

Two types of auditory sequence perception

RICHARD M. WARREN and JOHN M. ACKROFF
University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201

It has been suggested recently that there are two fundamentally distinct types of auditory sequence perception in man: (1) holistic pattern recognition (HPR), operating for component item durations from a few milliseconds up to about 200 msec; and (2) direct identification of components and their order (Direct ICO), requiring verbal encoding of names for constituent sounds and requiring item durations roughly 200 msec and above for extended sequences. The present study, using only the very first judgments from 795 untrained participants presented with recycled three-item sequences, provided data consistent with this dichotomous formulation. In addition, it appeared that separate bursts of a noise band generated on-line were treated as different components in HPR and could not be used for sequence matching; "frozen" noise bursts having identical microstructure were treated as the same component and permitted HPR. On-line noise bursts permitted Direct ICO, with naming based on long-term spectral characteristics of noise.

Until a few years ago, the literature on auditory temporal discrimination seemed relatively uncomplicated and tidy. Hirsh (1959) had found that trained subjects could name the order of two items such as a tone and a hiss having onset differences as short as 15 msec, and subsequent investigators reported comparable values for similar tasks (see Fay, 1966, for review). This resolving power seemed adequate to account for our ability to discriminate between permuted orders of sounds in speech and music.

In the late 1960s, the puzzling observation was made that the order could not be named within recycled sequences containing three or four non-related sounds such as hisses, tones, and buzzes each lasting 200 msec (Warren, 1968; Warren, Obusek, Farmer, & Warren, 1969). Naming at this item duration was possible for recycled sequences of four steady state vowels (Warren, 1968), and subsequent work with these vowel sequences indicated that the threshold for reporting order was 125 msec/item (Thomas, Hill, Carroll, & Garcia, 1970). Since the average duration of phonemes in speech is only 70 or 80 msec, how then can we discriminate between permuted orders of speech sounds? It seems that we are not required to identify individual phonemes and their order of occurrence to comprehend speech, but rather to recognize phoneme groupings. Using sequences much less familiar than speech, it

is possible to distinguish between permuted orders of very brief sounds: Warren (1974a) found that after only short training periods with permuted orders of novel sequences of sounds having durations from 400 down to 5 msec, listeners could recognize whether two sequences heard successively (either recycled or single-sequence presentation) had identical or permuted orders and could discriminate which one of two possible orders was heard when sequences were presented alone.

These and other related experiments have suggested that there are two basic types of temporal order discrimination (Warren, 1974a, 1976, Note 1). One of these we may call "direct identification of components and their order" (Direct ICO). As the name indicates, Direct ICO requires that a listener detect and name the constituent sounds of a sequence in their proper order without any prior information concerning their relative positions. It has a lower limit for item durations which, depending upon component sounds and response procedures, varies from about 125 to 600 msec for extended sequences containing several sounds (Warren, 1974b; Warren & Obusek, 1972; Warren et al., 1969). There is no effective upper limit of item duration for ICO. The other type, "holistic pattern recognition" (HPR), involves recognition of the entire pattern rather than individual components in their proper order. It has a lower limit of a few milliseconds per item (which merges with an overall perception of pitch and timbre at shorter item durations) and an upper limit of a few hundred milliseconds (which merges with Direct ICO) (Warren, 1974a, 1976, Note 1).¹ There is evidence that HPR not only is used for discriminating between permuted orders of brief sounds such as hisses, tones, and buzzes in the laboratory, but that it also serves as a basis for perception of speech and music (Warren, 1974a). It appears likely that

The research was supported by grants from National Science Foundation (GB-36986X and BMS73-06787), National Institutes of Health (HD07855), and University of Wisconsin-Milwaukee Graduate School to the first author and by a National Science Foundation Graduate Fellowship to the second author. Requests for reprints should be sent to Richard M. Warren, Department of Psychology, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201.

man's evolutionary development leading to the achievement of speech and music was built upon a preexisting capacity for HPR, since auditory HPR seems to be employed by infrahuman primates, cats, chinchillas, and birds (see Warren, 1976, for discussion). In contrast, Direct ICO appears to be limited to humans, requiring verbal categorization and storage of names, followed by retrieval of the names in the order stored.

The present series of studies attempts to extend the earlier work on HPR and ICO. A comparison was made using the same recycled sequences with instructions corresponding to either HPR or ICO. Using sequences containing noises which were either separate bursts from on-line generation or repeated statements of a single burst, it was possible to test whether HPR with noises required repetition of the same bursts. Throughout these studies, we used the somewhat inconvenient procedure of employing separate groups of subjects for each experimental condition, and restricting each subject to only a single judgment in order to minimize undesired practice effects and transfer of training with other stimuli and tasks.¹ Use of the single judgment procedure has some other advantages: replication of any part of an experiment is simple (not requiring repetition of all the training and prior judgments of the earlier study); performance as measured reflects ordinary preexperimental abilities (not involving special auditory skills established through training).

GENERAL METHOD

Stimuli

Each sequence consisted of three successive sounds repeated over and over in the same order without any pauses between items. Two types of sequences were used: (1) 2,500-Hz tone, 1,000-Hz square wave (low-pass filtered at 4,000 Hz with 32 dB/octave cutoff slope), wide-band noise (500-4,000 Hz with cutoff slopes of 32 dB/octave); (2) 2,000-Hz tone, high-band noise (one-third octave band centered at 5,000 Hz with 32 dB/octave cutoff slopes), and low-band noise (500-1,000 Hz band with 48 dB/octave cutoff slopes). Tones, square wave, and white noise (filtered to yield the experimental noise bands) were produced by electronic signal generators. The filter with 32 dB/octave slopes was General Radio Model 1952, and the filter with 48 dB/octave slopes was Rockland Model 1042 F. Grason-Stadler Series 1200 programming equipment was used for sequencing: component timing modules were calibrated with a quartz crystal time base; electronic switches were adjusted to desired rise/decay times by using a 10,000-Hz calibration tone and a Tektronix Model 564B oscilloscope with single-sweep storage. Rise/decay times were always 5 msec for item durations from 50 to 400 msec, 2.5 msec for 30-msec item durations, 1 msec for 10-msec durations, and 10 μ sec for 5-msec item durations. Listening was through matched TDH-49 headphones fitted with MX41/AR cushions, with all sounds at a level of 80 dB re 20 μ N/m² as measured by a Brüel and Kjaer Model 2204 sound-level meter with a 6-cc earphone coupler. Band-pass filtering from 500 to 5,000 Hz with cutoff slopes of 32 dB/octave was used before delivery to the headphones.

On-line generation of stimuli was used in Experiment 1 for naming of order within a sequence, and in Experiment 2 for

making same/different judgments with pairs of sequences having identical or permuted orders of items. On-line generation of stimuli could not be used in Experiment 3, since component noise bursts were repeated with identical waveforms. The sequences employed and their method of preparation will be described subsequently.

Subjects

A total of 795 subjects, recruited from the introductory psychology sections, were tested singly. Each received either \$1 or credit towards the course grade for participation, depending upon his or her wishes and the individual instructor's rules. All subjects served in only one experimental group and only their first judgments were used. None had ever participated in any other experiment on temporal order.

General Procedure

The subjects were tested individually in a double-walled audiometric room (IAC Model 1204A) in which both the subject and the experimenter were seated at a table. All subjects read the appropriate instructions (for either ICO or HPR judgments). Any questions concerning the tasks were answered by rephrasing the instructions. The subjects were allowed to listen for as long as they wished, and the sequence was terminated after they responded. The first sound heard in a sequence was always randomly determined. Each subject serving in a Direct ICO group attempted to name the order of the three components of a single recycled sequence by arranging three cards with appropriate names in the proper order. Earlier experiments had demonstrated a considerably lower threshold for card-ordering than for vocal reporting order (Warren & Obusek, 1972). Each subject participating in a HPR group switched back and forth between two three-item recycled sequences as often as they wished before deciding whether they were the same or different.

EXPERIMENT 1: DIRECT IDENTIFICATION OF COMPONENTS AND THEIR ORDER (DIRECT ICO)

Procedure

Subjects (six groups of 30 subjects) were given typewritten instructions informing them that they were to arrange three cards representing the three sounds of a repeating sequence in the order in which the sounds occurred, starting with whichever sound they chose. They were told that before listening to the recycled sequence, they could hear samples of each of the three sounds by pressing appropriately labeled buttons. When they believed they were familiar with the names of the sounds, they were to signal the experimenter who would then present the recycled stimulus. The subjects then listened to the stimulus as long as they wished and informed the experimenter when their arrangement of the cards was completed; this arrangement was then recorded and the sequence terminated. Each of the six groups received either: (1) the wide-band noise, tone, and square wave; or (2) the high-band noise, low-band noise, and tone, presented at one of the three following item durations: 100, 200, or 400 msec. There were two possible orders of items in each recycled three-item sequence, and each of these orders was heard by half the 30 subjects in a group. Only the very first judgment of each subject was used.

Results

The number of correct responses (total responses = 30) for each group is shown in Figure 1. The most probable chance score was 15 correct. A total of 20 correct responses was significantly better than chance at $p < .05$ (binomial expansion table),

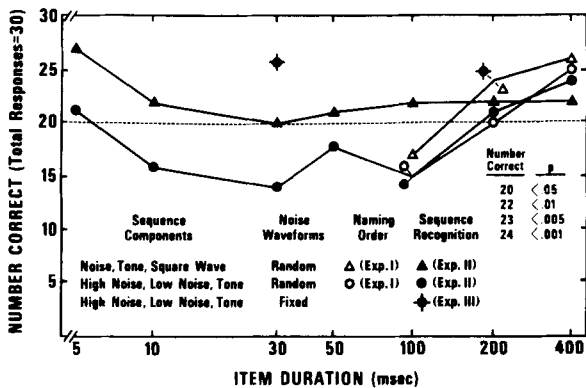


Figure 1. Accuracy of naming order of components and of recognizing whether two sequences had the same or permuted orders using recycled sequences. On-line generation of random noise bursts was used for naming in Experiment 1 and for recognition in Experiment 2. "Frozen" or fixed waveform noise bursts were used for recognition in Experiment 3, and the data shown were derived from Table 1 (see text for further description). The dashed horizontal line corresponds to the limit of performance significantly better than chance ($p < .05$).

and is represented by the horizontal dashed line in Figure 1. A decrease in Direct ICO (naming) accuracy occurred with a decrease in item durations: for the longest item duration (400 msec), each group had scores well above the chance level ($p < .001$) by binomial expansion; for the shortest item duration (100 msec), scores were not significantly above chance for either group; and for the intermediate duration of 200 msec, correct scores were significantly higher than chance for each group ($p < .001$ for sequences containing wide-band noise, tone, square wave; $p < .05$ for sequences containing high-band noise, low-band noise, and tone).

EXPERIMENT 2: HOLISTIC PATTERN RECOGNITION (HPR) INVOLVING ON-LINE NOISE BURSTS

Procedure

The 420 subjects (14 groups of 30 subjects) were given type-written instructions informing them that they would be listening to two sound channels, each consisting of a sequence of brief sounds repeated over and over without pause. All subjects were told that the arrangement of sounds in the two channels might be the same or different, and that they did not have to report the order of the sounds, or even what the component sounds were. For item durations from 100 through 400 msec, the subjects were told to report whether the two channels had their items in the same or different orders; for item durations from 5 through 50 msec (for which the component sounds were not as readily discriminable), they were told to report whether the two channels sounded exactly the same or different in any way. All subjects were told that they could listen to whichever channel they wished by using a control switch, and could change from one to the other as often as desired before they responded by calling out "same" or "different." Half received the pair with identical orders, and half the pair with permuted orders. Since any discriminable difference between the two channels could lead to the response of "different," considerable care was taken to ensure

that the two channels were matched in intensity and spectral response.

For sequences containing two noise bands, the "same-order" pair always had the succession: tone, high-band noise, low-band noise, tone, etc. For sequences containing one noise band, for the "same-order" pair, the succession was always: tone, square wave, broad-band noise, tone, etc. In order to avoid the possibility that a fixed alignment of items across the two sequences could serve as a cue to order, an "item-hold" procedure was used in which sequencing on one of the channels always stopped when the subject switched to the other (free-running) channel, and then recommenced when the subject switched back to the "item-hold" channel.

Results and Discussion

Figure 1 shows the number of the 30 same/different responses which were correct for each group. As in Experiment 1, the most probable chance score was 15 correct, and the statistical treatment of the data was the same as that used in the earlier experiment on ICO.

It can be seen in Figure 1 that same/different judgments in Experiment 2 using sequences containing tone, square wave, and noise band were more accurate than corresponding judgments involving tone, high noise band, and low noise band for all but the longest (400-msec) item durations. With sequences consisting of one noise band, judgments were at levels significantly above chance for every item duration, while with sequences consisting of two noise bands, judgments were no better than chance for over half of the item durations (10, 30, 50, and 100 msec). It should be noted that Experiment 1 had shown that for items lasting 200 or 400 msec it was possible to name the order of sounds with accuracy significantly above chance level with both sets of sequence components. The ability to name order would permit accurate same/different judgments without HPR for the two longest item-duration sequences in Experiment 2, and the similarity in performance in Experiments 1 and 2 is consistent with a common basis at these durations. Also, some subjects spontaneously named the order for each of the sequences while making their same/different judgments. For the shorter durations requiring HPR for same/different judgments, the consistently poorer performance with the two noise bands suggested a quite interesting possibility concerning the nature of HPR which led to Experiment 3.

Successive bursts of noise derived from the same generator and spectrum shapers are conventionally treated as samples of the same sound in experiments on sequence perception (Broadbent & Ladefoged, 1959; Fay, 1966; Hirsh, 1959; Neisser & Hirst, 1974; Warren et al., 1969). However, the microstructure and short-term running spectra are quite different for successive bursts of on-line noise, and it is possible that each burst is treated perceptually as a separate sound rather than an iteration of the same sound.

In contrast, bursts of a tone and a square wave maintain equivalent spectra and waveforms on repetition. Now let us consider our three-item sequences of Experiment 2: with one noise band, square wave, and tone, two of the three items are identical on repetition, and the possible orders of the sequences can be differentiated by whether the tone follows or precedes the square wave (the position of the noise is fixed by the other two components); with two noise bands and a tone, only the tone remains the same on repetition, and since the two noises vary in a random fashion each time they are presented, different orders of components might not be discriminable.

However, it is possible to produce sequences which restate all components (including those derived from noise) in acoustically identical form, and to arrange these components so that their orders are either identical or different. Experiment 3 employs such pairs of sequences, along with a variety of others derived from "frozen" noise bursts.

EXPERIMENT 3: HOLISTIC PATTERN RECOGNITION (HPR) INVOLVING REPETITION OF IDENTICAL ("FROZEN") WAVEFORMS

In Experiment 2, HPR involved on-line generation of the repeated sequences. Periodic gating of the output of signal generators produced acoustically matched restatements for the successive bursts of the tone and the square wave, but not for the filtered noises. In Experiment 3, three-item sequences consisting of the same sounds used in Experiment 2 were produced using a tape loop and/or a digital delay line to insure that each of the components within a sequence maintained an identical waveform for each restatement and to control waveform relations of components across the sequences to be compared.

Procedure

The component sounds and their orders were the same as those described in Experiment 2, as were the instructions to subjects. The difference from Experiment 2 was in the use of "frozen" waveforms rather than on-line generation of items for HPR, and the limitation of item durations to either 30 or 200 msec. A total of 13 groups of 15 subjects were used, and each subject made only one same or different judgment with a single sequence pair. The preparation of the "frozen" sequences was fairly complicated, and is described in the Appendix.

The stimuli with 30-msec item durations having "frozen" waveforms were too brief to provide tape loops of sufficient length. However, the digital delay line described in the Appendix could be used to repeat single statements of the entire 90-msec sequences by operating in a "recycle" mode with storage of 45 kbits.

As in Experiment 2, sequences were always presented in pairs containing the sounds in either same or permuted orders, and each pair had both sequences consisting of either: (1) tone, square wave, and broad-band noise; or (2) tone, high-band noise, and low-band noise. While each sequence within a pair had a "frozen" waveform repeated by the digital delay line, a choice

was available for the waveform relation for noises across sequence pairs. In some cases, each sequence of a pair was generated separately, and hence each had independent "frozen" waveforms for noise bursts. Other sequence pairs had identical waveforms for each component across sequences (matched waveforms were prepared both for same and for permuted item orders). All of the stimuli and testing conditions employed are listed in Table 1. As can be seen in the table, additional groups of 15 subjects each were used for some conditions to determine if results could be replicated.

Results and Discussion

Experiment 3 demonstrates that substitution of "frozen" noise bursts for the on-line bursts used in Experiment 2 can improve dramatically the accuracy of some same/different judgments.² The numbers of "same" and "different" judgments for each group of 15 subjects are shown in Table 1. It was not always permissible to combine judgments obtained with identical pairs and permuted-order pairs as was done for the on-line sequences of Experiment 2, since when each "frozen" noise burst was used for only one sequence of a pair (see Table 1), differences between bursts were confounded with order. However, when the same "frozen" noise bursts were used across sequences, order was the only variable distinguishing the paired sequences, and pooling of responses for same item-order and different item-order was possible. Pooling balanced for possible response bias in favor of "same" or "different" judgments. Two sets of pooled values were obtained from the data in Table 1, and each set appears in Figure 1 along with the corresponding data from the on-line conditions of Experiment 2.

As shown in Figure 1, in Experiment 2 only 14 out of 30 responses were correct for sequences containing tone and two noise bursts at 30-msec item durations, but a dramatic improvement occurred for comparable sequences with "frozen" waveforms in Experiment 3: 26 out of 30 responses were correct for corresponding conditions (Table 1 shows that scores before pooling for the original same item order, single waveform group and different item order, single waveform group were both 13 out of 15 correct). The difference between 14 and 26 correct in Experiments 2 and 3, respectively, is significant at $p < .005$ according to a 2 by 2 contingency table. For 200-msec items, it can be seen in Figure 1 that same/different judgments with the two noise bands and the tone were slightly (nonsignificantly) more accurate for "frozen" waveform sequences than for the on-line sequences of Experiment 2 (scores before pooling for the 200-msec groups with "frozen" waveforms are shown in Table 1). Turning to the sequences containing one noise band along with square wave and tone, same/different judgments were significantly more accurate than chance for on-line generation at 30 msec per item in Experiment 2, as shown in Figure 1, and Table 1 indicates that performance was

Table 1
Holistic Pattern Recognition for Pairs of Three-Item Recycled Sequences Containing "Frozen" Noise Bursts

Cross-Sequence Relation for Each Noise Burst	30-msec Items		200-msec Items
	Two Noise Bands ^a	One Noise Band ^b	Two Noise Bands ^a
	Judgments		
	Same/Different	Same/Different	Same/Different
Same Item Order, Single Waveform	13/2** 13/2r***	13/2**	13/2**
Same Item Order, Two Waveforms	3/12*	3/12*	
Permuted Item Order, Single Waveform	2/13** 1/14r***		4/11
Permuted Item Order, Two Waveforms	0/15*** 0/15r***	0/15*** 1/14r***	

^aSequences consisting of high-band noise, low-band noise, tone

^bSequences consisting of wide-band noise, square wave, tone

^rReplication using separate stimuli and different subjects

* $p < .05$; ** $p < .005$; *** $p < .001$ for significance of differences from chance level

very good when "frozen" waveforms of this duration were used for matching. The same "frozen" waveform was not used for both sequences of the different item-order pair (one noise-band, 30-msec item durations), so that each sequence of this pair differed in both item order and waveform, preventing a direct comparison with results in Experiment 2. However, comparable groups were used within Experiment 3 for the 30-msec item-duration sequences with one and with two noise bands, and it can be seen in Table 1 that equivalent high accuracy was attained with both types of sequences.

An examination of Table 1 shows that for the brief (30-msec) items, "same" judgments were made by listeners only when both item-order and waveform of the "frozen" noise bursts were matched across sequences. When item order differed with identical cross-sequence waveforms for each item (two noise-band pair), subjects judged the sequences to be different, as they did if the item order was the same but the waveform (and short-term spectrum) of the items varied across sequences (both two noise-band pair and one noise-band pair). Second groups of 15 subjects were used for four of the more interesting experimental conditions, and Table 1 shows that scores matched well for these attempts at replication.

SUMMARY AND GENERAL DISCUSSION

The observations made with unpracticed subjects in our experiments with three-item recycled sequences can be summarized as follows: (1) with

on-line sequences having item durations of 200 msec and above, performance for naming order and for matching orders was at levels significantly above chance for all sequences employed; (2) matching of orders of on-line sequences for the shortest item duration (5 msec) was easier than for durations from 10 through 200 msec; (3) for durations between 10 and 100 msec, matching orders of on-line sequences through holistic pattern recognition was significantly above chance for square wave, tone, and wide-band noise, but was not significantly above chance for high-band noise, low-band noise, and tone; (4) holistic pattern recognition for sequences with item durations from 10 through 100 msec required that at least two of the three components have fixed (either regular or "frozen") waveforms.

Let us consider each of these findings and their implications:

(1) Naming of order of items within sequences was significantly better than chance at 200 and 400 msec/item. This accurate identification of components and their order (Direct ICO) at long item durations in Experiment 1 could explain how the separate listeners used in Experiment 2 could tell whether the same slow sequences were in the same or permuted order without using holistic pattern recognition (HPR). There was supporting evidence that subjects did indeed use an ICO procedure for same/different judgments at long item durations: (a) scores were quite similar at 200 and 400 msec/item for Experiment 1 (naming order) and Experiment 2 (same/different judgments) as shown in Figure 1; (b) some subjects volunteered the information that they used identification of component orders for matching; and (c) experienced listeners (not used in the experiment) reported that it was extremely difficult (if not impossible) to ignore the obvious order of components in making same/different judgments.

(2) At the shortest item duration of 5 msec, entire 15-msec sequences were repeated almost 70 times/sec. This repetition rate is within the pitch range, and differences in timbre associated with the stimulus envelope could have been responsible for the relatively high accuracy of same/different judgments (see Green, 1971; Warren, 1974a).

(3) Interesting implications follow from the observations that same/different judgments for item durations from 10 through 100 msec were significantly above chance for all groups when sequence pairs with one noise were used, and that no groups performed at levels significantly above chance when the corresponding sequences containing two noises were used. First we have the observation that our untrained undergraduate students could make accurate HPR judgments listening to tone, square wave, and noise—an earlier study reporting successful HPR with similar sequences had employed graduate

students and staff who received special training with the sequences before being tested (Warren, 1974a). Since HPR was so accurate with sequences containing one on-line noise, why was it so much more difficult with sequences containing two on-line noises? This difference between performance with the two types of sequences would not be anticipated from the literature, and suggested the interpretation and experimental work described below.

(4) The unexpected differences in HPR for sequences with one and with two noise bands in Experiment 2 led to the hypothesis concerning sequence recognition tested in Experiment 3. Since recycled sequences containing two spectrally distinct noise bursts and a tone were generated on-line in Experiment 2, each of the two noise bursts had a different, randomly determined composition at every reappearance, while the tone burst contributed the same waveform for each repetition. It has been conventional in psychoacoustics to use a single term for noise bursts produced by a single generator and filtered in the same fashion, but such noise patterns are no more equivalent physically than are successive kaleidoscopic patterns seen through the same viewer. Thus, for the sequences containing high-band noise, low-band noise, and tone, there was no acoustic repetition for two of the three sounds, and so there was no pattern invariance within a sequence which could be used to judge equivalence for members of a sequence pair. The situation was rather different for the on-line generation of sequences containing only one noise band (wide-band noise, tone, and square wave). While the wide-band noise did differ randomly with each on-line capture, the tone and square waves always had the same regularly repeated waveforms. Even if the protean structure of the wide-band noise provided no help in achieving HPR within or across sequences, pairs with permuted orders had tone preceding square wave in one case and square wave preceding tone in the other. HPR judgments of same/different could be made accurately if based only on the order of tone and square wave.

There has been considerable evidence published that listeners can recognize a repeated noise burst when heard alone (rather than as part of a sequence containing other sounds as in our study). Guttman and Julesz (1963) observed that listeners could detect repetition of a recycled noise burst lasting up to 1 or 2 sec. Subsequent work has not only confirmed that recycling of segments derived from noise can be detected, but also that individual recycled noise bursts or noise-like patterns with differing short-term spectra and similar long-term spectra can be discriminated from each other and recognized (Pfafflin & Mathews, 1966; Pollack, 1972, 1975; Schubert & West, 1969). These observations would lead us to expect that repeating of short-term spectra

for the noise components used in Experiment 2 would improve discriminability. In Experiment 3, we used a digital delay line both to rearrange component items and to recycle 90-msec sequences containing 30-msec components with "frozen" waveforms (see Appendix). The discriminability of sequences improved greatly. While the fixed waveforms guaranteed the maintenance of fixed short-term spectra, we did not change waveforms while keeping short-term spectra constant, so that Experiment 3 cannot decide definitely whether the spectral constancy of components was responsible for the improved HPR. Chances are it was.²

Acoustically, the stimuli used in the present studies may be considered as a series of discrete sounds, but perceptually they appear to be treated as successions of separate sounds only at item durations of 200 msec and above. At shorter item durations, complex acoustic sequences seem to be recognized and differentiated as overall or holistic patterns. These statements concerning HPR seem to apply not only to sequences of unrelated sounds, but to speech and music as well (Warren, 1974a).

In this study, HPR was demonstrated for sequences with fixed envelopes. However, HPR can be somewhat flexible. Acoustic patterns of speech can remain intelligible at different speaking rates and over a range of internally consistent shifts in frequency and relative intensity of spectral components. Even chinchillas trained with speech can maintain learned discrimination between plosives despite changes in following vowels and in speakers (Kuhl & Miller, 1975). It appears possible that with practice, HPR could be accomplished by human subjects with sequences of arbitrarily selected sounds differing over some limited range of acoustic characteristics. There already has been a study measuring the latitude in component durations permitting HPR by untrained listeners with sequences of arbitrarily selected sounds (Warren, 1974b), and it also has been reported that training can extend the range of item durations permitting recognition of individual permuted orders (Warren, 1974a).

APPENDIX

Preparation of Sequences With "Frozen" Waveforms for Experiment 3

For the 200-msec items, sequence pairs contained high-band noise, low-band noise, and tone. The members of the "different" sequence pair were recorded simultaneously on a two-track Ampex 440B recorder at 15 ips. Each track had a repeated single statement of the 600-msec sequence, with the same waveforms but different orders on the two tracks. The interchanging of orders of the "frozen" noises was accomplished with the aid of Grason-Stadler Series 1200 programming equipment and a dual output Eventide Model 1745 digital delay line equipped

with built-in noise reduction system and anti-aliasing filters, and operating at a 50-kHz sampling rate with 10 bits per sample. The 9-in. tape loop was played back on a specially modified stereo Ampex AG 500 recorder, and both channels rerecorded on a four-track Ampex 440B recorder (which was also used for playback of the stimuli to the subjects). The preparation of the "same" stimulus pair was similar except that each track of the stimulus tape had the identical waveforms for each item in the same order. Pairs containing 30-msec items with the "same" order were relatively easy to produce. For same order and identical waveforms for both sequences of a pair, a single statement of the three successive items was recycled on the digital delay line with one channel offset by approximately half the total sequence duration (44 msec). Both channels were recorded at the same time on a four-track Ampex AG 440B recorder at 15 ips. To produce the same-order pairs with independent frozen waveforms for the corresponding noise bursts in each sequence, successive recordings were made on separate tracks. By changing the sequence programming when producing a second recycled sequence, different-order pairs with separate "frozen" waveforms were synthesized. All of the 30-msec item sequence pairs consisting of separate captures and independent "frozen" waveforms had a randomly determined cross-sequence item alignment.

A more elaborate procedure was required to produce identical waveforms in different orders for 30-msec-item high-band noise, low-band noise, tone recycled sequences. Both sequences of the pair were generated from a single statement of a three-item sequence recycled by the delay line and recorded on one track of an eight-channel Ampex Model 440B recorder. The component sounds were isolated by playback of the intact sequence, filtering off two of the sounds and rerecording the remaining sound on a separate track of the eight-track recorder. The 2,000-Hz tone was passed or rejected as needed by Brüel and Kjaer Model 2120 frequency analyzer set at 2,000 Hz as either a 1% band-pass filter or a band-stop filter (with a rejection of 60 dB one octave from the center frequency). A two-channel Rockland Model 1042F filter in ganged high-pass or low-pass mode (slopes of 96 dB/octave) was used to separate the component noises from each other. Each of the three separated sounds was then played back and recorded again on separate tracks of the eight-track recorder after passage through the delay line to shift the relative positions of the sounds as required. The individual components were then mixed, and the completed sequence rerecorded. In addition to this permuted item-order, single waveform stimulus pair, the same three items were used as components of an identical item-order, single waveform pair of sequences.

Considerable care was taken to ensure that the processing of stimuli did not introduce artifactual differences between members of a stimulus pair (which could cause extra judgments of "different"). The same number and types of processing steps were always used for both sequences of a pair. Matching of component waveforms was checked using oscillograph tracings, and spectral composition was checked using a spectrum analyzer (Rockland Model FFT 512/S). The high proportion of "same" judgments under appropriate conditions (see Table 1) indicates that artifactual perceptual differences were minimal.

All replications shown in Table 1 involved the same stimuli used with the original group, except for the permuted item-order, single waveform pair (the most complex preparation), which employed completely new sequences prepared as described above.

REFERENCE NOTE

1. Warren, R. M. Temporal resolution of auditory events. In D. A. Norman (Chair) *Perception of temporal order in hearing: Old pattern recognition problems in a new guise*. Symposium presented at the meeting of the American Psychological Association, Honolulu, September 1972.

REFERENCES

- BROADBENT, D. E., & LADEFOGED, P. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 1959, **31**, 1539.
- FAY, W. H. *Temporal sequence in the perception of speech*. The Hague: Mouton, 1966.
- GREEN, D. M. Temporal auditory acuity. *Psychological Review*, 1971, **78**, 540-551.
- GUTTMAN, N., & JULESZ, B. Lower limits of periodicity analysis. *Journal of the Acoustical Society of America*, 1963, **35**, 610.
- HIRSH, I. J. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 1959, **31**, 759-767.
- KUHL, P. K., & MILLER, J. D. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, 1975, **190**, 69-72.
- NEISSER, U., & HIRST, W. Effect of practice on the identification of auditory sequences. *Perception & Psychophysics*, 1974, **15**, 391-398.
- PFÄFFLIN, S. M., & MATHEWS, M. V. Detection of auditory signals in reproducible noise. *Journal of the Acoustical Society of America*, 1966, **39**, 340-345.
- PLOMP, R. Timbre as a multidimensional attribute of complex tones. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970.
- POLLACK, I. Memory for auditory waveform. *Journal of the Acoustical Society of America*, 1972, **52**, 1209-1215.
- POLLACK, I. Identification of random waveforms. *Journal of the Acoustical Society of America*, 1975, **58**, 1262-1271.
- SCHUBERT, E. D., & WEST, R. A. Recognition of repeated patterns: A study of short-term auditory storage. *Journal of the Acoustical Society of America*, 1969, **46**, 1493-1501.
- THOMAS, I. B., HILL, P. B., CARROLL, F. S., & GARCIA, B. Temporal order in the perception of vowels. *Journal of the Acoustical Society of America*, 1970, **48**, 1010-1013.
- WARREN, R. M. Relation of verbal transformations to other perceptual phenomena. In *Conference Publication No. 42, IEE/NPL Conference on Pattern Recognition*. Teddington, England: Institution of Electrical Engineers, 1968, (Suppl., 8 pp.).
- WARREN, R. M. Auditory temporal discrimination by trained listeners. *Cognitive Psychology*, 1974, **6**, 237-256. (a)
- WARREN, R. M. Auditory pattern recognition by untrained listeners. *Perception & Psychophysics*, 1974, **15**, 495-500. (b)
- WARREN, R. M. Auditory perception and speech evolution. *Annals of the New York Academy of Sciences*, 1976, in press.
- WARREN, R. M., & OBUSEK, C. J. Identification of temporal order within auditory sequences. *Perception & Psychophysics*, 1972, **12**, 86-90.
- WARREN, R. M., OBUSEK, C. J., FARMER, R. M., & WARREN, R. P. Auditory sequence: Confusion of patterns other than speech and music. *Science*, 1969, **164**, 586-587.
- WIGHTMAN, F. L., & GREEN, D. M. The perception of pitch. *American Scientist*, 1974, **62**, 208-215.

NOTES

1. In earlier experiments, it was found necessary to avoid providing experience with sequences differing slightly in the duration of their items: the ability to list items and their order (ICO) could be transferred in a series of short steps from longer to much briefer item durations through recognition of similarity in quality (HPR), without listeners having been able to perceive directly the individual components or their orders at the shorter durations (Warren, 1974a, b).

2. In order to determine the effect of changing the waveform while keeping the intensities of spectral components fixed, we passed a recycled "frozen" sequence with 30-msec items (tone,

high-band noise, low-band noise) through a phase shifter. Four experienced listeners switched back and forth between the unshifted sequence, and the sequence shifted through a variety of phase angles. Virtually no differences could be perceived by any of the listeners, despite great differences in waveforms as viewed on an oscilloscope. Thus, while changes in phase occasionally produce detectable differences in complex periodic sounds (see Plomp, 1970; Wightman & Green, 1974), phase changes are not readily detectable under our experimental conditions.

(Received for publications February 16, 1976;
revision accepted August 7, 1976.)