

Learning and Generalization of Auditory Temporal–Interval Discrimination in Humans

Beverly A. Wright, Dean V. Buonomano, Henry W. Mahncke, and Michael M. Merzenich

Keck Center for Integrative Neuroscience, University of California, San Francisco, San Francisco, California 94143-0732

The sensory encoding of the duration, interval, and order of different stimulus features provides vital information to the nervous system. The present study focuses on the influence of practice on auditory temporal–interval discrimination. The goals of the experiment were to determine (1) whether practice improved the ability to discriminate a standard interval of 100 msec bounded by brief 1 kHz tones from longer intervals, and, if so, (2) whether this improvement generalized to different tonal frequencies or temporal intervals. Learning was examined in 14 human subjects using an adaptive, two-alternative, forced-choice procedure. One hour of training per day for 10 d led to marked improvements in the ability to discriminate between the standard and longer intervals. The generalization of learning was evaluated by independently varying the spectral (tonal frequency) and temporal (interval) components of the stimuli in

four conditions tested both before and after the training phase. Remarkably, there was complete generalization to the trained interval of 100 msec bounded by tones at the untrained frequency of 4 kHz, but no generalization to the untrained intervals of 50, 200, or 500 msec bounded by tones at the trained frequency of 1 kHz. Thus, these data show that (1) temporal–interval discrimination using a 100-msec standard undergoes perceptual learning, and (2) the neural mechanisms underlying this learning are temporally, but not spectrally, specific. These results are compared with those from previous investigations of learning in visual spatial tasks, and are discussed in relation to biologically plausible models of temporal processing.

Key words: psychophysics; hearing; perceptual learning; auditory processing; temporal processing; human

The sensory encoding of temporal information such as the duration, interval, and order of different stimulus features provides vital information to the nervous system. This is clearly illustrated by the ever-increasing evidence of the importance of temporal cues in the perception of speech. The identification of individual consonant–vowel syllables correlates with the interval between air release and vocal cord vibration (e.g., “ba” versus “pa”; Lisker and Abramson, 1964), the duration of frequency transitions (e.g., “ba” versus “wa”; Liberman et al., 1956), and the silent time between consonants and vowels (e.g., “sa” versus “sta”; Dorman et al., 1979). Furthermore, prosodic cues such as pauses, the duration of speech segments, and speaking rate influence semantic content (Lehiste et al., 1976). Indeed, speech can still be understood even when the available cues are primarily temporal (Shannon et al., 1995), but not when the temporal cues are removed either by manipulations of the speech signal (Drullman et al., 1994a,b) or by impairments of temporal processing in the perceiver (Tallal and Piercy, 1973).

Most of the temporal cues involved in speech perception fall on a time scale of tens to hundreds of milliseconds. Relatively little is known about the neural mechanisms underlying temporal processing on this scale (for review, see Ivry, 1997). The processes used in these temporal tasks, however, must differ from those used

in spatial tasks. In a purely spatial task, all the relevant information is coded in the spatial pattern of active input fibers. For example, in a visual vernier acuity task, the stimuli to be discriminated are slightly displaced in space and, as a result, activate different spatial groups of photoreceptors in the retina. Similarly, in an auditory frequency discrimination task, the stimuli to be discriminated activate different spatial groups of hair cells on the basilar membrane. Performance in these tasks could therefore depend on detecting the activation of spatially distinct populations of peripheral neurons. In contrast, in a purely temporal task, all the relevant information is coded in the temporal pattern of the active input fibers. For example, in a temporal–interval discrimination task each temporal interval is often marked by a pair of stimuli that activate the same input fibers. Thus, performance on this task cannot depend directly on the detection of spatially distinct populations of peripheral neurons. The relevant information instead must be extracted from the temporal activity of the input fibers and coded in a form that can be used to solve the task.

The present study focuses on the influence of practice on the human discrimination of auditory temporal intervals in the millisecond time range. We undertook this project because very little is known about the learning of temporal tasks and because the psychophysical features of learning have the potential to provide insight into the neural encoding of temporal information. The specific goals of this experiment were to determine (1) whether practice improved the ability to discriminate between temporal intervals marked by tones of one frequency, and, if so, (2) whether this improvement generalized to intervals marked by tones of a different frequency or to intervals of different duration. The results show that the ability to discriminate longer temporal intervals from a 100 msec standard improves with practice and

Received Oct. 21, 1996; revised Feb. 19, 1997; accepted March 6, 1997.

This work was supported by the McDonnell-Pew Program in Cognitive Neuroscience, National Institutes of Health/NIDCD (R29-DC02997), the Office of Naval Research (N00014-96-1-0206), the Howard Hughes Medical Institute, and the Charles A. Dana Foundation. The comments of two anonymous reviewers helped to improve this paper. We thank Marshall Fong for technical assistance.

Correspondence should be addressed to Dr. Beverly A. Wright, Audiology and Hearing Sciences Program, Northwestern University, 2299 North Campus Drive, Evanston, IL 60208-3550.

Copyright © 1997 Society for Neuroscience 0270-6474/97/173956-08\$05.00/0

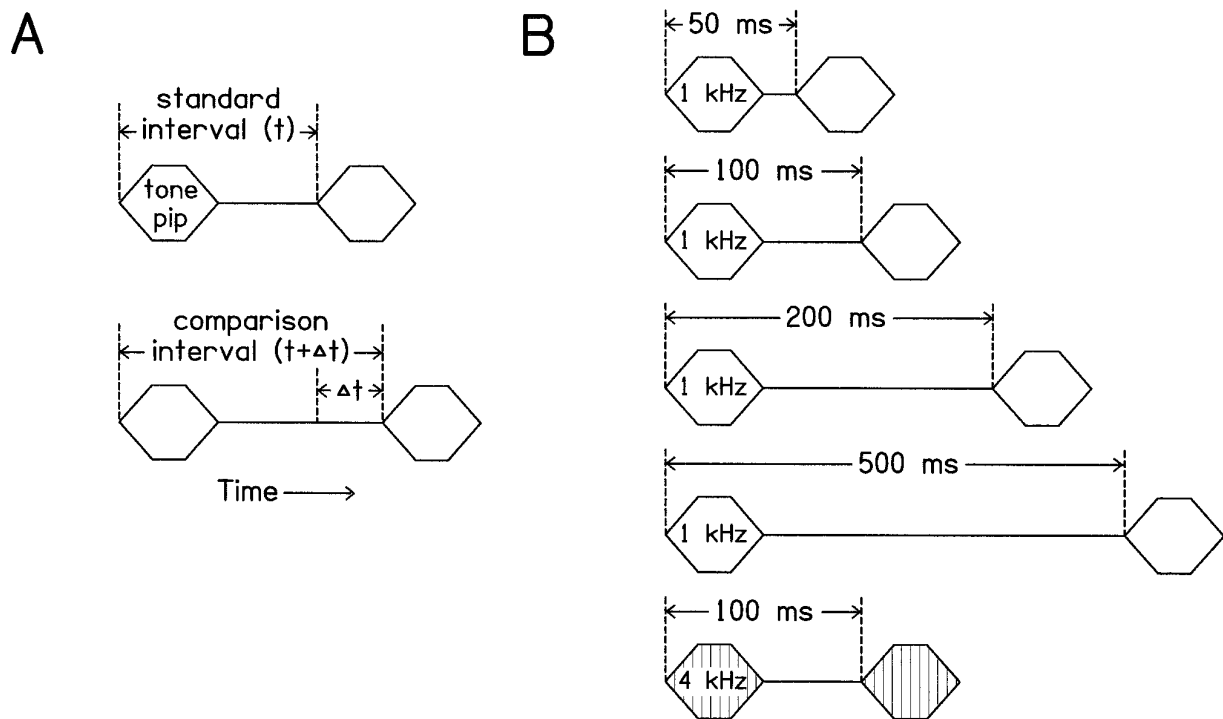


Figure 1. *A*, Schematic diagrams of the stimuli presented in the standard (*top*) and comparison (*bottom*) intervals in each two-interval forced-choice trial. *B*, Schematic diagrams of the standard intervals in the five conditions (*top* to *bottom*): 50, 100, 200, and 500 msec at 1 kHz and 100 msec at 4 kHz.

that this improvement generalizes to an untrained frequency, but not to untrained standard intervals.

MATERIALS AND METHODS

Subjects. Fourteen normal-hearing volunteers (10 females) between the ages of 21 and 46 years served as subjects. Subjects S4 and S3 were the second and third authors. The remaining subjects were paid for their participation. No subject had any previous experience in psychoacoustic or interval-discrimination tasks.

Stimuli. The task is illustrated schematically in Figure 1*A*. On each trial, two tone pips were presented in each of two observation periods. In one observation period, the tone pips were separated by a fixed temporal interval, t , referred to as the standard interval. In the other observation period, termed the comparison interval, the tone pips were separated by a temporal interval equal to t plus a variable, always positive, Δt . The comparison interval was randomly presented in either the first or the second observation period. The subject was asked to indicate which of the two observation periods contained the comparison (longer) interval by pressing a key on a computer keyboard. The response was not timed. The comparison interval was adjusted adaptively across trials (see Procedure) to determine the threshold Δt needed to discriminate the comparison from the standard interval.

Figure 1*B* shows schematic illustrations of the five listening conditions. The tone pips were presented at 1 kHz at standard intervals of 50, 100, 200, and 500 msec and at 4 kHz at a standard interval of 100 msec.

Each tone pip had a total duration of 15 msec, including 5 msec rise/fall ramps, and was always presented in zero phase. The temporal interval between the two pips was measured from the onset of the first pip to the onset of the second pip. The level of each pip was 86 dB sound pressure level. The onsets of the first tone pips in the first and second observation intervals were separated by 900 msec for standard intervals shorter than 500 msec and by 1250 msec for the standard interval of 500 msec. Each observation period was marked on a computer screen by a visual display. Visual feedback was provided after each trial.

The tone pips were digitally generated in the frequency domain using a digital-signal-processing board (TDT AP2) and a sampling period of 40 μ sec (25 kHz) for standard intervals shorter than 200 msec or 60 μ sec (16.7 kHz) for the standard interval of 500 msec. The tone pips were delivered to a 16-bit digital-to-analog converter (TDT DD1), followed by an anti-aliasing filter (8.5-kHz low-pass, TDT FT5) and an attenuator

(TDT PA4). They were presented through the left ear piece of Sennheiser HD265 headphones in circumaural cushions.

Procedure. The experiment consisted of a pretest, a training phase, and a post-test. In the pre- and post-tests, the threshold for interval discrimination was measured in all 14 subjects in three conditions: 100 and 200 msec at 1 kHz and 100 msec at 4 kHz. Thresholds were also measured in six of the subjects (S1, S2, S5, S9, S10, and S13) in two additional conditions: 50 and 500 msec at 1 kHz. The training phase occurred between the pre- and post-tests. It consisted of 1 hr of practice per day for a minimum of 10 d in the 100 msec at 1 kHz condition.

Within each 60-trial block, the comparison interval was adjusted adaptively by decreasing Δt after every three consecutive correct responses and increasing Δt after each incorrect response. The Δt values at which the direction of the change in Δt reversed from decreasing to increasing or increasing to decreasing, referred to as reversals, were noted. The first three reversals of each block of trials were discarded, and the 79% correct point on the psychometric function was estimated by taking the average value of the remaining reversals (Levitt, 1971). To ensure reliability, no estimate was calculated if there were fewer than four remaining reversals. Throughout this paper, threshold is expressed as the Δt needed to achieve 79% correct discriminations, divided by t —a Weber fraction. Thus, a threshold of 0.2 for a standard interval of 100 msec indicates that the subject could discriminate 120 msec from 100 msec intervals 79% of the time.

For subjects S6–S14, the step size was 10% of the standard interval until the third reversal and was 1% of the standard interval thereafter. For these subjects, the comparison interval at the start of each block of trials was always equal to 0 msec, forcing the subject to guess on the first trial. For subjects S1–S5, for all conditions, the step size was 3 msec until the third reversal and was 1 msec thereafter. For these subjects, the starting comparison interval was initially 20% of the standard interval plus 10 msec. This interval was adjusted after the first block in each condition during the pre- and post-tests, and between sessions during the training phase, to be the previous threshold plus 10 msec.

Five or six blocks (300–360 trials) were collected in each condition during the pre- and post-tests. During these tests, the conditions were all presented in random order for subjects S6–S14, but the trained condition was always presented first, followed by the other conditions in random order, for subjects S1–S5. Fifteen blocks (900 trials) were collected on each training day. Approximately 6% of the blocks from the pre- and

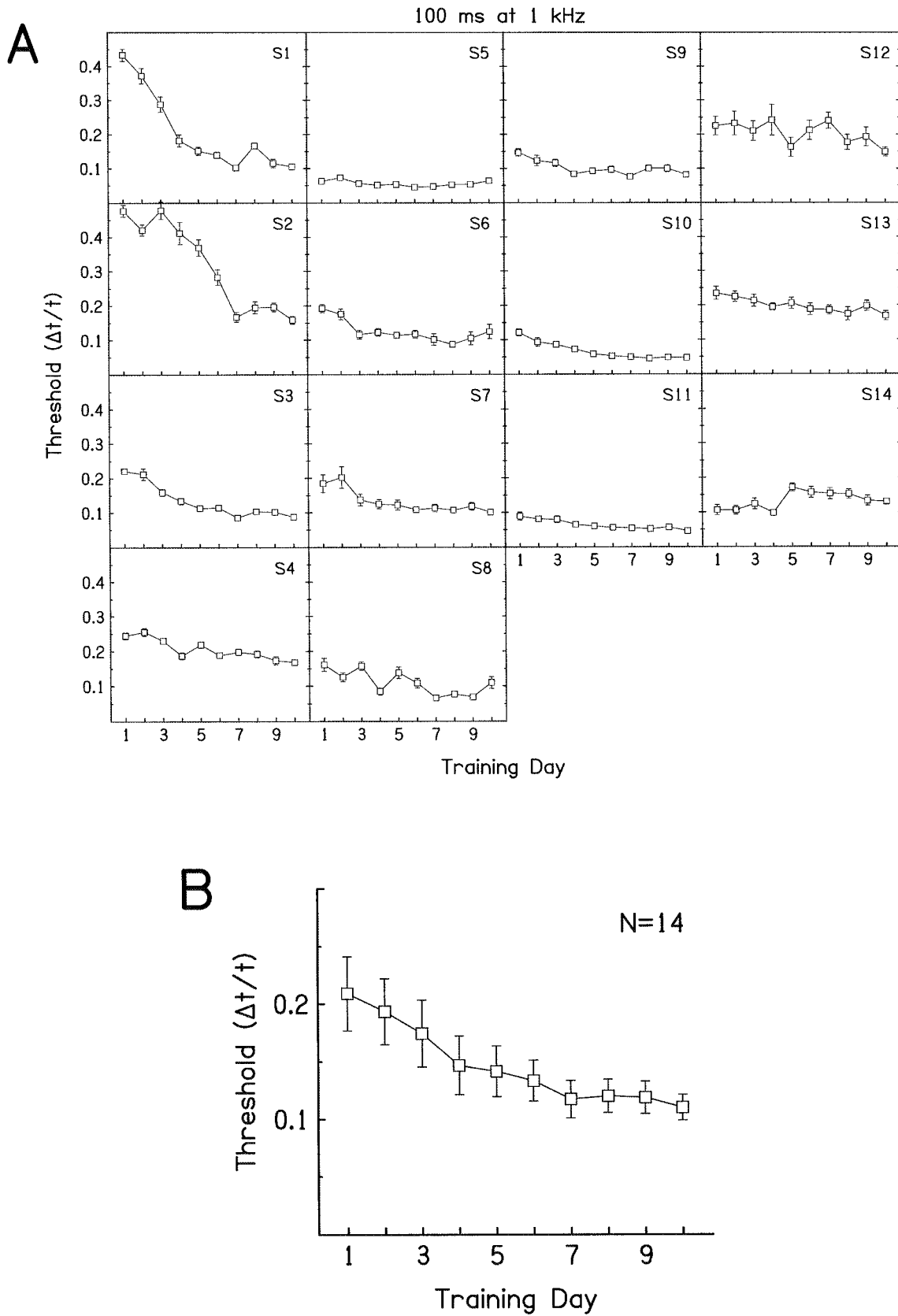


Figure 2. Learning results. The threshold for interval discrimination in the trained condition (100 msec at 1 kHz) as a function of the number of days of training. Threshold is expressed as the Δ millisecond value at threshold divided by the standard duration. Performance is shown for each individual subject in *A* and for the mean of all subjects in *B*. The error bars represent ± 1 SEM within-subjects (*A*) and across-subjects (*B*). Note that the ordinate range differs across *A* and *B*.

post-tests and 2% of the blocks from the training phase were discarded because of an insufficient number of reversals. Despite the differences in the procedures used with subjects S6–S14 and S1–S5, the results of the two groups were quite similar. Therefore, the data of both groups were combined in the analyses reported here.

RESULTS

The first purpose of this experiment was to determine whether practice could improve auditory interval discrimination. The results of each subject during the training phase are plotted in the 14 panels of Figure 2*A*. The open squares indicate the interval-discrimination threshold for the trained 100 msec at 1 kHz condition for the first 10 d of training. The error bars represent the SE within subjects. Figure 2*B* shows the mean performance of all 14 subjects (open squares) and the across-subject SE (error bars). The mean function follows a typical learning curve. A one-way ANOVA with repeated measures performed on the mean data revealed a statistically significant change in performance over training days ($F_{13,9} = 8.80, p < 0.0001$). Thus, auditory interval discrimination does improve with practice.

The second purpose of this experiment was to determine the generalization of learning in interval discrimination. The transfer of learning was evaluated using the data from the pre- and post-tests of only those subjects who showed significant learning during the training phase. An individual subject was judged to have shown significant learning if the results of that subject (1) showed a significant change ($p < 0.05$) in performance across the first 10 d of training according to a repeated-measures one-way ANOVA and (2) yielded a negative slope when fitted with a regression line. Eleven (S1–S11) of the fourteen subjects demonstrated learning by these criteria. Their data on the pre- and post-tests were analyzed to determine the generalization of learning.

Figure 3 shows the mean interval-discrimination thresholds across the qualifying subjects for the five conditions measured in the pre- (white bars) and post-tests (black bars). The data were analyzed using a two-way ANOVA on condition and time (pretest versus post-test), with time as a repeated-measures factor. There was no statistically significant main effect for condition, but both the main effect for time and the interaction between condition and time were statistically significant (Table 1). The interaction was further analyzed using F tests for simple effects (Brunner and Kintz, 1987). Thresholds differed between the pre- and post-tests only for the trained condition of 100 msec at 1 kHz and the untrained condition of 100 msec at 4 kHz. There were no significant differences for the untrained intervals, although they were tested at the trained frequency of 1 kHz (Table 2). Additionally, the threshold change between the pre- and post-tests for the untrained 100 msec at 4 kHz condition was the only one that was not significantly different from the change observed for the trained 100 msec at 1 kHz condition (Table 3).

It seems unlikely that the lack of generalization to the standard intervals of 200 and 500 msec resulted from the better pretest performance in those than in the other conditions. Two pieces of evidence support this view. First, there was also no generalization to the 50 msec standard interval, for which the mean pretest threshold was higher than that for the two 100 msec standard intervals. Second, for standard intervals in the range of 200 to 500 msec, highly trained subjects have thresholds around 0.06 (Getty, 1975; Divenyi and Danner, 1977), whereas the post-test thresholds of the present subjects were typically considerably poorer at 0.15. For comparison, for standard intervals near 100 msec, the highly

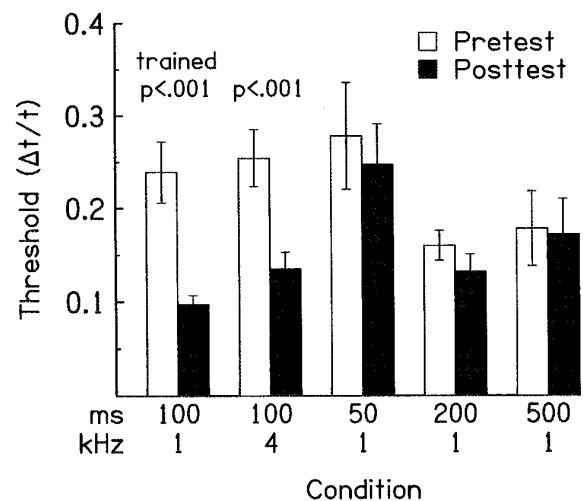


Figure 3. Generalization results. The mean threshold values in the five conditions measured before (white bars) and after (black bars) the training phase. The difference between the pre- and post-test thresholds in the 100 msec at 1 kHz condition represents the learning resulting from the training in that condition. The corresponding decrease in the 100 msec at 4 kHz condition reflects the generalization of that learning to an untrained frequency. The lack of significant differences between the pre- and post-tests in the remaining conditions indicates that the learning did not generalize to untrained temporal intervals. Threshold is expressed as the Δ millisecond value at threshold divided by the standard duration. The error bars represent ± 1 SEM across subjects. Only the data of subjects who individually demonstrated learning during the training phase are included. The data of 11 of the 14 subjects tested are represented in three conditions: 100 and 200 msec at 1 kHz and 100 msec at 4 kHz. The data of five of the six subjects tested are represented in the other two conditions: 50 and 500 msec at 1 kHz.

Table 1. Pre- versus post-test

	df	F	p
Condition	4,38	2.4	ns
Time	1,38	50.9	<0.001
Condition \times Time	4,38	6.1	<0.001

df, Degrees of freedom; ns, not significant.

Table 2. Pre- versus post-test in each condition

msec	kHz	df	F	p
100	1	1,38	42.2	<0.001
100	4	1,38	29.8	<0.001
50	1	1,38	1.0	ns
200	1	1,38	0.9	ns
500	1	1,38	0.0	ns

df, Degrees of freedom; ns, not significant.

Table 3. Threshold change between the pre- and post-tests in the trained versus each untrained condition

msec	kHz	df	F	p
100	4	1,38	0.5	ns
50	1	1,38	7.9	<0.01
200	1	1,38	13.5	<0.001
500	1	1,38	12.1	<0.005

df, Degrees of freedom; ns, not significant.

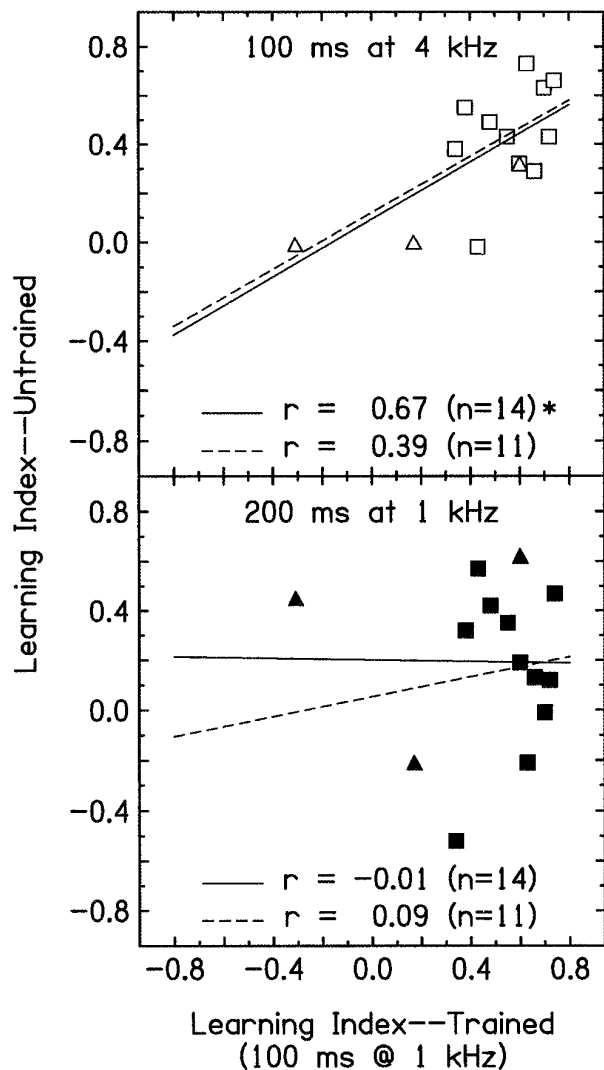


Figure 4. Correlation results. Correlation (r) between learning in the trained (*abscissa*) and each of two untrained (*ordinate*) conditions for all 14 subjects. Each *point* shows the learning index [(pre-post threshold)/pre-threshold] of an individual subject in the untrained 100 msec at 4 kHz (*top panel*) and 200 msec at 1 kHz (*bottom panel*) conditions plotted against the same measure in the trained 100 msec at 1 kHz condition. Different symbols mark the data of the 11 subjects classified as learners (*squares*) and the three subjects classified as nonlearners (*triangles*). The lines represent the least-squared error fits to the data of all 14 subjects (*solid lines*) or only the 11 learners (*dashed lines*). The asterisk in the *top panel* indicates the one statistically significant correlation.

trained subjects in previous reports, and our own subjects, all have thresholds near 0.1.

To examine the generalization of learning on a subject by subject basis, we calculated the correlations between the pre- and post-test improvements in the trained conditions and each of two untrained conditions. The results of all 14 subjects were individually converted into a learning index [(pre-post threshold)/pre-threshold] before performing the correlations. Each point in Figure 4 shows the learning index of an individual subject in the untrained 100 msec at 4 kHz (*top panel*) and 200 msec at 1 kHz (*bottom panel*) conditions plotted against the same measure in the trained 100 msec at 1 kHz condition. Different symbols mark the data of the 11 learners (*squares*) and 3 nonlearners (*triangles*). The solid

lines represent the least-squared error fits to the data of all 14 subjects. In that analysis, the learning index in the untrained 100 msec at 4 kHz condition correlated significantly with that in the trained 100 msec at 1 kHz condition ($r = 0.67, p < 0.01$). This indicates that the improvements observed in the untrained 100 msec at 4 kHz condition were related to learning on the trained condition for every subject and suggests that the same mechanism underlies the changes at 1 and 4 kHz. In contrast, there was no significant correlation between the learning index in the untrained 200 msec at 1 kHz condition and that in the trained condition ($r = -0.01, p = 0.96$), demonstrating the independence of the performance changes in those two conditions. The dashed lines show the least-squared error fits to the data of only the 11 subjects classified as learners (*squares*). In that analysis, there was no significant correlation between the trained condition and the 100 msec at 4 kHz ($r = 0.39, p = 0.25$) or the 200 msec at 1 kHz ($r = 0.09, p = 0.81$) conditions. However, the slope of the fitted line in the 100 msec at 4 kHz panel was very similar to that obtained with all 14 subjects. Figures 3 and 4 thus reveal that learning in interval discrimination transfers to the trained interval presented at an untrained frequency, but not to an untrained interval presented at the trained frequency.

DISCUSSION

Temporal perceptual learning

The present results show that (1) practice improves the ability of adult humans to discriminate longer temporal intervals from a 100 msec standard bounded by tones at 1 kHz, and (2) this learning is temporally, but not spectrally, specific. The average discrimination threshold for 14 naive subjects decreased from 0.21 to 0.11 during 10 d of training (Fig. 2). This learning transferred to performance on the trained 100 msec standard interval bounded by tones at the untrained frequency of 4 kHz, but not to the untrained standard intervals of 50, 200, or 500 msec bounded by tones at the trained frequency of 1 kHz (Figs. 3 and 4).

To our knowledge, these data represent the first systematic demonstration of perceptual learning and generalization on an interval discrimination task. Rammsayer (1994) previously reported that there was no significant learning on the same task using a 50-msec standard interval. His subjects, however, trained on 50 trials per day for 20 d, whereas the present subjects trained on 900 trials per day for at least 10 d. Differences in training time could therefore easily account for these discrepant results. Several other researchers have noted practice effects while examining other features of temporal processing (Michon, 1963; Hafter and Carrier, 1970; Penner, 1976).

Temporal Versus Spatial Perceptual Learning

Most investigations of perceptual learning have examined the discrimination of visual spatial patterns (Sagi and Tanne, 1994), including vernier acuity (McKee and Westheimer, 1978; Beard et al., 1995; Fahle et al., 1995; Poggio et al., 1995; Saarinen and Levi, 1995), spatial frequency discrimination (Fiorentini and Berardi, 1980, 1981), and orientation discrimination (Vogels and Orban, 1985; Karni and Sagi, 1991). Two main conclusions can be drawn from these studies. First, performance on visual spatial tasks improves with practice. Most visual spatial learning occurs over the course of 1 to 2 weeks of practice, with performance changing most rapidly early in the training period (Karni and Sagi, 1993). Second, learning in visual spatial tasks seems to be specific to the

trained task (Sagi and Tanne, 1994). For example, subjects trained to discriminate the orientation of a texture region in one visual hemifield showed no improvement when tested with perpendicular stimuli in the trained location or stimuli of the trained orientation in the opposite visual hemifield (Karni and Sagi, 1991). Thus, at least some spatial perceptual learning is spatially specific.

The present results extend the observation of perceptual learning to the temporal domain. Just as in visual spatial learning, temporal-interval discrimination improved with practice, and most of the learning occurred early in the 10 d of training. Also as in visual spatial learning, the improvements in interval discrimination showed some specificity to the trained task. However, unlike spatial learning, the temporal learning observed here was highly temporally, but not spatially (spectrally) specific.

A more comprehensive comparison of spatial and temporal learning would require an examination of generalization in the temporal domain after spatial perceptual learning. For instance, it is apparently not known if learning to discriminate the visual orientation of a line transfers to cases in which the line is presented at untrained durations or velocities. A demonstration of temporal generalization after spatial perceptual learning, combined with the present report of spatial generalization after temporal learning, would suggest that perceptual learning is specific only to the trained domain.

Models of Temporal Processing

Separate neural mechanisms seem to be involved in processing temporal information on the microsecond, millisecond, and second time scales. For example, a low frequency sound is localized by measuring the difference in the time of arrival of the sound at the two ears. These interaural delays range from 0 to $\sim 650 \mu\text{sec}$ in humans and thus are an order of magnitude shorter than those typically involved in speech. It appears that interaural delays are measured by a specialized neural circuit in the brainstem that combines a range of axonal path lengths with sharply tuned coincidence detectors to form an array of delay lines (Jeffress, 1948; Carr, 1993). At the other extreme, time estimation tasks generally require the discrimination of time intervals longer than 500 msec. Pharmacological and cognitive manipulations indicate that the as yet unknown neural mechanisms involved in time estimation may differ from those used for the discrimination of shorter intervals. Benzodiazepines and the D2 antagonist remoxipride make temporal discriminations more difficult for 1 sec, but not for 50 msec, standard intervals (Rammsayer, 1992, 1993). Similarly, increasing cognitive load by requiring subjects to simultaneously perform a visual task interferes with the ability to discriminate auditory intervals at a standard duration of 1 sec, but not of 50 msec (Rammsayer and Lima, 1991).

The present study focuses on temporal processing in the intermediate time range of from tens to hundreds of milliseconds. The encoding of temporal information on that time scale has been variously attributed to energy integration, internal clocks, neural delay lines, and the encoding of temporal patterns as spatial ones via time-dependent neuronal properties.

In one proposed model of interval discrimination, temporal intervals are represented by the total energy integrated over the interval. According to this idea, the first tone of a pair marking a temporal interval produces activity that decays with a certain time constant such that the total activity produced by the pair of tones depends on the interval between them. Short intervals produce a

high level of evoked activity and long intervals a low one. This model predicts that varying the intensity of either the first or second tone of the pair would degrade temporal processing, but that does not seem to be the case (e.g., Creelman, 1962; Allan and Kristofferson, 1974; Divenyi and Danner, 1977; Rammsayer, 1994). Thus, interval discrimination on the millisecond scale appears to depend on the direct measurement of the temporal interval.

One popular proposal of the mechanism underlying the direct measurement of temporal information is that the passage of time is encoded by an internal clock (Creelman, 1962; Treisman, 1963; Church, 1984). Internal clocks are hypothetical neural pacemakers that generate pulses recorded by a counter. Temporal intervals are measured by the number of registered pulses. The simplest clock models hold that the clock should be shielded from external factors to ensure reliability, and, thus, these models do not predict the perceptual learning reported here. A modified clock model in which variability in the clock pulses produces suboptimal performance could attribute the perceptual learning to a reduction in this variability, but would not be consistent with the observed interval-specific generalization of learning. One more complex clock model consists of two stages (Treisman et al., 1990); the first stage produces pulses at a fixed frequency, and the second modulates the pulse frequency according to the task demands. Such a model could account for the present results if it were assumed that the learning occurred at the second stage, perhaps attributable to a reduction in variability, and that there were independent first-stage oscillators dedicated to different base frequencies. In partial support of these assumptions, Treisman et al. (1994) reported electroencephalographic evidence of multiple first-stage oscillators.

The direct measurement of temporal information also has been attributed to neural delay lines in which different elements of a network are assigned different time delays (Braitenberg, 1967; Tank and Hopfield, 1987; Banks and Margoliash, 1993). In a delay line (Jeffress, 1948), responses specific to a particular interval are generated by delaying on a path the response to the initial stimulus and summing it with the response to the second stimulus. Only intervals in which the two stimuli are separated in time by an amount equal to the neural delay generate a response. Different intervals are encoded by spatially distinct neurons. Cellular properties other than conduction delays might also function as delay lines. These properties include slow inhibitory postsynaptic potentials and rebound excitation (Sullivan, 1982; Margoliash, 1983; Jaffe, 1992), cell threshold (Anton et al., 1991), and intrinsic oscillations (Fujita, 1982; Miall, 1989).

Buonomano and Merzenich (1995) have previously proposed a biologically plausible model for temporal processing that differs from delay-line models in that it does not require a fixed array of time constants that span the time range of the task in question. This dynamic network model relies on time-dependent neuronal properties such as paired-pulse facilitation (PPF) and slow inhibitory postsynaptic potentials (IPSPs). PPF refers to a form of short-term plasticity in which the second of a pair of action potentials can produce a larger postsynaptic response. Slow IPSPs refer to GABA_B-mediated postsynaptic currents. The time course of both paired-pulse facilitation and slow IPSPs ranges from tens to a few hundred milliseconds, the range relevant to the intervals studied here. The model is conceptually simple. Consider a large population of interconnected excitatory and inhibitory neurons in which the excitatory synapses exhibit PPF and the inhibitory synapses produce both slow and fast IPSPs. The first pulse mark-

ing the beginning of a temporal interval will activate a subpopulation of neurons and trigger a series of processes including PPF and slow IPSPs. Therefore, even though the second pulse, marking the end of the interval, may be identical to the first, it will activate a different subpopulation of neurons because the network is in a different state as the result of the occurrence of the first pulse; some synapses will be facilitated and others will be inhibited. Computer simulations have indicated that such a model can perform interval discrimination tasks although the connectivity of the network is random and there is no fixed array of time constants. Recent experimental data (Buonomano et al., 1995) support this model by showing that PPF and slow IPSPs lead to the activation of subpopulations of neurons in a history- and context-dependent manner. Thus, it appears that networks of neurons are intrinsically capable of processing temporal information.

Both the delay-line and dynamic network models can account for the temporal perceptual learning observed here. Because both models translate temporal information into a spatial representation, they can attribute temporal perceptual learning to an increased sensitivity to differences in spatial activation. They can also explain the lack of generalization to untrained intervals by assuming that the learning is spatially specific. Finally, they can account for the spatial (spectral) generalization of learning by assuming that at some level of auditory processing inputs from different frequencies converge on a common site and that learning happens at or after this site.

Conclusions

The human psychophysical results reported here indicate that (1) auditory temporal-interval discrimination improves with practice, and (2) this learning is temporally, but not spectrally, specific. These results are similar to those observed in visual spatial learning in that both types of learning follow a similar time course and show some specificity to the trained task. The reported pattern of temporal learning and generalization is inconsistent with energy integration and single-stage internal clock models of temporal processing, but could be accounted for by delay-line, dynamic network, or two-stage internal clock models. The further refinement of current models awaits a more detailed understanding of the psychophysics and physiology of temporal processing.

REFERENCES

- Allan LG, Kristofferson AB (1974) Psychophysical theories of duration discrimination. *Percept Psychophys* 16:26–34.
- Anton PS, Lynch G, Granger R (1991) Computation of frequency-to-spatial transform by olfactory bulb glomeruli. *Biol Cybern* 56:407–414.
- Bankes SC, Margoliash D (1993) Parametric modeling of the temporal dynamics of neuronal responses using connectionist architectures. *J Neurophysiol* 69:980–990.
- Beard BL, Levi DM, Reich LN (1995) Perceptual learning in parafoveal vision. *Vision Res* 35:1679–1690.
- Braitenberg V (1967) Is the cerebellar cortex a biological clock in the millisecond range? *Prog Brain Res* 25:334–336.
- Bruning JL, Kintz BL (1987) *Computational handbook of statistics*. Glenview, IL: Scott, Foresman and Company.
- Buonomano DV, Merzenich MM (1995) Temporal information transformed into a spatial code by a neural network with realistic properties. *Science* 267:1028–1030.
- Buonomano DV, Hickmott PW, Merzenich MM (1995) Associative plasticity of polysynaptic inputs to CA neurons: Implications for temporal information processing. *Soc Neurosci Abstr* 21:1297.
- Carr CE (1993) Processing of temporal information in the brain. *Annu Rev Neurosci* 16:223–243.
- Church RM (1984) Properties of the internal clock. In: *Timing and time perception* (Gibbon J, Allan LG, eds), pp 566–582. New York: New York Academy of Sciences.
- Creelman CD (1962) Human discrimination of auditory duration. *J Acoust Soc Am* 34:582–593.
- Divenyi PL, Danner WF (1977) Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. *Percept Psychophys* 21:125–142.
- Dorman MF, Raphael LJ, Liberman AM (1979) Some experiments on the sound of silence in phonetic perception. *J Acoust Soc Am* 65:1518–1532.
- Drullman R, Festen JM, Plomp R (1994a) Effect of temporal envelope smearing on speech reception. *J Acoust Soc Am* 95:1053–1064.
- Drullman R, Festen JM, Plomp R (1994b) Effect of reducing slow temporal modulations on speech reception. *J Acoust Soc Am* 95:2670–2680.
- Fahle M, Edelman S, Poggio T (1995) Fast perceptual learning in hyperacuity. *Vision Res* 35:3003–3013.
- Fiorentini A, Berardi N (1980) Perceptual learning specific for orientation and spatial frequency. *Nature* 287:43–44.
- Fiorentini A, Berardi N (1981) Learning in grating waveform discrimination: specificity for orientation and spatial frequency. *Vision Res* 21:1149–1158.
- Fujita M (1982) An adaptive filter model of the cerebellum. *Biol Cybern* 45:195–206.
- Getty DJ (1975) Discrimination of short temporal intervals: A comparison of two models. *Percept Psychophys* 18:1–8.
- Hafer ER, Carrier SC (1970) Masking-level differences obtained with a pulsed tonal masker. *J Acoust Soc Am* 47:1041–1047.
- Ivry R (1997) The representation of temporal information in perception and motor control. *Curr Opin Neurobiol*, in press.
- Jaffe S (1992) A neuronal model for variable response latency. In: *Analysis and modeling of neural systems* (Eeckman FH, ed), pp 405–410. Boston: Kluwer.
- Jeffress LA (1948) A place theory of sound localization. *J Comp Physiol Psychol* 41:35–39.
- Karni A, Sagi D (1991) Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proc Natl Acad Sci USA* 88:4966–4970.
- Karni A, Sagi D (1993) The time course of learning a visual skill. *Nature* 365:250–252.
- Lehiste I, Olive JP, Streeter LA (1976) Role of duration in disambiguating syntactically ambiguous sentences. *J Acoust Soc Am* 60:1199–1202.
- Levitt H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49:467–477.
- Liberman AM, Delattre PC, Gerstman LJ, Cooper FS (1956) Tempo of frequency change as a cue for distinguishing classes of speech sounds. *J Exp Psychol* 52:127–137.
- Lisker L, Abramson AS (1964) A cross language study of voicing in initial stops: acoustical measurements. *Word* 20:384–422.
- Margoliash D (1983) Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J Neurosci* 3:133–143.
- McKee SP, Westheimer G (1978) Improvement in vernier acuity with practice. *Percept Psychophys* 24:258–262.
- Miall C (1989) The storage of time intervals using oscillating neurons. *Neural Comput* 1:359–371.
- Michon JA (1963) Studies on subjective duration. *Acta Psychol* 22:441–450.
- Penner MJ (1976) The effect of marker variability on the discrimination of temporal intervals. *Percept Psychophys* 19:466–469.
- Poggio T, Fahle M, Edelman S (1995) Fast perceptual learning in visual hyperacuity. *Science* 256:1018–1021.
- Rammsayer TH (1992) Effects of benzodiazepine-induced sedation on temporal processing. *Hum Psychopharm* 7:311–318.
- Rammsayer TH (1993) On dopaminergic modulation of temporal information processing. *Biol Psychol* 36:209–222.
- Rammsayer TH (1994) Effects of practice and signal energy on duration discrimination of brief auditory intervals. *Percept Psychophys* 55:454–464.
- Rammsayer TH, Lima SD (1991) Duration discrimination of filled and

- empty auditory intervals: cognitive and perceptual factors. *Percept Psychophys* 50:565–574.
- Saarinen J, Levi DM (1995) Perceptual learning in vernier acuity: what is learned? *Vision Res* 35:519–527.
- Sagi D, Tanne D (1994) Perceptual learning: learning to see. *Curr Opin Neurobiol* 4:194–199.
- Shannon RV, Zeng F-G, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270:303–304.
- Sullivan III WE (1982) Possible neural mechanisms of target distance coding in auditory system of the echolocating bat *Myotis lucifugus*. *J Neurophysiol* 48:1033–1044.
- Tallal P, Piercy M (1973) Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241:468–469.
- Tank DW, Hopfield JJ (1987) Neural computation by concentrating information in time. *Proc Natl Acad Sci USA* 84:1896–1900.
- Treisman M (1963) Temporal discrimination and the indifference interval: implications for a model of the “internal clock.” *Psychol Monogr* 77:1–31.
- Treisman M, Faulkner A, Naish PLN, Brogan D (1990) The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception* 19:705–743.
- Treisman M, Cook N, Naish PLN, MacCrone JK (1994) The internal clock: Electroencephalographic evidence for oscillatory processes underlying time perception. *Q J Exp Psychol* 47A:241–289.
- Vogels R, Orban G (1985) The effect of practice on the oblique effect in line orientation judgements. *Vision Res* 25:1679–1687.