

Chapter 8: Conclusions

A Solution to the Problem of Serial Order?

How then do we store and retrieve a sequence of items in the correct order? The argument in this thesis has been that, for short-term memory at least, order is stored by associating each item with the start and end of the sequence. The relative strengths of these associations provide an approximate code for the item's position in the sequence. This code is stored together with the item to form a position-sensitive token. The order is retrieved by reinstating the positional codes and using them to cue tokens in memory. The evidence for this argument is summarised below.

Summary of Thesis

Chapter 1 introduced three possible solutions to the problem of serial order: chaining theory, positional theory, and ordinal theory. Experiment 1 in Chapter 2 failed to find any evidence for chaining theory in immediate serial recall, whereas Experiments 2 and 3 in Chapter 3 found evidence supporting positional theory. This evidence took the form of positional errors, either as transpositions between groups that maintain their position within groups (Experiment 2) or intrusions between trials that maintain their position within a trial (Experiment 3). These errors cannot be explained by ordinal theory.

Chapter 4 examined three more specific models of serial recall and used meta-analyses to argue that none was sufficient to capture the complete pattern of errors in short-term, serial recall. A new positional model was developed in Chapter 5 (the Start-End Model, SEM) that could reproduce the complete pattern of errors. This model was based on positional theory, and demonstrated how items can be ordered by cuing with approximate positional codes.

Chapter 6 examined the nature of those positional codes in more detail: in particular, whether the codes represent absolute position, or position relative to the start and end of a sequence. Experiments 4 and 5 resembled Experiments 2 and 3, except that they used groups and lists of different lengths. The positional errors in these cases supported the notion of relative position, in agreement with the predictions of SEM.

Finally, Chapter 7 examined the effects of repeated items in sequences, which pose problems for most models of serial recall. The robust yet complex nature of these effects suggested several additional factors contribute to recall of repeated items. Nonetheless, the effects appeared best explained by SEM: in particular, with its assumption that items are stored as position-sensitive tokens.

The Start-End Model

It is important to distinguish theory and model in this thesis. The theory that memory for serial order utilises positional information, where that information is defined relative to the start and end of a sequence, is based on experimental results from tasks such as serial recall. The model, SEM, is a more specific implementation of this theory, which makes further assumptions about short-term memory and the serial recall process, in order to fit data quantitatively. The success of SEM in fitting present data supports the more general positional theory. Nonetheless, the validity of the theory does not depend on the success of the model; SEM may be refuted by future data without necessarily refuting the theory.

SEM was reviewed in Chapter 5, where it was discussed in relation to further aspects of short-term memory, and compared with other models. However, it is useful to step away from the details of experiments and models, and consider more general implications of a positional theory of memory for serial order.

Serial Order in Short-term Memory

The problem of serial order in the present thesis has been confined to short-term memory for a novel sequence of items. The example given in Chapter 1 was of holding an unfamiliar telephone number in memory long enough to dial it. The present solution, in terms of storing each digit with a positional code defined by start and end markers, and reinstating these codes during dialling, may not seem particularly intuitive. Several questions might be asked of such a solution. For example, is not generating and reinstating positional codes for each digit in the telephone number somewhat laborious? What are the start and end markers that define these codes? How is an end marker used if the length of the telephone number is unknown? How is the order of positional codes themselves reinstated?

The answers to these questions are interrelated. Firstly, generating and reinstating positional codes is probably an automatic, unconscious process. One does not necessarily have to think “a 5 was near the start of the telephone number and a 9 was near the end”. In other words, the start and end markers may represent intrinsic properties of short-term memory that are not open to introspection. This also makes their interpretation difficult (which is why their psychological definition has remained vague in the present thesis). In fact, it is unclear whether much elaboration of start and end markers is possible in psychological terms. For example, there may be one group of neurons in the brain whose activity is triggered by the first digit in the telephone number and decreases with each subsequent digit, and another group of neurons whose activity increases with each subsequent digit. What is the psychological interpretation of these neurons? The answer can only be simply that the activity of the first group represents proximity to the start of the telephone number and the activity of the second group represents proximity to the end of the telephone number. This is equivalent to saying the groups of neurons are start and end markers.¹

The question of how an end marker can grow towards the end a sequence, when the end is not known in advance, was addressed in Chapter 6, with one answer requiring only an approximate level of expectation for the end of the sequence. With respect to the present example however, two points are worth noting. Firstly, an end marker may not be employed. The order of digits in the telephone number may be stored with reference to a start marker only (similar perhaps to the start-of-list context proposed by Page & Norris, 1996b). SEM is still able to store order with a single marker. Secondly, one may split the telephone number into groups of a predetermined size (i.e., the number may be grouped subjectively). In this case, an end marker can be employed to help define positions within a group, because the end of a group is known in advance. Alternatively, coding of position relative to the end of a sequence might be achieved during cumulative rehearsal of the digits (Chapter 5). Thus prior knowledge of the length of a telephone number, though helpful, is not necessary, particularly if the number is grouped subjectively or rehearsed cumulatively.

1. Note that few other positional theories have adequate psychological interpretation of the positional information they assume (e.g., Lee & Estes, 1981; Nairne, 1991), except perhaps those that attribute this information to temporal oscillators (e.g., Burgess & Hitch, 1996b), which have problems explaining the results in Chapter 6.

Finally, there is the concern that a positional theory simply displaces the problem of ordering the digits of the telephone number to the problem of ordering the positional codes associated with each digit. However, the problem of ordering the positional codes can have a simpler solution than the problem of ordering the digits. For example, in the Houghton (1990) model, all that need be done is activate the start marker. This activation then decays in a fixed manner, and the activation of the end marker grows in a yoked manner. Positional codes are therefore ordered automatically, as a consequence of the simple dynamics of the start marker. Though the data in Chapter 6 argue against a yoked end marker, positional codes in SEM are also an automatic consequence of two simple functions, one decreasing from the start of a sequence and one increasing towards the end. Moreover, these functions can be a fixed property of short-term memory: Though the order of digits must be learned for each novel telephone number, the order of positional codes is invariant (analogous perhaps to a pegword mnemonic). Ordering positional codes is less of a problem than it might first appear.

In summary, the present account of our ability to remember a novel telephone number for a short period of time does not necessarily suffer from conflicting intuitions. Indeed, it is defensible with respect to data from short-term memory for similar sequences in the present experiments. The next question is whether this solution to the problem of serial order in short-term memory generalises to the problem of serial order in other aspects of memory.

Serial Order in Long-term Memory

The problem of serial order is fundamental to most aspects of memory. Chapter 1 gave two examples of serial order in procedural memory (ordering phonemes in speech) and in episodic memory (ordering life-events). Is positional theory appropriate for these examples?

Procedural Memory

Though positional theory is supported by data from short-term memory, it faces problems when applied to long-term, procedural memory. These problems were mentioned in Chapters 1 and 5, with the main problem being the interference problem: People can store and retrieve numerous sequences of the same basic elements, with little to no interference between such sequences. If those sequences were coded by a single start and end marker, much greater interference would be expected. In other words, the interference between groups in STM is

rarely found between chunks or motor programs in procedural memory (the possible exception of speech errors is discussed below).

One solution to the interference problem is that each sequence has its own, unique start and end marker (Houghton, 1990). However, for procedural memory at least, a simpler solution may be an ordinal one. Indeed, several ordinal accounts have been suggested along neurally-inspired dimensions (e.g., Grossberg, 1978; Nigrin, 1993; Page, 1994). For example, order might be stored in a primacy gradient of synaptic strengths between a neuron representing the sequence and others representing the sequence elements, a hypothesis which has some neurophysiological support (Granger, Whitson, Larson & Lynch, 1994). Because the synaptic strengths associated with one “sequence neuron” are independent of those associated with another, as many sequences can be stored, in principle, as there are free neurons. Order can be retrieved by activating a sequence neuron, which primes the “element neurons” in proportion to their associated synaptic strength, and activating each element neuron in a cyclic process of selection and suppression of the most primed neuron. Models based on these ideas have demonstrated long-term learning of temporal sequences that does not suffer from the interference problem (Nigrin, 1993; Page, 1994).

In addition to the interference problem, a further problem concerns memory for long sequences. Even ordinal models face problems with long sequences, given that most biological dimensions are finite (and noisy). Long sequences are best stored as a hierarchy of subsequences, with the hierarchical decomposition continuing until the smallest subsequences are within short-term memory span. This is consistent with experimental evidence. For example, Klahr, Chase and Lovelace (1983) found that latency profiles in probed recall of the alphabet showed marked discontinuities across certain letter pairs. These discontinuities suggested a chunk boundary.² For Americans, these boundaries were coincident with the phrase boundaries of the “alphabet song”, which is often used to teach the alphabet. However, though it is well known that the presence of such structure aids serial learning (e.g., Martin, 1974), there is a surprising dearth of adequate psychological models for such data.

2. That the same data can be fitted by a linear model with variable interitem strengths (Scharroo, 1994) is not surprising; what is required is a demonstration of how such a chaining model can solve the interference problem.

Models of speech production generally assume separate mechanisms for dealing with the order and content of an utterance (e.g., MacKay, 1982). The order is sometimes phrased in terms of syntax or schemata (Lashley, 1951). For example, some models assume a syllable schema (e.g., Dell, 1988), in order to explain speech errors such as spoonerisms, the transpositions of phonemes between words (e.g., “dear old queen” spoken as “*queer old dean*”). Though resembling positional errors in STM, these errors respect more specific syllabic constraints (Hartley & Houghton, 1996). Other models assume that spoonerisms are failures of an editing process that operates over a speech output buffer, rather than indicating syllabic coding in long-term memory per se (Levelt, 1989). In either case, speech errors clearly entail more than simply positional information and, as such, do not really constitute evidence for positional coding in long-term memory.

The assumption of schemata to order the content of an utterance does not directly address the problem of serial order however. The question remains as to how that order is represented. MacKay (1982) assumed a hierarchy of timing nodes, consistent with evidence for a binary hierarchy in speech production (Gordon & Meyer, 1987). However, the nature of such timing nodes is rarely specified. They may represent internal oscillators, as assumed for the rhythm found in the stress patterns of speech (Robinson, 1977). Such oscillators have also been assumed to store order in short-term memory (e.g., Brown et al., 1996; Burgess & Hitch, 1996a, 1996b), though Chapter 6 argued against these models. In general, more evidence is required to determine the relationship between serial order in procedural memory (and speech production in particular), and serial order in short-term memory.

Episodic Memory

Another example in Chapter 1 concerned the ordering of events in the past. This is an example of episodic memory, in which temporal order is fundamental (Tulving, 1983). Short-term memory can also be viewed as an a form of episodic memory. As Nairne noted: “When items are forgotten from memory lists, it is not the items themselves that are forgotten, but rather their occurrences in prior spatiotemporal windows.” (Nairne, 1991, p. 332).

In the case of autobiographical memory, considerable research has focused how people date past events. Friedman (1993) distinguished three main theories. His distance theory

coded time along a single dimension (e.g., strength of memory), resembling ordinal theory; his location theory coded time by relating it to temporal schemata (e.g., days of the week), resembling positional theory; and his relative theory coded the order of two events (e.g., that “X occurred before Y”), resembling chaining theory. In reviewing the evidence, he argued for a location theory where events are dated via temporal schemata, occasionally supplemented by memory for relative order (and in rare cases, knowledge of exact dates). A single continuum, along which events recede hazily into the past, may be more illusory than real.

The main evidence for temporal schemata comes from scale effects, where memory for fine temporal detail is superior to that for coarse temporal detail. For example, an event attributed to the wrong week can still be attributed to the correct day, and such memory is not explicable simply by guesses based on general knowledge (Friedman & Wilkins, 1985). These multiple scales (e.g., day-in-week, week-in-month) resemble those seen in short-term memory (e.g., positions of item-in-group, group-in-list; Chapter 5). Unlike short-term memory however, the dating of past events is more likely to involve indirect, reconstructive inferences (e.g., “that was on wash-day, Monday”; Larsen & Thompson, 1995), rather than direct positional codes. Nevertheless, in the case of generalisations such as “towards the start of the week” or “towards the end of the week” there may be an additional role for the type of start and end markers assumed here for short-term memory. These markers would allow judgements of the order of events in the same week whose exact days could not be inferred.

In laboratory tests of episodic memory, positional theory is clearly supported. People are not only able to judge the frequency (Hintzman, 1976), recency (Yntema & Trask, 1963) and duration (Block, 1982) of events, but also their position within temporal sequences. Toggia and Kimble (1976), for example, showed that approximate positional judgements were above chance even under incidental learning of long lists of 96 words. Such judgements not only show primacy and recency effects, but also positional errors between lists (Hintzman, Block & Summers, 1973). Even positions of repeated items can be judged accurately, and often independently (Hintzman & Block, 1971), supporting multiple representations of repeated items (as in SEM). Given that covert serial recall of such long lists is highly improbable, these data constitute strong support for positional information in episodic memory.

Hintzman, Block and Summers (1973) attributed positional judgements to contextual associations, distinguishing reinstatable and nonreinstatable contexts, as in SEM (Chapter 5). Block (1982) showed that judgements of relative duration of lists were affected by both intrinsic and extrinsic contextual change between lists, but judgements of position were not. These results can be explained if duration judgements are based on changes in nonreinstatable context (e.g., SEM's general context), whereas position judgements are based on reinstatable context (e.g., SEM's positional codes).

Some evidence suggests that knowledge of relative order may also aid episodic judgements. Tzeng, Lee and Wetzel (1979) used a "study-phase retrieval model" to explain why displaced rehearsals do not disrupt judgements of temporal order. Displaced rehearsals are rehearsals of items that are retrieved during the study of later items, and tend to improve judgements of their relative order. This is problematic for positional models, in which displaced rehearsals will, if anything, recode the retrieved and studied items in adjacent positions, and hence impair discrimination of their relative order. In the study-phase retrieval model, displaced rehearsals are used to encode the relative order of the two items. This explains why judgements of relative order in categorised lists are better for intracategory than intercategory judgements, even though the latter are better separated in time and position (Tzeng & Cotton, 1980). However, that the relative order of two items can be coded explicitly is unequivocal; how such codes could underlie all positional judgements is equivocal. Numerous codes for relative order would be necessary to judge positions accurately, and they could not explain positional errors (Hintzman, Block & Summers, 1973). Coding relative order during study-phase retrievals would appear optional rather than necessary or sufficient.

In summary, serial order in episodic memory does utilise positional codes. With the lists of random words employed in most laboratory tasks, the start and end of a sequence provide the only salient means with which to define position. Together with the notion of nonpositional general context, a theory based on start and end markers, like SEM, can explain the basic findings in episodic judgement tasks, including serial position effects and positional errors.³ With more meaningful events in autobiographical memory, temporal order is likely to

3. Indeed, an alternative reading of SEM is "Search of Episodic Memory" (cf. Raaijmakers & Shiffrin, 1981).

be supplemented by inferences based on temporal schemata, with occasional help from general knowledge and the explicit coding of relative order.

Other Memory

There are other examples of serial order that are not easily classified as procedural or episodic memory. For example, memory for the order of British Monarchs is perhaps better classed as semantic memory (assuming the order has not been learned by rhyme, as in the alphabet example above). This order is more likely to be reconstructed by a series of propositional facts (e.g., “William and Mary must have followed James II because they replaced his Catholic rule”); inferences that do not fall naturally into any of the chaining, positional or ordinal theories in Chapter 1. Indeed, it is arguable whether such reconstruction counts as true memory for serial order.

Another distinction often made is between implicit and explicit memory. Though not necessarily distinct memory systems (Schacter & Tulving, 1994), these memories differ in their access. Explicit memory refers to conscious recollection of past episodes, whereas implicit memory refers to nonconscious use of previously acquired information. These notions are distinct from procedural and episodic memory, though obviously related. There is some evidence to suggest that positional information is used only in explicit memory. For example, serial position effects arise in explicit but not implicit tests (Brooks, 1994) and in “remember” but not “know” responses (Jones & Roediger, 1995).

Positional information may also require explicit encoding (which is not necessarily precluded by the “incidental” learning conditions of the studies mentioned in Chapter 6). Nevertheless, it may still be possible to encode order information implicitly, through nonpositional means. This would be necessary to account for implicit learning of temporal sequences (e.g., Stadler, 1993), though this is a contentious issue (Shanks & St. John, 1994), and unlikely to apply to the Hebb effect (Chapter 5; Sechler & Watkins, 1991). In either case, the hypothesis that explicit encoding and retrieval is necessary to utilise positional information, but not order information, appears an interesting and testable hypothesis.

A complete review of other types of memory is beyond present concerns. The purpose of the above discussion is to suggest more than one solution to the problem of serial order in

memory. This suggestion is unusual, but not unprecedented (Ebenholtz, 1972; Frensch, 1994). One solution is clearly positional, and this applies not only to the short-term memory considered here, but episodic memory more generally. It may be contingent on explicit encoding and retrieval processes. Another solution is probably ordinal, providing a simpler means of storing and retrieving temporal order in procedural memory, with the potential to be acquired and expressed implicitly. Though these ideas remain speculative, the fundamental nature of serial order clearly warrants their further investigation.

Chapter Summary

The present thesis has demonstrated that positional information is utilised in short-term memory for serial order. More generally, such positional information appears common to episodic (explicit) memory. It may also underlie serial order in procedural or implicit memory, though other representations of order, such as simpler ordinal representations, should not be discounted. In either case, this chapter has confirmed the claims of Chapter 1, that addressing the problem of serial order in short-term memory provides a good starting point with which to address the problem of serial order more generally.