

# 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 immediate 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 multi-attribute 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 dissimi-

lar 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.<sup>1</sup>

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).<sup>2</sup> 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-

sion of others. Recency effects, therefore, must result from the fact that end-of-list memory traces possess more identifying features than prerecency traces, allowing for better item selection at the time of recall. In accordance with other models of the modality effect, it is assumed that recency items from auditory lists are unique primarily because they contain residual auditory information; more generally, these traces contain more modality-dependent features than do traces from prerecency items. The extra advantage for the last item follows because it is the only item that does not suffer from overwriting by subsequent input. When an interfering item occurs, depending on its functional similarity to the preceding item, both the modality-dependent and modality-independent features will be degraded. Because the trace contains a smaller percentage of modality-dependent features, recall of prerecency items, then, will be determined almost exclusively by the remaining modality-independent inner-voice components. As a result, serial recall of prerecency items should be less dependent on the particular presentation modality; for example, one should find similar performance patterns for auditory and visual lists on the early and middle serial positions. Of course, this is one of the defining characteristics of the standard modality effect.

Overall, this account of auditory recency does not differ substantially from previous accounts by Crowder and Morton (1969) and O. C. Watkins and M. J. Watkins (1980). Both accounts place the locus of the auditory advantage in residual auditory trace information (i.e., information that is reduced or absent in prerecency items). Similarly, the present model assumes that utilization of the residual, modality-dependent information (if present) will depend on its discriminability: if the existing features fail to provide predictive information about which member of the recall set is specified by the trace, then auditory recency will be absent or sharply reduced. For example, following the reasoning of Darwin and Baddeley (1974), because the acoustic cues for stop consonants, over time, may provide less discriminative information than that provided by vowels, less auditory recency should be obtained for immediate memory lists composed only of stop consonants (BAH, DAH, GAH) when compared with lists composed primarily of vowels (GAH, GOO, GEE; Crowder, 1971). Or, if the auditory lists are composed of items high in phonological similarity, then auditory recency should be reduced (M. J. Watkins, O. C. Watkins, & Crowder, 1974). Finally, for lists of homophones (PEAR, PARE, PAIR), auditory recency should be virtually absent (Crowder, 1978b). In these instances, modality-dependent components are available to aid recall, but recency is reduced because they do not provide unique information specifying a particular member of the recall set. The presence of residual modality-dependent features, then, provides the opportunity for recency, but does not assure recency. Other factors, like the similarity and, consequently, the discriminability among recall set members will also be critical determinants of final recall performance.

**Suffix effects.** Although the particulars of the stimulus suffix effect are beyond the scope of the present treatment, this framework assumes that the reduction in auditory recency that is found with an added auditory suffix can be explained with the overwriting assumptions described earlier. Briefly, the suffix, provided that it is similar to, and grouped with, the last list item, will reduce recency by eliminating the modality-dependent features that produce the recency advantage; a similar process, of course, occurs normally for all prerecency items. For auditory lists followed by visual suffixes, even if those suffixes are fully processed by the system (e.g., Morton & Holloway, 1970; Nairne & Crowder, 1982), no reduction in recency recall is expected because the visual suffixes will contain visual modality-dependent features that bear little similarity to the features controlling recency for the auditory list.<sup>3</sup> The specifics of the masking process in the auditory case have not been finally determined, but a detailed description of some possible mechanisms can be found in Crowder (1978a). In addition, the suffix is known to have other "across-the-board" effects (see Balota & Engle, 1981; Penney, 1985; Routh, 1976) that also will not be treated here. It is sufficient simply to note that suffixes may potentially interfere with both modality-independent and modality-dependent features, but only the latter are primarily responsible for the selective impairment of recency.

**Visual recency.** Serial recall of lists presented visually, in the absence of sound, typically produces little, if any, improvement in recall of the last item. Although virtually every theory of the modality effect can explain why auditory presentation enhances recency, relatively little theoretical attention has been given to explaining the poor performance that is found at the end of visual lists. Given the increasing variety of presentation modes that produce recency (e.g., mouthing, lipreading, tactile presentation), the data pattern for visual presentation may turn out to be the primary pattern requiring explanation (cf. Nairne & McNabb, 1985). Past interpretations of visual recency performance have tended to rely on the "special status" of auditory traces as a way of predicting poor performance for visual lists; that is, little recency is found for visual presentation because visual traces lack the special properties inherent to auditory traces (e.g., greater durability, O. C. Watkins & M. J. Watkins, 1980; greater temporal distinctiveness, Gardiner, 1983, and Glenberg & Swanson, 1986). Crowder and Morton (1969), for example, assumed that auditory traces were less subject to decay than visual traces because of the special structural properties of PAS; iconic memory, in contrast, decayed too rapidly to be of any benefit in recalling the last list item.

The present framework assigns no special status to auditory-based features of a trace. Auditory features are merely examples of modality-dependent features, and do not differ in kind from any other modality-dependent features, including, in particular, visual ones. Of critical interest, then, is the question of why visual presentation fails to generate much recency during immediate serial recall.

Because the last list item in the series is not followed by another item, there is every reason to expect that the visually based, modality-dependent features should remain intact to aid recall. One possibility centers on the fact that salient visual events typically do occur subsequent to the end of the list. The subject is performing in a visual world, and there is every reason to suppose that his/her visual fixations following the list have the potential to overwrite the modality-dependent features of the last list item. Such an argument gains added weight when one considers that very few salient auditory events follow list presentation (perhaps the squeak of a chair or the tap of a pencil); consequently, there is less potential for "extraexperimental" stimuli to interfere with auditory list items (for some similar ideas, see Glenberg, 1984).

There have been two attempts to test a hypothesis of this sort, and the results have been mixed. First, Hitch (1975) sought to reduce the potential masking effects of background illumination on the recall of a visual series by running subjects in the dark and by requiring oral recall. Although small visual recency effects were obtained under these conditions, they did not differ in magnitude from the effects found under conditions employing normal, well-lit background illumination. In a second study, using a procedure in which list items are separated by periods of distraction, Glenberg, Eberhardt, and Beldon (1987) found enhanced visual recency when the possibility of visual interference was reduced by using a limited-vision mask (goggles), auditory interitem distraction tasks, and oral recall. Although relatively large visual recency effects were found under these conditions, still larger effects were obtained for auditory presentation; thus, despite the apparent lack of visual interference, modality effects were still demonstrated.

A second accounting of recency performance (or the lack of) under visual presentation conditions appeals not to overwriting per se, but rather to the speech-based character of short-term memory. Because most psychologists believe that the short-term memory system (at least for hearing subjects) has evolved primarily as a vehicle for producing and interpreting spoken language, it is reasonable to argue that subjects may not be adapted to use visual components of a trace as critical cues for item selection at recall; rather, in the absence of auditory cues, subjects likely tend to rely on the recoded, modality-independent features (the inner voice) to reconstruct what stimulus items have been presented (for a similar argument, see Broadbent & Broadbent, 1981; Nairne & McNabb, 1985). This reasoning is somewhat reminiscent of the logic used by Shand and Klima (1981) in their proposal of a "primary linguistic code" underlying recency performance. According to their account, substantial recency effects result only when the presentation modality is consistent with the subjects' dominant coding format in short-term memory. Thus, for most subjects, presentation conditions closely tied to speech or language perception produce recency; for deaf subjects, the dominant coding format may be visually based, as in American Sign Language. In the present case, however,

I am simply suggesting that subjects may be more likely to use modality-dependent features when those features are compatible with primary linguistic input. Nevertheless, the demonstration of recency does not need to be limited to linguistically relevant input; under the right experimental conditions—namely, task demands that induce the subject to attend to modality-dependent features that are not linguistically relevant—significant recency effects may still be found.

Evidence consistent with the idea that task demands may be an important determinant of when modality-dependent features are used to benefit recency comes from a recent study by Kallman and Cameron (1987). These authors showed that movement of a visually presented stimulus could significantly enhance recency recall, but only if the movement was critical to identification of the to-be-recalled information. For example, in one experiment, subjects were shown successively presented rectangles that moved during presentation in one of four diagonal directions; the subjects' task was to recall the direction of movement for each of the presented rectangles (i.e., left, right, up, or down). Kallman and Cameron found significant visual recency effects under these conditions, but, more to the point, recency was reduced in a condition in which the movement direction (i.e., the word *left*, *right*, *up*, or *down*) was simultaneously placed inside the moving stimulus. Thus, to the extent that the task demands induced the subjects to attend to the movement itself, any remaining modality-dependent features were likely to be used beneficially in recall.

A similar interpretation can be applied to the work of Broadbent and Broadbent (1981), who showed significant recency effects using abstract visual stimuli. What is interesting about using an abstract matrix or line figure is that these stimuli contain visual features that are hard to name. One could argue, then, that these experimental task demands were optimal for inducing subjects to attend to the modality-dependent visual features because few, if any, modality-independent features were available from recoding (for a related finding, see Hines, 1975). A recent series of experiments by Campbell, Dodd, and Brasher (1983) provides further support for this hypothesis; again, significant visual recency effects were produced with unusual visual stimuli, in this case sequentially presented arrows or hand signs. Campbell (1986) also produced significant visual recency effects using pseudo-homophones (wunn, tooe, threa, etc.), which, similarly, may heighten the salience of visual features because of their unique spellings. Although the exact empirical nature of these visual recency effects is not well understood—for example, Campbell (1986) reported that visual recency can sometimes be disrupted by auditory suffixes—their basic demonstration is important because they allow one to dismiss the notion that recency effects are the exclusive by-product of "special" auditory, or language-based, features. It is quite possible to produce marked visual recency effects, provided that the experimental conditions encourage the subjects to make use of the modality-dependent features that are present in the

input. One could similarly argue that static, visual presentation of overlearned letter stimuli will not promote reliance on visual features; rather, under such conditions, subjects will tend to rely on the rich, modality-independent features of the trace that result from automatic verbal labeling.

To summarize, recall of visually presented lists, then, is normally based almost exclusively on the analysis of the modality-independent features formed via the inner voice. As each list item occurs, it interferes with the previous one, reducing recall. Because the proportion of modality-independent features is large, recall of preterminal items is determined primarily by an analysis of these inner-voice features, regardless of the modality of presentation. Visual recency effects are not typically found because (1) many salient visual events follow the last list item, and therefore overwriting of the modality-dependent features may occur, and (2) the system is probably not adapted to use any remaining visual components of the trace in recall because immediate memory is so typically involved in the comprehension and production of spoken language. This analysis predicts that recall of prerecency items (because of overwriting) will produce similar performance levels for auditory and visual presentation, whereas a recency advantage will accrue in the auditory case for the terminal list item.

**Inner-voice activities.** Many of the preceding arguments have been based on the idea that terminal list items are especially recallable because of residual modality-dependent features that have not been overwritten by subsequently occurring material. But a similar argument can be made for the modality-independent features of the terminal list items. Unless the last list item is followed by a stimulus suffix, it would seem necessary to assume that recency items contain more inner-voice components than do prerecency items, in addition to more modality-dependent features. Why then does this relatively rich supply of modality-independent features not produce enhanced recency for visually presented lists? This question is particularly applicable to visual lists because, as argued previously, recall of visual recency items tends to be based primarily on the inner-voice components of the trace. To deal with this problem, Nairne and McNabb (1985) introduced the idea that the "background activities" of immediate memory might potentially interfere with list recall in the same way that externally presented events produce interference (see Johnson & Raye, 1981). By the term *background activities*, we were referring to the constant cognitive activity that occurs during and immediately after list presentation, for example, subjects formulating strategies for retrieval, rehearsing list items, wondering about the end of the session, and so on. It was assumed that these activities were controlled by the inner voice and, as a result, were likely to interfere with previously encoded modality-independent features. Thus, *rehearsal* of a preterminal item, after the list has concluded, might degrade the modality-independent features of the terminal list item in much the same way that an exter-

nally presented suffix can degrade the modality-dependent features of the last list item.

Although this reasoning is speculative, the idea that the background activities of immediate memory interfere with the recall of experimentally presented list items can be extended to other domains. For example, as noted earlier, Frankish (1985) reported that temporal separation of nine-item lists into groups of three substantially improved later recall when compared with the recall of ungrouped lists. However, this improvement for grouped lists occurred most dramatically with auditory lists; grouping effects for silent, visually presented lists was marginal. If the recall of visually presented lists is based primarily on an analysis of modality-independent features, as the present framework assumes, then it follows that it ought to be difficult to get grouping effects with visual presentation. Temporal separation can effectively isolate the occurrence of modality-dependent features between groups (unless a suffix occurs in the interval) but not the presence of modality-independent features; that is, the background activities of immediate memory (e.g., rehearsal) will preclude temporally based protection from overwriting during the interval because these activities are highly similar in format to the list traces. The same mechanism that was proposed to reduce recency for visually presented lists, then, will operate during the interval separating the groups, eliminating any grouping advantage. With auditory lists, the modality-dependent features of the traces should not be interfered with during the interval by inner-voice activities and, thus, will become protected by virtue of the temporal isolation.

An argument of this sort might also be applied to a phenomenon known as the "long-term modality effect," which is derived from a paradigm originally developed by Whitten and Bjork (1972; Bjork & Whitten, 1974) to study long-term recency effects (also see Tzeng, 1973). In this procedure, pairs of to-be-remembered items are presented for study, but separated from one another by distractor-filled interpresentation intervals (e.g., solving arithmetic problems). The end of the list is then followed by another filled distractor interval prior to free recall. Not only are significant recency effects found in this procedure, but, more to the point, the obtained recency effects are larger for auditory than visual presentation (see Gardiner & Gregg, 1979; Glenberg, 1984).

Glenberg and Swanson (1986) attempted to account for these data by assuming that subjects use temporally based search sets in the long-term case to tap temporally coded memory representations. Recency differences between the modalities are assumed to reflect a more fine-tuned encoding of temporal information for auditory than visual presentation; that is, auditory presentation leads to a more temporally distinctive memory representation than does visual presentation. If it is true that subjects encode fine-grained information about time of presentation (which can be conceptualized as a kind of modality-dependent feature), then one might expect time of occurrence information to be associated with a wider temporal region for

visual presentation because of the abundance of inner-voice activity that occurs immediately prior to and after stimulus presentation. Whereas the occurrence of auditory traces, based on the analysis of modality-dependent (i.e., auditory) features, could be sharply contrasted with these inner-voice activities, a similar discrimination should prove difficult for visual presentation, leading to less precision in the representation of occurrence information.

### Lipread and Mouthed Modality Effects

Enhanced recency performance for a terminal list item should be found for any presentation modality that produces discriminable modality-dependent features provided that (1) those features are not overwritten prior to the initiation of recall and (2) subjects are adapted to use those features as discriminable cues for selecting an item from the recall set. This last requirement is assumed to be influenced by the subject's natural tendencies to sample certain types of trace features and by the particular task demands that are employed in an experiment. As a result, under the right set of circumstances, one can anticipate demonstrations of substantial recency effects for presentation modalities that do not employ sound. Of course, it is demonstrations of modality-like effects in the absence of auditory stimulation that have proven difficult for existing theories of the modality effect to explain (e.g., Crowder & Morton, 1969; O. C. Watkins & M. J. Watkins, 1980).

Of particular interest is a flurry of recent articles showing that auditory-like serial recall patterns, using standard lists composed of letters, digits, or words, can be produced when subjects silently lipread or mouth the items as they are presented. Spoehr and Corin (1978) showed that a suffix silently mouthed by the experimenter (and lipread by the subject) produced significant interference in recall of the recency portions of a list presented aloud; Nairne and Crowder (1982) showed a comparable result under conditions in which subjects repeated visually presented lists aloud and silently mouthed a visually presented suffix. More importantly, Campbell and Dodd (1980) produced a modality effect, better recency compared with a visual control, when subjects simply lipread the items presented silently by the experimenter (see also, Greene & Crowder, 1984); in addition, Nairne and Walters (1983) and Greene and Crowder (1984) showed a recency advantage when subjects themselves mouthed visually presented list items rather than merely read them. Because none of these manipulations directly involves sound, theories that appeal to the special status of auditory events (e.g., Crowder & Morton, 1969; Gardiner, 1983; O. C. Watkins & M. J. Watkins, 1980) cannot explain the data patterns that are produced.

To apply the present ideas to these results, it is necessary to account for why the modality-dependent features formed from lipreading or mouthing a list are apt to be used by the subject as discriminable cues in recall. In the case of lipreading, the modality-dependent features would appear to be visually based, although the subject may encode information about the temporal orderings of the lip

movements (see Campbell & Dodd, 1980); for mouthed input, it is conceivable that some form of articulatory feedback is encoded as part of the trace (see Crowder, 1983; Nairne & Walters, 1983). One possibility is that these presentation modalities force the subject to pay particular attention to visual or articulatory cues. Lipreading in the absence of accompanying sound is not a normal activity and, therefore, may produce protracted encoding of the visual features, thereby increasing their salience; a similar argument could be made for silent mouthing of visual stimuli. Under normal visual presentation conditions (e.g., a list of digits), it is likely that the input allows for immediate resolution of content and little attention to the surface features of the stimulus (e.g., the shape of the digit) is induced. Thus, the task requirements for lipreading and mouthing visual stimuli may increase the subject's tendencies to use modality-dependent features as discriminative cues in recall. In fact, any form of input that demands a difficult or protracted encoding may induce a greater reliance on modality-dependent features in recall and, as a result, produce greater relative recency performance (see Campbell et al., 1983).

Although an argument of this sort may be true in the general case, simple appeals to protracted encoding or to the task requirements induced by lipreading or mouthing visual stimuli may not be sufficient to explain the recency that is produced. Lipread, mouthed, and auditory input often appear to possess an inherent similarity that goes beyond reliance on modality-dependent features as critical components of recall. Most importantly, these forms of input interfere with one another in ways that cannot be predicted on the basis of an appeal to nominal similarity. Specifically, as mentioned earlier, mouthed suffixes interfere with auditory lists (Nairne & Crowder, 1982), as do lipread suffixes (Campbell & Dodd, 1980; Greene & Crowder, 1984; Spoehr & Corin, 1978). Conversely, auditory suffixes interfere with recency performance in lipread (Campbell & Dodd, 1980; Greene & Crowder, 1984) and mouthed (Nairne & Walters, 1983) lists. Such a range of findings is problematic for any theory of suffix interference that is based on nominal similarity because these modes of input are certainly quite dissimilar. What lipreading, mouthing, and audition do have in common, however, is that they all may be involved in the normal perception of language; that is, information about speech gestures may be processed by the same perceptual systems as auditory information. As discussed previously, there is considerable evidence to support the idea that visual information about lip movements can importantly influence the auditory perception that the subject experiences (e.g., MacDonald & McGurk, 1978). Thus, extending the ideas of Morton et al. (1981), Crowder (1983), and Greene and Crowder (1984), the final form of the modality-dependent features of lipread, mouthed, and auditory input may be similar functionally (and therefore interfere with one another) because they reflect the output of a language analyzing system that helps an individual decide what language-based information has actually been presented.<sup>4</sup>



It is important to stress, however, that the locus of recency effects does not reside in this language analyzer itself, as implied by Greene and Crowder's (1984) liberalized version of PAS; instead, recency effects are produced as a consequence of residual modality-dependent features that result as output from the analyzer. These features remain available to enhance recency, particularly for the last item, for the reasons cited earlier: (1) the last list item is not followed by interfering postlist material, and (2) our information processing system is probably well adapted to use modality-dependent features that accrue from language-relevant input. Appealing to the encoded, modality-dependent features of immediate memory rather than to the structural properties of the language analyzer is reasonable because it does not restrict explanations of recency effects to language-based input. That is, the present account is capable of explaining the range of visually based recency effects that have been demonstrated by Campbell and her colleagues (Campbell, 1986; Campbell et al., 1983).

### SUMMARY AND CONCLUSIONS

To summarize the main points of this framework, auditory and visual presentation modalities are assumed to lead to qualitatively different memory traces, which, in turn, are differentially susceptible to interference from subsequent input. The attributes making up these traces do not differ from one another in any fundamental mnemonic sense; that is, the modality-dependent features that result from auditory presentation are not considered to be stronger, more distinctive, or less subject to decay than visual or any other type of feature. In fact, contrary to the PAS model, the present framework does not even use decay as an operative source of forgetting. As a result, in the absence of interfering material, recency effects should be found after indefinite intervals (see O. C. Watkins & M. J. Watkins, 1980). To explain the advantage that auditory presentation can sometimes show in immediate serial recall, one needs to consider the possibility that the locus of the modality differences in recency performance lies not in auditory superiority per se, but rather in the failings of visual presentation. Visual features may be susceptible to further visual interference and, because of our normal tendency to rely on speechlike processing in short-term memory environments, less likely to be sampled as the critical discriminative cues in recall. However, most important to the ideas of the proposed framework, neither of these failings are absolute: To the extent that visual interference following list presentation can be reduced, or the salience of visual features enhanced, visual recency effects should emerge.

Finally, as mentioned in the introduction, the ideas expressed in this framework for interpreting recency do not, as a whole, constitute a theory of immediate memory in any formal sense. Many of the assumptions possess degrees of freedom that are, at this point, disquieting; yet, a similar criticism can be applied to each of the post-PAS

explanations of the modality effect that have been offered by researchers. For example, notions about changing-state stimuli (Campbell & Dodd, 1980) or primary linguistic codes (Shand & Klima, 1981) have never been stated with ringing precision. What separates the latter from mere musings, however, is the fact that each has led investigators into productive experimental realms. These proposals have proven testable at qualitative levels of analysis, and our knowledge about recency effects has been advanced as a result.

### REFERENCES

- AYRES, T. J., JONIDES, J., REITMAN, J. S., EGAN, J. C., & HOWARD, D. A. (1979). Differing suffix effects for the same physical suffix. *Journal of Experimental Psychology: Human Learning & Memory*, 5, 315-321.
- BALOTA, D. A., & DUCHEK, J. M. (1986). Voice-specific information and the 20-second delayed-suffix effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 12, 509-516.
- BALOTA, D. A., & ENGLE, R. W. (1981). Structural and strategic factors in the stimulus suffix effect. *Journal of Verbal Learning & Verbal Behavior*, 20, 346-357.
- BJORK, R. A., & WHITTEN, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, 6, 173-189.
- BOWER, G. H. (1967). A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 230-325). New York: Academic Press.
- BROADBENT, D. E., & BROADBENT, M. H. P. (1981). Recency effects in visual memory. *Quarterly Journal of Experimental Psychology*, 33A, 1-15.
- BROADBENT, D. E., VINES, R., & BROADBENT, M. H. P. (1978). Recency effects in memory as a function of modality of intervening events. *Psychological Research*, 40, 5-13.
- CAMPBELL, R. (1986). Remembering with impurity: When precategorical acoustic storage is not acoustic, what is it? In D. A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production*. London: Academic Press.
- CAMPBELL, R., & DODD, B. (1980). Hearing by eye. *Quarterly Journal of Experimental Psychology*, 32, 85-99.
- CAMPBELL, R., DODD, B., & BRASHER, J. (1983). The sources of visual recency: Movement and language in serial recall. *Quarterly Journal of Experimental Psychology*, 35A, 571-587.
- CONRAD, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, 55, 75-84.
- CONRAD, R., & HULL, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychonomic Science*, 10, 135-136.
- CONWAY, M. A., & GATHERCOLE, S. E. (1987). Modality and long-term memory. *Journal of Memory & Language*, 26, 341-361.
- CORBALLIS, M. C. (1966). Rehearsal and decay in immediate recall of visually and aurally presented items. *Canadian Journal of Psychology*, 20, 43-51.
- CRAIK, F. I. M. (1969). Modality effects in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, 8, 658-664.
- CRAIK, F. I. M., & KIRSNER, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology*, 26, 274-278.
- CROWDER, R. G. (1971). The sounds of vowels and consonants in immediate memory. *Journal of Verbal Learning & Verbal Behavior*, 10, 587-596.
- CROWDER, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- CROWDER, R. G. (1978a). Mechanisms of backward masking in the stimulus suffix effect. *Psychological Review*, 85, 502-524.
- CROWDER, R. G. (1978b). Memory for phonologically uniform lists. *Journal of Verbal Learning & Verbal Behavior*, 17, 73-89.
- CROWDER, R. G. (1983). The purity of auditory memory. *Philosophical Transactions of the Royal Society*, 302B, 251-265.

