Memory & Cognition
1988, 16 (4), 343-352

A framework for interpreting recency effects in immediate serial recall

JAMES S. NAIRNE
University of Texas at Arlington, Arlington, Texas

A descriptive framework is offered for the interpretation of recency effects in immediate serial recall. Basic to the framework is a distinction between two types of trace features: (1) modality-dependent features, which represent the perceptual qualities of presentation, and (2) modality-independent features, which result from the set of encoding operations known as the "inner voice." Recency and modality effects emerge because certain types of modality-dependent (i.e., language-based) features are typically not subject to postlist interfering events and are likely to be sampled as discriminative cues in recall. The framework is used to interpret problematic findings in the modality effect literature, such as the effects of visual presentation, lipreading, mouthing, and stimulus class on the recall of recency items.

The question of how we remember recently presented information as it recedes backward in time has occupied the attention of memory theorists for some time. The typical finding that items near the end of a list are remembered best has served as an empirical cornerstone for a number of memory issues; for example, the separation between primary and secondary memory stores has relied on variations in the recency effect in free recall for much of its empirical base (see Greene, 1986, for a review). The concern of the present article, however, is with the recency effect in immediate serial recall, a procedure in which subjects are presented with relatively short lists of items to be recalled immediately in the exact order of presentation. Of particular interest is the finding that the size of the recency effect in immediate serial recall is dramatically affected by the specific modality of stimulus presentation. The *modality effect*, as it is termed, refers to the fact that recency performance is enhanced for auditorally presented items relative to silent visual controls (Conrad & Hull, 1968; Corballis, 1966; Craik, 1969; Murdock & Walker, 1969; Murray, 1966).

Although the modality effect has been the subject of a great deal of empirical research over the past 20 years, its theoretical basis has been attributed nearly universally to the inherent mnemonic superiority of auditory, usually sensory, memory traces. The most popular of these accounts has been the Precategorical Acoustic Storage (PAS) model (Crowder & Morton, 1969) in which the auditory recency advantage is proposed to accrue because storage time in auditory sensory memory (PAS) exceeds that of visual, or iconic, sensory memory. The PAS model has proven capable of explaining a wide range of empirical results (see Crowder, 1976, 1978a, for reviews), although recent demonstrations of substantial auditory-like serial recall patterns with nonauditory stimuli that are lipread (Campbell & Dodd, 1980; Spoehr & Corin, 1978) or silently mouthed (Greene & Crowder, 1984; Nairne & Walters, 1983) have proven difficult for the theory to handle. The theoretical underpinnings of the modality effect are therefore at issue, and some new accounts have been proposed. For example, Campbell and Dodd (1980) suggested that changing-state stimuli, through unstated mechanisms, may determine when recency effects are obtained; alternatively, Shand and Klima (1981) proposed that recency effects will be found whenever to-be-recalled stimuli are presented in a format that is compatible with the subject's normal dominant coding format in short-term memory. Despite some local success with the reported data at hand, neither of these proposals has proven capable of explaining the broad range of presentation conditions that can affect recency.

The purpose of the present article is to propose a general theoretical framework in which recency effects and, in particular, the modality effect might be explained. My intention is to provide a more extensive set of assumptions than those of Campbell and Dodd (1980) and Shand and Klima (1981), although the discussion remains at a qualitative level of analysis. As a result, I am not proposing a formal theory of immedite serial recall, but rather a set of ideas relevant to the interpretation of recency. My goal is to account for a wide range of problematic findings in the modality effect literature and, as a consequence, to act as a spur for further empirical work. The article is divided into two major sections: The first describes the theoretical assumptions in some detail, and the second applies those assumptions to particular empirical domains.

This article was originally written in 1985 during the time that I was on leave at UCLA. I would like to thank Ruth Campbell, Sue Gathercole, Art Glenberg, Robert Greene, and Alice F. Healy for valuable criticisms of the manuscript at various points in its development. Requests for reprints should be addressed to James S. Nairne, Department of Psychology, Box 19528, University of Texas, Arlington, TX 76019.

SUMMARY OF THE FRAMEWORK

Immediate serial recall is viewed in this article as a reconstructive process involving the analysis of multiattribute memory traces. These traces, formed during list presentation, are conceived as bundles of features or attributes (e.g., Bower, 1967) that the subject tries to interpret at recall by comparing each trace with the set of possible recallable items; in most immediate serial recall experiments, this set is relatively constrained (e.g., the digits 1 through 9). As a continuous record of immediate experience, it is assumed that the temporal orderings of these traces are preserved, perhaps through the sort of positional coding scheme described by Estes (1972). Ordered recall is accomplished by accessing each trace in the order in which it was established, with successful item recall being determined by how well the features of a particular trace specify a member of the recall set to the exclusion of others. Recency effects are a by-product of the fact that end-of-list memory traces tend to possess more identifying features, thereby allowing for better item selection at the time of recall. Recency differences as a function of modality are explained by assuming two things. First, an overwriting process is proposed to occur in which the features of a trace are degraded by subsequent input; the amount of overwriting is influenced by the similarity of successive input (Broadbent & Broadbent, 1981) and by whether items are perceived by the subject as belonging to the same "group" (Frankish & Turner, 1984; Kahneman, 1973). Second, those trace features that survive overwriting will aid recall as long as (1) they provide discriminative information about an item (i.e., those features uniquely specify a member of the recall set) and (2) the system is adapted to use those features as discriminative cues in recall.

Features of the Traces

In the present framework, immediate memory traces that result from serial list presentation are described in terms of two major classes of features or attributes: (1) modality-independent features resulting primarily from the set of encoding operations classified as the "inner voice," and (2) modality-dependent features as determined by the presented input. Any trace complex can be described in terms of both feature types, but it is assumed that in most situations the number of modality-independent features will exceed the number of modality-dependent features.

Modality-independent features. Any list item, regardless of the modality of presentation, is most likely to be encoded into immediate memory in terms of some sort of speechlike code (e.g., Conrad, 1964). Attributes encoded via the inner voice, although speechlike, are not tied to any particular presentation modality. Most importantly, auditory and visual presentation of an item are assumed to produce memory traces with identical inner-voice codes (i.e., an ensemble of speechlike features). In that sense, the term *speechlike* is a misnomer, because inner-voice features, as a class, are deemed to be dissimilar to any physical features encoded as a result of comprehending spoken language, including the features formed from the subject's own outer voice. (For evidence relevant to the modality-independent nature of inner voice encodings, see Geiselman & Glenny, 1977.) Although under various task demands, one can expect other modality-independent features to be a part of the trace (e.g., semantic or imaginal; see Shulman, 1972), such features are assumed to play a minimal role in most studies of immediate recall.

Modality-dependent features. The second major class of trace features consists of those physical, intraitem features that are unique to the particular mode of presentation. For items that are presented aloud, one can assume that specific auditory features (i.e., those unique to the particular voice) are represented as part of the trace complex (see Craik & Kirsner, 1974; Geiselman & Glenny, 1977). Likewise, visually presented input should lead to visually based trace features (see Broadbent, Vines, & Broadbent, 1978), and tactile input should lead to the encoding of tactile attributes (Nairne & McNabb, 1985; M. J. Watkins & O. C. Watkins, 1974). It is important to stress, however, that although these features represent the physical aspects of presentation, they are more aptly described as physical features "encoded by the system." For example, top–down contextual variables may lead the subject to encode only a selection of the possible physical features present in the nominal stimulus. In a study by Ayres, Jonides, Reitman, Egan, and Howard (1979), subjects were presented with a complex, but ambiguous, WA sound, which they previously had been biased, through instruction, to interpret either as a unit of speech or as a sound made by a trumpet. Although the nominally presented stimulus was the same in the two conditions, one can assume that the biasing manipulation produced two different sets of trace features: It is likely that those subjects expecting the speech WA tended to represent the speechlike aspects of the presented stimulus, whereas the nonspeechlike features were selected by the subjects expecting a trumpet sound. Any subsequent recall differences between the two groups, then, would be attributed to differences in the composition of the traces, even though those traces were formed from the same nominal stimulus (for a similar argument, see Balota & Duchek, 1986; Morton, Marcus, & Ottley, 1981).

In addition to differential selection of available stimulus features, perceptual information processing mechanisms may also influence or determine the final coded format of the physical features that are selected. Of particular interest is a recent idea suggested by Morton et al. (1981), Crowder (1983), and Greene and Crowder (1984) that trace components resulting from input closely tied to speech or language perception may possess some inherent similarity to one another because they reflect the output of a language-analyzing system. Such a system, at least in hearing subjects, is designed to interpret spoken language as presented; toward that end, visual as well as auditory aspects of the input may be used to determine what the speaker actually heard (see Crowder, 1983, p. 261).

Evidence consistent with these ideas comes from two sources: First, MacDonald and McGurk (1978) showed that visual information about lip movements can importantly influence what a subject actually perceives (i.e., hears); second, a number of authors (e.g., Gardiner, Gathercole, & Gregg, 1983; Greene & Crowder, 1984; Nairne & Crowder, 1982; Nairne & Walters, 1983) showed that in immediate memory experiments lipreading or mouthing input often results in memory performance that mirrors the performance found for auditory input. It is suggested, therefore, that lipread, mouthed, and auditory events may produce similar modality-dependent features, perhaps by virtue of the operation of a speech-analysis system.[1]

Consequently, it is important not to confuse the notion of modality-dependent features with the idea of sensory features, because there can be some important differences. Modality-dependent features, although determined by the particular mode of presentation, do not necessarily reflect faithful representations of the sensory qualities present in the nominal stimulus. As just outlined, for example, presentation modes whose sensory qualities appear qualitatively different (e.g., lipreading and sound) may produce trace components with highly similar features. It is possible, therefore, that subjects, under certain task demands, can produce task-dependent features internally that bear little resemblance to the actual presented stimulus. The term *modality dependent* is used to describe this class of features because, in the majority of instances, there is a high correspondence between the physical features of the stimulus and the modality-dependent features of the trace complex. However, in a more general sense, it is proposed that subjects represent the perceptual qualities of presentation; the resulting features, then, may or may not be similar to the sensory qualities present in the stimulus as nominally presented.

Overwriting Assumptions

It is assumed that the probability of recall of an event is reduced whenever its features recur in a later event. This process is referred to here as *overwriting*, where the active features of an immediate memory trace are rendered functionally lost by subsequently occurring material. The term *overwriting* is meant to stand for a hypothetical psychological process (e.g., erasure), but no particular mechanism is assumed (see Crowder, 1978a, for a discussion of possible mechanisms).[2] The amount of overwriting obtained in a given situation is determined by two variables:

1. *Similarity*. The more similar the encoding of event B to the trace of a previously encoded event A, the more event B will overwrite, and consequently reduce the recall of, event A. Similarity is defined by the overlap between the features of the two traces, as encoded by the system (Tversky, 1977). Consider the case in which event A is presented aloud by a male speaker. The immediate memory trace for event A should then consist primarily of the modality-independent speechlike features created by the inner voice, in addition to the modality-dependent features that are specific to the particular male speaker. If event B is then presented by the same male speaker, one can expect both the modality-dependent and the modality-independent features to be overwritten (as determined by the amount of feature overlap between A and B). If event B is presented visually, or in a female voice, there should be little overwriting of the modality-dependent features; on the other hand, because the character of modality-independent features is not influenced by the particular presentation conditions, these features should be overwritten to the same extent as when events A and B are presented in the same male voice.

2. *Grouping*. The second variable to influence the overwriting process is event grouping: Event B is capable of altering the immediate memory trace of event A if and only if it is perceived as belonging to the same stimulus set as event A. This means that how a subject segments list items will importantly determine when overwriting will occur, even if two events are highly similar. Events A and B may possess many features in common but not interfere with one another if they are perceived as belonging to two different stimulus sets. Put in this way, similarity is viewed as a necessary condition for overwriting to occur, but probably not a sufficient one. Segmentation of list items, through temporal and perhaps other means, may functionally insulate an event from subsequent interference through overwriting.

What determines when two list events will be assigned to two different stimulus sets? First, in most immediate memory experiments, it appears that how items are temporally separated is an important factor in determining what items are grouped together. For example, Ryan (1969) showed that immediate memory for nine-item lists could be improved significantly when extended pauses were inserted after the third and sixth digits. Frankish (1985) showed a similar result for auditory lists, except that his data indicated that the most substantial improvement in recall occurred for the last item in a temporally separated group. This is exactly the result that one would predict if overwriting occurred primarily within, but not across, groups. Thus, for a nine-item list presented in groups of three, the first item in a group should be interfered with by the second, and the second item should be altered by the third; however, because the fourth item occurs in a temporally distinct segment, the third list item should remain relatively free from interference and should be easily recalled. Other evidence touching on the role of temporal factors in grouping comes from experiments on the stimulus suffix effect, a paradigm examining how recall of the last item in a list is affected by the occurrence of a redundant event (usually a cue to begin recall). It has been known for some time that the damaging effect of a suffix is reduced if the suffix is delayed for a second or two following list presentation (see Crowder, 1976). However, Frankish and Turner (1984) showed that it is not the absolute time period that is critical, but the relative time period. For example, a suffix can be quite ineffective in reducing recall of a terminal list item, even if it occurs within one second after the list ends, provided that the list items themselves are presented at a high rate (say, 10 digits per second). The results of the study suggest that the interfering effect of one item on another depends on whether those items are perceived as belonging to the same stimulus grouping.

Second, although similarity by itself cannot be used to predict when overwriting will occur, it is likely that similarity does play a role in how subjects group items together in memory. Thus, with all other factors held constant, similar items will more likely be perceived as belonging to the same stimulus set than dissimilar items. Indeed, as Greene (1985) argued, grouping by similarity (in this case, semantic similarity) may be a critical factor underlying performance in the continuous-distractor paradigm developed by Bjork and Whitten (1974). Thus, in most situations, grouping may turn out to be the only necessary mechanism for predicting when overwriting will occur; that is, similar items will tend to interfere with one another by virtue of the fact that they are grouped together.

Utilization of Trace Features

Given the preceding assumptions about trace features and their susceptibility to interference, we can now speculate on how those features, for a given trace complex, are used by the subject during recall. As stated earlier, selection of an item to recall is viewed as a process in which the subject uses the existing trace features to discriminate among the set of possible recallable items. The encoded features of a trace, provided they have not been overwritten, will help in this selection process as long as (1) those features provide discriminative information about a particular item (i.e., those features uniquely specify an item from the recall set) and (2) the system is adapted to use those features as discriminative cues.

Discriminability. Consider a case in which a subject is attempting to select an item from the recall set based on the analysis of a trace complex containing a collection of modality-independent features (A) and a collection of modality-dependent features (X). In the best case, both the A and the X features will, in combination, help the subject make his/her selection. However, it is easy to conceive of a number of situations in which one or the other class of features will be more or less effective. Obviously, if the modality-dependent features (X) have been largely overwritten by subsequent input, then recall will be based primarily on inner-voice features (A). Alternatively, if the retrieval query asks for modality-specific information (e.g., to identify only the items spoken by a male voice from a list containing items presented by both male and female speakers), then the A features, because they are independent of presentation modality, will be of limited value. Of particular interest here is the role that similarity among the members of the recall set—and, as a consequence, among the memory traces of presented items—can play in serial recall performance. The collection of features that define the trace complex (A and X) will be effective only if those features uniquely specify a member of the recall set. To the extent that the A features or the X features overlap with the defining features of other members of the recall set, those features will be less important determinants of recall.

To illustrate, consider an experiment by Crowder (1978b) in which subjects were asked to recall lists of homophones (e.g., PEAR, PAIR, PARE) following auditory presentation. Homophones represent the extreme case in which the modality-dependent features of the trace are useless in specifying a member of the recall set; any encoded information about how an item sounds is identical with the modality-dependent information contained in any other trace formed from list presentation. Thus, saying a homophone aloud may produce a "richer" trace complex, in the sense of adding modality-dependent auditory features, but this does not necessarily improve serial recall. Item selection at recall is not a simple function of how well the features of a trace match the defining features of a recall set member; rather, the features of a trace will be effective only if they match the representative features of a recall set member to the exclusion of other recall set members. Put in different language, trace features are effective in helping recall only if they are distinctive (Gardiner, 1983; Glenberg & Swanson, 1986).

Salience. Even though the presence of distinctive features should, in principle, improve recall, they are unlikely to do so unless the subject is prepared to use those features as discriminative cues in recall. For example, suppose that a subject is presented visually with a list of random digits and is asked to vocalize (or not) particular letters of the alphabet in accordance with list presentation (thus a subject might say "A" to the first digit that appears, no matter what it is, "B" to the second item, and so on). One can assume that each trace complex would then be a composite of modality-dependent and modality-independent features, but not all of those features can reasonably be expected to help recall. Specifically, modality-dependent auditory information about how the letter "A" was vocalized is unlikely to help the subject recall the first list item, even though those features may uniquely specify a member of the recall set episodically (compared with other immediate memory traces).

APPLICATION OF THE FRAMEWORK

The Standard Modality Effect

Auditory recency. Serial recall of lists presented aloud typically produces near-perfect performance on the last one or two list items. Superior recency performance in the auditory case is particularly striking because it sits in sharp contrast to the general trend toward increasing errors that is found for the early and middle serial positions, and no similar enhancement in recall is found when the same list items are presented visually, in the absence of auditory stimulation. This performance pattern—namely, improvement for auditory lists that is restricted to recency items—defines the standard modality effect.

As stated earlier, recall of an item from immediate memory is determined by how well the modality-dependent and modality-independent features of the trace specify a particular member of the recall set to the exclu-