Stress Grouping Improves Performance on an Immediate Serial List Recall Task

Carolyn Reeves, A. René Schmauder, and Robin K. Morris
University of South Carolina

Five experiments investigated whether perceptual patterning afforded by imposing a recurrent stress pattern on auditorially presented lists has a positive effect on list recall. The experiments also addressed whether the recall advantage reflected the salience that the stress pattern created for certain items or whether the recall advantage arose from the distinct grouping configurations that were produced by the stress pattern. The authors explored these issues by examining immediate serial-recall performance for spoken lists that either did or did not have a stress pattern imposed on them. Lists had an anapest or dactylic stress pattern or were monotone and consisted of two stimulus types, either digit names or common English nouns. Results showed that stress patterns enhanced serial-recall performance and that the recall benefit derived primarily from the perceptual grouping afforded by the stress patterns. Results also showed that the grouping benefit derived from stress patterning generalizes to monotone lists. In contrast, salience effects are attached to the stimulus per se and do not transfer.

Because spoken language extends over time, the auditory information conveyed by the stress, intonation, and pauses in spoken language must be retained briefly in memory to ensure that the information is processed and integrated before the stimulus input decays (see Cowan, 1995). The performance characteristics of memory for auditory information have been repeatedly investigated using the serial recall of auditorially presented lists (e.g., see Burgess & Hitch, 1999; Estes, 1972; Henson, 1998; Lee & Estes, 1981; Murdock, 1983; Nairne, 1990; Page & Norris, 1998a, 1998b). On the basis of the results of an immediate serial-recall task, Frankish (1989) argued that pauses in lists function like pauses in the prosodic structure of spoken utterances. Furthermore, Frankish speculated that memory for auditory stimuli has a central role in language processing because memory uses prosodic features to facilitate retrieval of speech stimuli. If Frankish is correct, then it is important to extend these findings beyond pause information to explore how other features that contribute to prosody in spoken language affect recall of auditorially presented information. Here, we report the results of five experiments that investigated how stress in auditory word lists affects the serial recall of those lists. Of interest was how stress patterning influences the listeners’ memory for the spoken list items and what role perceptual grouping and item salience play in that process.

Stress facilitates language processing at the word level, at the phrase level, and at the sentence level.1 At the word level, stress aids word recognition during early processing stages as auditory speech stimuli are segmented and identified (L. M. Slowiaczek, 1990, 1991). Also, stress provides patterns that assist the process of lexical access (Grossjean & Gee, 1987). Stress contributes to the salience of individual words, facilitating word identification and integration into the semantic, syntactic, and discourse structures (Ladefoged, 1982), as well as enabling recall of words from memory (Gow & Gordon, 1993). Stress also aids later stages of stimulus processing, as words are processed as phrases and sentences. Stress influences syntactic processing by improving predictability of the syntactic structure of potentially ambiguous constructions (Beach, 1991). Also, stress differentiates structures, such as appositions and parentheticals, in spoken language (Price, Ostendorf, Shattuck-Hufnagel, & Fung, 1991). By making certain words more prominent perceptually, stress indicates the importance of these words and thus emphasizes the importance of the message as a whole (Bolinger, 1989). Thus, there is evidence that stress may affect not only on-line processing but also memory for prose.

To address the role of stress patterning in memory, we investigated whether imposing prosodic patterns on verbal strings affects recall performance in an immediate serial-list-recall task. Because

1. In English, stress and intonation interact because pitch changes in intonation contours often cooccur with rhythmically stressed syllables (Johns-Lewis, 1986; Ladefoged, 1982). We specifically manipulated stress patterns in the present experiments. However, we acknowledge that the relation between intonation and stress made it possible that participants perceived an intonation pattern in the stress-patterned groups. Because intonation has effects on processing similar to stress (e.g., Bolinger, 1989; Brown, 1983; Coulthurd, 1992; Ladefoged, 1982; Lieberman, 1985; Nagel, Shapiro, & Navv, 1994; Speer, Crowder, & Thomas, 1993; Swertz, Collier, & Trefker, 1994; Swertz & Geluykens, 1994; Wells, 1986), we do not anticipate any negative influence on processing of perceived intonation contours.
list items are word stimuli, we view the list recall task as an auditory language processing task, albeit a simplified one. Furthermore, we hypothesized that if stress did improve recall for auditory stimuli, stress may function in two ways: by increasing the salience of particular items and/or by dividing and organizing the speech stream into perceptual groups in memory. These two mechanisms were explored in the present experiments.

Research on memory for auditory language stimuli is indexed by three effects that are relevant to the present experiments: the recency effect, the grouping effect, and the salience (or von Restorff) effect. The recency effect is defined as better recall for items occupying the last few positions in a list, particularly for the list-final item, than for items in list-medial positions (e.g., Corbalis, 1966; see also Balota & Engle, 1981). Recency effects were used in the present experiments to assist us in identifying effects of stress patterning on processing and memory of spoken language stimuli.

The grouping effect observed in list recall data is based on the Gestalt notion of perceptual grouping and was made prominent in the cognitive literature by Kahneman (1973) and Broadbent (1975). In the auditory list recall literature, the grouping effect is the enhanced recall of items that form perceptually salient groups. Kahneman (1973) and Kahneman and Henik (1981) proposed that a verbal list is perceived as a whole, that is, as a group, and they noted the particular recall advantage characteristic of items that occur at the end of groups. Studies of the grouping effect have consistently revealed enhanced recall in auditory word lists in which pauses grouped the lists (Broadbent, 1975; Broadbent, Cooper, Frankish, & Broadbent, 1980; Frankish, 1985, 1989; Hitch, Burgess, Towe, & Culkin, 1996; Ryan, 1969a, 1969b). Frankish (1985, 1989) argued that pauses assisted list recall by segmenting the list into perceptual groups, and he showed that the items that were group-final, that is, those immediately preceding a structural boundary created by the pause, received the greatest recall benefit. In Frankish (1995), in which grouping was achieved by the use of naturally occurring intonational contours and laboratory-generated pitch patterns, recall was strongly enhanced when the utterance consisted of naturally occurring intonation contours. However, for the laboratory-generated pitch patterns, better recall was observed only when the group-final item of the pitch pattern was accentuated. Therefore, Frankish (1995) concluded that the presence of a pitch pattern alone is not a sufficient condition for improved recall. The present experiments investigated whether the imposition of another prosodic feature, stress, could also benefit the recall of auditory lists by configuring the lists into groups.

The notion that stress can be used to group lists perceptually and define structural boundaries between groups is supported by the auditory nonspeech literature. Deutsch (1995) and Bregman (1978, 1981, 1990, 1993) have addressed nonspeech grouping mechanisms in relation to the Gestalt principles of proximity, similarity, good continuation, and closure. Of particular interest for this study is the contribution of pitch, a major component of stress, to perceptual organization. Pitch is integral to two types of auditory grouping processes: perceptual capture, which involves the grouping of simultaneous sound elements, and auditory stream segregation, which involves grouping of sequences of events (Bregman, 1990, 1993). Studies of auditory streaming and perceptual capture have shown that pitch proximity and pitch similarity are effective in the perceptual grouping of rapid series of tones (Bregman, 1990; Bregman & Campbell, 1971; Dowling, 1973).

We call the last effect observed in serial list recall that is germane to the present research the salience effect. Discussed in detail by Wallace (1965, who called it the von Restorff effect; see also Crowder, 1976; Fabiani & Donchin, 1995; Huang & Willie, 1979; Murdock, 1960; Neath, 1998), the salience effect occurs when an item that is conspicuous on some perceptual dimension is recalled better than other items in the same ensemble in learning and memory tasks. Although much of the work investigating the von Restorff effect has used visual stimuli, there is evidence that the effect can also be found with auditory word stimuli (Holmes & Arborgast, 1979).

Stress in English creates a succession of regular beats, some metrically strong and others metrically weak, with the strong beats arriving at uniform intervals. This pattern of regularly occurring stressed syllables or words embedded in a continuum of more frequently occurring unstressed syllables may lead to a salience (von Restorff) effect for the more isolated stressed items. These two levels of strong and weak pulses also occur in music perception, where they define the metrical structure of a musical sequence (Bigand, 1993; Povel, 1981). This metrical structure creates perceptual expectations that lead listeners to focus attention on metrically strong beats in the musical sequence (Jones, 1976; Jones & Yee, 1993). Theoretically, the auditory salience that stress affords to certain words in spoken strings by making them more distinctive could focus attention on these words, perhaps facilitating on-line processing and/or contributing to their better recall from memory.

The Present Studies

In the present experiments, we created lists for a serial-recall task either with a monotone pattern or with one of two stress patterns— anapest or dactyl—imposed. The anapest pattern consisted of two unstressed items followed by a stressed item, and the dactyl pattern consisted of a stressed item followed by two unstressed items. Serial-recall performance for the stress-patterned lists was compared with serial-recall performance for monotone lists to evaluate how stress patterns influenced recall of auditory lists. Across experiments, list items either were digit names or common English nouns. Manipulation of the stress pattern (anapest or dactyl) and stimulus type (digit names or common, frequent nouns) let us dissociate effects due to the salience of individual items from effects that derived from the configural grouping patterns in the list. In addition, in all experiments reported here, suffixes and tone (unsuffixed) controls were appended to lists to examine if and how the presence of a not-to-be-remembered item affected grouping and recall. Because the suffix did not alter the basic pattern of results, we report data from the suffix manipulation in the Appendix and do not discuss the suffix effects elsewhere in this report.

Our research extends previous work on prosodic grouping by investigating whether the prosodic feature stress aids recall for items in a verbal string. The use of stress patterns allows us to evaluate simultaneously the separate, but not necessarily mutually exclusive, contributions of four factors purported to affect the immediate serial recall of spoken lists. Three of these factors are potential causes of any recall benefit accruing from the imposition
of prosodic patterns on a list: perceptual grouping (Frankish, 1985, 1989, 1995; Ryan, 1969a, 1969b), the perceptual salience of individual list items, and rehearsal (Ryan, 1969b). The fourth factor purported to affect list recall involves the presence or absence of instructions to the participants to group the lists.

Perceptual grouping in auditory lists is accomplished by creating perceptual discontinuities in the acoustic string, either by inserting temporal gaps (Frankish, 1985, 1989; Ryan, 1969a, 1969b) or nonspeech sounds (Ryan, 1969a) between list items or by changing the acoustic–auditory dimensions of a relevant parameter for certain list items (Frankish, 1995). Any discontinuities that result from these insertions or changes can create the perception of structural boundaries, which are necessary to the perception of grouping. Stress is, by definition, relational and involves a change in one or more acoustic parameters in certain items in the stimulus list. In the lists used here, prosodic stress, like pauses (Frankish, 1985, 1989; Ryan, 1969a, 1969b) and intonation contours (Frankish, 1995), may provide structural boundaries by marking acoustic divisions between items, thereby creating distinct groups of perceptual information that give rise to a grouping effect.

Additionally, our research both extends and differs from previous studies of grouping because prior studies have not investigated whether the perceptual salience per se of an item can benefit its recall. In the pausal grouping studies (e.g., Frankish, 1989; Ryan, 1969a, 1969b), stimuli were created in monotone (i.e., with the same pitch, the same loudness, and the same relative duration), with pauses inserted between groups of items. In Frankish’s (1995) intonation study, intonation contours were created using pitch changes in Experiments 1–4, but no item or items were distinctive. In Experiment 5 of the same study, the final item in each group was spoken with accented pitch, presumably making the group-final item more distinctive or salient. However, salience per se was not manipulated in the Frankish (1995) study.

In the present experiments, anapest and dactyl stress patterns were imposed on lists, and stressed items (one per group, three per list) in a given ensemble were produced at a higher pitch, with greater intensity, and with increased duration relative to the unstressed stimulus items. It should be noted that in languages other than English, the acoustic cues signaling stress may differ (e.g., see Williams, 1986, for a discussion of stress in Welsh). Although the acoustic parameters of the stressed items in Experiments 1–5 of the present study may lead to enhanced recall for those items relative to the unstressed items (e.g., see Hall & Blumstein, 1977, on the differential effects of duration on recall), we take the position that it is the distinctiveness of the stressed items, which comes from the configurative nature of the stress patterns, that may result in a recall advantage. In our lists, stressed items occur less frequently than unstressed items (one stressed item vs. two unstressed items per group; three stressed items vs. six unstressed items per list), and medial stresses, which are free from primacy and recency effects (and hence free also from potential ceiling effects), are surrounded by unstressed items, producing an isolation effect for the stressed items. We theorized that this isolation may make the stressed items more distinctive or salient, giving them an advantage over the unstressed items at retrieval.

A third factor that contributes causally to the recall of lists, grouped or ungrouped, is rehearsal. For our studies, the issue was whether rehearsal is primarily responsible for the recall benefit in prosodically grouped lists. Ryan (1969a, 1969b) suggested that pauses may enhance recall in pause-grouped lists by allowing time for rehearsal. Frankish (1989) investigated this claim by manipulating pause durations in auditory lists and found that pauses enhanced list recall even when their durations were too short to permit rehearsal. Frankish concluded that pauses assisted list recall largely by segmenting the list into perceptual groups. In the present studies, there were no pauses inserted into the lists, and, therefore, rehearsal opportunities were the same (i.e., virtually nonexistent) for monotone lists as they were for stress-patterned lists.

The fourth difference between our work and previous grouping studies was in the instructions to participants. Frankish (1989) and Ryan (1969a) instructed participants to recall all lists by using a strategy of mentally grouping the stimuli by threes. Thus, the notion of grouping was introduced explicitly to participants in the Frankish and the Ryan studies prior to list presentation. Moreover, in the Frankish study, grouping boundaries were indicated on the response sheets. These overt references to grouping may have influenced participants’ perception of the list stimuli in the Ryan and the Frankish studies as well as the retrieval of the stimuli in the Frankish study (cf. Neath, Surprenant, & Crowder, 1993). In contrast, our instructions did not mention, and our response sheets did not indicate, any list segmentation or grouping. Therefore, any grouping effects we observed should initially reflect natural processing of stressed (grouped) versus monotone (ungrouped) lists, providing a stronger test of the grouping hypothesis.

The following hypotheses about the influence of stress patterns on list recall were investigated in Experiments 1–5. The stress hypothesis says that in a serial-list-recall task with spoken lists, recall accuracy will be higher for stress-patterned lists than for monotone lists. This could occur either because stress defines perceptual groups in spoken word lists (cf. Frankish, 1985, 1989; Kahneman & Henik, 1981) or because stress makes certain items more salient perceptually (Bolinger, 1989). To investigate these two possible effects of stress, we used two additional hypotheses, the salience and terminal position hypotheses, which make specific predictions about why and where stress would influence recall.

The salience hypothesis proposes that the perceptual prominence of the stressed items makes these items easier to recall (see Holmes & Arboagast, 1979). According to the salience hypothesis, stressed items should be remembered better than unstressed items, regardless of list position. In contrast, the terminal position hypothesis states that stress patterns organize list stimuli into groups because discontinuities between stressed and unstressed items produce structural boundaries. Structural boundaries give rise to the perception of groups, which in turn improves recall, especially at group-final serial positions (see Frankish, 1985, 1989, 1995; Kahneman, 1973; Kahneman & Henik, 1981). The terminal position hypothesis states that enhanced recall will be observed in stress-patterned lists at three terminal loci: end of group position, end of list position, and the final group in a list. When stress-patterned lists are compared with monotone lists, the terminal position hypothesis predicts better recall for the stress-patterned lists at end of group positions, which are absent in ungrouped lists, and at positions that comprise the final group of the list, also absent in ungrouped lists. However, at end of list positions that are perceptually terminal, the terminal position hypothesis predicts no recall benefit for the stress-patterned lists over the monotone lists because the list-final position in both conditions is obviously list final because it is followed by a recall cue (in the present study, a pure
tone) that is perceptually different from the list stimuli or by the silence of the response interval itself.

The use of anapest and dactyl stress patterns allowed us to test the stress hypothesis and ensured that if stress patterning improved list recall as predicted by the stress hypothesis, those recall benefits would be characteristic of stress patterning in general and not just of one type of stress pattern. Using a dactyl stress pattern allowed us to measure the effects of salience as distinct from the effects due to group-final position. The salience hypothesis predicted superior recall for the stressed and thus salient item in the group, which in the dactyl pattern was the group-initial item. The terminal position hypothesis predicted superior recall for the item in group-final position in the dactyl pattern. Because the group-final position was an unattested position in dactyl lists, the effects due to salience and those due to terminal position could be contrasted.

Both word stimuli and digit stimuli were used to investigate whether enhanced recall attributable to stress patterning generalized beyond a single stimulus type. Experiments 1 and 2 used monosyllabic digit names, and each list in Experiments 3 and 4 used frequent nouns drawn from a pool of nine nouns without replacement. The digit names used in Experiments 1 and 2 were members of a well-known group and were from the same semantic category (numbers), whereas the frequent nouns of Experiments 3 and 4 were not well-known members of a single group and did not belong to a single semantic category. The use of unrelated but frequent contentful nouns took Experiments 3 and 4 closer to actual auditory language processing than did the digit stimuli of Experiments 1 and 2.

Moreover, we were concerned that because everyday life requires people to remember long strings of numbers (telephone numbers, social security numbers, bank account numbers, etc.), some participants may have developed strategies for improving recall of digit strings (see Ericsson & Chase, 1982), especially in the presence of groups (e.g., recoding the sequence 6 1 5 4 2 0 as six-fifteen four-twenty). The use of the nouns, which could not be recoded like the digits, ensured that any patterns we observed reflected stimulus grouping and not a strategy of stimulus recoding.

In summary, the present experiments investigated how stress patterning of auditory information influenced recall of auditory word lists from memory. Experiment 1 compared recall of stress-patterned lists with that of monotone lists to examine whether the imposition of a stress pattern on auditorially presented lists influenced serial recall (the stress hypothesis) and why and where that influence occurred (the terminal position and the salience hypotheses). Experiment 2 tested the differing predictions of the terminal position and salience hypotheses by using a dactyl stress pattern to create a situation in which stressed positions and group-final positions did not coincide. Experiments 3 and 4 replicated Experiments 1 and 2 and evaluated the extent to which any observed differences in recall in Experiments 1 and 2 generalized from the digit stimuli to lists of frequent nouns. Experiments 1–4 also evaluated the role of rehearsal in grouping effects (Ryan, 1969a, 1969b) and investigated whether listeners use stress to segment list stimuli into groups in the absence of instructions to do so (cf. Frankish, 1989; Ryan, 1969a). Experiment 5 tested whether stress grouping, a naturally occurring phenomenon in many languages, would transfer to monotone lists and what the ramifications of such transfer, should it occur, would be.

Experiment 1

Experiment 1 tested the stress hypothesis by investigating whether serial-recall performance would be enhanced by the presence of an anapest stress pattern (− −) on lists consisting of nine digit names. The stress hypothesis predicted better recall for the stress-patterned lists compared with their monotone controls, but it did not specify at which serial positions the superior recall would occur.

Experiment 1 also tested the salience and terminal position hypotheses, which specified why and where stress would influence recall. The serial positions critical for evaluating predictions of the salience hypothesis in anapest lists were stressed positions. The serial positions critical for evaluating predictions of the terminal position hypothesis were group-final positions. In both instances, these were Serial Positions 3 and 6 in the stress-grouped lists. Also critical for evaluating the terminal position hypothesis were Positions 7 and 8, which were unstressed members of the final group in the stress-grouped lists. Serial Position 9, which was both stressed and list final, was also critical.

The salience and terminal position hypotheses both predicted better recall in anapest stress-patterned lists at Serial Positions 3 and 6 compared with monotone list controls. The salience hypothesis predicted better recall for Positions 3 and 6 because the items in these positions were stressed, whereas the terminal position hypothesis made the same prediction because these positions were group final.

The terminal position hypothesis also predicted better recall for items in Serial Positions 7 and 8 in anapest lists than in monotone control lists. Those positions were members of the final group in the anapest lists but not in the ungrouped monotone lists.

In the anapest lists, the salience and terminal position hypotheses made different predictions for recall of the list item at Serial Position 9 compared with monotone lists. The salience hypothesis predicted better recall for anapest lists compared with the monotone lists at Serial Position 9 because that item was stressed in the anapest lists. According to the terminal position hypothesis, when it became clear (by the tone that followed) to participants that the item at Serial Position 9 was the last item in the list, they would treat that item similarly in both anapest and monotone lists (by virtue of its list-final position). Thus, there was no predicted difference between these two conditions on the terminal position hypothesis.

Method

Participants. Forty-nine members of the University of South Carolina Psychology Department's human participant pool participated for course credit. Data from an additional 4 participants were discarded for failure to follow instructions. All participants were native speakers of American English, with normal hearing, established by self-report.

Materials. The digit names one through ten (with seven omitted because it is bisyllabic) were said aloud by a female speaker in both a stressed and an unstressed form, along with the two speech suffixes go (unstressed) and GO (stressed), and the word ready. The pitch of the stressed stimuli and suffix approximated the musical note B below middle C, and the pitch of the unstressed stimuli and suffix and the word ready approximated the musical note F# below middle C. The fundamental frequency ranged from 232 Hz to 253 Hz for the stressed items and from 175 Hz to 190 Hz for the unstressed items. The durations ranged from 307.71 ms to 433.38 ms for the stressed items and from 239.68 ms to 341.52 ms for the unstressed
items. The intensity of the stimuli and suffix ranged from 79.2 dB to 81.0 dB for the stressed items and from 75.0 dB to 79.4 dB for the unstressed items.

The stimuli were recorded in a soundproof booth by a microphone connected to a Marantz PMD-201 cassette recorder (Superscope Technologies; Aurora, IL). The stimuli were then digitized using a Soundslicer-Pro board (Creative Labs; Milpitas, CA) and an IBM PS2-77-48DSX2 computer. A pure tone with a frequency of 1000 Hz was generated with a pure tone generator and then digitized; this tone served as the control for the speech stimuli. The pure tone and the suffixes served as recall cues at test. To control for unwrapped coarticulation effects and to control the length of the interval between words, individual files were made for each digit name in its stressed and monotone (unstressed) form. Also, individual files were created for the two speech suffixes go and GO and for the tone control (see the Appendix for discussion of the suffix conditions; they are not discussed further in the body of this article). The intervals between digits, digits and suffix, and digits and tone were then set using the procedure for P-centering outlined by Morton, Marcus, and Frankish (1976). The perceived reading rate was approximately 2 digits per second, or 520 ms per digit. The lists were then transferred from the computer to audiotape. Five 9-item practice lists were created using the same procedure as the experimental lists, with each of the five experimental lists conditions represented in the practice lists.

For the experimental lists, 100 lists of nine digit names each were created. Digit names were selected without replacement in each list from the pool of nine digit names, but digit names were repeated across lists. List pattern was either monotone or anapest. Monotone lists were produced by randomly generating lists of nine unstressed stimuli. For randomly generated sequences of nine digits, the anapest pattern was created by selecting two unstressed stimulus items followed by a stressed stimulus item (x - x - x). This pattern organized the list into groups of three items, with each group-final item being stressed.

The resulting 100 stimulus lists were divided into 20 sets of 5 lists each, and the three suffix conditions, unstressed go, stressed GO, and tone control, were then assigned pseudorandomly within each set of 5 lists. The three suffix types and the two list patterns (anapest and monotone) formed the basis of five within-subjects conditions.

Procedure. The participants' task was to recall, in writing, as completely and as accurately as possible the list that they heard on each trial. Digit names were recalled as digits (2, 5, 9, etc.). Participants were instructed to wait for a recall cue, which would be either a tone or the word go given on the tape, before beginning to write their responses on the response sheets provided. Participants were told to recall serially from left to right without backtracking, and they were instructed to leave blank or, preferably, to guess at any stimulus item(s) they could not remember. An experimenter was present to monitor compliance. Participants were tested in groups of 15 or fewer. Five practice lists preceded the experimental trials at each testing session. The start of each trial was signaled by the word ready, followed after a 700-ms delay by the digit list and either a speech suffix or the tone control. Each digit list was followed by a 15-s recall interval and then the next trial began immediately. The stimulus tapes were played for the participants on a Sony TC-520CSC cassette recorder (Tokyo, Japan) connected to Sony speakers. The entire procedure, including instructions, practice lists, the test itself, and debriefing, took approximately 1 hr.

Results and Discussion

Experiment 1 tested predictions of the stress hypothesis and evaluated the validity of the terminal position and salience hypotheses. The data in Figure 1 show superior recall performance in stress-patterned lists compared with the monotone lists, especially at those positions predicted by terminal position hypothesis (Serial Positions 3, 6, 7, and 8). We only report and discuss the tone conditions here (see the Appendix, Table A1 for means from the suffix conditions).

The stress hypothesis predicted that imposition of a stress pattern would improve recall for spoken lists, and a 2 (list type; monotone vs. anapest) × 9 (Serial Positions 1–9) analysis of variance (ANOVA) confirmed that prediction. The analysis revealed a main effect of list type, with better mean recall for the anapest condition at 13.7 items than for the monotone condition at 12.4 items, F(1, 48) = 59.17, p < .0002, MSE = 12.19. There also was a reliable main effect of serial position (see Figure 1), F(8, 384) = 138.12, p < .0002, MSE = 13.00, and, most importantly, an interaction of list type and serial position, F(8, 384) = 7.59, p < .0002, MSE = 3.75. Experiment 1 also tested the salience and terminal position hypotheses. Bonferroni familywise error rates were used to evaluate these contrasts. For the tone conditions in Experiment 1, the critical p value was < .012.

The salience and terminal position hypotheses correctly predicted better recall for the anapest condition versus monotone control condition for items in Serial Positions 3 and 6, t(48) = 2.97, p < .01, and t(48) = 4.87, p < .001, respectively. Either because items in those positions were stressed or because the items were group-final, recall was better when a stress pattern was present in the lists.

The terminal position hypothesis also correctly predicted better recall for items in the final group of the list, namely for items at Serial Positions 7 and 8 in the anapest condition versus the monotone condition. The salience hypothesis made no such predictions. From t tests we found reliable differences at Serial Positions 7 and 8 for the anapest condition versus its monotone control, t(48) = 4.07, p < .001, and t(48) = 4.66, p < .001, respectively.

The salience hypothesis and the terminal position hypothesis made different predictions for recall of items at Serial Position 9. The salience hypothesis' prediction of better recall for Serial Position 9 in the anapest condition versus the monotone condition was not upheld. Figure 1 shows that recall for both the anapest and monotone conditions was virtually identical. This was predicted by the terminal position hypothesis because in both lists, the tone recall cue made the final list item obviously list final. However, the ability to detect effects of salience at Position 9 may have been compromised by the generally high level of recall at that position so we would not want to draw strong conclusions on the basis of this data point alone.

The stress and terminal position hypotheses were strongly supported by the results from Experiment 1: Stress patterning improved recall at Serial Positions 3, 6, 7, and 8. It is important to note that these grouping effects were observed even though participants were not overtly instructed to use grouping strategies (cf. Frankish, 1989), and there were no temporal gaps to provide time for rehearsal. At the serial positions where the salience hypothesis and the terminal position hypothesis made predictions that differed, Serial Positions 7, 8, and 9 for the anapest and monotone lists, the terminal position hypothesis received stronger support than did the salience hypothesis.

Experiment 2

Experiment 2 was designed to compare unique predictions of the salience and terminal position hypotheses by using a dactyl stress pattern. With the dactyl pattern, the stress fell on the group-
initial position (/ - -). Stressed positions and group-final positions did not coincide in dactyl lists, so the salience hypothesis and the terminal position hypothesis made unique predictions.

As in Experiment 1, the stress hypothesis predicted better recall performance for dactyl-patterned lists than for monotone lists in Experiment 2. The salience hypothesis predicted better recall for the stressed items in dactyl lists versus monotone lists. In the dactyl lists, stressed items occurred at the group-initial Serial Positions 1, 4, and 7. Thus, Experiment 2 allowed us to examine the effects of stressed items that were not group final on recall performance.

The terminal position hypothesis predicted better recall in the dactyl lists versus the monotone lists for the group-final Serial Positions 3 and 6. In contrast with Experiment 1, Serial Positions 3 and 6 in the dactyl lists were unstressed, permitting us to look at effects due solely to group-final location. As in Experiment 1, the terminal position hypothesis also predicted better recall for Serial Position 7 in dactyl than in monotone lists because items in Position 7 were members of the final group in dactyl-patterned lists. The salience hypothesis also predicted good recall for Position 7 because it was stressed in dactyl groups. The terminal position hypothesis predicted better recall at Serial Position 8 in the dactyl conditions than in the monotone conditions because Position 8 was a member of the final group in the patterned lists. At Serial Position 9 the salience and terminal position hypotheses made the same prediction of no reliable difference in recall between monotone and dactyl lists. The salience hypothesis made this prediction because in both monotone and dactyl lists, Serial Position 9 was unstressed. As in Experiment 1, the terminal position hypothesis predicted equivalent recall because Position 9 was clearly list-final in both dactyl and monotone lists.

**Method**

*Participants.* Fifty participants took part in Experiment 2. Data from an additional 10 participants were excluded from the analyses reported here.

*Figure 1.* Experiment 1: Number of digit names recalled (out of 20) at each serial position in an immediate serial-recall task.
because those participants failed to follow instructions. All other participant characteristics were the same as in Experiment 1.

Materials. The lists were the same as in Experiment 1, except that the stress pattern was dactylic. The dactyl pattern consisted of one stressed stimulus item followed by two unstressed stimulus items ('-'). Thus, in contrast to the anapest pattern used in Experiment 1, stress was group-initial in the dactyl pattern. Stimulus construction and list construction were accomplished as in Experiment 1.

Procedure. All aspects of the procedure were the same as for Experiment 1.

Results and Discussion

The results of Experiment 2 are presented in Figure 2 (see the Appendix, Table A2 for means from the suffixed conditions). They indicate superior recall at Serial Positions 6, 7, and 8 in the stress-patterned lists compared with the monotone lists. This provides strong support for the stress and terminal position hypotheses, in the absence of unambiguous support for the salience hypothesis.

The stress hypothesis predicted better recall for the dactyl lists than for their monotone controls, and this was observed. A 2 (list type: monotone vs. dactyl) × 9 (serial positions) ANOVA indicated a reliable main effect of list type, with better mean recall for the dactyl condition (13.8 items) than for the monotone condition (12.8 items), $F(1, 49) = 18.08, p < .0002, MSE = 11.10$. There also was a reliable main effect of serial position (see Figure 2), $F(8, 392) = 98.90, p < .0002, MSE = 11.59$, and a reliable interaction of list type and serial position, $F(8, 392) = 6.37, p < .0002, MSE = 3.56$. Thus, the stress hypothesis was supported.

Experiment 2 also tested the salience and terminal position hypotheses. The Bonferroni familywise error correction was used, and the critical $p$ value for this experiment was <.012. The salience hypothesis predicted better recall in dactyl lists than in monotone lists for the stressed items at Serial Positions 1, 4, and 7. As Figure 2 shows, stressed Positions 1 (where ceiling effects were observed) and 4 in dactyl lists were not recalled better than in monotone lists. Better recall was observed at stressed Serial Posi-

![Figure 2. Experiment 2: Number of digit names recalled (out of 20) at each serial position in an immediate serial-recall task.](image-url)
tion 7 for the dactyl condition versus the monotone condition, \(r(49) = 5.84, p < .001\). Thus, the results at Position 7 were the only data consistent with predictions of the salience hypothesis. However, these results were also predicted by the terminal position hypothesis because Position 7 is a member of the final group. The lack of reliable effects at Position 4, a stressed position, and the confounding of stress and group-initial position at Serial Position 7 raised doubt as to whether the salience alone of stressed items in the dactyl lists is what caused the recall effects at Position 7. These results are consistent with the results of Experiment 1.

The terminal position hypothesis predicted better recall for the dactyl-patterned lists at serial positions that were group-final (Serial Positions 3 and 6). The results at Serial Position 3 did not differ as a result of stress, with nearly identical results in the dactyl and monotone conditions. However, at Serial Position 6, better recall was observed for the dactyl versus the monotone condition, \(r(49) = 4.01, p < .001\).

As in Experiment 1, the terminal position hypothesis also predicted better recall for Serial Positions 7 and 8, which were members of the final group in the list. From \(t\) tests we found reliable differences for the dactyl condition versus the monotone condition at Serial Position 7, as just reported, and at Serial Position 8, \(r(49) = 3.41, p < .001\).

The terminal position hypothesis predicted good recall for list-final Position 9 in dactyl lists, but because Position 9 was also list-final in monotone lists, no difference in recall was predicted or observed across the two conditions.

Experiment 2 provided support for the stress hypothesis and for the terminal position hypothesis. All predictions of the terminal position hypothesis were upheld, with the exception of Serial Position 3, where predicted recall differences were not observed. Because the one prediction unique to the salience hypothesis was not upheld (see results for Position 4), we conclude that salience alone did not lead to significantly better recall in stressed than in monotone lists.

The results of Experiment 2 and those of Experiment 1 showed that participants segmented patterned lists without explicit instructions to group during recall (cf. Frankish, 1989). These results also showed that rehearsal is not necessary for perceptual grouping effects to occur (cf. Ryan, 1969a, 1969b). Together, the results of these two experiments show that stress patterns enabled list segmentation and led spontaneously to grouping of list stimuli, as did pauses (Frankish, 1989) and intonation patterns with an accented item in group-final position (Frankish, 1995).

In Experiments 1 and 2, we used lists using number labels as stimuli. In Experiments 3 and 4, we investigated whether the results would generalize to lists with word stimuli that were not members of the same semantic category and were perhaps not as familiar as the number names used in Experiments 1 and 2.

Experiment 3

Experiment 3 replicated Experiment 1 and investigated whether the recall advantage resulting from the imposition of a stress pattern on spoken lists of digits would extend to spoken lists of frequent English nouns. The design of Experiment 3 was identical to that of Experiment 1, except that it used a set of nine nouns that were sampled without replacement within each list from the pool of nine. The nouns differed significantly from the digits used in Experiments 1 and 2. The nouns were not members of a single semantic category, as the digits were; the nouns were not well learned as a group, whereas the digits were; the nouns were rich in both denotation and connotation, whereas the digits were more abstract; and the nouns could not easily be recoded, whereas recoding the digits (e.g., recoding the sequence 1 6 in a list to the larger unit sixteen instead of coding it as the individual units one and six) was a fairly simple process. The use of unrelated, but frequent, semantically content-laden nouns made the stimuli in Experiment 3 more like normal language than were the digit stimuli of Experiments 1 and 2.

Method

Materials and procedure. The nine words used in Experiment 3 were boy, chair, dog, hill, key, side, moon, note, and land. The mean frequency of these words was 192 words per million (wpm), with a range of 63 to 476 wpm (Francis & Kucera, 1982). Stressed and unstressed versions of these words were recorded by the same female speaker as in Experiments 1 and 2 in a soundproof booth. The pitch of the stressed stimuli and suffix approximated the musical note middle C, whereas the unstressed stimuli and suffix and the word ready approximated the musical note G below middle C. The fundamental frequency of the stimuli ranged from 247 Hz to 258 Hz for the stressed words and from 176 Hz to 204 Hz for the unstressed words. The durations ranged from 365.53 ms to 458.46 ms for the stressed words and from 301.41 ms to 398.68 ms for the unstressed words. The intensity of the stimuli and suffix ranged from 73.9 dB to 79.3 dB for the stressed words and from 71.0 dB to 73.7 dB for the unstressed words.

Stimuli were prepared for presentation using the same procedures as for the digits in Experiments 1 and 2. The reading rate was approximately 505 ms per word. Eighty lists of nine items were created by randomly sampling from the group of nine words, without replacement within a nine-item list; words were repeated across lists. The anapest pattern was created by using two unstressed-version stimulus items followed by a stressed-version stimulus item (\(-/-\), as in Experiment 1. The response interval was 18 s, giving participants sufficient time to write out the words. All other aspects of the materials and procedure were as in Experiment 1.

Participants. Fifty participants took part in Experiment 3. Data from an additional 6 participants were excluded from analyses reported here for failure to follow instructions. All other participant characteristics were the same as in Experiment 1.

Predictions. Predictions were the same as for Experiment 1.

Results and Discussion

Experiment 3 data are depicted in Figure 3. They reveal an advantage in recall performance in Serial Positions 3, 6, and 7 for the stress patterned lists and thus provide support for the stress and terminal position hypotheses but, again, no unambiguous support for the salience hypothesis.

The stress hypothesis predicted that recall would be better for the anapest lists than for the monotone lists. A 2 (list type: monotone vs. anapest) \(\times\) 9 (serial positions) ANOVA provided strong support for the stress hypothesis. There was a reliable main effect of list type, with better recall accuracy for the anapest condition (7.92 items) than for the monotone condition (7.14 items), \(F(1, 49) = 26.03, p < .0002, MSEE = 5.32\). Also (see Figure 3), there was a reliable main effect of serial position, \(F(8, 392) = 137.56, p < .0002, MSEE = 10.27\), and a reliable interaction of list type and serial position, \(F(8, 392) = 5.55, p < .0002, MSEE = 3.45\). Experiment 3 also evaluated the salience and terminal position hypotheses using a series of planned pairwise com-
parisons and the Bonferroni correction, with a critical $p$ value of <.01.

The salience and terminal position hypotheses correctly predicted better recall for the anapest lists versus the monotone lists for items in Serial Positions 3 and 6, $t(49) = 3.07, p < .01$, and $t(49) = 5.36, p < .001$, respectively. Items in these positions in stress-grouped lists were both stressed and group-final and were recalled well (see Figure 3).

The terminal position hypothesis also predicted better recall for the items in the final group of the list, namely for items at Serial Positions 7 and 8. From $t$ tests we found reliable differences for items at Serial Position 7 for the anapest condition versus its monotone control, $t(49) = 4.23, p < .001$, but for Serial Position 8, the data failed to show the predicted recall differences for the anapest condition over the monotone condition.

The salience and terminal position hypotheses made different predictions for the recall of items at Serial Position 9. The salience hypothesis predicted better recall at stressed Serial Position 9 in the anapest condition versus its monotone control, but this difference was not reliable. Figure 3 shows that at Serial Position 9, both conditions were recalled almost equally well. The terminal position hypothesis correctly predicted no difference at Serial Position 9 because the presence of the nonspeech tone made it easy to identify the terminal list item as being the last item in the list in both anapest and monotone lists. One could again argue that this failure to find a difference at Position 9 was due to recency effects putting performance near ceiling. However, it is important to note that performance at Position 9 is lower in this experiment than in Experiment 1, leaving more opportunity to observe a processing difference if indeed there was one. Thus, it is not clear that the failure to find salience effects at Position 9 can be explained by ceiling effects. However, as in Experiment 1, we would not want to base our conclusions solely on the result obtained at Position 9. Instead, we would conclude that the terminal position hypothesis was generally supported by the results of Experiment 3 (with the exception of Serial Position 8). The salience hypothesis failed to
predict good recall at Serial Position 7, and its prediction of better recall for the anapest lists versus the monotone lists at Serial Position 9 was not observed.

Like Experiments 1 and 2, Experiment 3 found that the effect of stress patterning in a list was to group list stimuli and that this grouping influenced serial list recall. This finding complements Frankish’s (1989) finding that pauses grouped lists, even when time considerations did not allow for rehearsal. In contrast to Frankish’s work, where participants received explicit grouping instructions, Experiments 1–3 showed that grouping due to stress occurs even when participants receive no explicit instructions to group items at recall.

Of note was the difference in the overall level of recall for words compared with digits, as Figures 1 and 3 clearly show. We attribute this general lowering of the serial position curve for the word stimuli to the extra time needed to write out the words compared with the time needed to jot down the digits.

Experiment 4

In Experiment 4, a dactyl stress pattern (/ - -) was imposed on word lists like those used in Experiment 3. We sought confirmation that the better recall due to stress patterning that we observed in Experiment 2 generalized to word lists. As in Experiment 2, we examined Serial Positions 1 and 4 in the dactyl-patterned list for evidence that stressed positions that were neither group-final nor members of the final group would be recalled better due to the presence of stress.

Method

Materials and procedure. Experiment 4 used a dactyl stress pattern in lists of frequent words. Materials and procedure were the same as for Experiment 3, with the exception of the dactyl stress pattern, which was created as in Experiment 2.

Participants. Forty-eight participants took part in Experiment 4. Data from an additional 7 participants were excluded from the analyses reported below because the participants failed to follow instructions. All other participant characteristics were as reported in Experiment 1.

Predictions. The predictions were the same as for Experiment 2.

Results and Discussion

Experiment 4 results are depicted in Figure 4. There was an overall recall benefit for the stress-patterned lists, a finding that supports the stress hypothesis. Recall benefit was most obvious at Serial Positions 6 and 7, consistent with predictions of the terminal position hypothesis.

The stress hypothesis predicted, and we observed, better recall for dactyl lists (7.1 items) than for their monotone controls (6.5 items). A 2 (list type) × 9 (Serial Positions 1–9) ANOVA revealed reliable main effects of list type, \(F(1, 47) = 12.01, p < .0002\), \(MSE = 6.46\), and serial position, \(F(8, 376) = 146.7, p < .0002\), \(MSE = 11.17\). Also, there was a reliable interaction of list type and serial position, \(F(8, 376) = 6.07, p < .0002\), \(MSE = 2.34\). Planned pairwise comparisons were used to evaluate predictions of the salience and terminal position hypotheses at specific serial positions. For Experiment 4, the critical \(p\) value, corrected for familywise error, was <.012.

The salience hypothesis predicted better recall in dactyl lists than monotone lists for items at Serial Positions 1, 4, and 7. However, as seen in Figure 4, recall was not better for items at Position 1, where ceiling effects were observed, or at Position 4 in dactyl lists versus the monotone lists. At Serial Position 7, recall in the dactyl condition was better than recall in the monotone condition, \(t(47) = 7.30, p < .001\). However, better recall at Serial Position 7 for dactyl lists was also predicted by the terminal position hypothesis because this position was in the final group. This pattern of effects at Position 7 was also present in the results of Experiments 1 and 3 (see Figure 1 and Figure 3, respectively), where Position 7 was not stressed. Together with other evidence supporting the terminal position hypothesis, we concluded that the terminal position hypothesis provided a valid account of these effects at Position 7.

The terminal position hypothesis predicted better recall for items at unstressed Positions 3 and 6 in the dactyl-patterned lists because these were group-final positions. However, dactyl stress did not improve recall at Serial Position 3 relative to the monotone condition (see Figure 4). At Serial Position 6, there was better recall for the dactyl condition versus the monotone condition, \(t(47) = 2.84, p < .01\).

The terminal position hypothesis also predicted better recall for members of the last group in the dactyl lists versus in the monotone lists. There were reliable differences at Serial Position 7 for the dactyl condition versus the monotone condition (\(t\) tests reported above). However, at Serial Position 8, as in Experiment 3, there was no reliable difference in recall for items in stressed lists compared with monotone lists. In general, these results resembled the results of Experiments 1–3.

The terminal position hypothesis predicted good recall for Position 9 in dactyl lists, but because Position 9 was also list-final in monotone lists, good recall was predicted there as well. Thus, as in Experiments 1–3, no differences in recall for Serial Position 9 were predicted by the terminal position hypothesis across the stress-patterned and monotone conditions, and no differences were observed.

In summary, the predictions of the terminal position hypothesis for the stress-patterned lists of Experiment 4 were supported by the recall data for items at Serial Positions 6, 7, and 9 (but not at Serial Positions 3 or 8). The pattern of results we obtained in Experiment 4 is similar to the pattern obtained by Frankish (1989) using pause-grouped lists, indicating that perceptual grouping due to stress occurs at group boundaries even in the absence of instructions to group the lists, and in the absence of time to rehearse (cf. Ryan, 1969a, 1969b). The general lowering of the serial position curve seen in Experiment 3 for word stimuli was also observed for the word stimuli of Experiment 4.

Experiment 5

Experiment 5 was motivated by the observation that the monotone serial position curves of Experiments 1–4 were not smooth but were slightly scalloped. This slight scalloping effect is neither surprising nor is it unique to our work. Other research has shown evidence of spontaneous grouping in the absence of any overt grouping cues when list length exceeds span (see Henson, 1998). However, in our experiments, this spontaneous grouping may have been exaggerated by the generalization of stress-patterning effects.
to the monotone lists, given the randomized presentation order of the lists. Because language is highly configured, and because language learning and development is closely allied with auditory pattern recognition and pattern generalization (e.g., see Morgan, Meier, & Newport, 1987; Mehler, Jusczyk, Lamberta, et al., 1988; Fernald & Mazzie, 1991; Trehub & Trainor, 1993; Echols, 1993; Gerken, 1994; Harley, Howard, & Hart, 1995), we anticipated that the stress patterns we used in our experiments would generalize to the monotone lists as a natural part of the process of auditory language processing. However, the adoption of a grouping strategy that transferred to the monotone lists may have obscured the benefit to recall arising from item salience. This would be critical to evaluation of the results at Serial Position 4. This is the only position in dactyl-patterned lists that is both unaffected by ceiling effects (a medial position in the list) and that is predicted by the salience hypothesis alone to show a recall benefit (a stressed position that is not also a group-final or final group position).

Experiment 5 used monotone and dactyl stress-patterned lists that were blocked and counterbalanced to address two issues relevant to stress grouping and item salience: first, the extent to which grouping generalized, and second, the effect of item salience when the potential for generalized grouping effects was controlled. This generalization of stress pattern to monotone lists should be pronounced in the case where monotone lists followed presentation of dactyl lists. Of particular interest to us was the pattern of results at Serial Position 4 and whether that pattern was similar to the pattern of results obtained at other list-medial positions where better recall was observed in Experiments 1–4 for stress-patterned lists.

Method

Materials. Experiment 5 used a dactyl stress pattern in lists of digits. Materials were the same as for Experiment 2.

Procedure. The procedure was the same as for Experiment 1, with the exception that four practice lists instead of five were used, a tone was the sole recall cue, and presentation of list type was blocked instead of random. A block of 20 monotone lists and a block of 20 dactyl-patterned lists were created, with each participant receiving both blocks. Blocks were counter-

Figure 4. Experiment 4: Number of words recalled (out of 16) at each serial position in an immediate serial-recall task.
balanced to produce two orders of presentation, manipulated between participants. Each participant received one order of presentation (monotone first or dactyl first). Each testing session lasted for approximately 35 min.

Participants. Sixty-eight participants took part in Experiment 5. Data from 11 additional participants were excluded from the analyses reported below because the participants failed to follow instructions. All other participant characteristics were as reported in Experiment 1.

Predictions. The stress hypothesis predicted better recall for the dactyl lists than for their monotone controls. Because we were particularly interested in Serial Position 4, a medial position uncompromised by the ceiling effects associated with primacy (Serial Position 1) or recency (Serial Position 9), we limited our predictions to Serial Position 4 and other critical medial positions (3, 6, and 7). The terminal position hypothesis predicted better recall for the dactyl lists compared with the monotone lists at three medial serial positions, specifically group-final Serial Position 3 and 6 and at final group Position 7. The salience hypothesis predicted better recall for the dactyl lists compared with the monotone lists at two medial stressed positions, Serial Position 4 and Serial Position 7 (where the terminal position hypothesis also predicted superior recall).

Furthermore, we predicted that the benefit associated with the dactyl pattern would transfer to the monotone lists, producing better recall for the monotone lists when the order of presentation was dactyl first than when it was monotone first, as suggested by the pattern of results from Experiments 1–4.

Results and Discussion

In Experiment 5, the different pattern of results for Serial Position 4 versus Serial Positions 3, 6, and 7 showed that salience effects due to stress patterning were distinct from grouping effects due to stress patterning (see Table 1). The results of this experiment, together with the results of Experiment 2, showed that a dactyl stress pattern caused a grouping effect that transferred to monotone lists. Recall performance at medial positions critical to the terminal position hypothesis (3, 6, and 7) was enhanced for monotone lists that followed presentation of dactyl lists. Salience effects on recall performance were distinct from the effects of stress-patterned grouping on recall performance. There was a benefit observed for a stressed item in the dactyl list (Serial Position 4) but that benefit did not transfer to the monotone list.

A 2 (list type) × 2 (order of presentation) × 9 (serial position) ANOVA showed that dactyl stress patterning caused grouping effects that were transferred to monotone lists. As in the previous experiments, recall for the dactyl lists was better than recall for the monotone lists (see Table 1), and this was supported by a main effect of list type, \( F(1, 66) = 84.71, p < .0001, MSE = 1,792.00 \). Recall also differed depending on the position of the item in the list, with a main effect of serial position, \( F(8, 66) = 140.52, p < .0001, MSE = 1,526.00 \). There was no main effect of order of presentation (\( F < 1 \)). However, more importantly, order of presentation participated in a three-way interaction with serial position and list type, \( F(8, 528) = 7.27, p < .0001, MSE = 29.83 \), and there was a two-way interaction between order of presentation and list type, \( F(1, 1) = 28.63, p < .0001, MSE = 606.00 \). These interactions reflect the fact that performance on monotone lists was influenced by whether or not they were preceded by a stress-patterned list, but performance on the stress-patterned lists was relatively unaffected by order. Grouping effects were strongest in the stress-patterned lists, as evidenced by the interaction of list type and serial position, \( F(1, 8) = 39.65, p < .0001, MSE = 163.00 \). There was no interaction of serial position and order of presentation, \( F(1, 8) = 1.14, p > .3, MSE = 12.35 \).

The important interactions from the omnibus ANOVA were explored in more detail in a series of 2 (list type: monotone vs. dactyl) × 2 (order of presentation: monotone first vs. dactyl first) ANOVAs, one for each of four critical serial positions: 3, 4, 6, and 7. The pattern of the data at Positions 3, 6, and 7 is remarkably similar, whereas the pattern at Position 4 is distinctly different (see Table 1). At Positions 3, 6, and 7, when the monotone block followed the stress-patterned block, recall in the monotone lists was boosted to a level similar to that observed in the stress-patterned lists. The critical interaction of list type and order of block presentation was reliable for Serial Positions 3, 6, and 7, \( F(1, 66) = 19.49, p < .0001, MSE = 109.44; F(1, 66) = 38.00, p < .0001, MSE = 559.12; \) and \( F(1, 66) = 15.08, p < .0002, MSE = 136.00, \) respectively. In contrast, at Serial Position 4, there was no such interaction, \( F(1, 66) = 2.32, p > .1, MSE = 14.89 \).

At Serial Positions 3, 4, 6, and 7, there was a reliable main effect of list type, \( F(1, 66) = 24.94, p < .0001, MSE = 140.03; F(1, 66) = 15.71, p < .0002, MSE = 100.65; F(1, 66) = 85.24, p < .0001, MSE = 805.50; F(1, 66) = 116.46, p < .0001, MSE = 1,050.62, \) respectively. These effects of list type reflected the fact that recall was always better for dactyl lists than for monotone lists. Of importance, this was true whether the dactyl or monotone list was presented first (see Table 1). There was no effect of order of block presentation at any serial position tested.

Planned pairwise comparisons of recall of monotone lists preceding dactyl lists with recall of monotone lists following the dactyl lists revealed reliable differences at Serial Position 3, \( t(66) = 2.39, p < .02; \) at Serial Position 6, \( t(66) = 4.50, p < .01; \) and at Serial Position 7, \( t(66) = 2.54, p < .01. \) However, at Serial Position 4, there was no reliable difference, \( t(66) = 0.63, p < .53. \) These results indicate that the blocked presentation of the dactyl pattern induced a grouping strategy that applied to the monotone lists at serial positions relevant to grouping. The results at Serial Position 4, a position predicted by the salience hypothesis to show a recall benefit, suggest that item salience is associated with the item itself and did not generalize from the dactyl stress-patterned lists to the monotone lists.
In summary, in Experiment 5, the different pattern of results for Serial Position 4 versus Serial Positions 3, 6, and 7 showed clearly that effects of salience on recall due to stress patterning were distinct from the effects of grouping on recall due to stress patterning. Salience improved recall at Serial Position 4 for dactyl-patterned lists compared with the monotone lists (regardless of whether the dactyl lists were presented first or second), but the effects of salience, unlike the grouping effects, did not transfer to the monotone lists. Serial Position 7, which was predicted by both the salience and terminal position hypotheses to show a recall benefit, followed the pattern of results associated with the terminal position hypothesis and observed at Serial Positions 3 and 6. Thus, the data indicated that grouping contributed heavily to the better recall that was observed in grouped lists.

General Discussion

The stress hypothesis was strongly supported by Experiments 1–5: Presence of a stress pattern in a spoken list enhanced recall for the list. In Experiments 1–5, each stress-patterned condition showed reliably better recall than the monotone control condition.

The pattern of results from Experiments 1–5 indicate that stress patterns led to better recall largely by organizing the stimulus string into groups and not simply by making certain pieces of information conspicuous, salient, or distinctive (cf. Wallace, 1965). This conclusion was supported by the fact that the most robust and reliable effects in each of the five experiments reported were observed at serial positions that were group-final or members of the final group. To the extent that effects were observed at stressed positions that were not group-final, these effects were smaller and not as consistently reliable across experiments. Finally, in Experiment 5, distinct effects on recall were exhibited for salience versus grouping due to stress patterning.

The results of Experiments 1–5 suggested that whether conditions were randomized or blocked, the stress patterns transferred to the monotone lists with the result of enhanced recall for the monotone lists when dactyl lists preceded or were mixed with the monotone lists. Our results further indicated that when monotone and stress-patterned conditions were not blocked, a situation that approximates real-world language processing more closely than do blocked conditions, salience due to stress was not by itself sufficient to create reliable differences between stress-patterned and unpatterned (monotone) lists. Our data indicate that when presentation of stress-patterned and monotone conditions was randomized, stressed items per se were not recalled significantly better than their monotone counterparts. Instead, salience may have worked together with grouping effects to contribute to reliable differences where they were observed. When stress was combined with final group membership or group-final position, stimuli were recalled well. Serial Positions 3 and 6 in the anapest lists were both group-final and stressed, and Serial Position 7 in the dactyl lists was stressed and a member of the final group in the list. Items in all three of these serial positions were recalled reliably better across experiments than their counterparts in the monotone control conditions.

Clearly, stress patterning influenced serial-recall performance in a manner similar to pausal grouping (Frankish, 1985, 1989; Ryan, 1969a, 1969b) and intonational contours ending with an accent (Frankish, 1995) by improving retrieval of group-final items. However, a major difference between this study and those of Frankish (1989) and Ryan (1969a, b) was that both Frankish and Ryan instructed their participants to adopt a grouping strategy, whereas we gave no instructions to group. The results of the five experiments reported here indicated that the participants spontaneously segmented speech strings by using the acoustic–perceptual information afforded by stress patterns to indicate group boundaries and act as retrieval cues.

Our results also addressed the role of rehearsal in the better recall of pause-grouped lists (Frankish, 1989; Ryan, 1969a, 1969b). Ryan (1969a) argued that temporal grouping may aid recall primarily because pauses allow time for rehearsal. However, our data, together with Frankish's (1989) work, led us to reject Ryan's (1969a) rehearsal hypothesis as the primary explanation for the superior recall associated with temporal grouping. The stress-patterned lists we used contained no pauses, providing no opportunity for rehearsal. Thus, our results give compelling support to Frankish's argument that the primary role of prosody in verbal strings is to group the list, not to provide time for rehearsal.

Our results are also consonant with several of the major findings in the auditory nonspeech literature. Bregman (1990, 1993), among others, has shown that pitch can be used as a perceptual grouping mechanism, with group boundaries signaled by acoustic discontinuities or changes, and Deutsch and Feroe (1981) have shown that hierarchically encoded pitch series (series encoded according to Gestalt principles) result in better recall. Pitch was one of the parameters manipulated in our stress groups, and group boundaries were indicated by pitch changes. Moreover, our stress groups conformed to Gestalt principles and showed the recall benefit that accompanies pitch grouping.

The results obtained in these experiments also have implications for current models of serial recall. We consider four of these models: the feature model (Nairne, 1990), the start–end model (Henson, 1998), Burgess and Hitch's (1999) network model of the phonological loop, and the primacy model (Page & Norris, 1998b). The first three models address grouping, but the fourth model does not.

Nairne's (1990) feature model addresses how the temporal order information required for language understanding is retained. In the model, memory traces (or memory representations) are represented as feature vectors. Interference, defined as the overwriting of one memory trace by another, is the main cause of forgetting. Trace distinctiveness (lack of overlap in feature vectors) increases the likelihood of sampling the correct item, whereas feature similarity (overlap in feature vectors) is a necessary condition for the overwriting of an item by its successor. However, according to Nairne's model, a memory trace can overwrite another memory trace only if both traces are members of the same group. Moreover, because search sets in this model are limited to members of individual groups, grouping produces overall better recall, with primacy and recency effects in each group. Thus, Nairne's model can account for the overall enhanced recall observed for stress-patterned versus monotone lists in Experiments 1–5, and it can account for the more specific finding that the group final positions were the positions that showed the greatest benefit. However, the model would need to be modified to incorporate the distinct effects of salience and grouping observed in Experiment 5. This could be accomplished either by adding a representation of salience to the trace vectors or by altering their notion of distinctiveness to
include the effects of salience on memory for auditory language lists.

Henson's (1998) start–end model (SEM) of serial recall assumes that position in a verbal string is coded relative to the two most salient aspects of the string, its start and its end. The start marker is strongest at the beginning of the string, and its strength decreases toward the end. The end marker is weakest at the beginning of the string, and its strength grows toward the end of the string. This positional coding results in episodic tokens for each item in short-term memory. Like Nairne's (1990) feature model, Henson’s SEM incorporates grouping, and like the feature model, the SEM predicts primacy and recency effects in each group.

Burgess and Hitch's (1999) connectionist model extends Baddeley's (1986) phonological loop to include serial order and long-term learning. Item selection in serial order takes place at a lexical level, and the selection itself is made by a competitive queuing process. According to Burgess and Hitch, positional information is represented in short-term memory and is coded as a context signal, which, in effect, attaches serial position markers to each item. In grouped lists, the representation of serial order information reflects both the position of an item within the list and the position of the item within its group. The position of the item within the list is represented by a set of context nodes whose signal is modulated by a second set of context nodes connected to within-group positional information every time a pause is encountered. The pattern of activation in the second set of context nodes depends only on the position within the group, and these nodes reset with each pause (i.e., at the end of each group) and at the start of recall. Therefore, this model can readily explain primacy and recency effects in lists and in groups within a list. If the context nodes could be expanded to include context information about boundaries other than pauses, then the model may well be compatible with the results of the present research.

Unlike the other models we have discussed, Page and Norris's (1998a, 1998b) primacy model makes no provision for grouping, although the authors stated that the model will be amended to address grouping affects (Page & Norris, 1998b). In this model, decay, which occurs in the absence of cumulative rehearsal, is the primary factor in forgetting. The primacy model assumes that order information is stored directly through the graded activation of representations of list items. Because activation decays over time, earlier list items should be recalled better than later items, and this explains order effects. Noise is responsible for recall errors in the model. If we make an added assumption that stressed items have more active representations than unstressed items in the primacy model, then the model may easily account for the salience effects observed in Experiment 5. However, it is not clear to us whether or how stress patterning would affect the strength gradient in the primacy model, and, thus, it is difficult to see how this model could account for the grouping effects observed in Experiments 1–5 reported here.

The results of Experiments 1–5 are also pertinent to studies relating language processing and memory. M. L. Slowiaczek and Clifton (1980) studied the effects of blocking subvocalization on reading. They found that when subvocalization was blocked, reading comprehension, but not memory for individual word concepts, was affected. They noted that reading comprehension requires the integration of various types of information across clauses and sentences, and they theorized that subvocalization allows a prosodic restructuring of the written linguistic string to occur. According to Slowiaczek and Clifton, this prosodic restructuring is important because prosody, which is absent in written language, provides pattern information in the spoken language necessary to highlight important information and enable complex language processing. In this context, they referred specifically to the rhythmic aspect of prosody. They also proposed that the prosodic patterning of information creates a more durable memory structure, perhaps by facilitating rehearsal. The results of Experiments 1–5, along with Frankish's (1989, 1995) studies, support Slowiaczek and Clifton's contentions. Stress provides rhythmic patterns that group lists and result in better recall, with or without rehearsal. Experiment 5 shows that such grouping patterns can be used to restructure otherwise ungrouped lists, which include both spoken monotone lists and, presumably, written lists (in this context, sentences) devoid of any prosodic information.

The results of Experiments 1–5 may also be relevant to discourse structure. Clarke and Clarke (1978) have noted the strong, perhaps universal, tendency of languages to place "given" information before "new" information, and Bock and Irwin (1980) have suggested cogent reasons why this should be so. Accordingly, given information is placed closer to the beginning of an utterance, whereas new information comes closer to the end of the utterance. New information, because it is new, presumably must be retained in memory, however briefly, so that it can be integrated into the individual clause or sentence in which it appears and into the prose structure as a whole. We know from Experiments 1–5 that main effects of grouping are found toward the end of the list (i.e., sentence), where the new information is found. Thus, the better recall predicted by the terminal position hypothesis and so observed is consonant with the need for retention of new information predicted by the given–new information structure in spoken languages.

The finding that in English stress patterns, especially anapest patterns, produce better recall for linguistic strings than does monotone intonation may have implications for the syntactic structure of English phrases. Johnson (1965) and Rosenburg (1968) have shown that individual words in an utterance are recorded for storage into larger syntactic and associative-semantic structures (i.e., groups). Also, Jarvella (1970) has demonstrated in an immediate verbatim recall task that the likelihood of retrieval of an individual word is dependent on syntactic phrase structure and boundary considerations. In English and in other languages, such as German and Russian, phrases are usually constructed so that words bearing the largest semantic and syntactic load come last in the phrase. For example, in a prepositional phrase, the noun is phrase final, and in a verb phrase, the main verb is phrase final. In these phrases, and similar syntactic phrases, the end of group constituent is an open-class word (usually a noun or main verb) that carries a large syntactic and semantic load, and open-class words are the words most likely to receive stress. Therefore, by placing in group-final position the syntactically and semantically

---

3 It should be noted that the Nairne (1990), Henson (1998), and Burgess and Hitch (1999) models all predict primacy effects for individual groups. The results from Experiments 1–5 suggest that these models may overpredict primacy effects for stress-patterned grouping.
important words, the words most likely to receive stress in stress-timed languages, the phrase structure of English ensures that the salient information carried by these words will likely be available at retrieval.

In summary, the results of Experiments 1–5 have specific implications for our understanding of cognition in general and language processing in particular. The benefits of temporal grouping, or chunking information by the use of pauses, have been known since at least 1967 (Burgess & Hitch, 1999; Frankish, 1985, 1989; Hitch et al., 1996; Kahneman, 1973; Kahneman & Henik, 1981; Penney, 1978; Ryan, 1969a, 1969b; Wickelgren, 1967). Furthermore, the possibility that prosody serves as a grouping mechanism for speech was noted by Frankish (1989, 1995). Pauses are one type of grouping device used by the language system, but speakers regularly talk for a long time without pausing; however, the speech produced includes divisions and boundaries that are fully recognized by the perceiver (Bolinger, 1989). Indeed, explaining what aspects of the speech stream give rise to our perception of divisions and boundaries is one of the central problems facing researchers today. The results of our research, together with the earlier work on grouping, indicate that pauses and intonation are not the only prosodic mechanisms that serve to group speech, thereby aiding speech perception and enhancing memory. Language perception and retention are possible because additional prosodic devices like stress are present to help “chunk” (Miller, 1956) the utterance into segments and mark linguistic boundaries. The results of Experiments 1–5 show that grouping and the ensuing superior recall can be effectively achieved by stress, a feature that is more pervasive in spoken English than are pauses.

References


Appendix

Suffix Data: Experiments 1–4

<table>
<thead>
<tr>
<th>Serial position</th>
<th>Monotone-go</th>
<th>Anapest-go</th>
<th>Anapest-GO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.95</td>
<td>18.10</td>
<td>18.42</td>
</tr>
<tr>
<td>2</td>
<td>15.75</td>
<td>16.08</td>
<td>16.40</td>
</tr>
<tr>
<td>3</td>
<td>14.34</td>
<td>16.08</td>
<td>16.40</td>
</tr>
<tr>
<td>4</td>
<td>10.55</td>
<td>11.38</td>
<td>12.00</td>
</tr>
<tr>
<td>5</td>
<td>7.85</td>
<td>9.69</td>
<td>9.57</td>
</tr>
<tr>
<td>6</td>
<td>8.02</td>
<td>11.79</td>
<td>12.93</td>
</tr>
<tr>
<td>7</td>
<td>6.38</td>
<td>7.93</td>
<td>7.91</td>
</tr>
<tr>
<td>8</td>
<td>6.00</td>
<td>7.44</td>
<td>7.24</td>
</tr>
<tr>
<td>9</td>
<td>9.46</td>
<td>11.00</td>
<td>10.40</td>
</tr>
</tbody>
</table>

Table A2

Experiment 2: Digit Names, Dactyl Versus Monotone Pattern, Suffixed Conditions (Mean Number of Items Recalled)

<table>
<thead>
<tr>
<th>Serial position</th>
<th>Monotone-go</th>
<th>Dactyl-go</th>
<th>Dactyl-GO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.22</td>
<td>18.62</td>
<td>18.10</td>
</tr>
<tr>
<td>2</td>
<td>16.06</td>
<td>16.68</td>
<td>16.30</td>
</tr>
<tr>
<td>3</td>
<td>14.50</td>
<td>15.28</td>
<td>15.20</td>
</tr>
<tr>
<td>4</td>
<td>11.38</td>
<td>12.34</td>
<td>11.52</td>
</tr>
<tr>
<td>5</td>
<td>8.76</td>
<td>10.14</td>
<td>9.70</td>
</tr>
<tr>
<td>6</td>
<td>8.44</td>
<td>10.96</td>
<td>10.42</td>
</tr>
<tr>
<td>7</td>
<td>6.92</td>
<td>9.40</td>
<td>9.08</td>
</tr>
<tr>
<td>8</td>
<td>6.24</td>
<td>8.14</td>
<td>7.28</td>
</tr>
<tr>
<td>9</td>
<td>8.22</td>
<td>9.90</td>
<td>9.44</td>
</tr>
</tbody>
</table>

Table A3

Experiment 3: High Frequency Nouns, Anapest Versus Monotone Pattern, Suffixed Conditions (Mean Number of Items Recalled)

<table>
<thead>
<tr>
<th>Serial position</th>
<th>Monotone-go</th>
<th>Anapest-go</th>
<th>Anapest-GO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13.42</td>
<td>13.28</td>
<td>12.70</td>
</tr>
<tr>
<td>2</td>
<td>11.14</td>
<td>10.68</td>
<td>10.44</td>
</tr>
<tr>
<td>3</td>
<td>8.42</td>
<td>10.22</td>
<td>9.86</td>
</tr>
<tr>
<td>4</td>
<td>5.30</td>
<td>5.50</td>
<td>5.04</td>
</tr>
<tr>
<td>5</td>
<td>3.54</td>
<td>4.22</td>
<td>3.88</td>
</tr>
<tr>
<td>6</td>
<td>3.20</td>
<td>5.02</td>
<td>4.88</td>
</tr>
<tr>
<td>7</td>
<td>2.20</td>
<td>4.90</td>
<td>4.92</td>
</tr>
<tr>
<td>8</td>
<td>3.60</td>
<td>4.00</td>
<td>4.38</td>
</tr>
<tr>
<td>9</td>
<td>5.56</td>
<td>7.44</td>
<td>6.66</td>
</tr>
</tbody>
</table>

Table A4

Experiment 4: High Frequency Nouns, Dactyl Versus Monotone Pattern, Suffixed Conditions (Mean Number of Items Recalled)

<table>
<thead>
<tr>
<th>Serial position</th>
<th>Monotone-go</th>
<th>Dactyl-go</th>
<th>Dactyl-GO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13.39</td>
<td>13.04</td>
<td>12.81</td>
</tr>
<tr>
<td>2</td>
<td>10.56</td>
<td>9.83</td>
<td>9.77</td>
</tr>
<tr>
<td>3</td>
<td>7.89</td>
<td>8.83</td>
<td>8.64</td>
</tr>
<tr>
<td>4</td>
<td>4.56</td>
<td>4.16</td>
<td>4.60</td>
</tr>
<tr>
<td>5</td>
<td>2.68</td>
<td>3.10</td>
<td>2.54</td>
</tr>
<tr>
<td>6</td>
<td>2.00</td>
<td>2.85</td>
<td>3.14</td>
</tr>
<tr>
<td>7</td>
<td>1.93</td>
<td>4.29</td>
<td>3.62</td>
</tr>
<tr>
<td>8</td>
<td>3.02</td>
<td>3.95</td>
<td>3.39</td>
</tr>
<tr>
<td>9</td>
<td>5.54</td>
<td>5.70</td>
<td>5.45</td>
</tr>
</tbody>
</table>

Received April 23, 1999
Revision received March 30, 2000
Accepted April 15, 2000