Temporal coding in rhythm tasks revealed by modality effects

ARTHUR M. GLENBERG
University of Wisconsin, Madison, Wisconsin

and

MENACHEM JONA
Northwestern University, Evanston, Illinois

Temporal coding has been studied by examining the perception and reproduction of rhythms and by examining memory for the order of events in a list. We attempt to link these research programs both empirically and theoretically. Glenberg and Swanson (1986) proposed that the superior recall of auditory material, compared with visual material, reflects more accurate temporal coding for the auditory material. In this paper, we demonstrate that a similar modality effect can be produced in a rhythm task. Auditory rhythms composed of stimuli of two durations are reproduced more accurately than are visual rhythms. Furthermore, it appears that the auditory superiority reflects enhanced chunking of the auditory material rather than better identification of durations.

Progress has been made in understanding temporal coding from multiple perspectives. One way of thinking about temporal coding is in terms of when an event has occurred. This sort of coding, represented by work by Friedman (1987), Tzeng and Cotton (1980), and Winograd and Soloway (1985), typically involves the presentation and recall of words and other verbal materials. Temporal coding can also be thought of as a representation of the duration of events (e.g., Block, 1982; Ornstein, 1969). Finally, some investigations of temporal coding have focused on the perception and reproduction of rhythms (e.g., Martin, 1972; Povel, 1981). The major goal of this article is to build a bridge between these various perspectives.

We begin by demonstrating that a variable having important effects in recall, modality of presentation, plays an equally important role in the coding of rhythms. Previous work on the perception of patterns has revealed a type of modality effect. For example, Garner and Gottwald (1968) presented eight-element patterns composed of two different pitches or two different-colored lights. A number of measures indicated that the auditory patterns were easier to learn than the visual patterns (at least at fast presentation rates). Handel and Buffardi (1968) demonstrated similar effects. In those experiments, modality was confounded with a stimulus dimension. Auditory patterns were generated by sequencing two *pitches*, whereas visual

patterns were generated by sequencing lights of different *colors*. In the work reported here, the auditory and visual patterns are more comparable. In both modalities, the pattern is created by turning a single stimulus (tone or light) on or off and manipulating solely the durations of the stimuli. In preview, we have found that the coding of temporal patterns conveyed by an auditory signal is more sensitive to rhythmic structure than is coding of the same pattern conveyed by a visual signal. In addition, memory for auditory temporal patterns is superior to memory for visual temporal patterns. A second experiment demonstrates a constraint on these effects that helps to identify the processes used in temporal coding.

MODALITY AND TEMPORAL CODING

Many investigations of temporal coding have employed auditory stimuli exclusively (Deutsch, 1986; Povel, 1981; Povel & Essens, 1985). This strategy may reflect a choice dictated by convenience. However, there are reasons to believe that modality of presentation may have important effects, effects that point to limitations in current conceptions of temporal coding and point to new theoretical alternatives. Glenberg and Swanson's (1986) temporal-distinctiveness theory proposes that modality effects in recall tasks reflect a difference among modalities in temporal coding. In particular, they suggested (following Gardiner, 1983) that auditory stimuli result in more accurate coding of time of presentation (when events occur) than do visual stimuli. Furthermore, they specified that temporal information is used as a retrieval cue to aid the recall of the most recently presented information. In combination, these assumptions provide an explanation for the modality effect, that recall of recently presented

auditory information exceeds recall of recently presented visual information.

Three classes of findings support this temporal-coding interpretation of modality effects. First, memory of auditory events is more sensitive to temporal variables (e.g., the durations between items) than is memory for visual events (Glenberg & Swanson, 1986). Second, memory for order of events is affected by modality: Memory for order of auditory events exceeds memory for order of visual events (Glenberg & Fernandez, 1988; Greene & Crowder, 1988). Third, and most germane to the present discussion, is the finding that auditory rhythms are remembered better than are visual rhythms (Glenberg, Altman, Mann, Forman, & Procise, 1989).

In the experiments reported by Glenberg et al. (1989), subjects were presented sequences composed of short (e.g., 250 msec) and long (e.g., 750 msec) stimuli. These stimuli were either tones or visual signals. Immediate reproduction of the sequences was more accurate for the auditory stimuli than for the visual stimuli. This auditory advantage was relatively unaffected by signal-to-noise ratio, practice, and response mode (a sequence of button-presses in which one button represents short and the other long, or an analog reproduction of the actual temporal durations). To the extent that this rhythm-reproduction task taps temporal coding, the results demonstrate more accurate temporal coding for auditory than for visual stimuli.

On the face of it, the conclusion that temporal coding is more accurate for auditory than for visual stimuli does not fit well with direct investigations of the perception of duration (Crowder & Greene, 1987; also see Allan, 1979, for an earlier review). Those studies do not find strong modality differences. One resolution is to propose that Glenberg et al.'s (1989) results do not arise from modality differences in coding *duration*, but from modality differences in the coding of *ordinal relations* among the durations; that is, which durations follow which others. This view is consistent with the work of Povel (Povel, 1981; Povel & Essens, 1985). He proposed that coding of a rhythm involves two processes. One is imposing an isochronous beat (or clock) on the rhythm. (We will be using the terms "beat" and "beat interval" to signify a constant duration, not the onset of an event.) The beat is chosen so that when beats are aggregated and mapped onto the rhythm there are few beats that begin on unaccented portions of the rhythm. Once the beat is chosen, the pattern of individual elements (e.g., notes in a melody) within the beat are chunked. That is, a single symbol, the chunk, is used to represent several stimuli and their order.

EXPERIMENT 1

Experiment 1 provides a test between two alternatives. One alternative is that the auditory superiority in the reproduction of rhythms (Glenberg et al., 1989) reflects solely more accurate coding of the *duration* of auditory signals than of visual signals. The second alternative is

that the auditory superiority reflects a difference between the modalities in *chunking* by beats. Similarly to Povel and Essens (1985), suppose that temporal patterns are coded by first imposing an isochronous beat on the pattern and then chunking, with a single symbol, the events within a beat. We assume that the beat-based chunk preserves the number of stimuli within the beat, their relative durations (if not their metric properties), and their order. Suppose, furthermore, that chunking by beats is easier for auditory temporal patterns than for visual temporal patterns. Thus, representing auditory temporal patterns will require few symbols, because each symbol is a chunk representing several stimuli, whereas representing visual temporal patterns will require many symbols, because each symbol represents but a single visual stimulus. The greater use of chunking with auditory stimuli enhances maintenance of information in working memory, resulting in auditory superiority in sequence reproduction.

In Experiment 1, subjects tried to reproduce sequences composed of nine intermixed long (600 msec) and short (200 msec) stimuli. There were two important independent variables in the experiment. The first was the modality in which a sequence was presented. That is, all of the temporal durations in a sequence were carried by either an auditory or a visual stimulus. The second independent variable was designed to affect beat-based chunking of the temporal patterns. In the constant condition, the duration of the interstimulus interval (ISI) between successive stimuli in a sequence was always 200 msec. Thus, the duration of a short stimulus plus an ISI interval (200 + 200 msec) was exactly half of the duration of a long stimulus plus an ISI interval (600 + 200 msec), to facilitate chunking by 800-msec beats. In the variable condition, the ISI was randomly changed after each stimulus in the sequence. With a probability of .5, the ISI was 266 msec, and with a probability of .5, the ISI was 144 msec. Finally, in the split condition, the ISI after the first four stimuli was a constant 200 msec, whereas the ISI after the remaining stimuli varied randomly as in the variable condition.

We expected to find auditory superiority in the constant condition, replicating Glenberg et al. (1989). Performance in the variable condition should distinguish between the two accounts for this finding. Suppose that the auditory superiority reflects solely more accurate coding of the durations of the individual stimuli. Varying the duration of the ISI should not affect this process, so that we should find auditory superiority in the variable condition, too. However, suppose that the auditory superiority reflects the ease of imposing on auditory patterns a constant beat and the use of the beat to chunk successive durations. With the constant ISI, two short stimuli plus their ISIs (200 + 200 + 200 + 200) will exactly fill one 800-msec beat. Thus, beat-based chunking is possible and auditory superiority should be found. In the variable condition, however, two short stimuli and their ISIs may total, depending on the ISIs, 688, 800, or 932 msec. This variability precludes the use of a constant beat to chunk the

sequence, and thus performance in the auditory conditions should deteriorate to the level of performance in the visual conditions.

We included the split condition as a safeguard. Perhaps subjects require time to recognize the beat and to begin imposing it on the sequence. In this case, the first few beats in the split condition lure the subject into using the beat to encode the sequence, and then the introduction of the variable ISI disrupts this coding. Predictions for the split condition are the same as for the variable condition, except in the split condition the predictions only apply to reproduction of the end of the sequence.

Method

The presentation, timing, and scoring of all events was controlled by an Apple II+ computer. The auditory stimuli were produced by a fixed-frequency (926 Hz) square-wave oscillator used to drive a 2-in. loudspeaker (72 dB(A) at the ear). A diffused-lens, red LED, 5 mm in diameter, was mounted above the speaker, and the LED was used to present the visual stimuli. The speaker and the LED were approximately 70 cm from the subject, and they were placed atop a video monitor controlled by the computer.

Subjects were presented 12 blocks of sequences, and each block was composed of six sequences. The six sequences in a block represented one exemplar from each of the six conditions formed by the factorial combination of modality (auditory or visual) and ISI condition (constant, variable, and split). The order of the six sequences in a block was random. Also, the sequences in a block differed in the pattern of short- and long-duration stimuli, as described next.

The composition of each sequence was random within the following constraints. First, short stimuli always occurred in pairs. Thus, a pair of short stimuli along with the following ISIs would, in the constant condition, fall into a putative beat equal to the length of a long stimulus plus the following ISI. Second, no more than four stimuli of the same length could be presented successively. Third, a sequence contained a minimum of three transitions between runs of short and long stimuli.

Each sequence began with a ready signal that indicated the modality of the sequence but not the ISI condition. The ready signal was displayed on the video monitor for 1,500 msec. The ready signal was followed by a 1,000-msec blank screen. The appropriate sequence was then presented (using the loudspeaker or the LED), and it was followed by a 250-msec delay before the response probe was

displayed on the monitor. The probe consisted of an array of nine short blank lines. Subjects attempted to reproduce the sequence using two keys, one labeled "short" and the other "long." Each keypress resulted in a letter ("S" or "L") replacing one of the nine blank lines. A third key allowed the subject to erase a previous response. Pressing both the "short" and "long" keys indicated that the subject was satisfied with the reproduced sequence, and initiated presentation of a new sequence.

After providing their informed consent and listening to instructions, subjects practiced on four sequences, two of which were auditory and two visual. Orthogonally, two of the practice sequences were in the constant condition and two were in the variable condition. This practice was followed by the 12 blocks of trials.

Subjects. A total of 44 subjects (25 women, 19 men) participated. Of these, 24 were members of the University of Wisconsin summer community who were paid for their participation in the experiment, and 20 were enrolled in introductory psychology classes and received extra credit toward the course grade.

Results

Responses were scored by serial position. For each serial position, if the subject's response ("short" or "long") matched the true duration of the stimulus, a one was scored for that position, otherwise a zero was scored. The results of most interest are displayed in Figure 1. These results have been collapsed across subjects and blocks of trials.

Focus first on the results for the constant conditions. Note that performance on the auditory trials exceeds that for the visual trials; that is, there is a modality effect. Now consider performance in the variable condition. Here the modality effect is all but eliminated because of a drop in performance in the auditory condition. The results of the split condition are a combination of these two. In the initial four serial positions (before the introduction of the random ISI), performance in the auditory condition exceeds performance in the visual condition. Starting with the fifth serial position, however, there is a reversal for all but the last position. These results confirm the prediction of the chunking account of the rhythm modality effect. Apparently, the auditory advantage reflects chunking of the sequence, not more accurate identification of
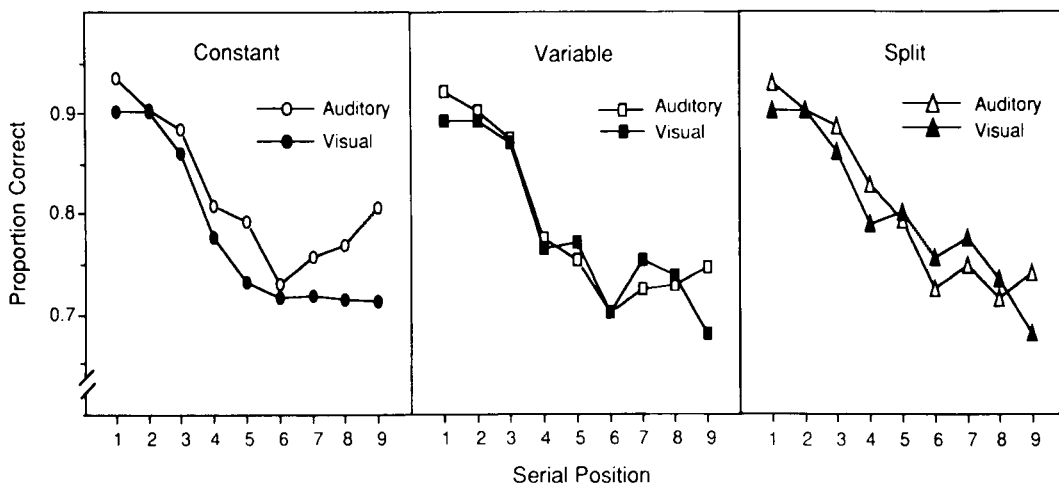

Figure 1. Proportion correct in the rhythm task used in Experiment 1.

the durations of individual sequences. There appears to be a residual auditory advantage in the final position for both the variable and split conditions. This may be noise, or as Balota and Engle (1981) claim, terminal and preterminal modality effects may reflect different processes.

The main observations were confirmed by a statistical analysis of the data using a Type 1 error rate of .05. Consider first an analysis of the data from the constant condition. There were significant main effects of modality $[F(1,43) = 7.411, MS_e = .058]$ and serial position $[F(8,344) = 38.405, MS_e = .019]$. The interaction, however, was not significant $[F(8,344) = 1.409, MS_e = .017, p = .19]$. In the variable condition, there was a significant main effect of position $[F(8,344) = 37.048, MS_e = .021]$, but the modality effect was not significant $(F < 1.0)$. Modality and position did interact $[F(8,344) = 2.763, MS_e = .016]$, but in a nonsystematic manner. That is, a simple effects analysis showed a significant auditory superiority in Positions 1 and 9 and a significant visual advantage in Position 7. The modality differences were far from significant at the other six positions. Finally, consider the analysis of the data from the split condition. The modality difference was not significant $(F < 1.0)$, but the main effect for position was significant $[F(8,344) = 41.425, MS_e = .019]$. The interaction of modality and position was also significant $[F(8,344) = 1.961, MS_e = .016]$. This interaction was systematic: Excluding the last position (because it may be special), the first four positions (constant condition) resulted in auditory superiority, whereas the next four positions (variable condition) reversed the effect $[F(1,43) = 4.284, MS_e = .027]$. (Considered alone, however, the reversal was not significant $[F < 1.0]$).

Discussion

Two points are worthy of note. First, there was a modality effect in the constant condition, and the effect was not confined to just the last few list positions. The reliable modality effect demonstrates that conceptions of temporal pattern encoding are deficient to the extent that they do not take modality into account.

Second, the pattern of effects is consistent with a greater use of chunking with the auditory items compared with the visual items. On this account, disrupting chunking using the variable ISI lowered performance on the auditory trials but not on the visual trials. Furthermore, when chunking was disrupted, there was very little difference between the modalities in the accuracy of coding durations (at least to the extent that accuracy of duration coding is measured by classification into the categories of long and short).

These results help to explain a puzzling inconsistency in the literature on modality effects and temporal coding. Crowder and Greene (1987) and Schab and Crowder (1989) studied memory for duration using the irregular-list technique. Subjects were presented a sequence of events (e.g., a permutation of the letters A-E) with various ISIs within the sequence. Following the sequence, the subjects were prompted with three letters defining two

adjacent temporal intervals and the subjects were to indicate which of the intervals was longer. Although performance was affected by the actual durations, little (Schab & Crowder) or no (Crowder & Greene) modality effect was found. These investigators took this null effect as evidence against Glenberg and Swanson's (1986) claim that temporal coding is more accurate for auditory than for visual stimuli. We now can appreciate one reason (for another, see discussion after Experiment 2) why the irregular-list technique does not produce any evidence of auditory superiority, whereas that evidence is commonplace in recall tasks and in the rhythm-reproduction task. The irregular-list technique is comparable to the random ISI condition of Experiment 1. In both cases, the development of chunking is disrupted by asynchronous presentations. That is, the auditory superiority in temporal coding is not (just) in perceiving durations; in fact, judging from the results in the variable condition, individual durations are perceived and remembered about equally well, regardless of modality.[1] Instead, the auditory superiority comes about when those durations can be chunked, as in the constant condition.

We now turn to a discussion of how that chunking might take place in the rhythm task. Our goals are modest in that we wish to provide an explanation that is sensitive to modality differences in chunking. This explanation must be combined with hypotheses of clock (beat) induction (e.g., Longuet-Higgins & Lee, 1982, or Povel & Essens, 1985) and recall (e.g., Glenberg & Swanson, 1986) to provide a complete account of the data.

Our hypothesis begins with the assertion that subjects use beats to chunk successive durations, perhaps because of previous experience with music and rhythms. Because chunking reduces the number of symbols needed to code a sequence, it enhances memory and reproduction of the sequence. We assume that analog duration codes are generated by an echoic system (but unlike in Crowder & Morton, 1969, this system is not reserved for linguistic stimuli). Thus, when the echoic system is stimulated for m msec (where $m <$ duration of echoic memory), an m-msec duration code is created. When the coded durations of successive stimuli (and their ISIs) match the duration of a beat, then the coded durations are effortlessly recoded as a chunk. These chunks code ordinal relationships (what follows what) and they are maintained in working memory. When the coded durations of successive stimuli do not match the duration of a beat, then the durations must be individually coded and maintained in working memory as separate symbols, one for each stimulus.

As an example, consider the coding of the sequence S, S, L, S, S, L, L, S, S, in which S indicates a short stimulus (200 msec plus a 200-msec ISI) and L indicates a long stimulus (600 msec plus a 200-msec ISI). Assume that the subject is attempting to code the sequence using a beat of 800 msec. The coded durations of the first two stimuli exactly fit one beat, and hence the durations can be recoded as one symbol, say, SS; the coded duration of the next stimulus, an L, would also be recoded as one

symbol, say, L, and so on. By the end of the sequence, the nine durations would be coded as six symbols, SS, L, SS, L, L, SS. If the subject were using a 1,600-msec beat, working memory would hold just three chunks, SSL, SSL, and LSS.

This sort of beat-based chunking will only work, however, if the coded durations of successive stimuli exactly match the beat duration (as in the example). At this point, we introduce the crucial assumption: Auditory stimuli have automatic access to the echoic system. Thus, an m-msec auditory stimulus results in a coded duration of m-msec. In contrast, visual signals do not have automatic access to the system. Instead, a subject-controlled process (e.g, humming to oneself) must be used to activate the system. Variability in the timing and strength of this subject-controlled activation degrades the analog representation of duration of the visual signal (e.g., an m-msec signal may be coded as a variable duration of less than m msec). This in turn degrades the system's ability to form symbolic, beat-based chunks.

Consider the coding of the example sequence when it is presented visually. To simulate inaccurate coding of durations, assume that visual stimuli are coded as about 10% less than their actual values. The first stimulus has a coded duration of 180 msec, the ISI is coded as 200 msec (for simplicity), the second stimulus is coded as 180 msec, and the second ISI is coded as 200 msec. The total coded duration is 760 msec, which does not fill the 800-msec beat. In this case, the subject could wait and try to add the next stimulus (the L) into the beat, but this will not do, because it is too long, with a coded duration of 720 msec. Because groups of coded durations do not correspond to one beat, the subject is forced to code each stimulus duration separately, resulting in nine symbols in working memory at the end of the sequence, and poor sequence reproduction. Note that subjects should have little difficulty discriminating short visual stimuli from long visual stimuli (e.g., a coded duration of 180 msec should be quite distinct from one of 720 msec). The problem is in packaging the stimuli. Furthermore, in the visual condition, it will not do to simply reduce the length of a beat by 10%. The reason is that we assume that the coding of duration for the visual modality is variable, not simply inaccurate by a constant (as in this example).

The difference between auditory and visual sequences in amount of chunking will have two consequences. First, sequences that are near memory-span length will be reproduced more accurately with auditory presentation than with visual presentation. That is, auditory sequences will be chunked to within memory span, whereas the number of symbols needed to code the visual stimulus will exceed memory span. At first glance, it may seem odd that subjects apparently do not effectively chunk the visual sequences; after all, the sequences are very simple and short stimuli always come in pairs. The model provides at least a partial answer. With auditory sequences, the chunks are formed relatively effortlessly, because two processes are presumed to be automatic. First, auditory

stimuli have automatic access to echoic memory, hence duration coding is automatic. Second, the coded durations will match the duration of a beat and will be chunked automatically. For visual stimuli, both of these processes require some conscious effort. Subjects must consciously stimulate echoic memory to code durations of visual stimuli. Because of variability in duration coding, the coded durations will not match the duration of a beat and will not be automatically chunked. Any chunking that does occur must be the result of additional effortful processing.

A second consequence of modality differences in chunking is that subjects should have difficulty imposing a beat on the visual sequences. Our informal observations are consistent with this prediction. It is relatively easy to tap one's foot in rhythm with the constant auditory sequences, and relatively difficult to do so with the visual sequences. There is also evidence from the pattern-perception literature consonant with this prediction. Subjects spend more time observing visual patterns (composed of two different-colored lights) than auditory patterns (composed of two different-frequency sounds) before attempting to reproduce the patterns (Garner & Gottwald, 1968; Handel & Buffardi, 1968).

The chunking model accounts for the major findings in Experiment 1. First, in the constant condition, there is auditory superiority in the reproduction of the sequences. According to the model, this finding reflects greater chunking of the auditory sequences compared with the visual sequences. Second, in the variable condition, reproduction accuracy of the auditory sequences is reduced. In our account, the variable ISI condition induces a sequence of durations that do not match the representation of the duration of a beat. Hence, there is a reduction in chunking and reproduction accuracy. Indeed, reproduction of the variable auditory sequences is at the same level as the constant visual sequences, an effect predicted by the model because both modalities are being processed in similar ways. Third, the variable condition causes little disruption in the reproduction of visual sequences. By the model, elements comprising visual sequences are coded individually even in the constant condition (due to variability in subject-controlled input to the echoic memory). Thus, the variable ISI cannot produce any further reduction in chunking and there is no further reduction in performance.

EXPERIMENT 2

We have attributed the auditory superiority found in the rhythm task to the automatic chunking of durations. Furthermore, we have proposed that there are limits on the duration of stimuli that will produce automatic chunking. Namely, the duration of a beat must be short enough to be represented in echoic memory, about 2 sec (Darwin, Turvey, & Crowder, 1972; Treisman, 1964). From this limitation, we can derive new predictions from our model of the coding mechanism. First, as the beat duration increases, we should observe a decrease in the difference

between auditory and visual rhythms. That is, with long beats (e.g., more than 2,000 msec), the subject should have difficulty representing a beat duration, precluding automatic chunking of the auditory stimuli. In this case, each stimulus (whether auditory or visual) will be represented in working memory as a separate symbol. Furthermore, increasing the beat duration should result in a decrease in accuracy in the auditory condition (because more symbols are required to represent the sequence), but little decrease in the visual condition (because each stimulus is already encoded as a separate symbol). However, this latter prediction may be countermanded by other factors. With a longer beat duration, subjects will have more time to consciously recode the individual symbols using a control process outside the purview of the model (as opposed to the automatic chunking discussed above). This control process should work equally well on symbols generated by auditory and visual signals. Thus, increasing the beat duration may result in improved performance, but not a difference between the modalities.

Method

Except for a few details noted next, the method was very similar to the constant condition of Experiment 1. The beat durations were 800 msec, 1,600 msec, and 2,400 msec. A beat duration was defined as the duration of a long stimulus plus the duration of the (constant) ISI, which was always one-third the duration of the long stimulus. Also, a short stimulus was always equal in duration to the ISI. For example, in the 1,600-msec condition, the long stimulus was 1,200 msec, the ISI was 400 msec, and the short stimulus was 400 msec. Subjects were randomly assigned to beat durations. Each subject practiced on four sequences (two auditory and two visual) and was then presented with 15 auditory and 15 visual sequences, which were randomly intermixed. These sequences were constructed using the same algorithm as used in Experiment 1. A total of 42 subjects (15 men, 27 women) participated in the experiment (14 in each beat-duration condition).

Results

The subjects' responses were scored as in Experiment 1. The data of most interest are in Figure 2. Note

first the results from the 800-msec condition. As in the constant condition of Experiment 1, there was a robust modality effect, and as in our previous work, the modality effect was found at all serial positions. In the 1,600-msec condition, the modality effect was more variable across the positions. Finally, in the 2,400-msec condition, there was little or no modality effect. These observations were confirmed by an analysis of variance. There were significant main effects for duration [$F(2,39) = 6.69$, $MS_e = .022$], position [$F(8,312) = 47.78$, $MS_e = .025$], and modality [$F(1,39) = 12.69$, $MS_e = .046$]. Importantly, there was a significant modality × duration interaction [$F(2,39) = 3.28$, $MS_e = .046$]. This interaction indicates that the size of the modality effect decreased with increasing beat durations. The only other significant effect was the duration × position interaction [$F(16,312) = 3.33$, $MS_e = .025$]. This interaction reflects the greater drop across serial position in the shorter beat-duration conditions compared with the long beat-duration condition.

Discussion

These results confirm the prediction that the modality effect will decrease with long beat durations. Apparently, when the beat duration is too long (greater than the duration of echoic memory), the automatic chunking responsible for the rhythm modality effect is reduced or eliminated.

One might argue that the absence of a modality effect in the 2,400-msec condition is simply a ceiling effect. Note, however, that there is a modality effect in both the 800-msec and 1,600-msec conditions at the initial serial positions where recall is over 90% accurate. Thus, it is unlikely that a modality effect in the 2,400-msec condition, in which performance near the end of the sequence hovers between 80% and 85%, is being obscured by ceiling effects. Furthermore, Garner and Gottwald (1968) report a similar interaction between rate and modality using observation time as the dependent variable. Their subjects observed visual patterns for longer than auditory patterns before attempting to reproduce the pattern. How-
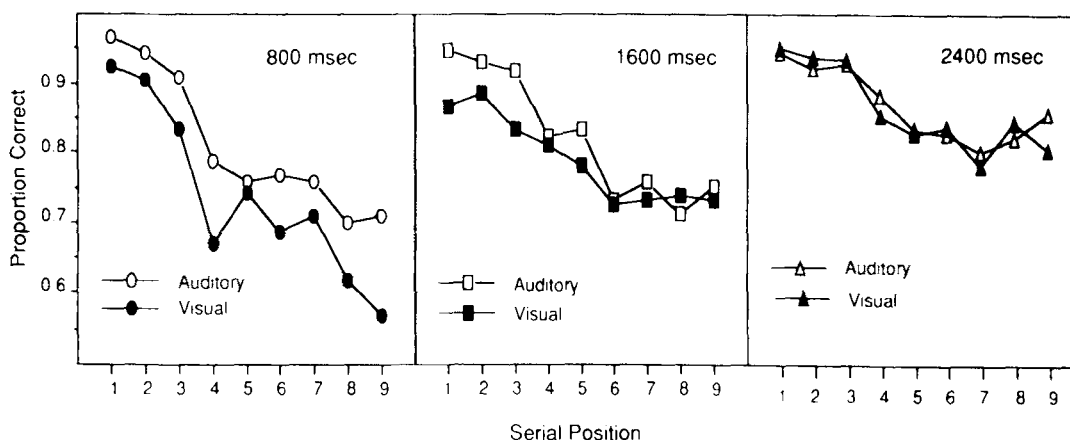

Figure 2. Proportion correct in the rhythm task used in Experiment 2.

ever, the difference in observation time decreased with slower presentation rates.

GENERAL DISCUSSION

The experiments generated three important empirical effects. First, there was a modality effect in the reproduction of rhythms: Auditory rhythms were reproduced more accurately than were visual rhythms. Second, reproduction of auditory sequences was disrupted by the variable ISI more than was reproduction of visual rhythms. Third, the auditory advantage decreased when long beat durations were used. All three of these findings are consistent with our chunking model.

The model can also account for a number of other features in the literature on time judgments and rhythm perception. Allan (1979, pp. 346-347) reviews effects of nontemporal information on time judgments, many of which are consistent with the model. First, filled intervals are judged as longer than unfilled intervals. Within the model, a filled interval (particularly an auditory filled interval) will automatically activate echoic memory, leading to accurate coding of duration. To code the duration of an unfilled interval, a more variable conscious-control process (e.g., humming to oneself) must be used to activate echoic memory. Second, the model accounts for the finding that a filled auditory interval is perceived as longer than a filled visual interval. With a visual interval, variability in the controlled onset of the activation of echoic memory tends to reduce coded duration. That is, the onset of activation may be delayed (reducing coded duration), but activation is unlikely to begin before the stimulus (which might otherwise have increased coded duration). Similarly, variability in attention to the stimulus during its presentation will tend to decrease coded duration. Third, discriminability of filled auditory intervals is greater than discriminability of filled visual intervals: Variability in the encoded duration of visual signals of constant duration should decrease discriminability.

Schab and Crowder (1989) derive the following generalization from their review of the literature on modality and time estimation: "For durations longer than a few seconds, and for responses more precise than absolute category judgments, we suggest that both duration experience and duration judgment are mediated by central processes and are relatively *independent* of peripheral sensory parameters ... For shorter intervals ... a different judgment process may be operating, one that perhaps relies more on sensory features in which duration and temporal patterning ... *are influenced* by the modality of presentation" (pp. 386-387, emphasis added). This generalization is consistent with the chunking model. The model only codes durations shorter than or equal to the maximum duration of echoic memory (perhaps 2 sec). Coding of those short durations should (and does) produce modality differences. Coding of longer durations is outside the purview of the model and hence may not be sensitive to modality.

Finally, the model provides a second reason why Crowder and his colleagues (Crowder & Greene, 1987; Schab & Crowder, 1989) have found little evidence for modality effects in judging durations using the irregular-list technique. In those experiments, subjects were presented temporal markers (e.g., five letters) that were separated by irregular durations. Subjects were asked to judge which of two durations (defined by three successive markers) was the longer. The finding of interest was that modality of the markers made little difference. According to the model, auditory markers will automatically activate duration coding in echoic memory, whereas visual markers require controlled activation. However, because the intervals are empty, the subject must continue to activate echoic memory using a controlled process (e.g., humming or counting) regardless of the marker modality. Thus, according to the model, for the great majority of the duration being encoded, the same process is used for both modalities and so no difference is expected.

The results from Experiments 1 and 2 point to a contribution of temporal (or ordinal) coding processes in rhythm modality effects, whereas other results (e.g., Glenberg & Swanson, 1986) point to a contribution of temporal coding to modality effects in recall tasks. Are these two modality effects produced by the same cognitive processes? At first glance, the answer appears to be no. After all, the modality effect found in the rhythm task extends throughout the list, whereas the modality effect in recall is usually confined to the end of a list. The modality effect in recall is often associated with a large recency effect, but there are only modest recency effects in the rhythm task. Finally, the modality effect in recall is not substantially affected by presentation rate, whereas the modality effect in the rhythm task is (Experiment 2).

On the other hand, the mere fact of auditory superiority in both tasks is unlikely to be coincidental. Furthermore, there is good reason to believe that the auditory superiority in both tasks is related to temporal factors. For example, Glenberg and Swanson (1986) demonstrated that recall of auditory material is more sensitive to temporal manipulations than is recall of visual materials, and the results of Experiment 1 demonstrate that memory for auditory sequences is more affected by manipulations of rhythm (constant versus variable ISI) than is memory for visual sequences. Also, we have unpublished evidence that the two modality effects are produced by the same cognitive processes. Namely, we have found that the size of the rhythm modality effect and the size of the recall modality effect correlate $(r = .66, n = 19, p < .01)$ across subjects. Our measure of the size of the modality effect was performance in the last auditory serial position from which was partialled performance in the last visual serial position. This measure avoids various statistical artifacts (see Chapman & Chapman, 1988), including individual differences in overall performance. Additional information on this technique may be found in Glenberg (1990).

What remains is to characterize the temporal process that is common to the rhythm and recall tasks. Our model

of the rhythm task specifies two levels of coding. First, durations are coded using a ratio scale in echoic memory. The second level of coding uses chunks stored in working memory. Because auditory rhythms are more compactly stored (chunked) in working memory, they can be reproduced more accurately. Note that this advantage reflects, in part, an ordinal coding mechanism. The information in a chunk codes which events followed which others, not the metric properties of the events. Interestingly, recent work on temporal coding in recall-like tasks has also pointed to storage of ordinal information. For example, Tzeng and Cotton (1980) and Winograd and Soloway (1985) have concluded that memory for time of occurrence depends on storing ordinal associations between items on a list. Similarly, using more naturalistic stimuli, Friedman (1987) has developed a hierarchical account of temporal coding. In his account, memory for time of occurrence does not derive from a special time code but from relationships among events, some of which explicitly refer to times (the wash is done on Sundays, the lawn is mowed in the summer). Thus, these different research programs have reached a similar conclusion about temporal coding: memory for time of occurrence depends on coding ordinal relations among events rather than on special temporal codes. Perhaps, then, a more accurate characterization of this work is that it is investigating ordinal memory, rather than temporal memory with metric properties.

Our model of the rhythm task provides an explanation for why temporal (ordinal) coding is superior for the auditory mode: Auditory stimuli automatically activate duration coding in echoic memory, which, in turn, facilitates chunking. Why should temporal (ordinal) coding be enhanced for auditory material in the recall task? A recent formulation by Penney (1989) suggests an answer. Penney proposes that auditory and visual information are processed in separate streams. The auditory stream results in automatic activation of an A code and a P code. Penney characterizes the A code as a rich and relatively durable (perhaps up to 60 sec) sensory code. The P code is a phonological code in working memory. In contrast, the sensory code generated by visual stimuli is ephemeral, and a control process is needed to generate the P code for visual stimuli. Importantly, Penney reviews a large body of data indicating that the A code results in strong associations between successive items. Thus, Penney's auditory stream and the A code provide just the mechanism needed to produce better ordinal coding of auditory than of visual information.

The positive correlation between the size of the modality effect in the rhythm task and the recall task can now be understood at three levels. First, modality effects in both tasks reflect better "temporal" coding for auditory stimuli than for visual stimuli. Second, "temporal" codes are based on ordinal relations in both tasks. Third, these ordinal relations arise from the operation of sensory coding systems. For the recall task, the sensory system may be Penney's A code, and for the rhythm task, the sen-

sory system may be duration coding in echoic memory. Furthermore, these sensory systems may be identical. That is, the chunking model may be taken as a specification of how Penney's A code works in the domain of rhythm coding. Determining exactly how to integrate Penney's A code and the chunking model, and what theoretical benefits may result, is a task left for the future.

REFERENCES

ALLAN, L. G. (1979). The perception of time. *Perception & Psychophysics*, **26**, 340-354.

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.

BLOCK, R. A. (1982). Temporal judgments and contextual change. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 530-544.

CHAPMAN, L. J., & CHAPMAN, J. P. (1988). Artifactual and genuine relationships of laterality scores to accuracy. *Psychological Bulletin*, **104**, 127-136.

CROWDER, R. G., & GREENE, R. L. (1987). On the remembrance of times past: The irregular list technique. *Journal of Experimental Psychology: General*, **116**, 265-278.

CROWDER, R. G., & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics*, **5**, 365-373.

DARWIN, C. J., TURVEY, M. T., & CROWDER, R. T. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*, **3**, 255-267.

DEUTSCH, D. (1986). Recognition of durations embedded in temporal patterns. *Perception & Psychophysics*, **39**, 179-186.

FRIEDMAN, W. J. (1987). A follow-up to "scale effects in memory for the time of events": The earthquake study. *Memory & Cognition*, **15**, 518-520.

GARDINER, J. M. (1983). On recency and echoic memory. *Philosophical Transactions of the Royal Society of London*, **302**, 267-282.

GARNER, W. R., & GOTTWALD, R. L. (1968). The perception and learning of temporal patterns. *Quarterly Journal of Experimental Psychology*, **20**, 97-109.

GLENBERG, A. M. (1990). Common processes underlie enhanced recency effects for auditory and changing-state stimuli. *Memory & Cognition*, **18**, 638-650.

GLENBERG, A. M., MANN, S., ALTMAN, L., FORMAN, T., & PROCISE, S. (1989). Modality effects in the coding of rhythms. *Memory & Cognition*, **17**, 373-383.

GLENBERG, A. M., & FERNANDEZ, A. (1988). Evidence for auditory temporal distinctiveness: Modality effects in order and frequency judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 3-15.

GLENBERG, A. M., & SWANSON, N. C. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 3-15.

GREENE, R. L., & CROWDER, R. G. (1988). Memory for serial position: Effects of spacing, vocalization, and stimulus suffixes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 740-748.

HANDEL, S., & BUFFARDI, L. (1968). Pattern perception: Integrating information presented in two modalities. *Science*, **162**, 1026-1028.

LONGUET-HIGGINS, H.C., & LEE, C. S. (1982). The perception of musical rhythms. *Perception*, **7**, 115-128.

MARTIN, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, **79**, 487-509.

ORNSTEIN, R. E. (1969). *On the experience of time*. Harmondsworth, U.K.: Penguin.

PENNEY, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, **17**, 398-422.

POVEL, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 3-18.

POVEL, D.-J., & ESSENS, P. (1985). Perception of temporal patterns. *Music Perception*, 4, 411-440.

SCHAB, F. R., & CROWDER, R. G. (1989). Accuracy of temporal coding: Auditory-visual comparisons. *Memory & Cognition*, 17, 384-397.

TREISMAN, A. M. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning & Verbal Behavior*, 3, 449-459.

TZENG, O. J. L., & COTTON, B. (1980). A study-phase retrieval model of temporal coding. *Journal of Experimental Psychology: Human Learning & Memory*, 6, 705-716.

WINOGRAD, E., & SOLOWAY, R. M. (1985). Reminding as a basis for temporal judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 11, 262-271.

Forthcoming *Memory & Cognition* Articles

The following is a list of forthcoming *Memory & Cognition* articles that are currently in press. They are given in approximate order of acceptance. Each entry includes the name and address of the author with whom to communicate for further prepublication information.

"Mental rotation of compound stimuli: The effects of task demands, practice, and figural goodness" by L. Paquet (Dept. of Psychology, Carleton Univ., Ottawa, ON, Canada K1S 5B6)

"Phonological assembly in reading: Lexical contribution leads to violation of graphophonological rules" by R. Peereman (Lab. de Psychologie Experimentale Univ. libre de Bruxelles, Avenue Buyl 117, B-1050 Brussels, Belgium)

"Effects of background music on the remembering of filmed events" by M. Boltz, M. Schulkind, & S. Kantra (M.B., Dept. of Psychology, Haverford College, Haverford, PA 19041)

"Spacing judgments as an index of integration from context-induced relational processing: Implications for the free recall of ambiguous prose passages" by L.D. Stern, R.G. Dahlgren, & L.L. Gaffney (L.D.S., Dept. of Psychology, Eastern Washington Univ., Cheney, WA 99004)

"A bias in favor of the positive response to high-frequency words in recognition memory" by Y. Hoshino (Dept. of Psychology, Tokyo Metropolitan Univ., Minamiosawa 1-1, Hachioji, Tokyo 192-03, Japan)

"Can we have a distinctive theory of memory?" by S.R. Schmidt (Dept. of Psychology, Middle Tennessee St. Univ., Murfreesboro, TN 37132)

"Forgetting in recognition memory with and without recollective experience" by J.M. Gardiner & R.I. Java (J.M.G., Memory & Cognition Research Group, Dept. of Social Sciences, City University, Northampton Square, London EC1V 0HB, U.K.)

"Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training" by D.R. Needham & I.M. Begg (I.M.B., Dept. of Psychology, McMaster Univ., Hamilton, ON, Canada L8S 4K1)

"Memory conjunction errors: Miscombination of stored stimulus features can produce illusions of memory" by M.T. Reinitz, W.J. Lammers, & B.P. Cochran (M.T.R., Dept. of Psychology, Boston Univ., 64 Cummington St., Boston, MA 02215)

"Structure and strategy in encoding simplified graphs" by D.J.

Schiano & B. Tversky (D.J.S., MS 262-3, NASA Ames Research Ctr., Moffett Field, CA 94035)

"Music-dependent memory in immediate and delayed word recall" by W.R. Balch, K. Bowman, & L.A. Mohler (W.R.B., 109 Eiche Library, Pennsylvania State Univ., Altoona, PA 16601)

"Mood-congruent and -incongruent learning" by M. Rinck, U. Glowalla, & K. Schneider (M.R., Dept. of Psychology, Justus-Liebig-University Giessen, Otto-Behaghel-Strasse 10/F, D-6300 Giessen, Germany)

"Automaticity and the detection of speech" by J.W. Mullennix, J.R. Sawusch, & L.F. Garrison (J.W.M., Dept. of Psychology, Wayne State Univ., 71 W. Warren St., Detroit, MI 48202)

"Autobiographical fluency: A method for the study of personal memory" by B.H. Dritschel, J.M.G. Williams, A.D. Baddeley, & I. Nimmo-Smith (B.H.D., MRC Applied Psychology Unit, 15 Chaucer Rd., Cambridge CB2 2EF, U.K.)

"Size effects in visual recognition memory are determined by perceived size" by B. Milliken & P. Jolicoeur (B.M., Dept. of Psychology, Univ. of Waterloo, Waterloo, ON, Canada N2L 3G1)

"Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure" by K.E. Stanovich & A.E. Cunningham (K.E.S., Ontario Inst. for Studies in Education, 252 Bloor St. West, Toronto, ON, Canada M5S 1V6)

"Output and retrieval interference in the missing-number task" by J.A. Hadley, A.F. Healy, & B.B. Murdock, Jr. (A.F.H., Dept. of Psychology, Univ. of Colorado, Campus Box 345, Boulder, CO 80309)

"Context availability and the recall of abstract and concrete words" by P.J. Schwanenflugel, C. Akin, & W.M. Luh (P.J.S., Dept. of Educational Psychology, 325 Aderhold Hall, Univ. of Georgia, Athens, GA 30602)

"Repetition priming with Japanese Kana scripts in word-fragment completion" by S. Komatsu & M. Naito (S.K., Faculty of Education, Kagawa Univ., 1-1 Saiwai-cho, Takamatsu, Kagawa 760, Japan)