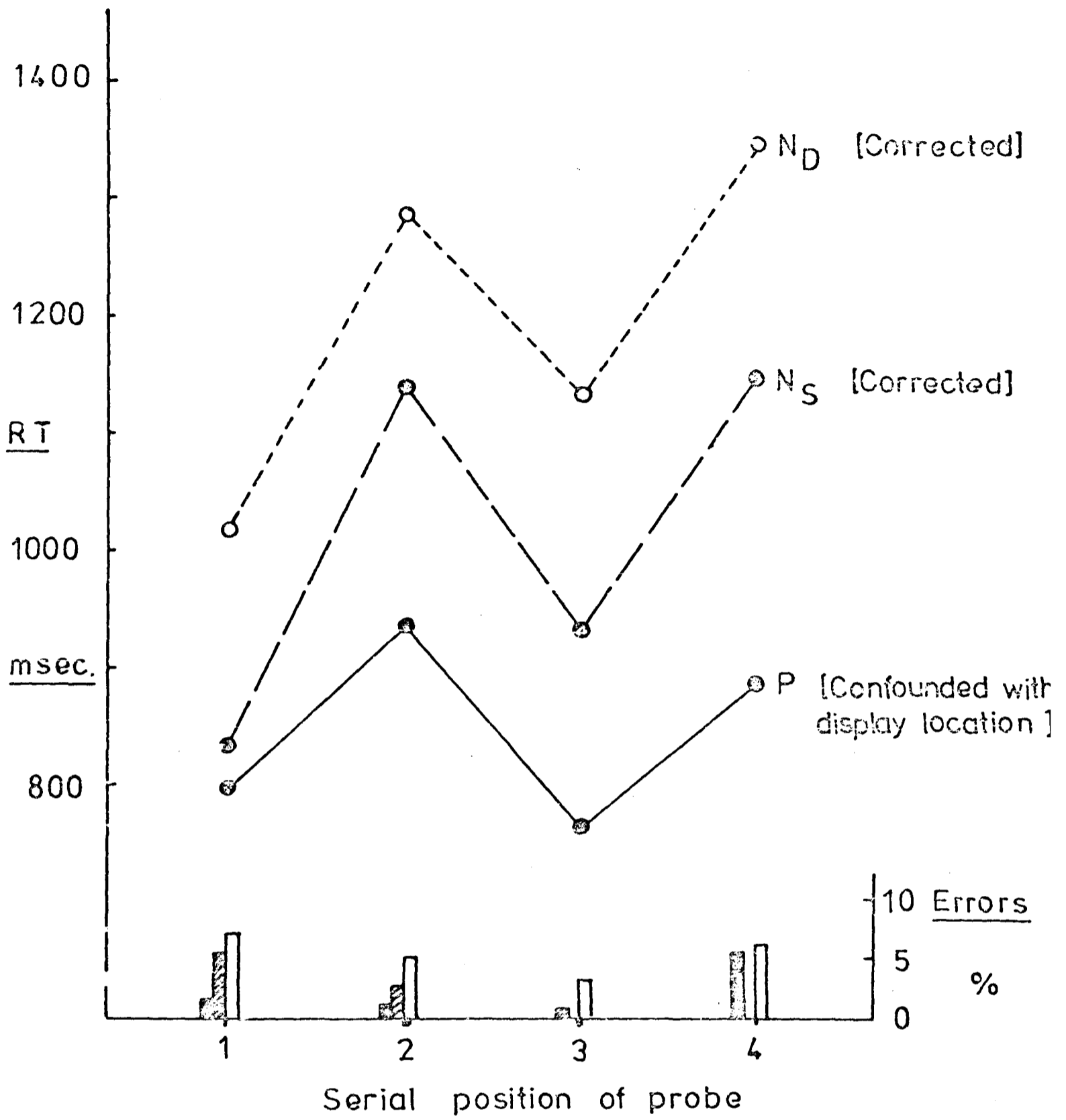
FIG. 6.3b. EXP. Vb - IRn.

"Corrected" serial position, Condition N.

FIG. 6.4



EXP. Vb-CR. Interaction of serial position and display location - Cond. N.



EXP. Vb - CR.

FIG. 6.5

CHAPTER VII. Immediate recognition and recency.

Introduction.

7.1 EXPERIMENT VI. Immediate recognition and recency.

7.11 Method.

7.12 Results.

7.13 Discussion.

Summary.

CHAPTER VII. Immediate recognition and recency.

The five experiments reported in the previous two chapters have provided a firm basis for preferring the hypothesis of two forms of post-categorical active memory to any available single-store hypothesis. It has been shown that the provision of a position cue facilitates contextual recall, and that acoustic similarity influences both contextual recall and probed location; neither factor had comparable effects on item recognition. The nature of the effects on the former group of tasks, whose common characteristic is that they require reference to stored order information, were largely in accord with the premise that the order of items is represented in a non-associative short-term store with the properties proposed in Section 1.4 for the response buffer. However, the lack of effects on item recognition provides only negative evidence on the nature of the representations mediating this task. In the remaining experiments to be reported, I attempted to investigate this question more directly, in relation to the theoretical issues raised in Section 3.1. In the present chapter, I shall describe a somewhat ambitious experiment on the dynamics of the representations underlying immediate recognition. In the following chapter, the question of their format will be considered.

7.1. EXPERIMENT VI. Immediate recognition and recency.

In Section 3.11 (to which reference may be necessary) five models of immediate recognition were outlined and discussed in some detail. On the basis of evidence available in the literature, I concluded tentatively in favour of the trace-strength inspection model elaborated from that proposed by Baddeley & Ecob (1970). According to this model, recently-presented items are represented at permanent content-addressable representations (which I have identified with Morton's 'logogens') as item-traces whose strength declines as a decelerating function of the time/items since presentation (the 'recency function'); RT is an increasing, but otherwise unspecified, function of the difference between the strength of the 'item-trace' addressed by the probe, and a decision criterion. That both the 'trace strength models have a 'forgetting axiom' - the assumption of non-linear decay - distinguishes them sharply from the scanning models, which include no explicit assumptions about the dynamics of the representations. Predictions based on this axiom are therefore crucial and may be tested by manipulations of recency (see Section 3.125). In Exp. VI I attempted to manipulate independently the recency both of negative probes and of the items of the memory set. The points to which the experiment was directed may be summarised under the following headings:

1) The recency of the negative probe. In Section 3.125(ii) I pointed out that the effects of manipulating the recency of the negative probe provide a crucial test between the models.

a) Models I, II and parallel scanning models: According to any model which posits a search, serial or parallel, of items only of the

positive set, the recency of a negative probe should have no effect on the search process. No effect on the slope of the set-size function is therefore expected. Should there be an effect on the intercept? It might be argued that the more recent the probe, the more rapidly it should be encoded, so that RT should be shorter for a recent negative probe through a reduction in intercept. On the other hand, it might conceivably be claimed that the subject might not expect a recently presented item to appear as a negative probe, so that RT could be longer for a recent probe. Either way, there should certainly be no interaction with memory set-size.

b) Models III and V. The trace strength inspection model clearly predicts that, provided the subject does not know in advance what kind of probe he is to receive, a more recent negative probe will take longer to reject since its strength will be closer to the criterion. Also, since the criterion will cut off a larger piece of the tail of the negative items' distribution if they are recent, more errors will be made the more recent the negative probe. Whether there will be an interaction with set-size or not must depend on the nature of the decision function. If differences in strength near the criterion have more effect on decision time than differences far from the criterion, then the magnitude of the effect should increase with set-size. So a monotonic interaction of recency and set size, while not necessary, is certainly possible. As was pointed out in Section 3.11, as regards the manipulation of recency, Model V may be considered merely a rather cumbersome discrete-state version of Model III, and makes the same predictions.

c) Model IV. According to this model, negative decisions are made by default; since the strength of a negative item is thus not inspected, it cannot influence the decision, and no effect of recency

on negative RT is predicted. However, there should still be more errors when the probe is recent, since its prior 'advantage' in strength will make it more likely to exceed the criterion K before the criterial time T_c is up.

The evidence available to date is restricted to two studies; those of Zechmeister (1971) and Darley & Arabin (1972), who both found that it took longer to reject recent negative probes than probes which had not occurred before in the experiment. The effect increased with set-size in both cases. The means chosen for manipulating recency in both these studies was somewhat extreme. Memory sets were drawn without replacement from a large set of words so that each was used only once. Negative probes were drawn either from immediately preceding trials or had never occurred before. Thus non-recent probes could be rejected on the basis of whether or not they had occurred in the experiment; in fact Darley & Arabin do interpret their finding in terms of a familiarity-judgement-or-search version of Model V. Zechmeister's experiment also has the disadvantage that recency is a between-subjects variable - so that there are problems of criterion change to muddle the issue and only two set sizes were used. In the present experiment, the means chosen to manipulate recency was rather more subtle. Memory sets were chosen from a set of 16 consonants. A 'recent' negative probe on trial n was defined as one of the four letters presented prior to the probe of trial n-1. A 'novel' negative probe was an item which had not occurred as one of the ten letters presented prior to and including the probe of trial n-1. Since all 16 letters were presented equally often, and very frequently, expectancy effects would be expected to be minimal, and recency was not confounded with occurrence in the experiment.

2) Recency of the memory set items. According to the two trace strength models, decay of trace strength is non-linear and decelerating. The slopes both of the set-size function and of the serial position function (provided rehearsal is controlled) are functions of the rate of decay. For a given set size, the criterion is held to be constant, and the non-linearity of the recency function may be seen in that of the serial position curve. As regards the set-size function, while the non-linearity of the recency function need not result in a non-linear set-size function if, for instance, prior items influence the initial increment of strength (see the discussion on p.103a), the slope of the set-size function will still depend on the rate at which strength decays as a function of time and/or subsequent items. Thus if presentation of the probe is delayed, while rehearsal is prevented, rate of decay should have slowed, and the slope of the set-size function and the serial position function should both be smaller than when test is immediate. This prediction is tested in the present experiment.

What predictions do the scanning models make? As they are stated by Sternberg (1966, 1969a), no particular effect of a small delay on slope is predicted, provided that the capacity of active memory is not exceeded. If one were to introduce an assumption that items in active memory decayed, then the degradation resulting from a probe delay should increase comparison time (cf. Sternberg, 1967b) and the slope should increase.

The means chosen for delaying the probe while preventing rehearsal was as follows. In one condition (I), as soon as the memory set had been presented, an interpolated vowel was presented for the subject to name as fast as he could; his response triggered the presentation of the probe. In condition N, the probe was presented immediately

after the memory set. Thus in condition N there was no delay, and in condition I there was a totally filled interval of about $\frac{1}{2}$ sec.. As it turned out, the interval was probably on the short side. Nevertheless, the portion of the recency curve involved should be displaced by the delay away from the point at which decay is fastest (as a result, perhaps of the 'sensory component' - Section 3.131). A pilot study seemed to indicate that the delay was large enough to produce a marked change in slope in the direction predicted by the trace strength hypothesis.

3) First-item primacy effects. It will be recalled (Section 3.124 if it is not) that though marked and appropriately non-linear recency effects have typically been obtained when rehearsal is controlled (see also Exp. II and Exp. V) there is frequently visible in the serial position data a small primacy effect: mean RT is often faster when the first item is probed than when the second is probed. It was suggested in Section 3.124 that this was an acquisition effect: i.e. that the first item received a greater initial increment in trace strength than the others due to lack of prior activity in the sensory feature-analysers. This may be tested by preceding the first item with an irrelevant prefix. This should share physical features with the memory set items so as to activate the appropriate feature-analysers, but should not be a member of the stimulus ensemble from which the memory set items are chosen. In this experiment the memory set was always preceded by an irrelevant vowel, but never the same vowel that appeared as the interpolated vowel in Condition I.

4) There were several other predictions derived in Section 3.12 from the various models, which the present experiment was not specif-

ically designed to test, but to which the data may nevertheless be relevant. The reader is referred

a) to the discussion of the relation between RT variance and set size in Section 3.122.

b) to the prediction derived from Model III concerning the relation between the ratio of the slopes of the positive and negative RT set size functions (the 'slope ratio') and the way in which the different classes of errors should increase with set size (Section 3.121).

Thus, in Exp. VI, memory lists of one to four consonants were presented, each list length in a separate block of trials, at a brisk rate, preceded always by a vowel. A recognition probe was presented either immediately (condition N) or after an interval filled precisely by the naming of a vowel (condition I). Half the negative probes were 'recent', half 'novel'. Since the double task was quite demanding and since the RT differences expected were not large, I made a point of using subjects who were proven veterans of at least one experiment of mine, and they were given two full sessions of practice before the data presented here were recorded.

7.11. Method.

Eight experienced subjects each served for two practice and four experimental sessions. In each session, they were given one block of trials under condition I and one under condition N. Each block was divided into four sub-blocks, one for each memory list length m . Each sub-block was preceded by a display informing the subject of the value of m for that sub-block. It began with two dummy trials, later discarded, which were followed by 2^4 experimental trials. A complete block therefore contained 10^4 trials and the subject was

permitted no break between trials. Each block lasted for approximately 13 minutes, and a five minute rest was allowed between blocks.

Design. The order of the two delay conditions and of the sub-blocks within each condition was balanced both within and across subjects. Each of the 2^4 possible orders of sub-blocks was combined once with each of the 2 orders of conditions in the 48 sessions of the experiment. These were allocated so that:

1) in sessions 3-6, the experimental sessions, the order of conditions and of sub-blocks was perfectly balanced by means of Latin squares both across each session and within each subject over the four sessions.

2) in sessions 2, 5 and 6, the order of conditions and sub-blocks was the reverse of that in sessions 1, 3 and 4 respectively. Data from sessions 3 and 6 and from sessions 4 and 5 were combined in the analysis to give two observations per cell; thus the contribution of practice effects to the error variance was minimised.

The memory lists and probes were chosen randomly from the set of 16 consonants: CDFGHJKLNPRSTXYZ, with the following constraints:

1) There were no repeats in a list.

2) Each serial position was probed equally often within each sub-block.

3) Positive and negative probes in a sub-block were equally frequent.

4) Half the negative probes in a sub-block were 'recent', half 'novel'.

A novel negative probe was one which had not occurred as one of the ten consonants presented prior to the present trial. A recent negative probe was one of the four consonants presented prior to

the probe of the previous trial. For the purposes of this definition, the same consonant presented as list item and as probe counted as two items.

5) The probe from the previous trial was never presented on the present trial. Recency, was, it should be noted, defined in terms of the number of intervening items. Since the intertrial interval was constant, and memory set-size blocked, the average difference in recency in terms of time between novel and recent negative probes varied with set size. The actual values (to the nearest sec.) are shown in the following table:

Set size (m) =	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	
Recent probes were presented on average:	11	10	9	8	sec.
Novel probes had not been presented for:	33	27	23	23	sec.

- prior to the probe of the present trial

The irrelevant prefix preceding the memory set was chosen randomly from the vowels AEIOU, but was never the same as the interpolated item, in condition I. The interpolated vowel was chosen randomly from the same five vowels, with the constraint that each vowel was used approximately equally often within each sub-block.

Display. The background display served as a reference frame for two locations separated horizontally by about $1^{\circ}20'$ of visual angle. Approximately $2^{\circ}30'$ beneath the centre of these was a third location. The locations were each indicated by four points at the corners of a square with the display location at its centre.

A trial: was initiated by the onset of the background display. This was followed after 500 msec. by the prefix and m letters of

the memory set presented one after the other in the top left-hand location, each for 400 msec. followed by a gap of 100 msec..

Condition N. The probe was presented in the bottom location 100 msec. after the offset of the last item. The subject pressed the right-hand index key if he thought the probe item was a member of the memory set, the left-hand index key if not. A feedback display consisting of a tick if the response was correct, a filled square if not, then followed for 1 sec. succeeded by a $2\frac{1}{2}$ sec. inter-trial interval.

Condition I. 100 msec. after the offset of the last item, an interpolated vowel was presented in the right-hand top location until the subject named it. He was required to preface the vowel name with the phoneme /b/, so as to reduce the variability due to voicing onset. His response triggered the voice key which caused the probe item to be presented immediately in the bottom location. The sequence of events was then as in condition N, save that the inter-trial interval was only 2 sec., thus equalising the probe-probe interval across delay conditions. Following any trial on which the vowel naming task was unsuccessful, as a result of error or voice-key failure, the experimenter was able, without interrupting the sequence of trials, to instruct the computer to ignore that trial.

Instructions. "1) Please do not attempt to rehearse or group the items - at any moment attend only to the item most recently presented.

2) The first item in every list will always be a vowel; it is therefore completely irrelevant to the task, but nevertheless please treat it as you do the other members of the list - it is intended to control for any advantage the first item gets because of nothing preceding it.

3) In condition I, the task of naming the interpolated vowel is intended to take up all your attention, so please concentrate on naming it as fast and smoothly as you can.

4) Please try and respond as rapidly and consistently as you can, while avoiding errors."

Subjects claimed to understand and comply with these instructions. The double task in condition I proved difficult for the first trials, but a steady rhythm was soon achieved and subjects made effectively no errors on the vowel naming part of the task.

7.12. Results.

In order to check that the duration of the delay in condition I did not change with set-size, the mean and pooled standard deviation of the vowel-naming RTs were computed separately for each set-size, as shown in TABLE 7.1. The largest difference between means is 10 msec. which is no more than 2.3% of the total delay. It therefore seems safe to assume that the delay introduced by the vowel naming task in condition I was comparable over different set-sizes.

Mean item recognition error rate was 3.8% overall. Correct RTs were pooled across sessions 3 and 6 and across sessions 4 and 5, giving two observations per trial-type per subject. Fig. 7.1 shows the set-size functions for each combination of response type and delay condition. The correct RT data were subjected to a four-way analysis of variance: Response type x Conditions (I or N) x Set size x Subjects (random). Each of the 2 observations per cell was based on 24 trials.

1) The effects of set-size on mean RT. There is an effect of set-size significant at the $p < 0.001$ level ($F(3,21) = 39.2$). An

analysis of trend shows this to be attributable to a highly significant linear trend ($p < 0.001$, $F(1,21) = 113.3$) which accounts for 96.5% of the variance due to this factor. There are, however, signs of non-linearity in the data (see Fig. 7.1) but the quadratic trend does not reach significance ($F(1,21) = 3.04$). However, if positive and negative RTs are submitted separately to equivalent analyses, the data for the positive RTs show a quadratic trend significant at the $p < 0.025$ level ($F(1,21) = 6.2$) and the linear trend accounts for only 91.7% of the variance. There is no quadratic trend in the negative RTs ($F < 1$). The non-linearity of the positive RT function may be seen to result from fast RTs when $m = 1$ (Fig. 7.3).

2) The effects of delay. Set-size functions for the two delay conditions (combining response types) are shown in Fig. 7.2. The overall effect of delay is significant ($p < 0.025$, $F(1,7) = 8.6$) but this absolute difference is not very meaningful; what interests us is whether there is an interaction between delay Condition and Set size. Such an interaction is visible in the data; the slope of the set-size function is smaller in Condition I than in condition N. The direction of the interaction is that predicted by the trace strength inspection model. If linear regression lines are fitted to these data, the slopes are 30.1 msec./item when there was a delay (Condition I), 42.2 msec./item when there was not. (In view of the departure from linearity of the set-size function, the regression coefficient must be regarded as no more than a useful descriptive statistic.) Unfortunately, although the slope difference is quite large compared to some reported in the literature, the interaction between Conditions and Set-size fails to reach significance ($F(3,21) = 1.88$). This is due to a large interaction of this interaction with subjects ($p < 0.001$,

$F(21,128) = 3.43$). In Figs. 7.5a, b the Conditions x Set-size interaction is shown separately for individual subjects. It may be seen that while the data from five subjects show the interaction quite clearly, the data from subjects 2, 3 and 8 do not. The same may be seen in TABLE 7.2, in which regression slopes computed for individual subjects' data are shown.

3) Response type. Positive RTs were faster than negative RTs by an average of 67 msec. ($p < 0.001$, $F(1,7) = 74.4$). There is an interaction between Response Type and Set-size ($p < 0.05$, $F(3,21) = 3.82$) due presumably to the non-linear trend in the positive RT set-size function already mentioned. As may be seen from TABLE 7.2, the negative/positive slope ratio for individuals ranges from 0.35 to 1.5 though the two functions are effectively parallel in the group data. Response type also interacts with Conditions ($p < 0.05$, $F(1,7) = 7.2$) as may be seen in Fig. 7.1.

No other effects approach significance in the above analysis. In order to assess the effects of the recency of the negative probe, the correct negative RT data were submitted to a four-way analysis of variance: Conditions (I,N) x Set size x Recency ('novel' versus 'recent') x Subjects (random). Each of the two observations per cell was based on 12 trials.

4) The effects of recency of the negative probe. The results of the analysis are very clear. There are highly significant effects of Recency ($p < 0.001$, $F(1,7) = 32.4$), Set-size ($p < 0.001$, $F(3,21) = 33.1$) and their interaction ($p < 0.001$, $F(3,21) = 16.7$). No other effects approach significance. The interaction is shown in Fig. 7.4.

Recent probes were rejected more slowly and less accurately than novel probes, and the difference increased monotonically with set size.

Positive RT data were submitted, separately for each set-size ($m = 2, 3$ and 4) to three-way analyses of variance: Conditions (I, N) x Serial Position x Subjects (random). The data, together with negative RT data, are summarised in Fig. 7.6.

5) The effects of serial position. For each set size, there is a significant effect of serial position (for $m = 2$, $p < 0.025$, $F(1,7) = 8.98$; for $m = 3$, $p < 0.001$, $F(2,14) = 16.1$; for $m=4$, $p < 0.001$, $F(3,21) = 12.1$). The attempt to eliminate a primacy effect by the use of an irrelevant prefix seems to have been successful, save for $m = 3$, condition N, where RT was slightly faster when the first item was probed than when the second was probed. In general, the serial position data exhibit monotonic recency, accelerating with serial position. Only one subject (S_3) deviated from this pattern. Her serial position data are shown in Fig. 7.8. Her data show general increases in RT with serial position. However, her error rate was unusually high and included a lot of fast anticipations. She was, moreover, one of the subjects who failed to show a Conditions x Set size interaction. It may be seen that the distribution of her errors with respect to serial position was markedly different in the two delay conditions. I have no idea what she was doing.

6) The interaction of serial position and conditions. In all cases the slope of the serial position function was greater for Condition N than for Condition I, as predicted by the trace strength inspection theory. However, only in the case of $m = 2$ is the Serial

position x Conditions interaction significant (for $m = 2$, $p < 0.001$, $F(1,7) = 19.2$; for $m = 3$, n.s., $F(2,14) = 2.3$; for $m = 4$, n.s., $F(3,21) = 2.23$).

7) RT variance and set size. Since the models outlined in Section 3.11 make differential predictions about the relationship between RT variance and set-size (see Section 3.122) RT variance was computed, separately for each set-size, response type (separately for the two kinds of negative probe) and subject. The mean variances are depicted in Fig. 7.7. The variance of positive RTs shows a monotonic but not very linear increase with set-size. The behaviour of negative RT variance appears to depend on whether the probe was novel or recent. If the probe was novel, the data show no overall increase in variance with increasing set-size. On the other hand, if the probe was recent, negative RT variance may be seen to increase with set-size at a rate considerably in excess of that for positive RT variance. However, it should be mentioned that individual data were highly variable, and I did not think a statistical test worth doing.

8) Relationship between slope ratio and errors. It was suggested in Section 3.131 that, according to Model III, if the slopes of the positive and negative set size functions were parallel (as they are here) then all other things being equal, false negatives should increase faster with set size than false positives. But Fig. 7.3 shows that the opposite is the case. However it may be seen from Fig. 7.4 that this is due only to the increase in false positive frequency when the probe is recent; when the probe is novel, the rate of increase is slightly smaller for false positives than for false negatives. All other things are therefore not equal, as the

interaction between recency and set-size has already suggested.

9) Relationship between negative and slowest positive responses. In Section 3.124(ii) it was pointed out that, according to Model IV, the 'default' model, the longest mean positive RT in the serial position curve should be shorter than the mean negative RT (however novel the probe). Ideally, to test this, response allocation and handedness should be carefully balanced, which was not the case in this experiment. However, it may be noted that although 7 of the 8 subjects were right-handed, and all responded positively with their right hand, mean negative RT for novel probes was still as fast or faster than the longest positive RT when $m = 2, 3$ and 4 . (See Fig. 7.6.)

7.13. Discussion.

1) Recency of the negative probe. The clearest findings from Exp. VI were:

a) subjects rejected a 'recent' negative probe more slowly and less accurately than a 'novel' one.

b) this effect interacted strongly with set-size such that the difference between the two classes of probe increased with set-size.

I have argued that any model postulating that only the items of the memory set are scanned can account for the first of these results only by the ad hoc assumption that a recent negative probe was encoded more slowly than a novel one. But there should then be no interaction with set-size. Might this interaction be accounted for in terms of an effect on encoding of time-based rather than item-based recency? For it may be seen from the table in Section 7.11

that a 'recent' negative probe was on average more recent in terms of time the larger the set size. However, this was also true of 'novel' probes, and in fact the average difference in time of occurrence between the two classes of probe actually decreased with set-size. A scanning model could therefore be accommodated to the interaction found only by an assumption of the general form: there was a tendency to encode a negative item more slowly the more recently (in terms of time) it had occurred, provided that it had occurred no more than about 11 secs. ago. I find this implausible, particularly since a difference in recency between 11 sec. and 8 sec. would have had to generate an average encoding time difference of 60 msec..

Another possible amendment to a scanning model which would enable it to account for (a) is the addition of a 'displacement axiom'. That is, it might be assumed that the subject does not start each trial with a clear 'active memory', but that items on the present trial merely displace earlier items, so that there may linger in 'active memory' items from previous trials; if one of these items is presented as a probe, the difficulty of sorting out the present memory set items from past items may lead to a longer RT or an error. However, aside from the complications this axiom would introduce into the standard predictions of the scanning models, it also predicts an interaction of recency and set-size in the wrong direction. For, the smaller the set-size, the fewer items from previous trials should be displaced from 'active memory', so that the effect of recency of the negative probe should decrease with set-size, which is the opposite of what in fact occurred.

It should also be noted that were one to postulate a short-term version of Theios' (1972) model for the constant-set recognition paradigm (Section 2.232), according to which a self-terminating scan of both positive and negative items is performed in an order

determined by their recency, an analagous objection to the last would apply with respect to the direction predicted for the interaction; moreover, according to such a model, recent negative probes should have been rejected faster than novel probes.

The trace-strength default model (Model IV) is also counter-indicated by finding (a) since no effect of negative probe recency on RT is predicted.

It seems, then, that this result, which confirms and extends those of Zechmeister (1971) and Darley & Arabin (1972) critically favours Model III, the trace-strength inspection model, or a mixed model (Model V). How is the interaction to be accounted for? There are at least two non-exclusive possibilities.

(i) It is possible that time since presentation as well as number of intervening items is important as a determinant of amount of decay (cf. Wickelgren, 1970b). Given that decay may be assumed to be non-linear and to a floor, it may be that the difference in strength produced by the dependence of time-based recency on set-size in this experiment (see table, Section 7.11) applied only to 'recent' negative probes. Not only does this assumption account for the interaction (b), but it may also explain the greater increase in false positive responses relative to false negatives with increasing set-size for recent probes - see Section 7.12 (9). For, the larger the set-size, the nearer to the criterion the strength distribution of recent negative items would be, relative to that of novel negative items; hence, more errors would be made.

(ii) The effects of recency on negative RT depend critically on the nature of the decision function. If differences in strength near the criterion have a greater effect on decision time than equivalent differences far from the criterion then the effect of a given manipulation of recency must depend on set-size in the manner

found, since the greater the set-size the lower the subject should set his criterion (see Fig. 3.1). The effect upon RT of a lower criterion should be greater for 'recent' negative probes than for 'novel' negative probes, which is what was found. This interpretation of interaction (b) is supported by the finding of a marked increase in RT variance with set-size for recent but not for novel negative probes - see Section 7.12 -(7). As set-size is increased and the criterion lowered, the trace strengths of recent negative items will move into that section of the decision function where small differences in strength produce large differences in decision time: an increase in the variability of decision time and hence RT is an inevitable consequence.

2) Recency of the memory set. The attempt to manipulate the recency of the memory set items with a filled probe delay must be accounted only a partial success. As predicted by Model III, the group data exhibit a reduction in slope in both the set-size and all three serial position functions following a filled delay of about $\frac{1}{2}$ sec.. However, three out of the eight subjects failed to show the effect, which therefore does not reach significance (save in the case of the Serial Position x Conditions interaction, when $m = 2$). Whether this failure is due to a genuine strategy difference, or merely to the fact that the effect is small in comparison to the normal range of random variability is difficult to determine. Only S_3 exhibited a markedly deviant pattern of performance in other respects - see Section 7.12 (5) and Fig. 7.8. Her error rate data might be interpreted as suggesting that she was attempting in some way to compensate for the effects of the filled delay. The only characteristic common to all three subjects that sets them apart from

the other five is that they were the three most experienced in RT experiments. S_3 and S_8 also had particularly fast vowel-naming RTs.

3) Serial position effects. The primacy effect commonly shown by the first item of the memory set did indeed seem to be largely abolished by the use of an irrelevant prefix having some physical features in common with items of the memory set. It is therefore probable that the primacy effect is, as suggested, an acquisition effect. The general form of the serial position curves is otherwise in accord with those obtained in Exps. II and V, exhibiting the monotonic and accelerating recency effect predicted by the trace strength models. Moreover, the fact that the shape of the curve was not changed by a delay of 550 msec. filled with visually similar material interposed between memory set and probe would seem to rule out an attempt to accommodate an exhaustive scanning model to the recency effects by supposing a special strategy of sensory matching for the last item (see Section 3.124). That the slope of the curve was reduced may however be interpreted as a reduction in rate of decay due partly to the partial loss of the 'sensory component' of trace strength (Section 3.131).

4) Other aspects of the data. Considering only 'novel' negative probes, negative RT variance did not appear to increase with set-size as positive RT variance did, and false positives increased more slowly than false negatives. Both these findings are predicted by the trace strength inspection model (and by Model V). Both statements cease to apply when the negative probe was recent. This has already been argued to relate to the interaction between set-size and the effects of recency.

Summary.

In Exp. VI, lists of from one to four consonants, chosen from an ensemble of 16, were presented visually and sequentially at a rate of 2 items/sec.. Each list was preceded by a visually similar but irrelevant vowel. Lists were followed by a recognition probe which was presented either immediately or after a delay determined by the time taken by the subject to name a visually-presented vowel (about $\frac{1}{2}$ sec.). The recency of negative probes was systematically manipulated. The main findings were:

- A 'recent' negative probe was rejected significantly more slowly and less accurately than a 'novel' negative probe.
- This effect increased significantly with set-size.
- The slopes of the group set-size function and the serial position functions were all reduced by the interpolation of a filled delay between list and probe. This effect did not however reach significance since it was not shown by three of the eight subjects.
- Pure monotonic accelerating recency functions were exhibited by the serial position data (save for a small primacy effect in the delay condition (I) when $m = 3$).
- A significant departure from linearity was found in the positive RT set-size function.
- Positive RT variance increased non-linearly with set size. Negative RT variance showed no overall increase with set size for 'novel' probes but a marked increase for 'recent' probes.

These findings were held to be compatible only with the trace strength inspection model (Model III, Section 3.11) or a mixed model (Model V, Section 3.11). These two models are indistinguishable in terms of present evidence, but the former is to be preferred on

the grounds of parsimony. The interaction between set-size and negative probe recency, and the differential behaviour of 'novel' and 'recent' negative RT variance could be accounted for by assuming a non-linear decision function, such that changes in strength near the criterion result in larger changes in decision time than equivalent differences more distant from the criterion.

The experiment needs replicating. A longer delay interval should be used in order to increase the likelihood of obtaining a significant change in slope. Also, it would be worth attempting to control recency of the negative probe in terms of time as well as in terms of number of intervening items. This could be done either by varying inter-trial interval with set-size, or by varying set-size randomly from trial to trial, or both. It might even be possible to manipulate time-based and item-based recency independently, and thus determine their relative importance.

TABLE 7.1

Set size (m) :	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	
Mean RT:	453.0	455.2	460.7	463.2	
Pooled s.d.	63.3	49.4	52.8	54.5	msec.

TABLE 7.1 Experiment VI - IRn. Mean and pooled standard deviation of the RT for naming the interpolated vowel in Condition I, shown separately for each set size.

TABLE 7.2

Response:	<u>YES</u>		<u>NO</u>		<u>Mean</u>	
Condition	<u>I</u>	<u>N</u>	<u>I</u>	<u>N</u>	<u>I</u>	<u>N</u>
Subject:						
S1	47.6	69.4	40.0	83.5	43.6	76.3
S2	33.2	31.1	33.8	24.5	33.4	27.9
S3	27.1	24.3	28.7	18.9	27.9	21.7
S4	21.6	33.3	16.1	46.4	18.8	39.7
S5	16.3	33.3	5.7	29.2	10.9	31.1
S6	21.2	69.4	31.5	49.9	26.8	59.5
S7	18.8	31.7	24.7	36.1	21.8	33.9
S8	58.2	50.1	57.7	45.6	58.0	47.6
Mean:	30.3	42.8	29.8	41.6	30.1	42.2
Mean intercept:	402.8	335.4	454.1	421.4	428.5	378.8

TABLE 7.2 Experiment VI - IRn. Slopes of the regression lines (in msec./item) fitted to the individual set-size functions; slope and intercept of the regression lines fitted to the group data.

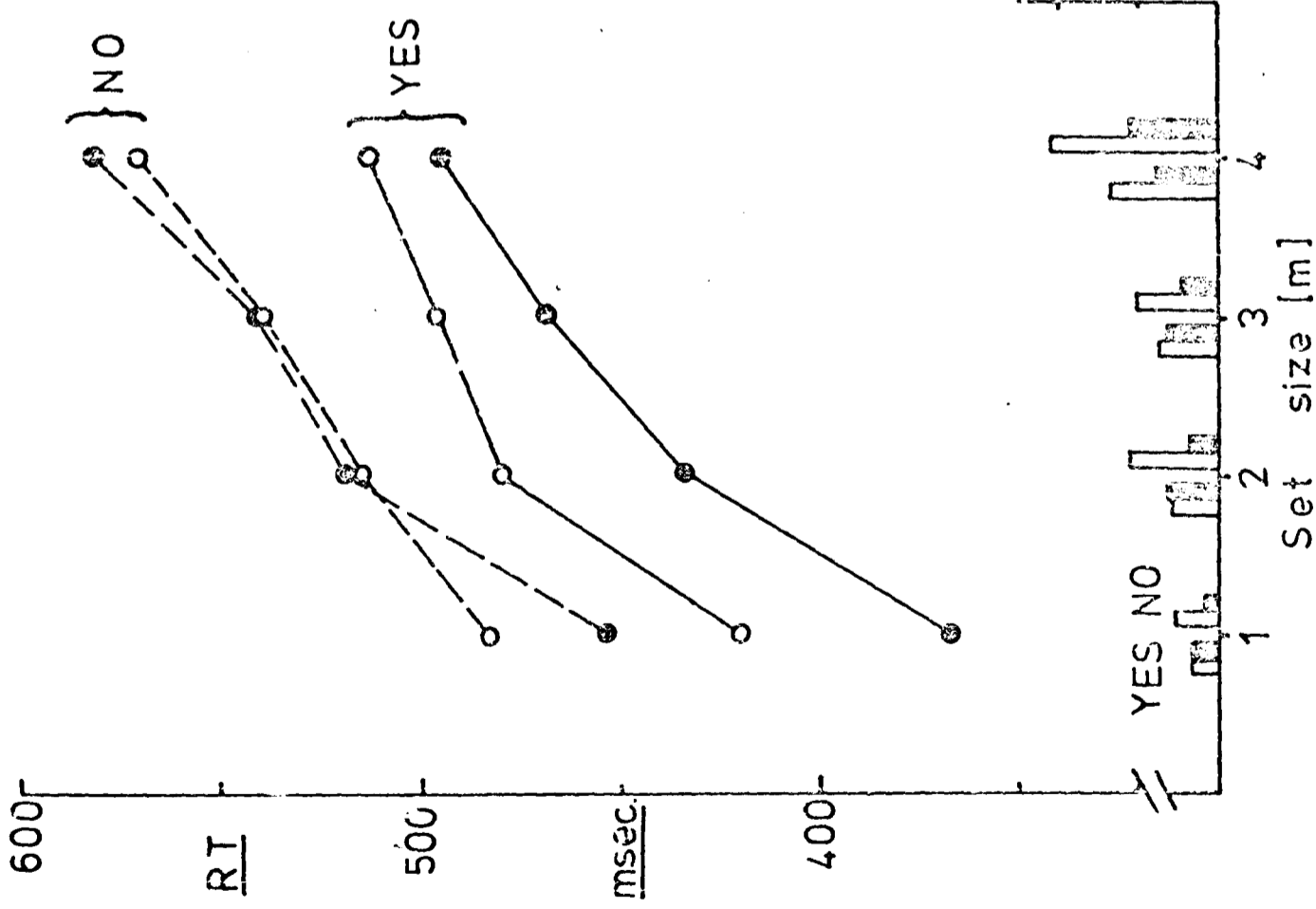


FIG 7.1 EXP.VI IRn. Set size functions.

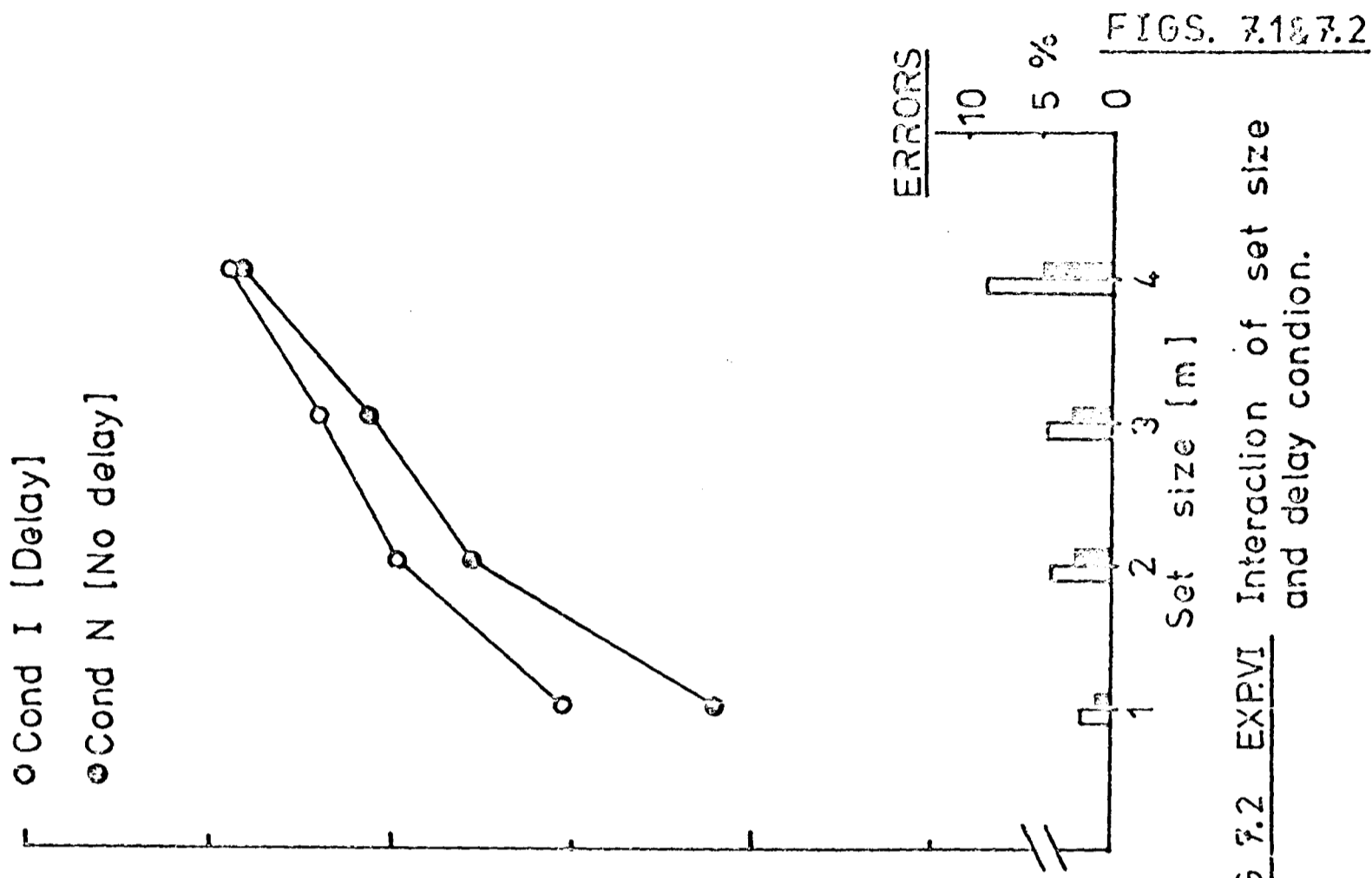


FIG 7.2 EXP.VI Interaction of set size and delay condition.

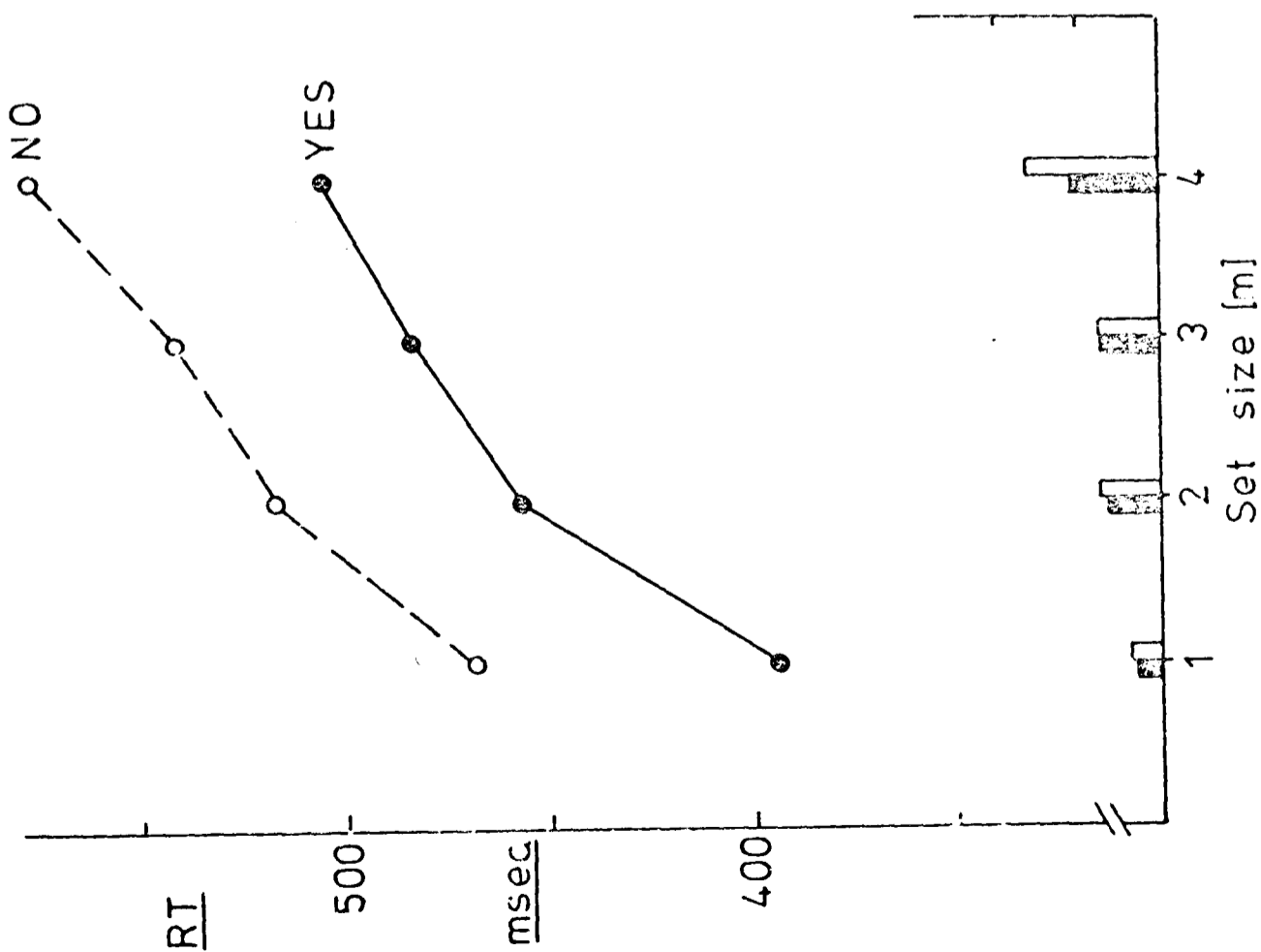


FIG. 7.3 EXP. VI - IRn. Interaction of set size and response type.

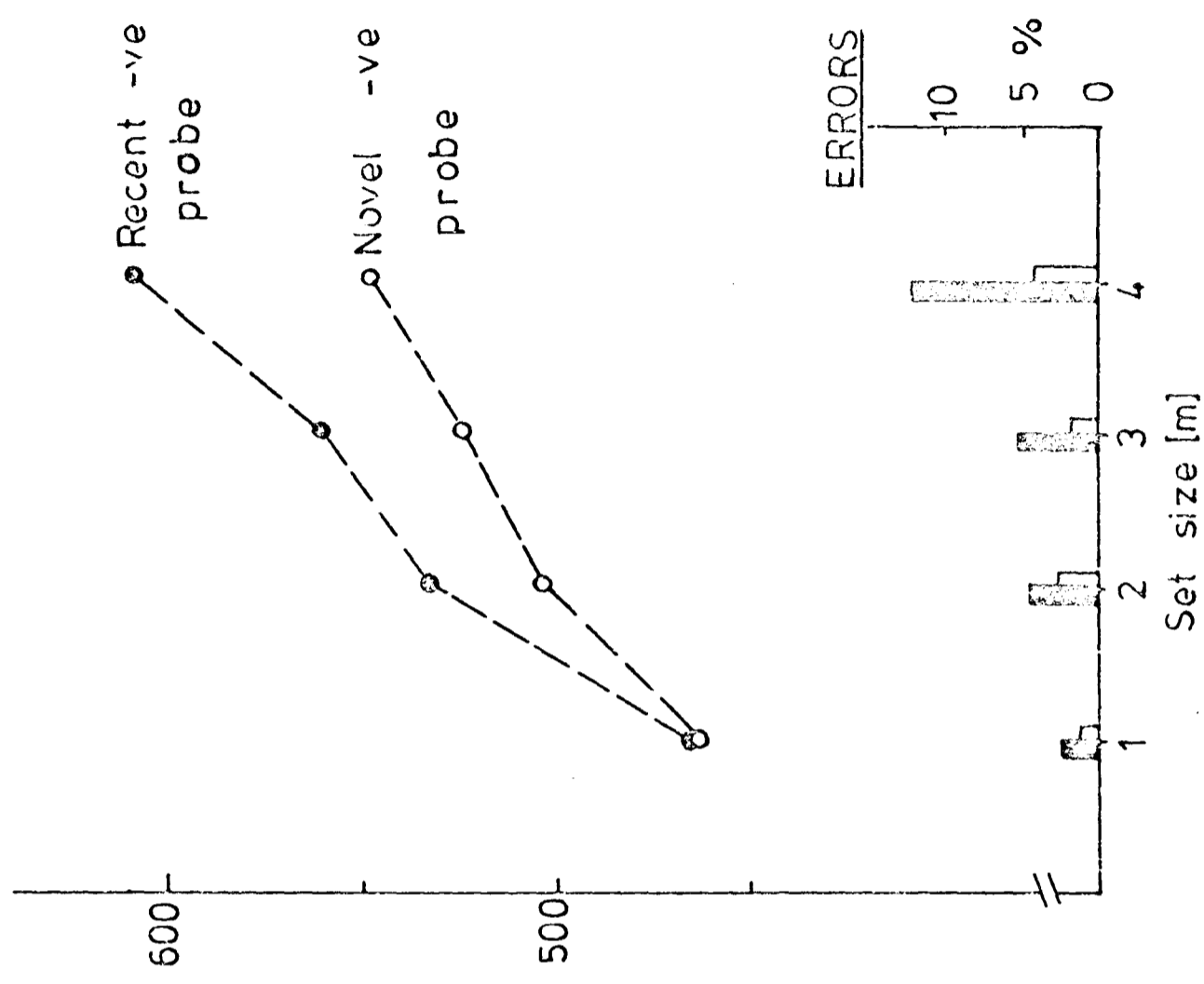
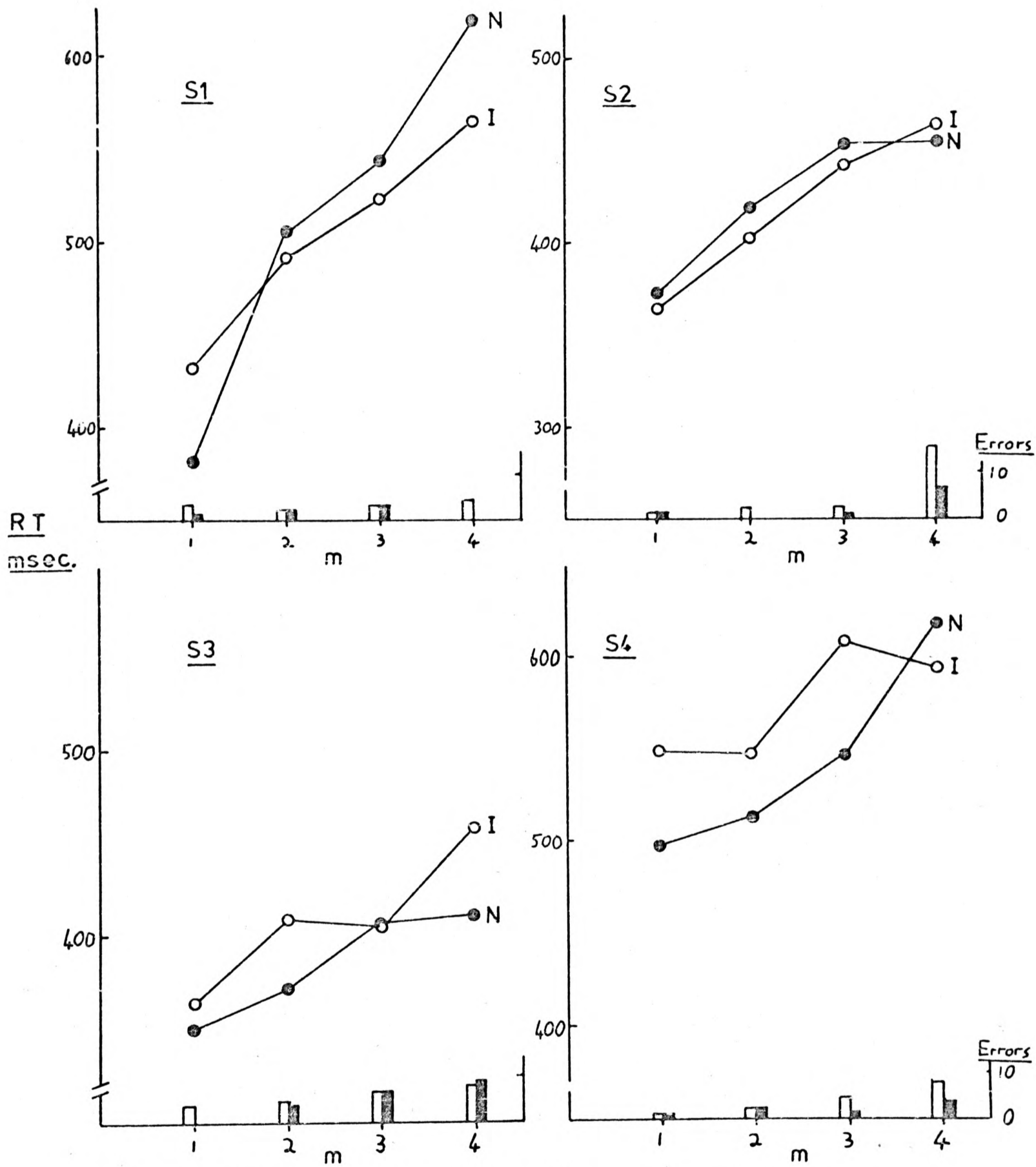
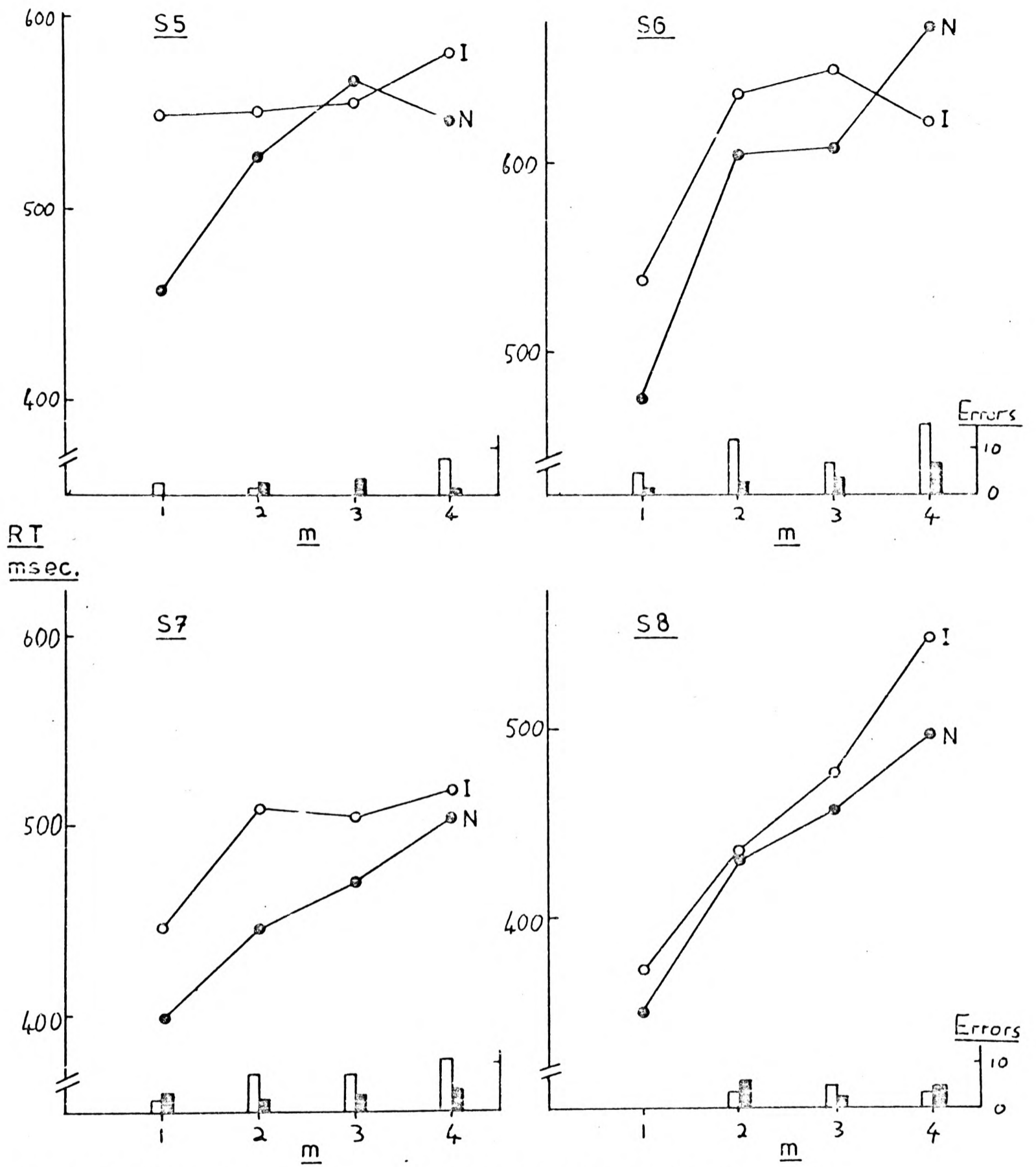


FIG. 7.4 EXP. VI. Effect of the recency of the negative probe.



EXP. VI - IRn. Interaction of delay condition and set size
- individual data.

FIG. 7.5a



EXP. VI - IRn. Interaction of delay condition and set size
- individual data.

FIG. 7,5b

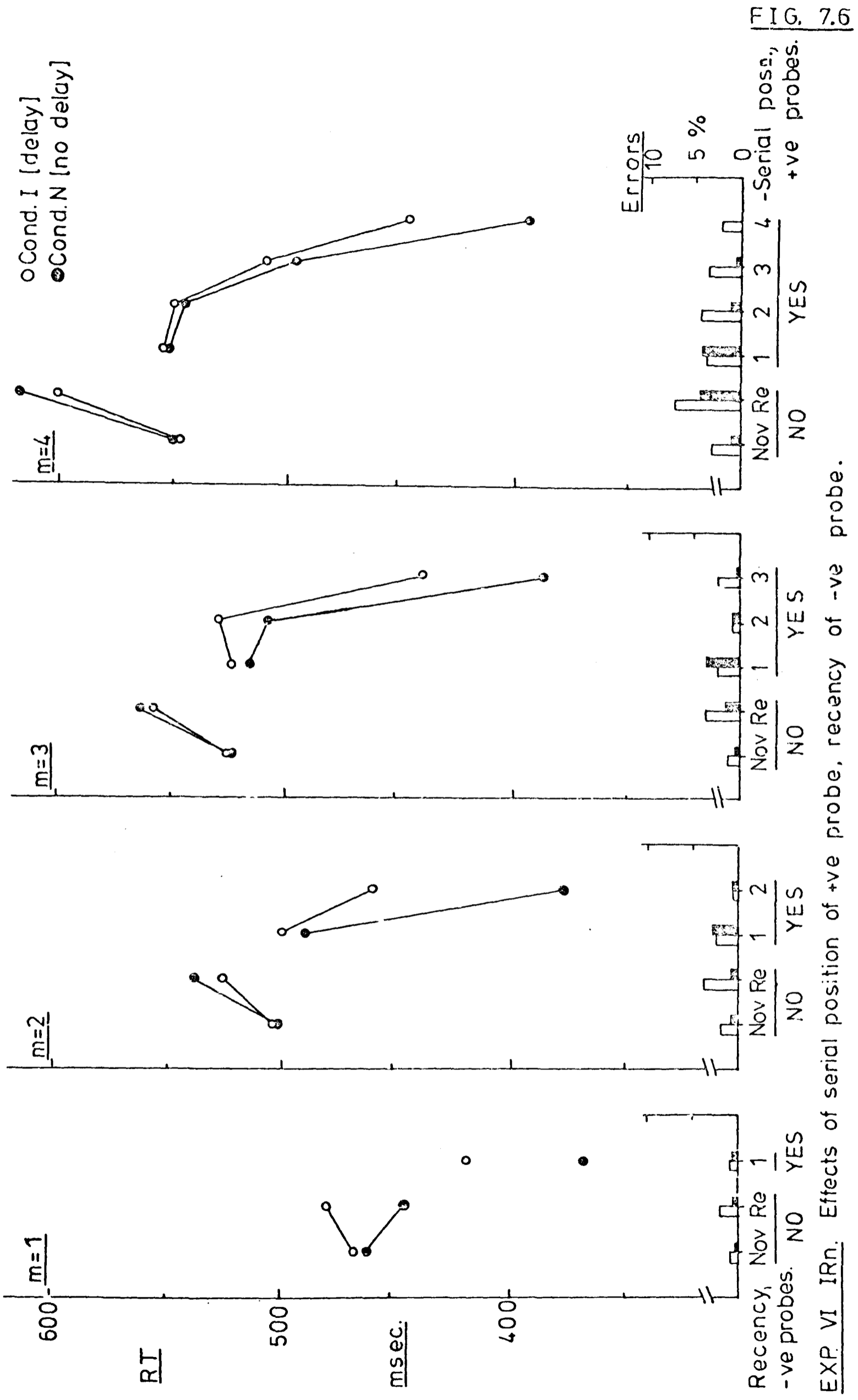
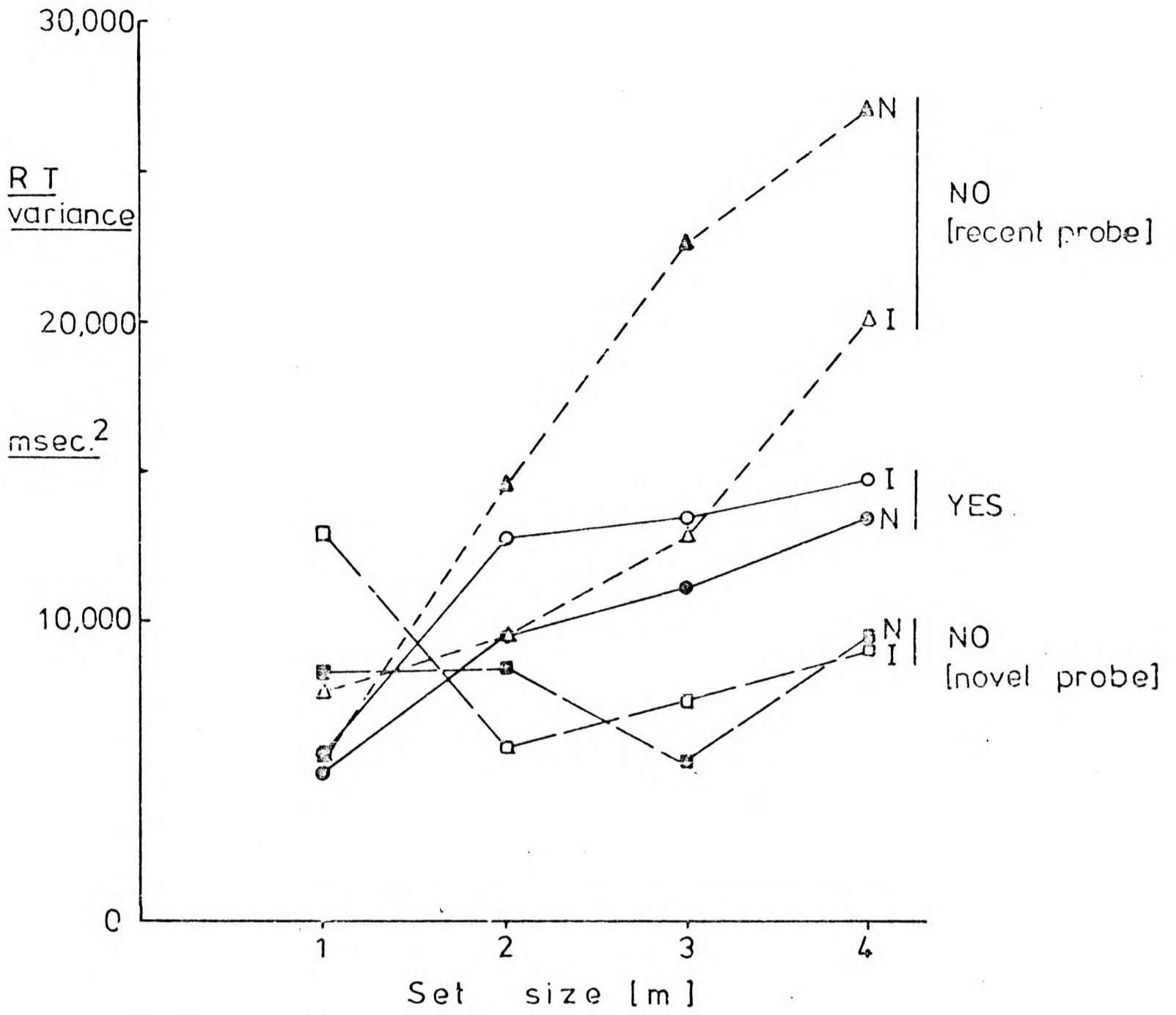


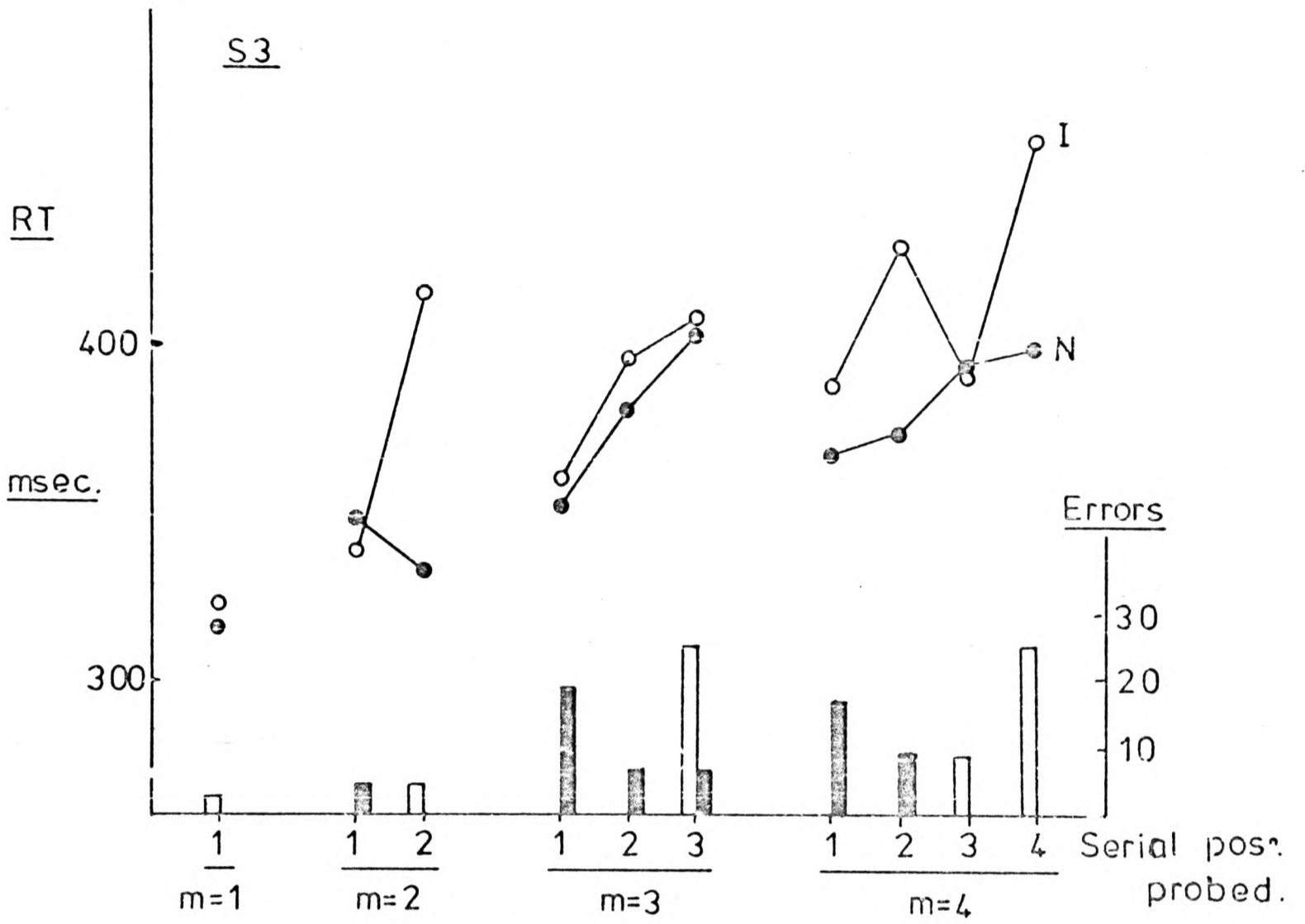
FIG. 7.6

EXP VI IRn. Effects of serial position of +ve probe, recency of -ve probe.



EXP. VI-IRn. RT variance.

FIG. 7.7



EXP. VI-IRn.

Subject 3 - serial position data.

FIG. 7.8

CHAPTER VIII. Immediate recognition and semantic class.

Introduction.

8.1 EXPERIMENT VII. Specifying the semantic subset: Mk I.

8.11 Method.

8.12 Results.

8.13 Discussion.

8.2 EXPERIMENT VIII. Specifying the semantic subset: Mk II.

8.21 Method.

8.22 Results.

8.23 Discussion.

Summary.

CHAPTER VIII. Immediate recognition and semantic class.

From the dynamics of the 'item-traces' mediating item recognition we turn in this chapter to the question of their format. This issue was reviewed in Section 3.13, and various alternative possibilities were there outlined. It was pointed out that a representation having the properties of a 'logogen' was particularly compatible with a trace strength hypothesis, while a scanning model imposed less of a theoretical constraint in this respect. The evidence reviewed led me to conclude that the logogen model accounted most fully for the evidence available in the literature, provided that it included the postulate of a 'sensory component' of trace strength (Section 3.131). This was held to facilitate IRn of probes presented either in the same physical form as in the memory set, or, for one or two items only, in the form expected by the subject - the latter facilitation resulting from a restricted capacity for the efferent generation of sensory traces.

The evidence obtained in experiments reported in previous chapters was also entirely compatible with the logogen model. Exp. II showed that subjects were able to perform the IRn task without adverse effects from within-list acoustic similarity. Experiments IV and V showed no facilitatory effects on IRn of a cue specifying the relevant position(s) in the memory set sequence. Both results were to be expected if logogens are content-addressable and amodal or conceptual representations. A logogen is, by virtue of its connections (via what Morton, 1970, calls the 'cognitive system') to other logogens, a semantic representation. Hence, while similarity of physical format should not affect IRn, semantic similarity may. In Section 3.132 (ii) I reviewed some studies on the effects of the semantic related-

ness of members of the memory set which were readily interpretable in these terms. It was also remarked that experiments need to be done on whether there are interfering effects of semantic similarity. However, we consider here a somewhat separate issue. While specification of the spatio-temporal coordinates of the relevant portion of the memory set is not expected to facilitate access or decision, specification of the semantic features could in appropriate circumstances have such an effect. To put it another way, we have seen that it does not help the subject to tell him where and when the relevant subset of the memory set was presented; we now ask whether it helps to tell him which subset is relevant, in terms of its semantic class.

The two experiments reported in this chapter must be regarded as exploratory, and some account of their historical context will be helpful. Together with Exp. IV they were conducted as part of a series of experiments in which I attempted to test the hypothesis that IRn involved a search process. The rationale common to the experiments was: if the subject is obliged to search active memory to determine whether an item is present in the memory set, and if the memory set is divided into two or more subsets, then specifying which subset is relevant to the decision could facilitate performance by abbreviating the necessary search; in particular, even when list length is constant, we would expect RT to depend on the size of the subset in the same way that for homogenous lists it depends on the size of the set. Of course, failure to find an effect of 'specifying the subset' in a single experiment would not disprove the search hypothesis. For the subset may have been specified in terms of an inappropriate attribute, or the attribute may have been cued in a

manner which allowed the subject inadequate time to process and act upon the cue. Such considerations suggested the experimental strategy of trying a variety of means of cueing the subsets, and a variety of attributes for distinguishing them, and moving on to others if negative results were obtained. In the first experiment, here reported as Exp. VII, memory sets of a constant size contained letters and digits in various proportions. The probe was also either a letter or a digit. After four subjects had been run, there was no sign of an interaction of probe class and list composition suggestive of an increase in RT with subset-size. The experiment was therefore terminated and a similar experiment, reported here as Exp. VIII was designed with changes intended to increase the availability of a 'directed search' strategy. The changes appeared to be crucial: as will be shown, RT increased with the size of the subset. At the time this appeared to constitute confirmation of the hypothesis that the subject not only searched but could perform a 'directed search' in IRn. However, if I may anticipate the discussion, the mounting evidence against scanning models for IRn has lately obliged me to reconsider this conclusion.

8.1. EXPERIMENT VII. Specifying the semantic subset: Mk I.

Five-item lists composed of from 0-5 digits, the remaining items being letters, were presented visually in a fixed display location at a rate of 600 msec. per item. After a delay of 700 msec. a probe item which was, unpredictably, a letter or a digit, was presented for recognition. Subjects pressed one key if the probe was in the list, another if it was not. The question of interest is whether RT depends on the size of the subset of memory set items of the same class as the probe item.

8.11. Method.

Four subjects each served for four sessions; data from the first session were discarded. In each session there were two blocks of 120 trials. Each block lasted about 20 minutes. Subjects were encouraged to have a short break half-way through a block and were given a five-minute rest between blocks.

Design. All six possible list compositions were used. There were four 'mixed' list types, containing both digits (D) and letters (L): 1D/4L, 2D/3L, 3D/2L, 4D/1L, and two 'pure' list types: 0D/5L, 5D/0L. In any block the correct response was positive on half the trials, negative on the rest. The 60 positive trials in a block included one example of each of the possible combinations of mixed list type, probe type and serial position probed, and two of each of the combinations of pure list type and serial position probed. The 60 negative trials included five examples of each of the possible combinations of list type and probe type. New blocks were generated for each session. The memory set items for each trial

were chosen randomly (without replacement) from the digits 1-9 and the letters A-J (excluding I). The positions occupied by letters and digits other than the probed item was also entirely random.

A trial: There was no background display. A trial was initiated by the appearance for 600 msec., at a location just to the left of the centre of the screen, of a warning arrow. 100 msec. after its offset, and at a locus just to its right, there followed the five items of the memory set, one after the other at the same location; they were each displayed for 500 msec., separated by a 100 msec. gap. 100 msec. after the offset of the last item, and to its right, a second warning arrow was displayed for 500 msec., followed after a delay of 100 msec. by the probe item, again to the right. The right index finger was used for a positive response, the left for a negative response. The response terminated the probe display, no feedback was given and another trial followed after 3 sec.. Feedback on errors and mean RT was given at the end of a block.

Instructions: Subjects were asked to respond as fast as possible while avoiding errors. No instructions were given concerning rehearsal.

8.12. Results.

False negative responses were made on 6.5% of positive trials, false positives on 4.0% of negative trials. Correct RTs were separately pooled over blocks 1-3 and blocks 4-6 to check for practice effects. Since there appeared to be none, the data from all sessions are summarised in Fig. 8.1, in which mean RT and error rate for each combination of list type, probe type and response is shown. Data from 'pure' and 'mixed' lists were subjected to separate analyses.

The correct RT data from the 'mixed' conditions were submitted to a five-way analysis of variance (mixed model): List type x Probe type x Response x Practice x Subjects (random). The single observation in each cell was based on 15 trials. Two effects reach significance. There is a main effect of Response ($p < 0.01$, $F(1,3) = 54.0$); positive responses were on average 58 msec. faster than negative responses. There is also an interaction between Response and Probe Type ($p < 0.025$, $F(1,3) = 23.5$): as may be seen from Fig. 8.1, the difference between positive and negative RT was greater when the probe was a letter.

1) There is little sign of an interaction between List type and Probe type ($F(3,9) = 2.7$) to suggest that RT increased with the size of the relevant subset. The absence of such an effect may be seen more clearly in Fig. 8.2, in which RT and errors are plotted against the size of the subset probed, without respect to its class. If the subject were searching only the relevant subset, we would expect RT to increase monotonically from $s = 1$ to $s = 4$. (Note, however, that there is some sign that errors increased in this fashion.)

2) The only other effect to approach significance in this analysis is the main effect of List type ($F(3,9) = 3.6$). Fig. 8.1 suggests that this is the result of two tendencies. Overall, RT increased slightly as the number of letters in the memory set increased. In addition, RT increased the less homogenous the list was.

3) The negative RT data from the 'pure' lists also suggest that the subjects made little use of the semantic class of the probe. These data were subjected to a four-way analysis of variance: List type x Probe type x Practice x Subjects (random). If subjects were

using the semantic class of the probe to reject probes of the opposite class, we should expect an interaction between List type and Probe type. But there is no such interaction ($F(1,3) = 1.2$) and the only effect to reach significance is that of Probe type ($p < 0.05$, $F(1,3) = 10.3$), letters being rejected slower than digits.

4) Serial position data from the positive trials are shown in Fig. 8.3. Positive RT data from the mixed lists were submitted to a five-way analysis of variance: Serial position probed x List type x Probe type x Practice x Subjects (random). Serial position is the only effect to reach significance ($p < 0.001$, $F(4,12) = 9.98$). A clear monotonic recency effect is obtained. That positive RT is shorter for pure than for mixed lists is probably due to no more than their greater homogeneity, as indicated above.

8.13. Discussion.

The RT data provide no grounds for supposing that subjects were making use of the semantic class of the probe to restrict a search process to only those memory set items sharing that class. Only the error data show a hint of such an effect, in that false negatives increased somewhat with subset-size. Even on those pure list trials when the probe could be rejected on the basis of semantic class alone, there was little evidence that it was.

What is responsible for the negative result? Perhaps it takes so long to extract the semantic class of the probe that it would be highly inefficient to wait until it is available to indicate the subset to be searched. On the other hand, available evidence suggests that classification of a character as a letter or digit is fast and may take place prior to its identification: for instance, Brand (1971) who, as it happens, in one of her experiments used the same apparatus

and display characters as me, found in a visual search task that even relatively unpracticed subjects took no longer to find¹ a target specified as 'any digit' than they did to find a particular digit. Nevertheless, it was decided to indicate the semantic class of the probe in the next experiment by means of an extra cue - display location. Another possibility is that a directed search strategy requires the subject to have 'tagged' the memory set by semantic class at input, and that the presentation rate used was too fast to permit this. Accordingly, in Exp. VIII, a slower presentation rate was employed. In addition trials of a particular composition were blocked so that subjects would know in advance the structure of the list.

There is one difficulty. Subsequent to the experiments reported in this chapter, Kaminsky & Derosa (1972) published details of an experiment which, though similar in crucial respects to the present one, gave a different result. They used six-item lists, presented at a brisk rate, containing two, three or four digits, the remaining items being letters. They also included a base-line condition: pure lists consisting of 2-6 letters or digits, probed only with items of the same class. Trial types occurred in a random sequence. The main differences from the present experiment were a) the probe delay was about 3 sec. and b) a condition was included in which the subject was warned of the probe class 2 sec. before the probe onset. When the subject was warned of the probe class, there was a straightforward increase in RT with subset-size, equal in slope to the set-size function obtained from the base-line condition. But this is unsurprising according to almost any hypothesis: if the subject knew the class of the probe 2 sec. before the probe, he was able to rehearse only the items of the relevant subset. Our problem is that even when the

¹in a background of letters.

subject was not warned of the probe class, there was still an increase in RT with subset-size, which is contrary to the present results. However, the increase in RT was markedly non-linear and much smaller in slope than that found in the base-line condition. Most important, less than half the subjects showed the effect (and the slope of the subset-size function for one of them was so great - 100 msec./item - as to suggest a search of the response buffer). Nevertheless, given the nature of the hypothesis, the fact that even some of the subjects showed an effect poses a problem. However, given the long probe delay, there exists an alternative explanation. Suppose that on those trials when no warning was given, the subjects continued to rehearse semantic classes selectively. If, on average, they gave as much rehearsal time to each subset, then the items of smaller subsets would on average be more recent at test than those of larger subsets. Hence on a trace strength hypothesis, RT would show some increase, though probably not very much, with subset size. One other feature of Kaminsky & Derosa's data argues against a 'directed search' hypothesis. When the subsets were grouped - i.e. items of the same class were adjacent, RT was longer than when they were random. Yet grouping would be expected to provide an extra dimension of organisation facilitating a directed search.

There are three other features of the data from Exp. VII that merit brief mention. Firstly, though rehearsal was not explicitly forbidden, the serial position data show the by now familiar monotonic recency effect. This is further evidence against a search strategy (see Section 3.124) and implies that subjects were either not rehearsing at all (given the reasonably fast presentation rate and short probe delay) or, as one subject claimed, rehearsing the whole list

just once during the 700 msec. probe delay. Secondly, there was evidence that performance on letters was inferior to that on digits; this ties in with the fact that memory span for digits is greater than that for letters, and that the slope of the set-size function in IR_n is greater for letters than for digits (Cavanagh, 1972). The latter, according to trace strength theory, implies a greater decay rate for letters' item-traces. Thirdly, performance deteriorated with increasing heterogeneity of the memory list. This almost certainly relates to Broadbent & Gregory's (1964) finding that with presentation rates of the present order, recall of lists in which letters and digits alternated was worse than recall of lists in which the classes were grouped. It may be either an acquisition effect, or relate to a generalisation of trace strength between semantically similar items (see Section 3.132 - iii).

8.2. EXPERIMENT VIII: Specifying the semantic subset: Mk II.

This experiment differed from Exp. VII for reasons already mentioned, in the following respects:

1) Digits and letters were always presented in different spatial locations, so that display location was available as a cue for semantic class.

2) Presentation rate was reduced to 1 item/sec..

3) The different list types were presented in separate blocks.

In addition, the following changes were made:

4) Four-item lists were used.

5) Probe delay was varied randomly between three values and a visual noise mask was presented during it.

6) Subjects were instructed not to rehearse.

8.21. Method.

Five subjects each served for three hour-long sessions. The first was a practice session and data from it were discarded. In each session, a subject was given five blocks each of 60 trials. A short rest was allowed between blocks.

Design. Each subject underwent each list type condition, of which there were five, in every session. The order of conditions in the practice session was fixed at 4D/OL, OD/4L, 1D/3L, 3D/1L, and 2D/2L. The order for the first experimental session was determined by a 5 x 5 Latin square and reversed in the second experimental session. Under each condition a block of 60 trials was given, of which the first 12 were 'warm-up' trials, subsequently discarded. Of the remaining 48, half were positive, half negative. Their composition was as follows:

'Mixed' lists, positive trials: there were three trials for each combination of probe type (D or I) and serial position probed. Each of the possible orders of non-probed items occurred equally often. The three values of probe delay used were allocated randomly to these orders so that overall they were balanced in this respect as well as over probe type and serial position.

'Mixed' lists, negative trials: each combination of probe type and delay was tested four times; order of items was balanced over delays as nearly as possible.

'Pure' lists, positive trials: each combination of probe type and probe delay was tested twice.

'Pure' lists, negative trials: each combination of probe type and probe delay was probed four times.

On each trial, digits were sampled randomly (without replacement) from the set 2-9 and letters from the set A-H.

Procedure. A trial consisted of the following sequence of events. A background display, which consisted of a horizontal line and remained visible throughout the trial, was displayed for 1 sec.. The four items of the memory set were each displayed for 750 msec., separated by intervals of 250 msec.. Letters were always displayed above the line, digits below. The two display locations were separated by approx. $1\frac{3}{4}^{\circ}$ of visual angle. 250 msec. after the final item, a rectangular field of dynamic visual noise was displayed for one of three durations; 250, 500 or 750 msec.. The mask was centred on the line and subtended about $1\frac{3}{4}^{\circ}$ horizontally and $4\frac{1}{2}^{\circ}$ vertically; subjects were told to fixate a blob at its centre. 250 msec. after the mask, the probe item was presented. Subjects responded with their right index finger for a positive response, the left for a

negative response. The response terminated the display and an interval of 3 sec. followed before the next trial. No feedback was given until the end of the block.

Thus, presentation rate was 1 sec./item and the probe delay was equally often $\frac{3}{4}$, 1 or $1\frac{1}{4}$ sec..

Instructions. Subjects were asked to respond as fast as they could while avoiding errors. They were also instructed not to rehearse. (Some subjects in fact found this difficult, given the long probe delay and slow presentation rate.) They were also told, "In the pure list conditions, there will still be some probes of the opposite class: to these you should be able to key 'no' very easily. In the mixed list conditions, it is not logically necessary, if you are given, say, a letter probe, to refer to your memory for the digits. You may find this helps you." These instructions are clearly biased towards a 'directed search' strategy.

8.22. Results.

The data were analysed in a similar fashion to those of Exp. VII save that practice was not included as a variable. Also, any RT more than three standard deviations from the mean of a (probe type x list type x subject) cell was discarded: 1.3% of all RTs were rejected in this way. False negative responses were made on 6.4% of positive trials, false **positives** on 3.9% of **negative** trials.

The mean correct RT for each combination of list type, probe type and response is shown in Fig. 8.4. The correct RT data from the mixed conditions were submitted to a five-way analysis of variance: List type x Probe type x Response x Delay x Subjects (random).

Positive RTs were faster than negative RTs by 37 msec. ($p < 0.05$, $F(1,4) = 8.2$). RT was shorter for a letter than for a digit probe, but this effect fails to reach significance ($F(1,4) = 5.3$).

1) Of the remaining effects, none approach significance save the interaction of interest, that between List type and Probe type, which is significant at the $p < 0.01$ level ($F(2,8) = 10.5$). As may be seen in Fig. 8.4, RT increased as the size of the relevant subset increased. Regression lines fitted to the data show that the rate of increase is 16 msec./item for digits, 27.8 msec./item for letters. The nature of this effect is more clearly seen in Fig. 8.5, in which RT is plotted as a function of the size of the relevant subset.

2) It also appears that subjects made use of semantic class (or display location) to reject negative probes of the opposite class in pure lists. A four-way analysis of variance was performed on the correct negative RTs from pure lists: Probe type x List type x Delay x Subjects (random). Aside from Probe type, letters being slower than digits ($p < 0.05$, $F(1,4) = 10.9$), only the interaction of Probe type and List type reaches significance ($p < 0.01$, $F(1,4) = 49.6$). As may be seen in Figs. 8.4, 8.5 and 8.6, probes were rejected very much faster when they differed in class from all the items of the memory set.

3) To look at serial position effects, a four-way analysis of variance was performed on data from the mixed condition positive trials: List type x Probe type x Serial position x Subjects (random). Though a very slight recency effect is discernible, Serial Position

does not reach significance ($F(3,12) = 2.8$). Also, in this analysis, the Probe type x List type interaction fails to reach significance ($F(2,8) = 3.47$). No other effects approach a significant level. A four-way analysis was also performed on the pure list positive RTs: List type x Serial position x Delay x Subjects (random). No effects approach significance. Serial position data are summarised in Fig. 8.6.

4) It may be seen from Figs. 8.5 and 8.6 that positive RTs were shorter when the lists were pure than when they were mixed.

8.23. Discussion.

The data clearly show that for mixed lists, RT increased systematically with the size of the subset whose members were of the same class as the probe item. On the face of it, this result favours the hypothesis that subjects were able to pursue a 'directed search' strategy. Moreover, that the slopes of the positive and negative subset-size functions are parallel must presumably on these grounds be interpreted as indicating that the subset was searched exhaustively. That positive RT was shorter for pure than for mixed lists, even though all four items shared the semantic class of the probe must, on the 'directed search' hypothesis, be attributed to extra time being required for the process of classifying the probe and for 'selecting' a subset.

However, as I indicated in the introduction to these experiments, several factors have led me to re-evaluate these conclusions. The experiments reported in Chapter 6 showed no sign that subjects were able to perform a strategy of 'directed search' when the subset was

defined in terms of ordinal position, when both temporal and spatial cues were given. It would be strange if, in a store which must code order, subjects could restrict a search to a subset defined by semantic class, but not by temporal position. Moreover the evidence reviewed in Section 3.1, and my own data, from Exp. VI in particular, have provided strong evidence for rejecting a scanning model for IRn. We must therefore consider whether the present result may not after all be interpreted in terms of a trace strength hypothesis. One feature of the data which might lead one to suspect an interpretation in terms of a directed search is the values derived for the scanning rate. At 16 msec./item the value for digits indicates a scanning rate twice that reported in all the studies reviewed by Cavanagh (1972). The value of 27.8 msec./item for letters is also faster than that obtained in nearly all the studies covered by Cavanagh's review. Moreover, the individual data were somewhat variable. Though all subjects showed a Probe type x List type interaction in the appropriate direction, its shape differed from subject to subject. It will also be recalled that the interaction failed to reach significance when positive RTs alone were analysed.

How can the data be interpreted in terms of trace strength theory?

1) In this experiment, in contrast to the last, the different list types were blocked rather than presented in a random order. Consider a block in which the memory set on every trial contains one digit and three letters. The mean of the distribution of the strengths of the negative digits on every trial will be lower than that for letters, since fewer have occurred recently. The optimal criterion placing will therefore be lower on the strength continuum for digits than for letters. Hence the interaction between list type

and probe type may be explained by supposing that the subject is able to use the cue to select the criterion appropriate to the class of the probe. If he does in the present example the average difference between the strength of the probe item and the criterion will be greater (and the RT less) for both responses when the probe is a digit than when it is a letter. When the list contains three digits and one letter, the inverse will hold. The obvious test of this explanation would be to repeat the experiment with trials of different list composition occurring randomly within a block.

2) The effect may be enhanced by a second factor. Subjects complained that it was difficult not to rehearse, and admitted to doing so some of the time. This was borne out by two features of the data: there were practically no effects of serial position, which is what one would expect as a result of a combination of rehearsal and variable probe delay (Section 3.124); there was no effect of probe delay. If subjects did rehearse, then the explanation offered earlier in this chapter for the behaviour of some of the subjects of Kaminsky & Derosa (1972) may apply here also. If subjects allocate on average equal rehearsal time to each subset, then the strength of list items from the smaller class will on average be the higher at presentation of the probe. This possibility could be avoided by stricter control of rehearsal, under conditions of fast presentation and immediate test.

3) One subject's comment suggests a third possibility. He claimed that when the list contained a single item of one class, he visualised that item, while 'saying the others to himself'. Thus a 'sensory component' of trace strength, resulting from Posner-type 'generation' (Section 3.131) may have contributed to the shorter RT for single items. Of course, one would expect the use of a visual

noise mask to restrict the effectiveness of any 'generation', but the efficiency of the mask is unknown, and an interval of 250 msec. elapsed between its offset and the probe, which might have allowed enough time to generate a sensory trace. A generation strategy of this sort could be prevented by giving an interpolated task, as in Exp. VI.

Why should positive RT be shorter when the lists were pure? In fact, if negative probes of the opposite class are not considered, it is only positive RTs which were markedly shorter for pure than for mixed lists. That false positives were particularly frequent for pure lists implies that the decision criterion was pitched somewhat lower on the strength continuum. In addition, there may also be a facilitative effect of homogeneity such as was observed in Exp. VII.

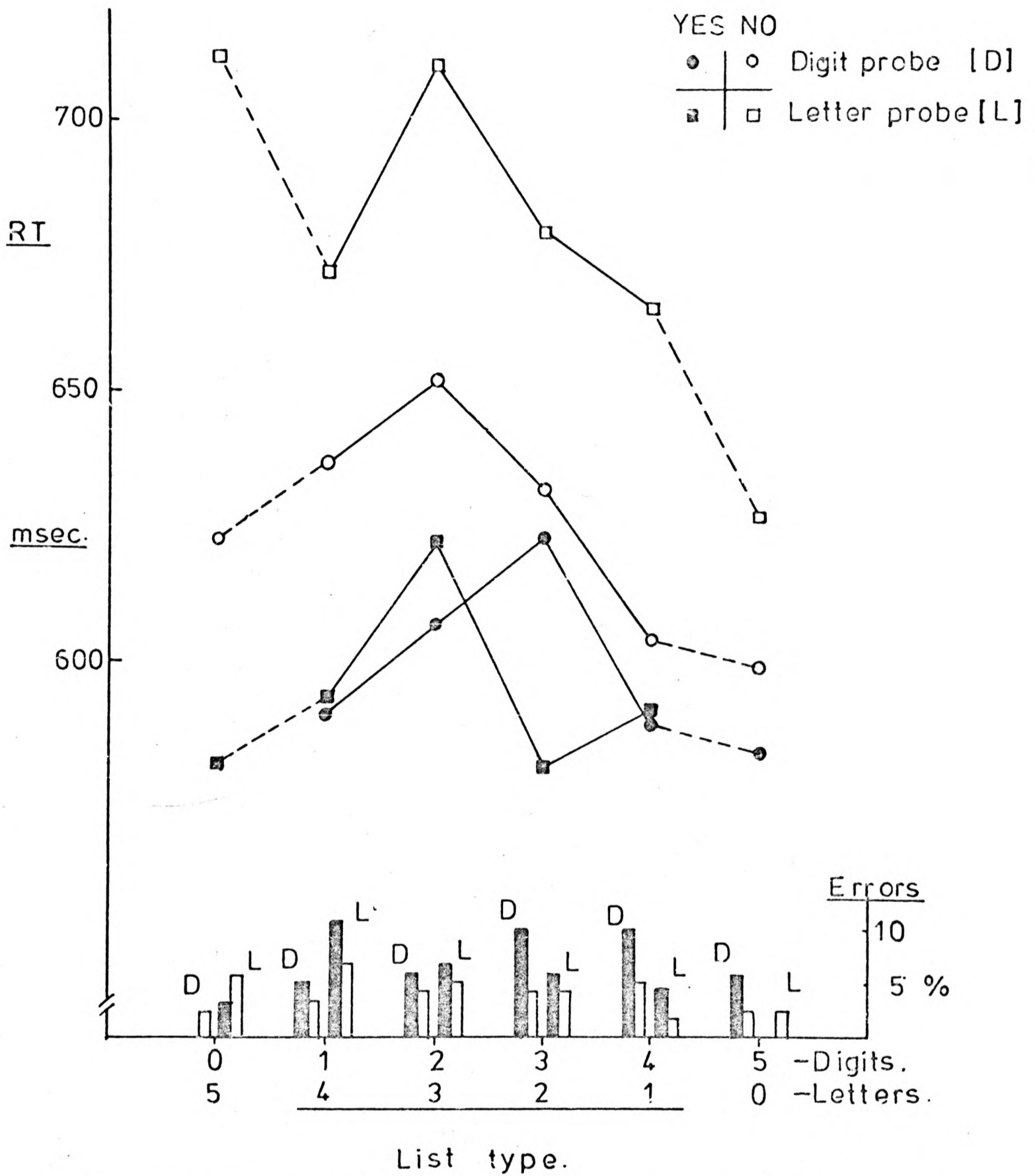
Lacking further data, and in the light of the several possible (if baroque) interpretations of the data in terms of a trace strength hypothesis, I am obliged to conclude that the present experiments, and that of Kaminsky & Derosa (1972) are not decisive with regard to the question of when, how and why IRn RT is dependent on the size of a subset defined by semantic class. I have suggested above ways of changing the experiment so as to remove effects predicted by a trace strength hypothesis without making a 'directed search' strategy impossible. The question must therefore wait upon the outcome of further experiments. Perhaps more importantly, experiments are also required to investigate whether there are interfering effects of semantic similarity on item recognition.

One final point: is there justification for assuming that results obtained with letters and digits can be generalised to other semantic

classes? Results obtained in dichotic memory experiments provide some evidence that there is. Gray & Wedderburn (1960) showed that subjects were as likely to report dichotically presented pairs of words in an order determined by semantic constraints, even when this involved alternating across ear of presentation, as they were to report ear by ear. Broadbent & Gregory (1964) replicated this result with unrelated words of the same class, and also for digits and letters. The latter authors also showed that the deleterious effect of alternating class, when presentation was monaural, applied no less to other word classes (like colour-words) than it did to digits and letters.

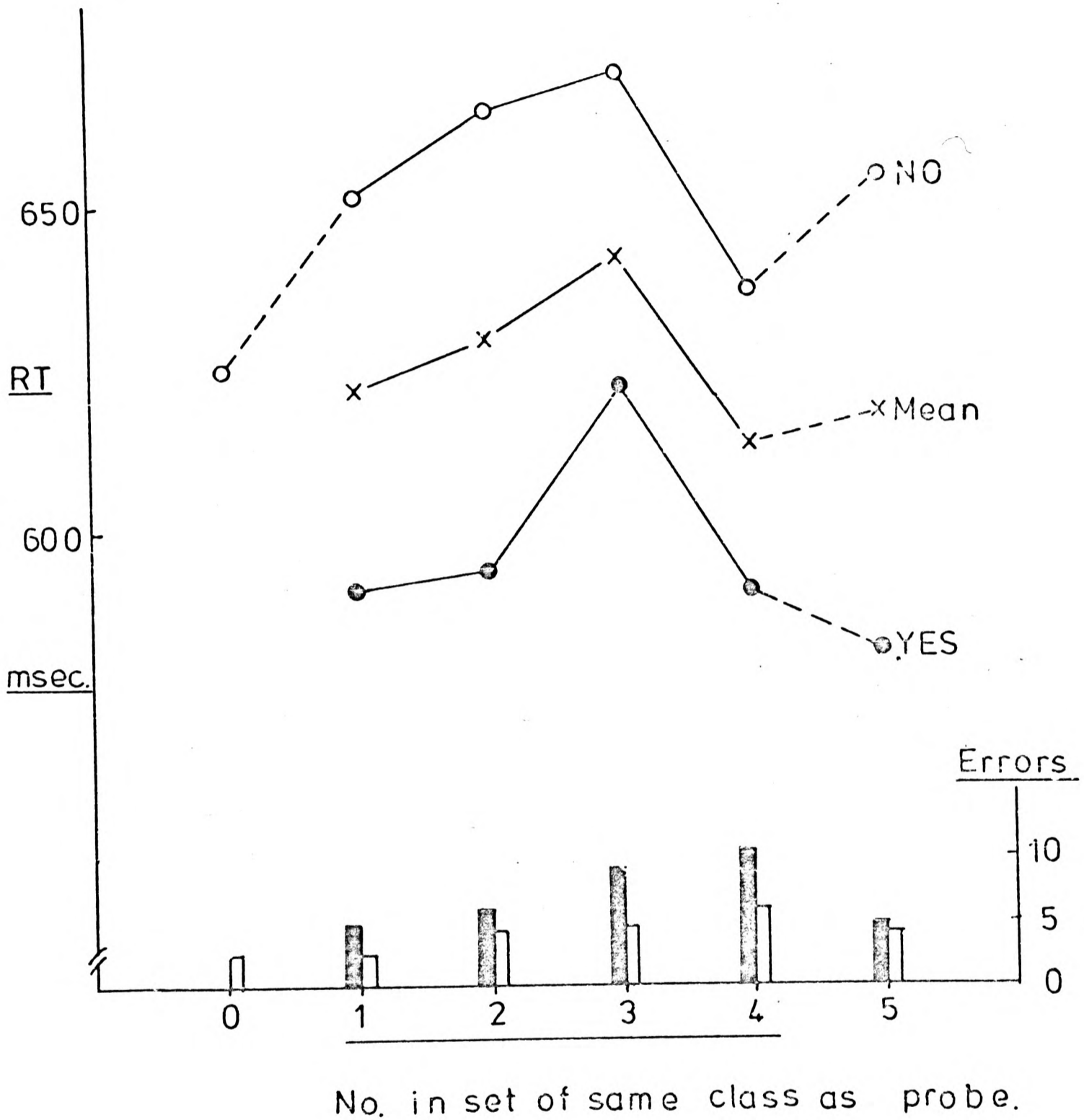
Summary.

In this chapter were reported two exploratory experiments on the role of semantic class in IRn. In both experiments, mixed memory sets of digits and letters were presented and subjects were tested with a probe item that was unpredictably a digit or a letter. The question of interest was the nature of the dependence, if any, of RT on the size of the subset of the same class as the probe. In the first experiment, in which five-item lists were employed, RT did not increase with the size of the subset probed, but merely showed a decrease as the homogeneity of the memory set was increased. It was, however, thought possible that it might take too long to extract the semantic class of the probe for it to be useful. In the second experiment, in which four-item lists were used, an additional cue for semantic class - display location - was provided. In addition, trials with memory sets of a given composition were grouped in blocks and presentation rate was reduced. This time, RT systematically increased with the size of the relevant subset. Though there existed a prima facie case for interpreting this result, and some data of Kaminsky & Derosa (1972) in terms of a 'directed search' hypothesis, it was argued that the data could also be interpreted in terms of trace-strength theory, provided that the subject is able to adjust his decision criterion on the basis of the semantic class indicated for the probe. In addition, two ways were suggested in which the average strength at test of items of the smaller subset items might be enhanced: differential rehearsal, and a 'sensory component'. Means were suggested for controlling these possibilities in future experiments.



EXP. VII - IRn. Effects of list composition and probe type.

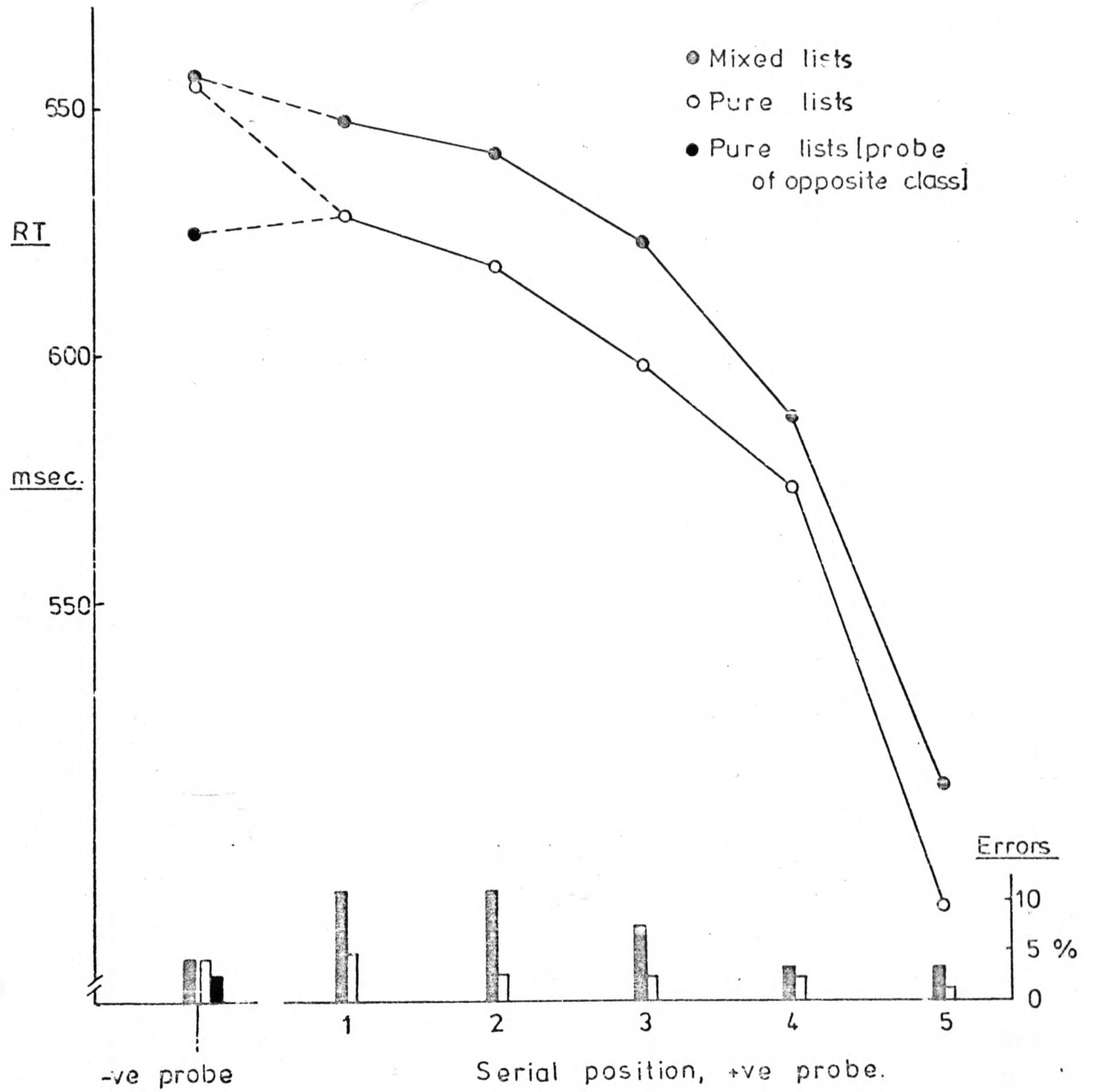
FIG. 3.1



EXP. VII - IRn.

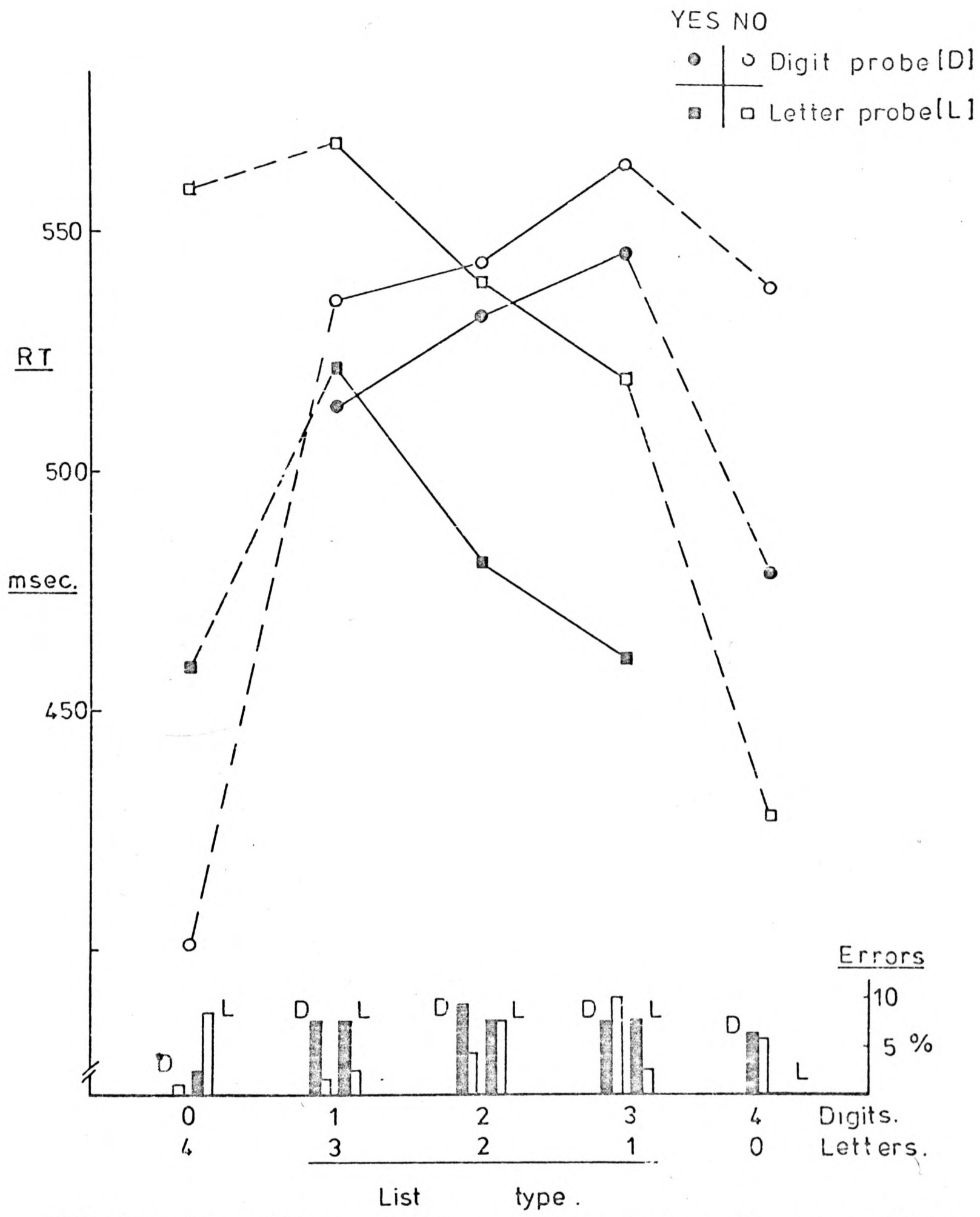
Subset-size functions.

FIG. 8.2

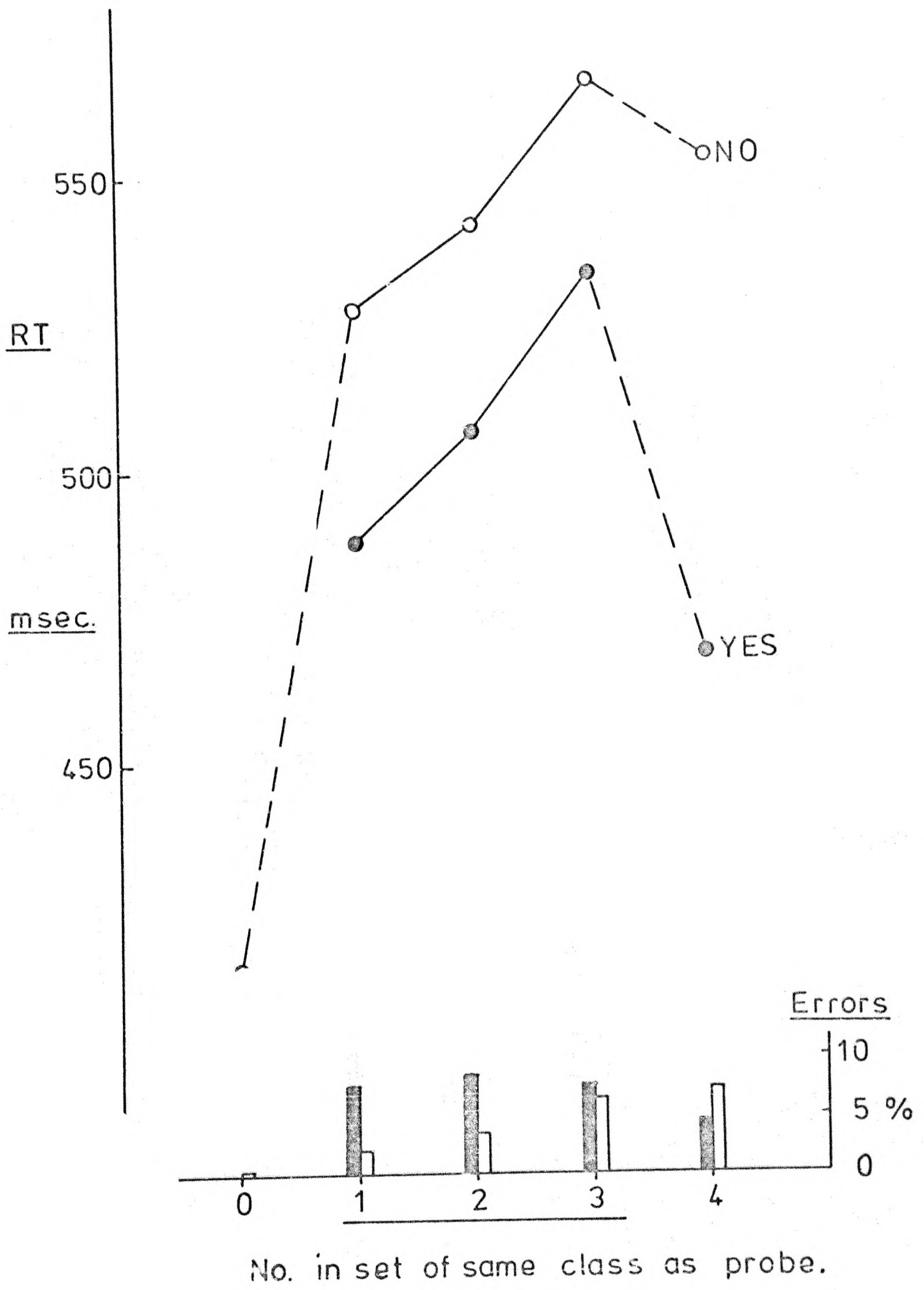


EXP. VII-IRn

FIG. 8.3

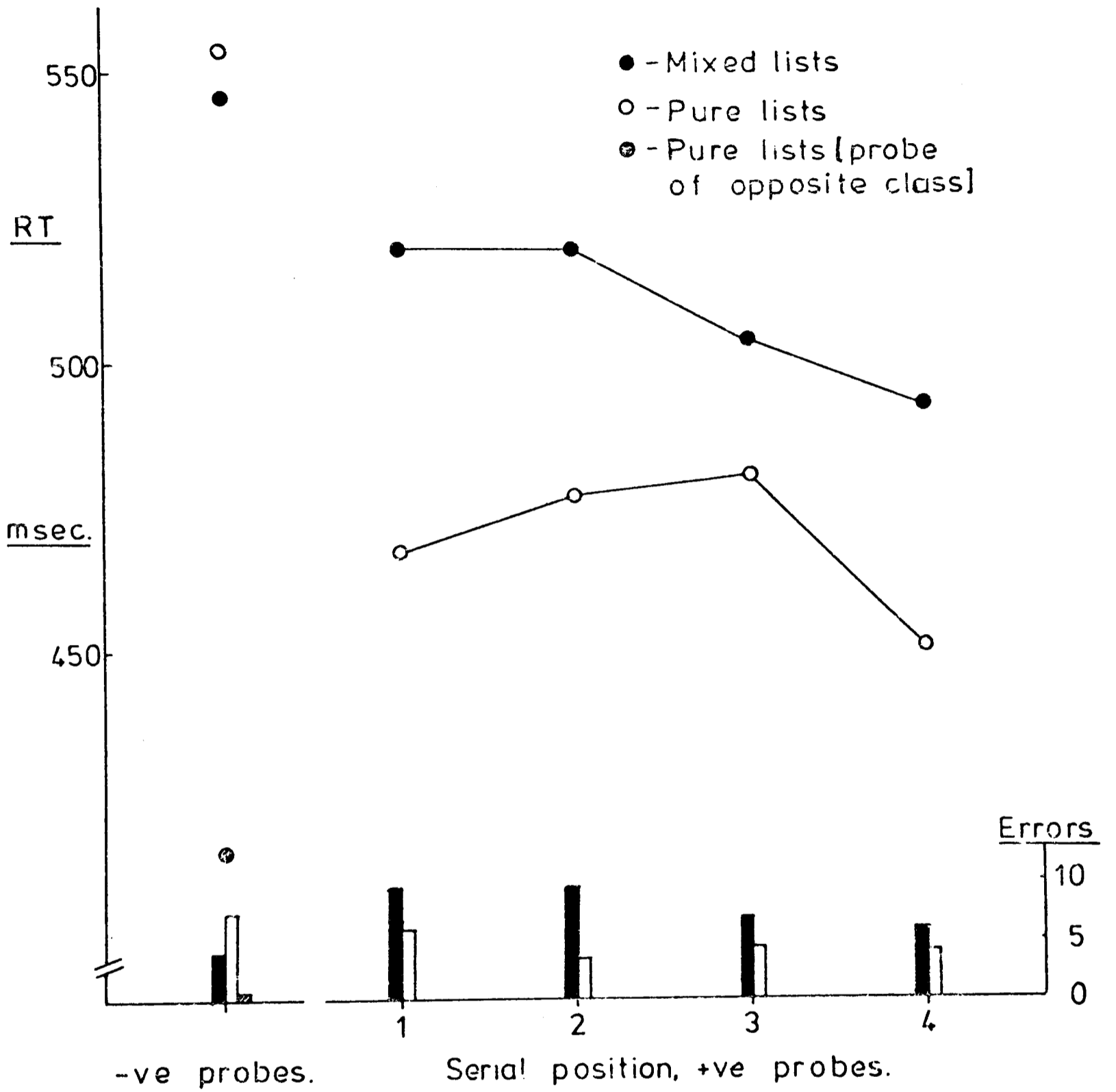


EXP. VIII - IRn. Effects of list composition and probe type. FIG. 8.4



EXP. VIII - IR_n. Subset-size functions

FIG. 8.5



EXP. VIII - IR_n.

FIG. 8.6

CHAPTER IX. Concluding comments, speculation and plans.

9.1 Theory.

9.2 Experiments.

9.3 What is active memory?

CHAPTER IX. Concluding comments, speculation and plans.

In view of the summaries provided in each chapter and in the abstract, I do not propose to burden the reader with a further detailed review of the experiments and their implications. In what follows, bearing in mind that the value of a piece of research is to be found as much in the questions it generates as in the answers it yields, I have tried to look forwards and sideways as well as backwards.

9.1. Theory.

In the first part of this thesis, I had three aims, two general and one specific. The first was to try and show that at least some of the inconsistencies and dissonances which seem inevitable if one tries to interpret the existing evidence on short-term verbal memory in terms of a single post-categorical storage mechanism may be resolved by the hypothesis of two retention mechanisms in post-categorical active memory, the logogen system and the response buffer proposed by Morton (1970) and amended in Section 1.4. Secondly, I tried to show that the same resolution works both for evidence from the short-term forgetting studies reviewed in Chapter 2 and for results obtained with the memory-scanning paradigms covered in Chapter 3, and suggest that this fact may supply the unifying link between the two sources of evidence that common sense demands. The specific aim was to explore in detail the models available to account for performance in Sternberg's item recognition paradigm and their predictive power with respect to evidence already available in the literature. It proved extremely difficult to account for this

evidence in terms of either pure scanning theory; nor was Model IV, the trace-strength 'default' model, strikingly successful. It was however argued that the evidence was compatible with the trace strength inspection model and in Ch. 7 it also proved possible to interpret my own data in these terms.

These several aims made it necessary to review a large number of published studies. I am conscious that in many cases the attention given to individual experiments was superficial. It seems an inevitable consequence of an attempt at global synthesis that one should ask the question, 'Does it fit?' more often than one asks 'What are the possible explanations available to account for this single finding?'. I take some comfort from the thought that the argument is structured in part as a mosaic rather than a linear progression. Should one or two tiles turn out to have been insecurely glued, the picture as a whole need not be scrapped. There was also a deliberate restriction of the discussion to behavioural evidence from experiments on short-term verbal memory in normal subjects. There are a number of theoretical questions related to the view put forward in this thesis of which brief mention should be made, though their detailed exploration is beyond the scope of this thesis (and the present competence of its author).

1) The relationship of the model to theories of the production and the perception of speech. The response buffer has been spoken of as being clearly located on the output side of the system, as an ordered buffer store of potential responses coded as a string of phonemes or articulatory commands. Because the acoustic waveform emitted by a speaker is largely continuous, there have been doubts cast on the idea that there are any discrete internal representations

corresponding to phonemes. But in what do articulatory commands then consist? Is there in fact a need for a pre-execution buffered store of such commands? Some models of speech production definitely require such a store, others are indifferent. (See, e.g., MacNeilage 1970, for a review of some of these issues.)

As for speech perception, nothing has been said about the mechanism whereby the appropriate logogens are activated by the speech stream as it is transcribed by the receptors. Hasler (1973)¹ has argued that there is a necessity for an ordered phonemic store prior to word identification. If this argument were to prove valid, would it not be unparsimonious to propose two ordered phonemic stores, one for perception, one for response? Perhaps it would then be necessary to think of the response buffer as a storage system that could alternate between two modes, input and output. However, the fact that one can listen and talk simultaneously would pose problems for such an account. Finally, it is interesting that it is possible to shadow (slowly) 'gobledgook' or a foreign language with phonemes similar to those of English. Does this imply that there is a direct output from phonemic (or other) 'sub-logogens' to the response buffer?

2) Comprehension. I attempted in Section 2.31 to argue that the conventional estimates of the capacity of 'primary memory' were in fact estimates (though in some cases probably under-estimates) of the capacity of the response buffer, and that recently-activated logogens might remain available as a source of linguistic information long after such estimates of capacity would suggest if they were taken to apply to active memory as a whole, provided that the appropriate semantic cues were available for access to logogens. Secondly, the postulate that output from logogens to the response buffer

¹in an unpublished undergraduate research project entitled 'Phonemes and Space', Oxford University, 1973.

is not automatic allows the possibility that in the perception of continuous prose, function words might well not be entered into the response buffer, thus increasing its effective capacity. Thirdly, the idea that since inactive memory is content-addressable, items in any component of active memory may serve as 'addresses' to regions of inactive memory, provides a further means by which the reader or hearer may retain contact with recent context. However, a rather more profound analysis of the active memory requirements of comprehension is obviously needed.

3) Attention. That there may be selection of logogen outputs for entry to the response buffer does not imply that there are no attentional limitations prior to logogens in the pattern-recognition system (cf. Treisman, 1969). Nevertheless, I suggested that in Sperling's task (Section 2.34) activation by the display elements of the appropriate logogens might occur effectively in parallel, and that the major limit on performance was the rate at which representations could be established in the response buffer, the latter being necessary for the recall of items. It is possible that a similar explanation might be applied in the case of dichotic memory; this might explain the partial dependence of the advantage of ear-by-ear report on whether order errors are scored (see Broadbent, 1971, p.161) though interpretation is complicated by the relatively long duration of PAS. If I am right in thinking that immediate recognition is mediated by trace strength at logogens, then an obvious experimental approach to these problems is to use a recognition paradigm and measure RT. This should escape the 'read-out' bottleneck and lay bare any pre-logogen limitations on divided attention.

4) Non-verbal immediate memory. The work reported in this thesis has been restricted to immediate memory for verbal items. In Section 3.132(v) it was considered problematic that reasonably typical results were obtained in the Sternberg immediate recognition paradigm even when initially unfamiliar visual figures were used as stimuli. Moreover, evidence has recently been presented (Kroll et al., 1970; Phillips & Baddeley, 1971) which implies the survival in appropriate conditions of short-term visual representations for much longer periods of time than have hitherto been thought possible (save in the case of individuals capable of eidetic imagery). At present, it is not possible to say whether such results may be interpreted in terms of pre-categorical visual storage, maintained perhaps by a capacity for rehearsal or central regeneration of sensory traces of the sort mentioned briefly in Section 3.131, or whether a more radical theoretical solution is required. If there is a non-sensory short-term visual store, it is hard to see why there should be so little evidence for its use in ordinary short-term memory tasks, even when the availability of verbal coding is restricted (by, for instance, suppressing compatible articulation - Levy, 1971 - see Section 2.35).

5) Neuropsychology. It is only very recently that brain-damaged patients have been found whose consequent behavioural disorders may be identified as specific to short-term as opposed to long-term memory tasks. Warrington and associates (Warrington & Shallice, 1969, 1972; Warrington & Weiskrantz, 1973) have studied one such patient, KF, in some detail. KF is particularly interesting, in that his existence may prove difficult to account for in terms of our model as it stands. KF has considerable difficulty in repeating what he has just heard, and shows signs of a mild nominal dysphasia. With

auditory presentation, his digit span is very small, and on the Brown-Peterson distractor task, his performance is poor and the normal rapid decay component virtually absent. However, his performance on long-term memory tasks is comparable to that of normal subjects. In Section 1.11 this evidence was mentioned in support of the active/inactive memory distinction and as implying that representation is at least some component of active memory was not a prerequisite of entry into inactive memory. So far, so good. What may present some difficulty for us is that KF's performance on short-term memory tasks is not nearly so severely impaired when presentation is visual. Warrington & Shallice (1972) argue that this implies the existence of a post-iconic visually-coded short-term store for verbal material. In terms of our model, if the damage is at or subsequent to the logogens, there should be no modality difference. But if the damage is prior to the logogens, i.e. in the auditory analysis → logogens pathway, we would expect difficulties in identifying spoken words. Yet KF is, according to Warrington & Shallice (1969) able to identify single spoken words and detect target words in a sequence when presentation is auditory and fairly fast. One possibility that merits further investigation is whether perhaps the order in which speech elements (at the distinctive feature level) become available to the logogens may be garbled as a result of the damage. This might not greatly impair KF's ability to identify single short words, detect short words in a string or identify words in a slow sequence as in the long-term memory tasks, but would hinder the identification either of short words in a fast sequence or of longer words. It would certainly be interesting to know whether KF's performance on the Sternberg IRn task is normal with auditory presentation of several items. If it were, in terms of our model this would certainly rule out damage up

to or including the logogens. The deficit suffered by KF that is common to both modalities of presentation is more readily interpretable in terms of our model. Warrington & Shallice (1969) showed that there was no motor speech defect. But it remains possible that the capacity of KF's response buffer is severely reduced or that his capacity for sub-vocal rehearsal is disrupted by damage to the response buffer -> logogens pathway.

Clearly KF and other similar patients merit further study, and the problem of the modality difference emphasises the importance of determining the nature of short-term visual retention, as mentioned under (4).

9.2. Experiments.

The evidence presented in the four experimental chapters converges upon the model described in Section 1.4. along three avenues. Firstly, we have seen that two experimental factors, within-list acoustic similarity and the provision of a position cue with the probe had clearly contrasting effects on tasks requiring reference to stored order information (contextual recall and probed location) and tasks which do not (item recognition), even though care was taken to match experimental parameters across tasks (Exps. I- V). Secondly the nature of the effects on the former group of tasks favoured the hypothesis that the order/position of items is consulted primarily via their representation in a non-associative temporally-ordered store, whose contents are a sequence of phonemically-coded items. Thirdly, while the absence of effects of acoustic similarity and a position cue on IRn told us relatively little about the representations mediating that task, Exp. VI provided more positive evidence. It was shown that manipulations of the recency of negative probes and of the items of the memory set influenced IRn RT in a manner compatible with the hypothesis that IRn is mediated primarily by inspection of the (decaying) trace-strength at a representation to which the probe provides direct access (Model III, Section 3.11). Moreover, the results appear incompatible with an exhaustive or self-terminating scanning model, and with a version of the trace-strength hypothesis which assumes that negative decisions are made by default. Only Exp. VIII yielded any difficulties for our model, in that the results seemed to suggest that in IRn, a directed search strategy might be possible; but it was argued that the results could also be interpreted in terms of the trace strength inspection model.

Ways of modifying this experiment to make it more decisive were suggested in Ch.8. Also the manipulation of the recency of the memory set in Exp. VI did not produce a statistically significant effect, though its direction was that predicted by the trace-strength inspection model, and means were suggested in Ch.7 whereby a more marked and reliable effect might be obtained.

The research strategy suggested by the logogens/response buffer distinction as applied to active memory seems therefore to have yielded dividends. It has proved possible to tease apart the two forms of representation by measuring reaction times in tasks which test memory for a recent short sequence of verbal items by probing different attributes of the sequence. But this is only a beginning. There is no shortage of experiments waiting to be done. The most crucial need at present is for experiments comparing the effect on IRn and CR (and the other tasks) of experimental factors which do influence item recognition. So far, direct comparisons of IRn and CR have shown systematic effects of particular variables on CR, none on IRn; a demonstration of a dissociation in the opposite direction is therefore needed. Among the factors which may serve this purpose are the following:

- 1) Semantic similarity. Logogens are (by virtue of their associative links to other logogens) semantic representations, and there should be a 'generalisation' of trace strength between semantically related logogens. It should, in IRn, take longer to reject, but less time to accept, a probe when the memory set contains a close synonym of the probe item than when it does not. No effect of semantic similarity on response buffer search is expected, so provided that effects of semantic similarity at input or during rehearsal are adequately controlled, CR should be unaffected by this variable.

Ideally, memory set and probe should be presented in different modalities, lest the subject make use of physical features of the words.

2) Recency of the memory set. In Exp. VI, it was shown that a filled delay between list and probe reduced the slope of the IRn set-size function (at least for some subjects). A comparable experiment using the CR task would be expected to show an opposite effect. A filled delay would be expected to result in degradation of the representations in the response buffer, slower read-out and thus an increase in the slope of the CR set-size function.

3) Rate of presentation. I have suggested (Section 2.34, 3.334) that the response buffer has a limiting rate of input much lower than that of the logogen system. During some pilot studies intended primarily to determine optimal rates of presentation for the control of rehearsal, I obtained some informal data in line with this suggestion. At the highest rate of presentation I tried for CR - 400 msec./item - error frequency already showed a marked increase over that found at lower rates. Yet performance in the IRn task at this same rate was, if anything, faster than and as accurate as that at lower rates. Even at almost double this rate - 220 msec./item - though there was by now an increase in error frequency, IRn performance was still remarkably good. At this latter rate, to perform the task was a bizarre experience. One felt that one was responding totally automatically and more or less randomly. One was unable to recall the entire sequence, and it is of particular interest that one underestimated the number of items in the memory set. Subjects accused me of presenting four-item lists when they were expecting five items. This phenomenal loss of items which are nevertheless adequately represented to mediate IRn would certainly bear following up, since in addition to present concerns, it carries the implication

that transfer of items from the logogen system to the response buffer may conceivably be involved in temporal ordering and numerosity judgements in perception. Given the theoretical and practical problems inherent in the use of RTs when error rate is high, it may be preferable to use a computer-controlled threshold-estimation technique, to determine maximum rates of presentation for a given level of accuracy on different probe tasks.

4) Interpolated irrelevant items. One way to disrupt a trace-strength strategy for IRn would be to present irrelevant items, from the same stimulus ensemble, between memory set and probe. If an accuracy pay-off is applied subjects should be driven to search the response buffer even in IRn, leading to radical changes in, for instance, serial position effects. Since a search of the response buffer is already required in CR, any effect of interpolated items should be comparable to that of adding extra items to the memory set, and no radical change in the pattern of results would be expected.

9.3. What is active memory?

Ideas about the nature of memory have on the whole been implicitly or explicitly stimulus-oriented. The assumption that memory is a record of (sensory) experience is very old. Plato compared memory to a wax tablet on which experience was engraved and Aristotle considered the act of remembering to consist in perceiving a present impression. Augustine referred to the 'great cave of memory' in which images were laid away to be brought forth when there was need of them. The theme is also very deeply rooted in the empiricist tradition. Locke's description of memory as 'the storehouse of our ideas' (where for 'idea' one may read 'image' without doing much violence to the sense), the power of the mind to revive perceptions it has once had, is implicit also in the philosophy of Hume, James Mill, and even William James. Yet we are creatures of action, possessing skills, and must also have response memories; as Miller, Galanter & Pribram (1960) have it, to the Image corresponds the Plan. By response memories, I do not mean representations of kinaesthetic experience, importantly related to response memories as these are, but representations of action patterns, motor commands, action schemata. The most obvious example is the internal representation of a learned skill, like riding a bicycle. But the importance of response memories extends beyond 'automatic' motor skills. For instance, accounts of cognitive development, both Soviet and Swiss, have emphasised the role of the development of 'action-schemata' in mediating a child's discovery of the structure of the world. Recent research on the crucial role of organisation in long-term memory has suggested that what subjects do to the material, overtly or covertly, is at least as important a determinant of whether

and how it will be remembered as what the experimenter and the material do to the subjects. Even that archetype for the idea of memory as a quasi-sensory record - the classical rhetor skilled in the Art of Memory (Yates, 1966) - had to remember which way he must next turn in his progress through the theatre or temple so vividly depicted in his mind's eye.

The same sensory bias appears to have influenced ideas about short-term memory. In the last fifteen years, a great deal of work has been devoted to the search for behavioural evidence demonstrating the existence of active or primary memory as a temporary storage mechanism to be distinguished from that involved in long-term memory. The early work emphasised the sensory characteristics of this labile form of representation, such as decay and acoustic coding. However, the discovery of modality-specific forms of sensory storage of rather brief duration, subject to overwriting or decay, has tended to push the box with 'primary memory' written on it further towards the centre of the large black box representing the central nervous system. In fact, it has become something of a problem to know where to put it in relation to the senses and long-term memory; it has also proved difficult to specify to what level its contents have been analysed and in what form they are represented. So far, however, the possibility that there exist short-term response memories which may play a role in short-term memory tasks has with a few honourable exceptions received little emphasis. Yet some form of temporary storage for the responses one is about to emit is probably essential. A phrase, a sequence of musical notes, or the actions involved in writing a sequence of graphic characters take some time to emit once they have been centrally determined, and their emission seems in some sense 'automatic'. While it is possible that each component of the

response is activated directly from the relevant representation in inactive memory, it is, as Lashley (1951) argued to such good effect, highly unlikely. It is characteristic of behavioural action patterns that they are serially ordered; that they are sometimes very fast suggests that the later component elements of a response sequence must be assembled prior to the initiation of the first. It is therefore entirely plausible that there exist short-term response memories which hold a buffered sequence of outputs which may subsequently be executed in order, freeing the central processes to monitor outcomes and prepare future responses. It may seem semantically deviant to call a mechanism a memory when it represents events which have not yet (overtly) occurred. However, in this thesis I have laboured to show that the response buffer for speech, originally hypothesised by Morton (1970), may serve to represent perceptual events recoded as potential responses. More than this, use of the response buffer in most short-term memory tasks is essential because of the requirement that the serial order of recent events be retained, and because it is possible to recirculate its contents to the logogen system - the process of sub-vocal rehearsal.

If a serially ordered response buffer for speech does indeed play the major role in short-term verbal memory tasks that I have argued, a speculative but conceptually simple answer to the question, 'What is active memory?' may now be given. It may no longer be necessary to postulate a rather mysterious 'primary memory' mechanism separate from the sensory and response systems and from inactive memory. Active memory may consist of no more than dynamic traces at the permanent structural units in the brain representing, respectively, sensory attributes, the motor components of responses, and the concepts or category states mediating between them. This may be

spelled out in a little more detail. We may distinguish three components of active memory as a whole:

1) Precategorical sensory stores. These are modality-specific, of relatively brief duration, non-associative (in Wickelgren's, 1972 sense - see Section 2.11) and subject to overwriting or decay. There may, as suggested earlier (Section 3.131) be some capability for generation of sensory traces in the absence of sensory input, by 'efferent' activation from the conceptual representations. This capability (which constitutes a form of rehearsal) may also enable the life of stimulus-generated sensory traces to be usefully extended, though the capacity and nature of this process have yet to be determined for the different modalities. The organization of the sensory stores is appropriate to the modalities they serve. Precategorical acoustic storage (Crowder & Morton, 1969) must be organized so as to preserve the temporal order of speech elements. The organization of the sensory store specific to the visual modality appears on the other hand to be primarily spatial in character. (This is not to say that there is no retention of temporal information in iconic memory; for instance, the recent unpublished work of A.M. Treisman and associates suggests that direction of movement may be coded in the icon.)

2) Response buffer stores. The response buffer hypothesised for speech is a limited-capacity non-associative store which holds an ordered sequence of articulatory commands for a duration rather longer than that characteristic of the precategorical sensory stores. Its contents may be fed back to the conceptual representations, which Morton calls logogens, for identification, or emitted as overt responses. Responses which have entered it may be lost primarily through displacement by subsequent entries, although some decay may also

occur. It is probable that there exist response buffers specific to response modalities other than speech. If their contents may also be recirculated back to conceptual representations (i.e. rehearsed) then it is possible that they too play a part in the performance of some short-term memory tasks. Note that some response modalities are spatially organised, which raises the possibility that response buffers specific to them may have a function in the retention of spatial information.

3) The third component consists of active traces at representations of the amodal concepts or category-states, or at the pathways, mediating between the sensory and response modalities (or between sensory and sensory, or response and response modalities). Wickelgren speaks of such representations as being in the 'conceptual modality'. In his sense this form of storage is associative: the representations may be directly accessed given the appropriate stimulus or response attributes. This kind of active trace seems to be destroyed in the absence of rehearsal both by the passage of time and by the activation of other traces. We have considered only logogens, the representations of words, but there are probably representations of non-verbal category-states which may function in the same way.

All of these forms of storage may be involved in immediate memory for a short sequence of verbal items. Each may be destroyed by a distracting task which prevents rehearsal. But they differ in the attributes they retain, in their modes of access, and in the manner in which the traces they retain are lost. Discovery of the properties of the different components of active memory requires the use of short-term memory tasks which selectively tap the individual

components. I have tried to indicate some of the ways in which this may be achieved.

- AARONSON, D. (1967). Temporal factors in perception and short-term memory. Psychol. Bull., 67, 130-144.
- AARONSON, D. (1968). Temporal course of perception in an immediate recall task. J.E.P., 76, 129-140.
- AARONSON, D. and STERNBERG, S. (1964). Effects of presentation rate and signal-to-noise ratio on immediate recall. Paper read at a meeting of the Eastern Psychological Association.
- ALLPORT, D.A. (1973). Word recognition and the recognition buffer. Paper presented at a meeting of the Experimental Psychology Society, University College, London. January 1973.
- ANDERS, T.R. and LILLYQUIST, T.D. (1971). Retrieval time in forward and backward recall. Psychon. Sci., 22, 205-206.
- ANDERSON, J.R. and BOWER, G.H. (1972). Recognition and retrieval processes in free recall. Psychol. Rev., 79, 97-123.
- ANDERSON, N.S. (1960). Post-stimulus cuing in immediate memory. J.E.P., 60, 216-221.
- 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.), Advances in the Psychology of Learning and Motivation Research and Theory. Vol. 2. Academic Press: New York.
- AUDLEY, R.J. (1971). Tutorial review: some observations on theories of choice reaction time. Paper read at Attention and Performance IV: Boulder, Colorado.
- AVERBACH, E. and CORRIELL, A.S. (1961). Short-term memory in vision. Bell Systm. Tech. J., 40, 309-328.

- BADDELEY, A.D. (1964). Immediate memory and the 'perception' of letter sequences. Q.J.E.P., 16, 364-347.
- BADDELEY, A.D. (1966a). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. Q.J.E.P., 18, 362-365.
- BADDELEY, A.D. (1966b). The influence of acoustic and semantic similarity on long-term memory for word sequences. Q.J.E.P., 18, 302-309.
- BADDELEY, A.D. (1968). How does acoustic similarity influence short-term memory? Q.J.E.P., 20, 249-264.
- BADDELEY, A.D. (1970). Estimating the short-term component in free recall. B.J.P., 61, 13-15.
- BADDELEY, A.D. (1972). Human memory. In: P.C. Dodwell (Ed.), New Horizons in Psychology. 2. Penguin: Harmondsworth, England.
- BADDELEY, A.D., CONRAD, R. and HULL, A.J. (1965). Predictability and immediate memory for consonant sequences. Q.J.E.P., 17, 175-177.
- BADDELEY, A.D. and ECOB, J.R. (1970). Reaction time and short-term memory: A trace strength alternative to the high-speed exhaustive scanning hypothesis. Technical Report No. 13. Center for Human Information Processing, University of California at San Diego.
- BADDELEY, A.D. and PATTERSON, K. (1971). Relation between long-term and short-term memory. Brit. Med. Bull., 27, No. 3, 237-242.
- BADDELEY, A.D., SCOTT, D., DRYNAN, R. and SMITH, J.C. (1971). Short-term memory and the limited capacity hypothesis. B.J.P., 60, 51-55.

- BERTELSON, P. (1965). Serial choice reaction-time as a function of response versus signal and response repetition. Nature, 206, 217-218.
- BOWER, G.H. (1967). A multi-component theory of the memory trace. In: K.W. Spence and J.T. Spence (Eds.), Advances in the Psychology of Learning and Motivation Research and Theory. Vol. 1. New York: Academic Press.
- BOWER, G.H. (1970). Organizational factors in memory. Cognitive Psychol., 1, 18-46.
- BRAND, J. (1971). Classification without identification in visual search. Q.J.E.P., 23, 178-186.
- BRIGGS, G.E. and BLAHA, J. (1969). Memory retrieval and central comparison times in information processing. J.E.P., 79, 395-402.
- BROADBENT, D.E. (1958). Perception and Communication. London: Pergamon Press.
- BROADBENT, D.E. (1966). Two-state threshold model and rating scale experiments. J. Ac. Soc. Amer., 40, 244-245.
- BROADBENT, D.E. (1970). Review Lecture: Psychological aspects of short-term and long-term memory. Proc. Roy. Soc. Lond. B., 175, 333-350.
- BROADBENT, D.E. (1971). Decision and Stress. London: Academic Press.
- BROADBENT, D.E. and GREGORY, M. (1964). Stimulus set and response set: The alternation of attention. Q.J.E.P., 16, 309-317.
- BROWN, J. (1958). Some tests of the decay theory of immediate memory. Q.J.E.P., 10, 12-21.
- BROWN, J. (1959). Information, redundancy, and the decay of the memory trace. In: D.V. Blake and A.M. Uttley, (Eds.), Proceedings of the Symposium on Mechanisation of Thought Processes. London. H.M.S.C.

- BRYDEN, M.P. (1967). A model for the sequential organisation of behaviour. Canad. J. Psychol., 21, 37-56.
- BURROWS, D. (1972). Modality effects in retrieval of information from S.T.M. Percept. & Psychophys., 11, 365-372.
- BURROWS, D. and MURDOCK, B.B. Jnr. (1969). Effects of extended practice on high speed scanning. J.E.P., 82, 231-237.
- BURROWS, D. and OKADA, R. (1971). Serial position effects in high speed memory search. Percept. & Psychophys., 10, 305-308.
- CAVANAGH, J.P. (1972). Relation between the immediate memory span and the memory search rate. Psychol. Rev., 79, 525-530.
- CERMAK, L.S. (1972). Human memory, research and theory. New York: Ronald.
- CHASE, W.G. and POSNER, M.I. (1965). Paper presented at Mid-Western Psychological Association Meeting. (See: Posner, 1967, for details.)
- CHASE, W.G. and CALFEE, R.C. (1969). Modality and similarity effects in short-term recognition memory. J.E.P., 81, 510-514.
- CHECKOSKY, S.F. (1971). Speeded classification of multi-dimensional stimuli. J.E.P., 87, 383-388.
- CHI, M.T.H. and CHASE, W.G. (1972). Effects of modality and similarity on context recall. J.E.P., 96, 219-222.
- CLARKE, F.R. (1957). Constant ratio-rule for confusion matrices in speech communication. J. Ac. Soc. Amer., 29, 715-720.
- CLIFTON, C. Jnr., and BIRENBAUM, S. (1970). Effects of serial position and delay of probe in a memory scan task. J.E.P., 86, 69-76.
- COLLINS, A.M. and QUILLIAN, M.R. (1972). Experiments on semantic memory and language comprehension. In: L.W. Gregg (Ed.), Cognition in Learning and Memory. New York: Wiley.

- CONNOR, J.M. (1972). Serial and parallel encoding processes in memory and visual search. J.E.P., 96, 363-370.
- CONRAD, R. (1959). Errors of immediate memory. B.J.P., 50, 349-359.
- CONRAD, R. (1960). Serial order intrusions in immediate memory. B.J.P., 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. (1964). Acoustic confusions in immediate memory. B.J.P., 55, 75-84.
- CONRAD, R. (1965). Order error in immediate recall of sequences. Journ. Verb. Learn. Verb. Behav., 4, 161-169.
- CONRAD, R. (1967). Interference or decay over short retention intervals? Journ. Verb. Learn. Verb. Behav., 6, 49-54.
- CONRAD, R. (1970). Short-term memory processes in the deaf. B.J.P., 61, 179-195.
- CONRAD, R. (1972). Short-term memory in the deaf: a test for speech coding. B.J.P., 63, 173-180.
- CONRAD, R. and HULL, A.J. (1964). Information, acoustic confusion and memory span. B.J.P., 55, 429-432.
- CORBALLIS, M.C. (1967). Serial order in recognition and recall. J.E.P., 74, 99-105.
- CORBALLIS, M.C., KIRBY, J. and MILLER, A. (1972). Access to elements of a memorised list. J.E.P., 94, 185-190.
- CORCORAN, D.W.J. (1971). Pattern Recognition. Harmondsworth, England: Penguin.
- CRAIK, F.I.M. (1970). The fate of primary memory items in free recall. Journ. Verb. Learn. Verb. Behav., 9, 143-148.
- CRAIK, F.I.M. (1971). Primary memory. Brit. Med. Bull., 27, No. 3, 232-236.

- CRAIK, F.I.M. and LOCKHART, R.S. (1972). Levels of processing: A framework for memory research. Journ. Verb. Learn. Verb. Behav., 11, 671-684.
- CROSSMAN, E.R.F.W. (1960). Information and serial order in immediate memory. In: C. Cherry (Ed.), Information Theory. London: Butterworths.
- CROWDER, R.G. and MORTON, J. (1969). Pre-categorical acoustic storage (PAS). Percept. & Psychophys., 5, 365-373.
- CUMMING, G. and COLTHEART, M. (1969). Short-term memory: Are item and position information stored independently? Psychon. Sci., 15, 216-218.
- DARLEY, C.F. and ARABIE, P. (1972) I am grateful to Saul Sternberg for details of this study, a description of which is now available in: ATKINSON, R.C. HERMANN, D.T. and WESCOURT, K.T. (1973) Search processes in recognition memory. Tech. Report No. 204, Institute for Mathematical Studies in the Social Sciences, Stanford University, June 1971.
- DARLEY, C.F., KLATZKY, R.L. and ATKINSON, R.C. (1972). Effects of memory load and reaction time. J.E.P., 96, 232-234.
- DEROSA, D.V. (1969). Transformation on sets in short-term memory: set size reduction by deletion. J.E.P., 82, 415-426.
- DEROSA, D.V. and BAUMGARTE, E.R. (1971). Probe digit recall of items from temporally organized memory sets. J.E.P., 91, 154-158.
- DEROSA, D.V. and BECKWITH, M. (1971). Retrieval of information from organized memory sets. Psychon. Sci., 23, 177-179.
- DEROSA, D.V. and MORIN, R.E. (1970). Recognition reaction time for digits in consecutive and non-consecutive memorized sets. J.E.P., 83, 472-479.
- DONALDSON, W. (1971). Retention of order and item information. J.E.P., 90, 293-296.
- DONALDSON, W. and GLATHE, H. (1969). Recognition memory for order and item information. J.E.P., 82, 557-560.

- DONDERS, F.C. (1969). Over de snelheid van psychische processen. Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool, 1868-69, Tweede reeks, II, 92-120. Translation by W.G. Koster In: W.G. Koster (Ed.), Attention and Performance II. Acta Psychol., 1969, 30, 412-31.
- DUMAS, D., GROSS, E., CHECKOSKY, S.F. (1972). Effects of attribute probability in a memory search task. J.E.P., 93, 327-332.
- EBBINGHAUS, H. (1885). Über das Gedächtnis. Duncker & Humblot: Leipzig.
- EBENHOLTZ, S.M. (1963). Position mediated transfer between serial learning and a spatial discrimination task. J.E.P., 65, 603-608.
- EICHELMAN, W.M. (1970). Stimulus and response repetition effects for naming letters at two response-stimulus intervals. Percept. & Psychophys., 7, 94-96.
- ELLIS, S.H. and CHASE, W.G. (1971). Parallel processing in item recognition. Percept. & Psychophys., 10, 379-384.
- FALMAGNE, J.C. (1965). Stochastic models for choice reaction time with applications to experimental results. J.M.P., 12, 77-124.
- FISHER, D.F., MONTY, R.A. and GLUCKSBERG, S. (1969). Visual confusion matrices: fact or artifact? J. Psychol., 71, 111-125.
- FORRIN, B. and MORIN, R.E. (1969). Recognition times for items in short and long-term memory. In: W.G. Koster (Ed.), Attention and Performance II. Acta Psychologica, 30, 126-141.
- FOSS, D.J. and DOWELL, B.E. (1971). High speed memory retrieval with auditory presented stimuli. Percept. & Psychophys., 9, 465-468.
- FOZARD, J.L. & YNTEMA, D.B. (1966). The effect of repetition on the apparent recency of pictures. Paper presented at the meeting of the Eastern Psychological Association, New York.

- FREUND, R.D., LOFTUS, G.R. and ATKINSON, R.C. (1969). Applications of multiprocess models for memory to continuous recognition tasks. J.M.P., 6, 576-594.
- FUCHS, A.H. (1969). Recall for order and content of serial word-lists in short-term memory. J.E.P., 82, 14-21.
- FUCHS, A.H. (1971). Response latency and serial position in short-term memory. Psychon. Sci., 22, 75-76.
- GLANZER, M. and CUNITZ, A.R. (1966). Two storage mechanisms in free recall. Journ. Verb. Learn. Verb. Behav., 5, 351-360.
- GLANZER, M., GIANUTSOS, R., and DUBIN, S. (1969). The removal of items from short-term storage. Journ. Verb. Learn. Verb. Behav., 8, 435-437.
- GLASSMAN, W.F. (1972). Sub-vocal activity and acoustic confusions in S.T.M. J.E.P., 96, 164-169.
- GOLDRING, S.L. (1969). The effect of presentation and test mode on short-term retention of words. Technical Report No. 68. Center for Cognitive Learning, University of Wisconsin.
- GORFEIN, D.S. and JACOBSON, D.E. (1972). Proactive effects in short-term recognition memory. J.E.P., 95, 211-214.
- GRABOI, D. (1971). Searching for targets: the effects of specific practice. Technical Report No. 9. Center for Human Information Processing, University of California at San Diego.
- GRAY, J.A. and WEDDERBURN, A.A. (1960). Grouping strategies with simultaneous stimuli. Q.J.E.P., 12, 180-184.
- GUMENIK, W.E. (1969). Effects of articulatory activity and auditory, visual and semantic similarity on the short-term memory of visually presented paired associates. J.E.P., 82, 70-74.
- HARVEY, N. (1973). Coding across modalities. Doctoral thesis in preparation, University of Oxford.

- HAWKINS, H.L. and HOSKING, K. (1969). Stimulus probability as a determinant of discrete choice reaction time. J.E.P., 82, 435-440.
- HAYS, W.L. (1963). Statistics. New York: Holt, Rinehart and Winston.
- HEBB, D.O. (1949). Organisation of Behaviour. New York: Wiley.
- HENDERSON, L. and MATTHEWS, M.L. (1971). Perception and memory loss of item and order information in S.T.M. Percept. & Psychophys., 9, 231-233.
- HINRICHS, J.V. (1970). A two-process memory-strength theory for judgement of recency. Psychol. Rev., 77, 223-233.
- HINRICHS, J.V. and BUSCHKE, H. (1968). Judgement of recency under steady-state conditions. J.E.P., 78, 574-579.
- HINTZMAN, D.L. (1967). Articulatory coding in short-term memory. Journ, Verb. Learn. Verb. Behav., 6, 312-316.
- HINTZMAN, D.L. (1969). Recognition time, effects of recency, frequency and the spacing of repetitions. J.E.P., 79, 192-194.
- HITCH, G.J. (1972). Organisation and retrieval in immediate memory. Unpublished Doctoral thesis, University of Cambridge.
- HODGE, D.C. (1962). Legibility of a uniform strokewidth alphabet. J. Engin. Psychol., 1, 34-46.
- HUNTER, I.M.L. (1964). Memory. Harmondsworth, England: Penguin.
- HYMAN, R. (1953). Stimulus information as a determinant of reaction-time. J.E.P., 45, 188-196.
- INGLEBY, J.D. (1969). Decision-making processes in human perception and memory. Unpublished doctoral thesis: University of Cambridge.
- JAHNKE, J.C. (1969). The Ranschburg Effect. Psychol. Rev., 76, 592-605.
- JAHNKE, J.C. (1972). Distribution of repeated and non-repeated target elements and short-term recognition memory. J.E.P., 96,

- JAMES, W. (1890). The Principles of Psychology. New York: Henry Holt & Co.
- JENSEN, A.R. and ROHWER, W.D. Jnr. (1965). What is learned in serial learning? Journ. Verb. Learn. Verb. Behav., 4, 62-67.
- JUOLA, J.F. and ATKINSON, R.C. (1971). Memory-scanning for words versus categories. Journ. Verb. Learn. Verb. Behav., 10, 522-527.
- JUOLA, J.F., FISCHLER, I., WOOD, C.T. and ATKINSON, R.C. (1971). Recognition time for information stored in long-term memory. Percept. & Psychophys., 10, 8-14.
- KAMINSKY, C.A. and DEROSA, D.V. (1972). Influence of retrieval cues and set organization on short-term recognition memory. J.E.P., 96, 449-454.
- KANTOWITZ, B.H., ORNSTEIN, P.A. and SCHWARTZ, M. (1972). Encoding and immediate serial recall of consonant strings. J.E.P., 93, 105-110.
- KAY, H, and POULTON, E.C. (1951). Anticipation in memorising. B.J.P., 42, 34-41.
- KENNEDY, R.A. (1968). Response latency in the serial learning of short lists. B.J.P., 59, 1-5.
- KENNEDY, R.A. and HAMILTON, D. (1969). Time to locate probe items in short lists of digits. Amer. J. Psychol., 82, 272-275.
- KEPPEL, G. and UNDERWOOD, E.J. (1962). Proactive inhibition in short-term retention of single items. Journ. Verb. Learn. Verb. Behav., 1, 153-161.
- KINTSCH, W. (1970). Learning, Memory and Conceptual Processes. New York: Wiley.
- KINTSCH, W. and BUSCHKE, H. (1969). Homophones and synonyms in short-term memory. J.E.P., 80, 403-407.

- KIRSNER, K. and CRAIK, F.I.M. (1971). Naming and decision processes in short-term recognition memory. J.E.P., 83, 149-157.
- KIRSNER, K. (1972). Naming latency facilitation: an analysis of the encoding component in recognition RT. J.E.P., 95, 171-176.
- KLATZKY, R.L. and ATKINSON, R.C. (1970). Memory scans based on alternative test stimulus representations. Percept. & Psychophys., 8, 113-117.
- KLATZKY, R.L. and ATKINSON, R.C. (1971). Specialization of the cerebral hemispheres in scanning for information in short-term memory. Percept. & Psychophys., 10, 335-338.
- KLATZKY, R.L., JUOLA, J.F. and ATKINSON, R.C. (1971). Test stimulus representation and experimental context effects in memory scanning. J.E.P., 87, 281-288.
- KLATZKY, R.L. and SMITH, E.E. (1972). Stimulus expectancy and retrieval from short-term memory. J.E.P., 94, 101-107.
- KONORSKI, J. (1961). The Physiological Approach to the Problem of Recent Memory. In: J.F. Delafresnay, (Ed.). Brain Mechanisms and Learning. Oxford: Basil Blackwell.
- KONORSKI, J. (1967). Some new ideas concerning the physiological mechanisms of perception. Acta Biologicae Experimentalis, 27, 147-161.
- KRISTOFFERSON, M.W. (1972a). Effects of practice on character classification performance. Canad. J. Psychol., 26, 54-60.
- KRISTOFFERSON, M.W. (1972b). When item recognition and visual search functions are similar. Percept. & Psychophys., 12, 379-384.
- KRISTOFFERSON, M.W. (1972c). Types and frequency of errors in visual search. Percept. & Psychophys., 11, 325-328.

- KROLL, N.E.A., PARKS, T., PARKINSON, S.R., BIEBER, S.L. and JOHNSON, A.L. (1970). Short-term memory while shadowing: recall of visually and aurally presented letters. J.E.P., 85, 220-224.
- KRUEGER, L.E. (1970). Effect of stimulus probability on two-choice reaction time. J.E.P., 84, 377-379.
- KRUEGER, L.E. (1971). Effect of direction of sequential presentation and redundancy on short-term recognition memory. Percept. & Psychophys., 9, 121-124.
- LAMING, D.R.J. (1968). Information Theory of Choice Reaction Times. New York: Academic Press.
- LANDAUER, T.K. (1962). Rate of implicit speech. Percept. & Mot. Skills., 15, 646.
- LASHLEY, K.S. (1951). The problem of serial order in behaviour. In: L.A. Jeffress, (Ed.), Cerebral Mechanisms in Behaviour: The Hixon Symposium. New York: Wiley.
- LAUCHERY, K.R. and PINKUS, A.L. (1966). Short-term memory: Effects of acoustic similarity, presentation rate and presentation mode. Psychon. Sci., 6, 285-286.
- LEVY, B.A. (1971). Role of articulation in auditory and visual short-term memory. Journ. Verb. Learn. Verb. Behav., 10, 123-132.
- LIVELY, B.L. (1972). Speed/accuracy trade-off and practice in a memory search task. J.E.P., 96, 97-103.
- LIVELY, B.L. and SANFORD, B.J. (1972). The use of category information in a memory search task. J.E.P., 93, 379-385.
- LOESS, H. (1964). Proactive inhibition in short-term memory. Journ. Verb. Learn. Verb. Behav., 3, 362-368.

- McCLEOD, P.D., WILLIAMS, C.E. and BROADBENT, D.E. (1971). Free recall with assistance from one and from two retrieval cues. B.J.P., 62, 59-65.
- McNICOL, D. (1970). Progressive decay of the memory trace. Psychon. Sci., 20, 196.
- McNICOL, D. (1971a). The origins of transpositions in immediate memory. Aust. J. Psychol., 23, 9-17.
- McNICOL, D. (1971b). The confusion of order in short-term memory. Aust. J. Psychol., 23, 77-84.
- McNICOL, D. (1971c). Short-term recall of messages containing repeated items. Aust. J. Psychol., 23, 241-251.
- MACNEILAGE, P.F. (1970). Motor control of serial ordering of speech. Psychol. Rev., 77, 182-196.
- MADIGAN, S.A. and McCABE, L. (1971). Perfect recall and total forgetting: a problem for models of S.T.M. Journ. Verb. Learn. Verb. Behav., 10, 101-106.
- MANDLER, G. (1972). Organization and recognition. In: E. Tulving and W. Donaldson (Eds.), Organization and Memory. New York: Academic Press.
- MARCEL, A.J. (1970). Some constraints on sequential and parallel processing and the limits of attention. In: A.F. Sanders (Ed.), Attention and Performance III. Acta Psychologica, 33, 77-92.
- MASSARO, D.W. (1970). Preperceptual auditory images. J.E.P., 85, 411-417.
- MATTHEWS, M.L. and HENDERSON, L. (1970). Fast presentation rate and the recall of item and order information. Nature, 226, 374-376.

- MAYZNER, M.S. and SCHOENBERG, L. (1964). Single letter and digram frequency effects in immediate serial recall. Journ. Verb. Learn. Verb. Behav., 3, 397-400.
- MELTON, A W. (1963). Implications of short-term memory for a general theory of memory. Journ. Verb. Learn. Verb. Behav., 2, 1-21.
- MELTON, A.W. and IRWIN, J.M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. Amer. J. Psychol., 53, 173-203.
- MERICKLE, P.M., COLTHEART, M. and LOWE, D.G. (1971). On the selective effects of a patterned masking stimulus. Canad. J. Psychol., 25, 264-279.
- MILLER, G.A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol. Rev., 63, 81-97.
- MILLER, G.A. and NICELY, P.E. (1955). An analysis of perceptual confusions among some English consonants. J. Ac. Soc. Amer., 27, 338-352.
- MILLER, G.A., GALANTER, E. and PRIBRAM, K.H. (1960). Plans and the Structure of Behaviour. New York: Holt, Rinehart and Winston.
- MILNER, P.M. (1961). A neural mechanism for the immediate recall of sequences. Kybernetik, 1, 76-81.
- MORIN, R.E., DEROSA, D.V. and STULZ, V. (1967). Recognition memory and reaction time. In: A.F. Sanders (Ed.), Attention and Performance I. Acta Psychol., 27, 298-305.
- MORIN, R.E., DEROSA, D.V. and ULM, R. (1967). Short-term recognition memory for spatially isolated items. Psychon. Sci., 9, 617-618.

- MORTON, J. (1964a). A preliminary functional model for language behaviour. International Audiology, 1964, 3, 216-225.
(Reprinted in R.C. Oldfield and J.C. Marshall (Eds.), Language. Harmondsworth, England: Penguin.)
- MORTON, J. (1964b). A model for continuous language behaviour. Language and Speech, 7, 40-70.
- MORTON, J. (1964c). The effects of context upon speed of reading, eye movements and eye-voice span. Q.J.E.P., 16, 340-354.
- MORTON, J. (1968). Repeated items and decay in memory. Psychon. Sci., 12, 219-220.
- MORTON, J. (1969a). The interaction of information in word recognition. Psychol. Rev., 76, 165-178.
- MORTON, J. (1969b). Categories of interference: Verbal mediation and conflict in card-sorting. B.J.P., 60, 329-346.
- MORTON, J. A functional model for memory. In: D.A. Norman (Ed.), Models of Human Memory. New York: Academic Press.
- MORTON, J. and BROADBENT, D.E. (1967). Passive versus active models, or is your homunculus really necessary?
In: W. Wathen-Dunn (Ed.), Models for the Perception of Speech and Visual Form. Cambridge, Massachusetts: M.I.T. Press.
- MORTON, J. and HOLLOWAY, C.M. (1970). On the absence of cross-modal suffix effects in S.T.S. Q.J.E.P., 22, 167-176.
- MORTON, J., CROWDER, R.G. and PRUSSIN, H.A. (1971). Experiments with the stimulus suffix effect. J.E.P., 91, 161-190.
- MOSS, S.M. and SHARAC, J. (1970). Accuracy and latency in short-term memory: evidence for a dual retrieval process. J.E.P., 84, 40-46.
- MURDOCK, B.B. Jnr. (1961). The retention of individual items. J.E.P., 62, 618-625.

- MURDOCK, B.B. Jnr. (1963a). Short-term retention of single paired-associates. J.E.P., 65, 433-443.
- MURDOCK, B.B. Jnr. (1963b). Interpolated recall in short-term memory. J.E.P., 66, 525-532.
- MURDOCK, B.B. Jnr. (1965). Signal detection theory and short-term memory. J.E.P., 70, 443-447.
- MURDOCK, B.B. Jnr. (1966). The criterion problem in short-term memory. J.E.P., 72, 317-324.
- MURDOCK, B.B. Jnr. (1967). Recent developments in short-term memory. B.J.P., 58, 421-433.
- MURDOCK, B.B. Jnr. (1968a). Serial order effects in S.T.M. J.E.P., (Monograph Supplement No. 42), 76, 1-15.
- MURDOCK, B.B. Jnr. (1968b). Response latencies in immediate memory. Q.J.E.P., 20, 79-82.
- MURDOCK, B.B. Jnr. (1971). A parallel process model for scanning. Percept. & Psychophys., 10, 289-291.
- MURDOCK, B.B. Jnr. (1972). Short Term Memory. In: G.H. Bower and J.T. Spence (Eds.), The Psychology of Learning and Motivation. Vol. 5. New York: Academic Press.
- MURDOCK, B.B. Jnr. and VOM SAAL, W. (1967). Transpositions in short-term memory. J.E.P., 74, 137-143.
- MURDOCK, B.B. Jnr. and WALKER, K.D. (1969). Modality effects in free recall. Journ. Verb. Learn. Verb. Behav., 8, 665-676.
- MURRAY, D.J. (1966). Vocalization-at-presentation and immediate recall, with varying recall methods. Q.J.E.P., 18, 9-18.
- MURRAY, D.J. (1967). The role of speech responses in immediate memory. Canad. J. Psychol., 21, 263-275.
- MYERS, C.S. (1911). A text book of experimental psychology, with laboratory exercises. Cambridge.

- NEISSER, U. (1954). An experimental distinction between perceptual process and verbal response. J.E.P., 47, 399-402.
- NEISSER, U. (1963). Decision time without reaction time: experiments in visual scanning. Amer. J. Psychol., 76, 376-385.
- NEISSER, U. (1964). Visual search. Scientific American, 210, (June), 94-102.
- NEISSER, U. (1967). Cognitive Psychology. New York: Appleton-Century-Crofts.
- NICKERSON, R.S. (1966). Response times with a memory-dependent decision task. J.E.P., 72, 761-769.
- NICKERSON, R.S. (1972). Binary-classification reaction time: A review of some human information-processing capabilities. Psychon. Monogr. Suppl., 4, No. 17 (Whole No. 65.).
- NORMAN, D.A. (1966). Acquisition and retention in short-term memory. J.E.P., 72, 369-381.
- NORMAN, D.A. (1968). Toward a theory of memory and attention. Psychol. Rev., 75, 522-536.
- NORMAN, D.A. (1969a). Memory while shadowing. Q.J.E.P., 21, 85-94.
- NORMAN, D.A. (1969b). Memory and attention. New York: Wiley.
- NORMAN, D.A. (Ed.) (1970). Models of Human Memory. New York: Academic Press.
- NORMAN, D.A. (1973). What have the animal experiments taught us about human memory? In: J.A. Deutsch (Ed.), The Physiological Basis of Memory. New York and London: Academic Press.
- NORMAN, D.A. and WICKELCREN, W.A. (1965). Short-term recognition for single digits and pairs of digits. J.E.P., 70, 479-489.

- NORMAN, D.A. and WAUGH, N.C. (1968). Stimulus and response interference in recognition experiments. J.E.P., 78, 551-559.
- NORMAN, D.A. and WICKELGREN, W.A. (1969). Strength theory of decision rules and latency in short-term memory. J.M.P., 6, 192-208.
- NORMAN, D.A. and RUMELHART, D.E. (1970). A system for perception and memory. In: D.A. Norman (Ed.), Models of Human Memory. New York: Academic Press.
- OKADA, R. (1971). Decision latencies in short-term recognition memory. J.E.P., 90, 27-32.
- PAIVIO, A. (1971). Imagery and Verbal Processes. New York: Holt, Rinehart and Winston.
- PETERSON, L.R. (1967). Search and Judgement in Memory. In: B.J. Kleinmuntz (Ed.), Concepts and the Structure of Memory. New York: Wiley.
- PETERSON, L.R. and PETERSON, M.J. (1959). Short-term retention of individual verbal items. J.E.P., 58, 193-198.
- PETERSON, L.R., JOHNSON, S.T. and COATNEY, R. (1969). The effect of repeated occurrences on judgements of recency. Journ. Verb. Learn. Verb. Behav., 8, 591-596.
- PETRUSIC, W.M. and DILLON, R.F. (1972). Proactive interference in short-term recognition and recall memory. J.E.P., 95, 412-418.
- PEW, R.W. and GARDNER, G.T. (1965). Unpublished data, University of Michigan. Reported in D.F. Fisher, R.A. Monty, S. Glucksberg. (1969). J. Psychol., 71, 111-125.
- PHILIPS, W.A. and BADDELEY, A.D. (1971). Reaction-time and short-term visual memory. Psychon. Sci., 22, 73-74.
- POSNER, M.I. (1964). Rate of presentation and order of recall in immediate memory. B.J.P., 55, 303-306.

- POSNER, M.I. (1967). Short-term memory systems in human information processing. In: A.F. Saunders (Ed.), Attention and Performance I. Acta Psychol., 27, 267-284.
- POSNER, M.I. (1969). Abstraction and the process of recognition. In: J.T. Spence and G. Bower, (Eds.), The Psychology of Learning and Motivation. Vol. 3. New York: Academic Press.
- POSNER, M.I. and ROSSMAN, E. (1965). Effects of size and location of informational transforms upon short-term retention. J.E.P., 70, 496-505.
- POSNER, M.I. and KONICK, A.F. (1966). On the role of interference in short term retention. J.E.P., 72, 221-231.
- POSNER, M.I. and KEELE, S.W. (1967). Decay of visual information from a single letter. Science, 158, 137-139.
- POSNER, M.I. and MITCHELL, R.F. (1967). Chronometric analysis of classification. Psychol. Rev., 74, 392-409.
- POSNER, M.I., BOIES, S.J., EICHELMAN, W.H. and TAYLOR, R.L. (1969). Retention of visual and name codes of single letters. J.E.P. Monogr., 79, (1, pt. 2).
- POSTMAN, L. (1961). The present status of interference theory. In: C.N. Cofer (Ed.), Verbal Learning and Verbal Behaviour. New York: McGraw-Hill.
- POSTMAN, L. (1964). Short-term and incidental learning. In: A.W. Melton (Ed.), Categories of Human Learning. New York: Academic Press.
- RABBITT, P.M.A. (1967). Learning to ignore irrelevant information. Amer. J. Psychol., 80, 1-13.
- RABBITT, P.M.A. (1968). Repetition effects and signal classification strategies in serial choice-response tasks. Q.J.E.P., 20, 232-240.

- RABBITT, P.M.A. (1971). Times for analysing stimuli and selecting responses. Brit. Med. Bull., 27, 259-265.
- RAYMOND, B. (1969). Short-term storage and long-term storage in free recall. Journ. Verb. Learn. Verb. Behav., 8, 567-574.
- REICHER, G., LIGON, E. and CONRAD, C. (1969). Interference in short-term memory. J.E.P., 80, 95-100.
- REITMAN, J.S. (1970). Computer Simulation of an Information Processing Model of Short-Term Memory. In: D.A. Norman (Ed.), Models of Human Memory. New York: Academic Press.
- REMINGTON, R.J. (1969). Analysis of sequential effects in choice reaction times. J.E.P., 82, 250-257.
- ROSS, J. (1970). Extended practice with a single character classification task. Percept. & Psychophys., 8, 276-278.
- RUMELHART, D.E. (1970). A multicomponent theory of the perception of briefly exposed visual displays. J.M.P., 7, 191-218.
- RUNDUS, D.J. (1970). An analysis of rehearsal processes in free recall. Technical Report No. 159. Institute for Mathematical Studies in the Social Sciences, Stanford University, California.
- RYAN, J. (1969). Temporal grouping, rehearsal and short-term memory. Q.J.E.P., 21, 148-155.
- SANDERS, A.F. and VAN BORSELEN, J.W. (1966). Continuing memory and information processing. J.E.P., 71, 844-848.
- SCARBOROUGH, D.L. and STERNBERG, S. (1967). Processing items and their order in sequential visual displays. Paper presented to the Eastern Psychological Association, Boston.
- SELFRIDGE, O.G. (1959). Pandemonium: A Paradigm of Learning. In: D.V. Blake and A.M. Uttley (Eds.), Proceedings of the Symposium on Mechanisation of Thought Processes. London: H.M.S.O

- SEYMOUR, P.H.K. (1973). A model for reading, naming and comparison. B.J.P., 64, 35-49.
- SHULMAN, H.G. (1970). Encoding and retention of semantic and phonemic information in S.T.M. Journ. Verb. Learn. Verb. Behav., 9, 499-508.
- SHULMAN, H.G. and GREENEBURG, S.N. (1971). Perceptual deficit due to division of attention between memory and perception. J.E.P., 88, 171-176.
- SIMPSON, P.J. (1972). High-speed memory scanning: stability and generality. J.E.P., 96, 239-246.
- SMITH, E.E. (1967). Effects of familiarity on stimulus recognition and categorization. J.E.P., 74, 324-332.
- SPERLING, G. (1960). The information available in brief visual presentations. Psychol. Monog., 74, No. 11. (Whole No. 498).
- SPERLING, G. (1963). A model for visual memory tasks. Human Factors, 5, 19-31.
- SPERLING, G. (1967). Successive approximations to a model for short-term memory. In: W.G. Koster (Ed.), Attention and Performance I, Acta Psychol., 27, 285-292.
- SPERLING, G. and SPEELMAN, R.G. (1970). Acoustic similarity and auditory short-term memory. In: D.A. Norman (Ed.), Models of Human Memory. New York: Academic Press.
- STERNBERG, S. (1966). High speed scanning in human memory. Science, 153, 652-654.
- STERNBERG, S. (1967a). Retrieval of contextual information from memory. Psychon. Sci., 8, 55-56.
- STERNBERG, S. (1967b). Two operations in character recognition: some evidence from reaction-time measurements. Percept. & Psychophys., 2, 45-53.

- STERNBERG, S. (1969a). Memory-scanning: Mental processes revealed by reaction-time measurements. Amer. Sci., 57, 421-457.
- STERNBERG, S. (1969b). The discovery of processing stages: Extensions of Donder's method. In: W.G. Koster (Ed.), Attention and Performance II, Acta Psychol., 30, 276-315.
- STERNBERG, S. (1973). Memory scanning: new findings and current controversies. Fourth Sir Frederick Bartlett Lecture. Delivered at a meeting of the Experimental Psychology Society at University College, London in January, 1973,
- SWANSON, J.M. and BRIGGS, G.E. (1969). Information processing as a function of speed versus accuracy. J.E.P., 81, 223-229.
- SWANSON, J.M., JOHNSEN, A.M. and BRIGGS, G.E. (1972). Recoding in a memory search task. J.E.P., 93, 1-9.
- SWENSSON, R.G. and EDWARDS, W. (1971). Response strategies in a two-choice reaction task with continuous cost for time. J.E.P., 88, 67-81.
- THEIOS, J. (1972). Reaction time measurements in the study of memory processes: theory and data. Technical Report No 72-2. Wisconsin Mathematical Psychology Program, University of Wisconsin.
- THEIOS, J. and SMITH, P.G. (1970). Sequential dependencies in choice reaction times. Technical Report No. 70-8. Wisconsin Mathematical Psychology Program, University of Wisconsin.
- THEIOS, J., SMITH, P.G., HAVILAND, S.E., TRAUPMAN, J. and MOY, M.C. (1971). Memory scanning is a serial, self-terminating process. Technical Report No. 71-1. Wisconsin Mathematical Psychology Program, University of Wisconsin.
- TITCHENER, E.B. (1905). Experimental Psychology: Instructor's manual, quantitative. New York.

- TREISMAN, A.M. (1964). Monitoring and storage of irrelevant messages in selective attention. Journ. Verb. Learn. Verb. Behav., 3, 449-459.
- TREISMAN, A.M. (1969). Strategies and models of selective attention. Psychol. Rev., 76, 282-299.
- TREISMAN, A.M. and GEFEN, G. (1967). Selective attention: Perception or response? Q.J.E.P., 19, 1-16.
- TULVING, E. and ARBUCKLE, T.Y. (1963). Sources of intra-trial interference in immediate recall of paired-associates. Journ. Verb. Learn. Verb. Behav., 1, 321-334.
- TULVING, E. and PATKAU, J.E. (1962). Concurrent effects of contextual constraints and word frequency on immediate memory and learning of verbal material. Canad. J. Psychol., 16, 83-95.
- TULVING, E. and PEARLSTONE, Z. (1966). Availability versus accessibility of information in memory for words. Journ. Verb. Learn. Verb. Behav., 5, 381-391.
- TULVING, E. and COLOTLA, V.A. (1970). Free recall of trilingual lists. Cognitive Psychol., 1, 86-98.
- TULVING, E. and MADIGAN, S.A. (1970). Memory and verbal learning. Ann. Rev. Psychol., 21, 437-484.
- TULVING, E. and THOMPSON, D.M. (1971). Retrieval processes in recognition memory: effects of associative context. J.E.P., 87, 116-124.
- TULVING, E. and DONALDSON, W. (Eds.). (1972). Organisation of Memory. New York: Academic Press.
- UNDERWOOD, B.J. (1969). Attributes of Memory. Psychol. Rev., 76, 559-573.

- WARRINGTON, E.K. (1971). Neurological disorders of memory. Brit. Med. Bull., 27, No. 3, 243-247.
- WARRINGTON, E.K. and SHALLICE, T. (1969). The selective impairment of auditory verbal short-term memory. Brain, 92, 885-896.
- WARRINGTON, E.K. and SHALLICE, T. (1972). Neuropsychological evidence of visual storage in short-term memory tasks. Q.J.E.P., 24, 30-40.
- WARRINGTON, E.K. and WEISKRANTZ, L. (1973). Short- and long-term memory deficits in man. In: J.A. Deutsch (Ed.), The Physiological Basis of Memory. New York and London: Academic Press.
- WATTENBARGER, B.L. (1968). Speed and accuracy set in visual search. Paper presented at a meeting of the Mid-Western Psychological Association, Chicago.
- WATTENBARGER, B.L. and PACHELLA, R.G. (1972). The effect of memory load on R.T. in character classification. Percept. & Psychophys. 12, 100-102.
- WAUGH, N.C. (1969). The effects of recency and repetition on recall latencies. In: W.G. Koster (Ed.), Attention and Performance II, Acta Psychol., 30, 115-125.
- WAUGH, N.C. (1970). Retrieval time in S.T.M. B.J.P., 61, 1-12.
- WAUGH, N.C. and NORMAN, D.A. (1965). Primary memory. Psychol. Rev., 72, 89-104.
- WAUGH, N.C. and NORMAN, D.A. (1968). The measure of interference in primary memory. Journ. Verb. Learn. Verb. Behav., 7, 617-626.
- WEBER, R.J., CROSS, M. and CARLTON, M. (1968). Searching circular sequences. J.E.P., 78, 588-592.
- WEBER, R.J. and CASTLEMAN, J. (1969). Metered memory search. Psychon. Sci., 16, 311-312.

- WEBER, R.J. and BLAGOWSKY, J. (1970). Metered memory search with implicit and explicit scanning. J.E.P., 84, 343-348.
- WEISKRANTZ, L. (1970). A long-term view of short-term memory in psychology. In: G. Horn and R.A. Hinde (Eds.), Short-Term Changes in Neural Activity and Behaviour. Cambridge, England: Cambridge University Press.
- WELFORD, A.T. (1968). Fundamentals of skill. London: Methuen.
- WICKELGREN, W.A. (1965a). Acoustic similarity and retro-active interference in short-term memory. Journ. Verb. Learn. Verb. Behav., 4, 53-62.
- WICKELGREN, W.A. (1965b). Distinctive features and errors in short-term memory for English vowels. J. Acou. Soc. Amer., 38, 583-588.
- 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.E.P., 70, 102-108.
- WICKELGREN, W.A. (1965e). Short-term memory for repeated and non-repeated items. Q.J.E.P., 17, 14-25.
- WICKELGREN, W.A. (1966a). Phonemic similarity and interference in short-term memory for single letters. J.E.P., 71, 396-404.
- WICKELGREN, W.A. (1966b). Distinctive features and errors in short-term memory for English consonants. J. Acou. Soc. Amer., 39, 388-398.
- WICKELGREN, W.A. (1966c). Short-term recognition for single letters and phonemic similarity of retro-active interference. Q.J.E.P., 18, 55-62.
- WICKELGREN, W.A. (1966d). Associative intrusions in short-term memory. J.E.P., 72, 853-858.

- WICKELGREN, W.A. (1967a). Rehearsal grouping and hierarchical organisation of serial position cues in immediate memory. Q.J.E.P., 19, 97-102.
- WICKELGREN, W.A. (1967b). Exponential decay and independence from irrelevant associations in short-term memory for serial order. J.E.P., 73, 165-171.
- WICKELGREN, W.A. (1968). Testing two-state theories with operating characteristics and a posteriori probabilities. Psychol. Bull., 69, 126-131.
- WICKELGREN, W.A. (1969a). Auditory or articulatory coding in verbal short-term memory. Psychol. Rev., 76, 232-235.
- WICKELGREN, W.A. (1969b). Context-sensitive coding, associative memory, and serial order in (speech) behaviour. Psychol. Rev., 76, 1-15.
- WICKELGREN, W.A. (1970a). Multi-trace strength theory. In: D.A. Norman (Ed.), Models of Human Memory. New York: Academic Press.
- WICKELGREN, W.A. (1970b). Time, interference and rate of presentation in short-term memory for items. J.M.P., 7, 219-235.
- WICKELGREN, W.A. (1972). Coding, Retrieval and Dynamics of Multi-Trace Associative Memory. In: L.W. Gregg (Ed.), Cognition in Learning and Memory. New York: Wiley.
- WICKELGREN, W.A. and NORMAN, D.A. (1966). Strength models and serial position in short-term recognition memory. J.M.P., 3, 316-347.
- WICKELGREN, W.A. and WHITMAN, P.T. (1970). Visual very short-term memory is non-associative. J.E.P., 84, 277-281.
- WICKELGREN, W.A. and BERIAN, K.M. (1971). Dual trace theory and the consolidation of L.T.M. J.M.P., 8, 404-417.

- WICKELGREN, W.A. and NORMAN, D.A. (1971). Invariance of forgetting rate with number of repetitions in short-term recognition memory. Psychon. Sci., 22, 363-364.
- WILCOX, L. and WILDING, J.M. (1970). Effect of item similarity on the speed of memory search. Nature, 227, 1152-1153.
- WINGFIELD, A. and BRANCA, A.A. (1970). Strategy in high-speed memory search. J.E.P., 83, 63-67.
- WINOGRAD, E. (1968). List differentiation as a function of frequency and retention interval. J.E.P., 76, Monog. Suppl. 1-18.
- WOLF, T.M. and JAHNKE, J.C. (1968). The effects of intra-serial repetition on short-term recognition and recall. J.E.P., 77, 572-580.
- WOODWARD, A.E. and MURDOCK, B.B. Jnr. (1968). Positional and sequential probes in serial learning. Canad. J. Psychol., 22, 131-138.
- WOODWORTH, R.S. (1938). Experimental Psychology. New York: Holt, Rinehart and Winston.
- YATES, F.A. (1966). The Art of Memory. London: Routledge and Kegan Paul.
- YELLOTT, J.I. Jnr. (1967). Correction for guessing in choice reaction time. Psychon. Sci., 8, 321-322.
- YNTEMA, D.B. and TRASK, F.P. (1963). Recall as a search process. Journ. Verb. Learn. Verb. Behav., 2, 67-74.
- YOUNG, R.K. (1968). Serial Learning. In: T.R. Dixon and D.L. Horton (Eds.), Verbal Behaviour and General Behaviour Theory. New Jersey: Prentice-Hall.
- ZECHMEISTER, E.B. (1971). Short-term recognition memory for words: why search? J.E.P., 89, 265-273.

