

Perceptual-Learning Evidence for Separate Processing of Asynchrony and Order Tasks

Julia A. Mossbridge,¹ Matthew B. Fitzgerald,¹ Erin S. O'Connor,¹ and Beverly A. Wright^{1,2}

¹Department of Communication Sciences and Disorders, Northwestern University, Frances Searle Building, Evanston, Illinois 60208, and ²Northwestern University Institute for Neuroscience, Chicago, Illinois 60611-3010

Normal perception depends, in part, on accurate judgments of the temporal relationships between sensory events. Two such relative-timing skills are the ability to detect stimulus asynchrony and to discriminate stimulus order. Here we investigated the neural processes contributing to the performance of auditory asynchrony and order tasks in humans, using a perceptual-learning paradigm. In each of two parallel experiments, we tested listeners on a pretest and a posttest consisting of auditory relative-timing conditions. Between these two tests, we trained a subset of listeners ~1 h/d for 6–8 d on a single relative-timing condition. The trained listeners practiced asynchrony detection in one experiment and order discrimination in the other. Both groups were trained at sound onset with tones at 0.25 and 4.0 kHz. The remaining listeners in each experiment, who served as controls, did not receive multihour training during the 8–10 d between the pretest and posttest. These controls improved even without intervening training, adding to evidence that a single session of exposure to perceptual tasks can yield learning. Most importantly, each of the two groups of trained listeners learned more on their respective trained conditions than controls, but this learning occurred only on the two trained conditions. Neither group of trained listeners generalized their learning to the other task (order or asynchrony), an untrained temporal position (sound offset), or untrained frequency pairs. Thus, it appears that multihour training on relative-timing skills affects task-specific neural circuits that are tuned to a given temporal position and combination of stimulus components.

Key words: asynchrony; order; relative timing; temporal perception; auditory processing; human perceptual learning

Introduction

Each of our sensory systems provides us with an ongoing stream of information representing individual perceptual events. One critical step in making sense of this information is to place these events relative to each other in time. In this paper, we investigate learning on two such relative-timing tasks in the auditory system: asynchrony detection and order discrimination. In an auditory asynchrony-detection task, the judgment is whether the beginnings or endings of the frequency components of a sound are synchronous or asynchronous. Sensitivity to this distinction helps us, for example, to separate sounds that arise from different sources (Bregman et al., 1994; Darwin, 1997; Carlyon, 2004) and to distinguish between some speech sounds (e.g., “pa” vs “ba”) (Lisker and Abramson, 1964). In an order-discrimination task, the judgment is whether the onset or offset of a given sound component precedes or follows that of another. The ability to differentiate presentation order enables us to distinguish, for instance, musical melodies (e.g., ascending vs descending scales) and words (e.g., “back” vs “cab”) (Broadbent and Ladefoged,

1959; Hirsh, 1959). Here we report that multihour training improves performance on auditory asynchrony-detection and order-discrimination tasks at sound onset, but that this learning does not generalize between tasks (asynchrony vs order), to an untrained temporal position (sound offset), or to untrained frequency pairs. This lack of generalization reveals key characteristics of the neural circuitry engaged by training on auditory asynchrony and order judgments.

Although there is some previous evidence that training improves performance on auditory relative-timing tasks, this improvement has only been documented for order discrimination, and the generalization of this learning from trained to untrained conditions has not been investigated. Most reports examining performance on auditory relative-timing tasks include data only from trained listeners, implying that experience leads to improvement [asynchrony (Hirsh and Sherrick, 1961; Summerfield, 1982; Parker, 1988; Zera and Green, 1993a,b; Tillmann and Bharucha, 2002) and order (Hirsh, 1959; Hirsh and Sherrick, 1961; Swisher and Hirsh, 1972; Wier and Green, 1975; Miller et al., 1976; Pisoni, 1980; Pastore et al., 1982, 1988; Pastore, 1983; Kelly and Watson, 1986)]. For order-discrimination tasks, such improvements have been either specifically mentioned without supporting data (Broadbent and Ladefoged, 1959; Divenyi and Hirsh, 1974; Pastore, 1983; Kewley-Port et al., 1988; Pastore and Farrington, 1996) or reported with supporting data, including changes in performance during the training period (Nickerson and Freeman, 1974; Warren, 1974; McFarland et al., 1998), com-

Received May 26, 2006; revised Oct. 17, 2006; accepted Oct. 18, 2006.

This work was supported by the National Institutes of Health/National Institute for Deafness and Other Communication Disorders. We thank Karen Banai, Julia Huyck, Jeanette Ortiz, Andy Sabin, Yuxuan Zhang, and two anonymous reviewers for providing helpful comments on previous drafts of this paper.

Correspondence should be addressed to Beverly A. Wright, Department of Communication Sciences and Disorders, Northwestern University, Frances Searle Building, 2240 Campus Drive, Evanston, IL 60208. E-mail: b-wright@northwestern.edu.

DOI:10.1523/JNEUROSCI.2254-06.2006

Copyright © 2006 Society for Neuroscience 0270-6474/06/2612708-09\$15.00/0

parisons between the performance of trained and untrained listeners (Barsz, 1996), or improvement from pretraining to post-training (Merzenich et al., 1996). In each of these cases of documented learning, listeners improved significantly on a trained ordering task, in which two or more consecutive sounds (pure or complex tones and/or noises) were arranged in different orders, and listeners identified specific instances of (Warren, 1974; Merzenich et al., 1996) or discriminated between (Nickerson and Freeman, 1974; Barsz, 1996; McFarland et al., 1998; Takahashi et al., 2004) these orders. However, there is no similar documentation of learning on auditory asynchrony detection, nor has generalization to untrained conditions been explored for either auditory relative-timing task.

Here we investigated learning on both auditory asynchrony-detection and order-discrimination tasks, focusing on the generalization of this learning to several untrained conditions. Each untrained condition was similar to the trained condition, except for the parameter(s) being explored: the task (asynchrony vs order), the temporal position of the judgment (onset vs offset), and the frequencies of the compared sounds. Our basic assumption was that generalization of learning from a trained condition to an untrained condition occurs if and only if the neural circuitry modified during training also influences performance on the untrained condition (Karni, 1996; Karni and Bertini, 1997; Wright et al., 1997; Wright and Fitzgerald, 2001; Karmarkar and Buonomano, 2003). Given this assumption, we reasoned that the pattern of improvement across the trained and untrained conditions would inform us about the nature of the neural circuitry affected by training. For example, if the same circuitry were affected by training on both asynchrony and order judgments, and this circuitry also were to process both temporal positions and all frequency pairs, then learning should generalize to all untrained conditions. At the other extreme, if training were to affect circuitry that processes only a single task and each of the stimulus parameters separately, then training should induce learning only on the trained condition.

We conducted two parallel training experiments. In each experiment, trained listeners were tested on several conditions in a pretest and then were trained over multiple sessions on one of the pretest conditions: either asynchrony detection or order discrimination at sound onset. After training, these listeners were retested on a posttest that was identical to the pretest. Controls in each experiment were tested on the same pretests and posttests as the trained listeners to which they would be compared but were not trained on any task. In both experiments, exposure to the pretest induced significant learning in controls. More importantly, for both groups of trained listeners, multihour training resulted in additional improvement on their respective trained conditions, but this learning did not generalize to any untrained condition. The specificity of this learning suggests that the circuitry affected by multihour training separately processes each relative-timing task, temporal position, and frequency pair.

Materials and Methods

Experimental design. We collected data in two experiments referred to as the asynchrony and order experiments. Both experiments consisted of a pretest, a training phase, and a posttest. In each experiment, in the pretest and posttest, we measured performance on six related relative-timing conditions. These conditions were the same within each listener but differed between the experiments. In the training phase, which occurred between the pretest and posttest, we measured thresholds in a subset of randomly chosen listeners, referred to as trained listeners, for ~1 h/d for 6–8 d on either asynchrony-onset detection or order-onset discrimination, using tones at 0.25 and 4.0 kHz. The remaining listeners, referred to

as controls, participated only in the pretest and posttest. In the asynchrony experiment, the pretest and posttest were separated by an average of 10.3 d for the trained listeners and 10.6 d for the controls. In the order experiment, the separation was, on average, 10.1 d for both the trained listeners and controls. Finally, to assess the stability of posttest performance, we administered the posttest to a subset of the trained listeners a second time, 4–8 weeks after the original one.

Listeners. Forty-five paid participants (34 females) between the ages of 17 and 31 years served as listeners. All reported normal hearing. No listener had previous experience with psychoacoustic tasks. There were six trained listeners and seven controls in the asynchrony experiment, and a separate group of 14 trained listeners and 18 controls in the order experiment. To determine whether listeners could follow instructions and perform normally on a psychoacoustic test, each listener performed one to two 30-trial blocks of a simple tone-detection task. Only individuals who passed this screening were used as listeners in these experiments.

The results reported for the order experiment were actually obtained in two separate order-training experiments. The pretest and posttest conditions differed slightly between the two experiments, but two conditions were shared between these pretests and posttests. In both experiments, trained listeners practiced exactly the same order-onset condition. Because pretest and posttest performance on the two shared conditions did not differ significantly between the trained listeners in the two experiments or between the controls in the two experiments, we combined the data to form a single trained group and a single control group.

Conditions. In all conditions, in each of two observation periods, we presented listeners with a standard stimulus and a signal stimulus. The signal occurred randomly in either the first or second observation period. The listener indicated which of the observation periods contained the signal by pressing a key on a computer keyboard and received visual feedback as to whether the response was correct or incorrect after every trial throughout the experiment.

The two trained conditions differed only in the relative-timing task. In the trained “asynchrony onset at 0.25 and 4 kHz condition,” listeners had to determine whether two tones began at the same time or at different times (Fig. 1A). This was the trained condition in the asynchrony experiment. Both standard and signal stimuli were composed of tones at 0.25 and 4 kHz that ended simultaneously. In the standard, the two tones began simultaneously. In contrast, in the signal, the two tones began asynchronously, with one tone leading the other by some temporal disparity (Δt). The leading tone was selected at random for each signal presentation. In the trained “order onset at 0.25 and 4 kHz condition,” the listener had to discriminate the temporal order of the onsets of the two tones (Fig. 1B). This was the trained condition in the order experiment. In the standard, the onset of the 0.25 kHz tone led that of the 4 kHz tone, and, in the signal, the onset of the 4 kHz tone led that of the 0.25 kHz tone. The temporal disparity between the tone onsets (Δt) was the same in both the standard and signal stimuli.

The untrained conditions on the pretests and posttests fell into three categories. First, to determine whether learning generalized between the asynchrony-onset and order-onset tasks, in each experiment, we tested the trained condition from the other experiment. Second, to determine whether learning generalized from stimulus onset to offset, in each experiment, we tested the offset versions of both the trained and untrained tasks. In the signal for the “asynchrony offset at 0.25 and 4 kHz condition,” the two tones ended at different times, with the lagging tone selected at random for each signal presentation (Fig. 1C). Similarly, in the “order offset at 0.25 and 4 kHz condition,” the offset of the 0.25 kHz tone lagged that of the 4 kHz tone in the standard, and the offset of the 4 kHz tone lagged that of the 0.25 kHz tone in the signal (Fig. 1D). Third, to determine whether learning generalized to untrained frequencies, in each experiment, we tested several conditions identical to the trained condition but with untrained frequency pairs. In both the asynchrony and order experiments, these frequency pairs included 0.75 and 1.25 kHz as well as 0.5 and 1.5 kHz. Additionally, order-onset discrimination was also tested at two more untrained frequency pairs (0.25 and 1.25 kHz as well as 3 and 4 kHz), both of which shared one frequency in common with the trained condition.

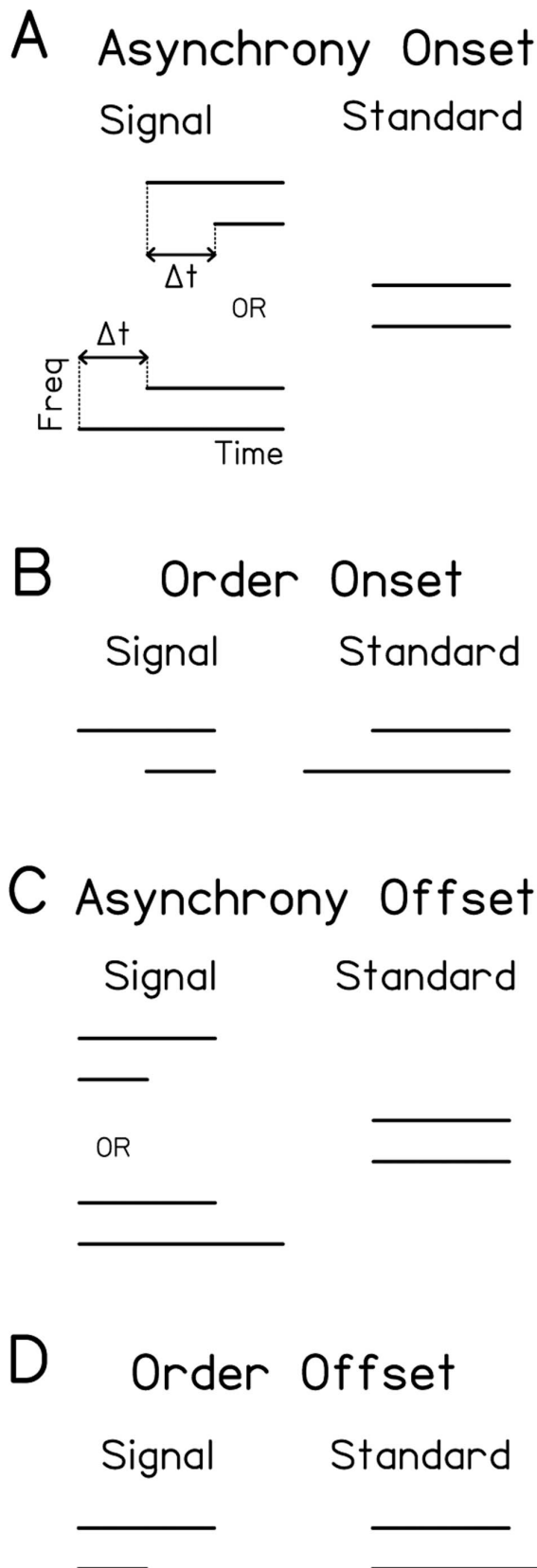


Figure 1. Relative-timing conditions. Schematic diagrams of the signal and standard observation periods in each two-interval forced-choice trial for four relative-timing conditions: asynchrony onset (**A**), order onset (**B**), asynchrony offset (**C**), and order offset (**D**). All stimuli consisted of two-tone complexes in which the duration of the higher frequency tone was always 500 ms. The frequencies of the tones depended on the condition parameters (see Fig. 3 and Materials and Methods). Freq, Frequency.

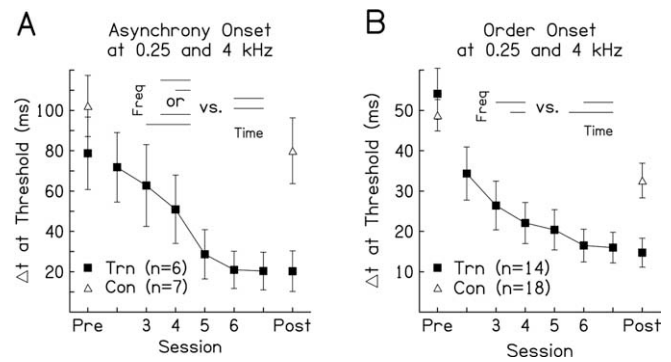


Figure 2. Improvement over the training sessions. Mean learning curves for listeners trained on asynchrony-onset detection (**A**) or order-onset discrimination (**B**) with tones at 0.25 and 4.0 kHz. Each learning curve shows the average thresholds of trained listeners on the trained condition (Trn) on the pretest (Pre), each training session (Session), and posttest (Post; filled squares); the pretest and posttest thresholds of controls (Con) are also shown (open triangles). Error bars show ± 1 SEM across listeners. Schematics of the trained conditions are at the top of each panel. Freq, Frequency.

Stimuli. In all conditions, each tone was gated on in zero phase at 70 dB sound pressure level and had 10 ms raised cosine rise/fall times. The total duration of the high-frequency tone was always 500 ms, and the duration of the low-frequency tone depended on Δt , following the basic stimulus design used in previous investigations of auditory relative-timing performance (Hirsh, 1959; Pisoni, 1980; Parker, 1988; Zera and Green, 1993a). Δt was measured from the onset of the first tone to the onset of the second tone in the onset conditions and from offset to offset in the offset conditions. For all conditions, the offsets of the higher-frequency tones in the two observation periods were separated by 1000 ms.

The tones were generated digitally using a digital-signal-processing board (TDT AP2; Tucker-Davis Technologies, Gainesville, FL) and delivered to two 16-bit digital-to-analog converters (TDT DD1; Tucker-Davis Technologies), followed by an anti-aliasing filter (8.5 kHz low pass, TDT FT5; Tucker-Davis Technologies) and two attenuators (TDT PA4; Tucker-Davis Technologies). We then sent the tones through a headphone buffer (TDT HB6; Tucker-Davis Technologies) and into the left earpiece of Sennheiser HD265 headphones in circumaural cushions.

Procedure. Within each 60 trial block, we adaptively adjusted Δt to estimate threshold by decreasing Δt after every three consecutive correct responses and increasing Δt after each incorrect response. The Δt values at which the direction of change reversed from decreasing to increasing or vice versa are referred to as reversals. The step size was 10 ms through the third reversal and was 2 ms thereafter. After discarding the first three reversals, the Δt at the 79.4% correct point on the psychometric function (threshold) was estimated by taking the average value of the greatest even number of remaining reversals as long as four reversals remained (Levitt, 1971). When there were fewer than four reversals, no threshold estimate was obtained on that block. On the pretest, we set the initial Δt at 60 ms for the first trial of each condition, and, during the training phase, we set it at 10 ms above the average threshold obtained from the previous session performed by that particular listener. Throughout the experiment, if there were too few reversals to obtain a threshold estimate for a given block of trials, the initial Δt value was adjusted for the subsequent blocks.

In each condition on the pretests and posttests, we collected four to five threshold estimates (240–300 trials). Listeners completed one condition before moving to the next, and condition order was determined by a randomized Latin-square design. During the training sessions, we collected 12 threshold estimates (720 trials) from each trained listener per day.

Results

Improvement over training sessions

Multihour training on asynchrony and order judgments at sound onset resulted in significant improvement over the training sessions on both group and individual measures of performance (Fig. 2A,B). On average, the two groups of trained listeners

learned significantly on the condition on which they were trained, as indicated by one-way ANOVAs on mean thresholds with repeated measures over the training sessions (asynchrony trained, $F_{(5,25)} = 5.56$, $p = 0.001$; order trained, $F_{(5,25)} = 9.86$, $p < 0.001$). Average thresholds after training were slightly better than those reported previously for groups of listeners trained in making asynchrony and order judgments at sound onset with similar stimuli [asynchrony, 20 ms here vs ~ 50 ms for a 0.8 and 3.7 kHz pair (Parker, 1988); order, 14 ms here vs ~ 20 ms for a 0.25 and 4.8 kHz pair (Hirsh, 1959)].

Individually, nearly all of the trained listeners improved significantly on the condition on which they were trained. The thresholds of all six of the asynchrony-trained listeners and of 12 of the 14 order-trained listeners changed significantly across training sessions according to a one-way ANOVA and yielded a significant negative slope when fitted with a regression line (Wright et al., 1997), indicating that these listeners learned over the training phase. We performed all analyses both with and without the two remaining nonlearners and reached the same statistical conclusions; therefore, the results presented are based on the data from all listeners.

Improvement on pretest and posttest conditions

In both experiments, controls improved from pretest to posttest across all conditions, and the trained listeners learned more than controls only on their respective trained conditions (Fig. 3*A,B*). The 2.5 h pretest induced significant learning in controls, who did not receive multihour training, in both the asynchrony and order experiments. For both groups of controls, separate time (pretest vs posttest) $\times$ condition ANOVAs with repeated measures on time each revealed a significant main effect of time (asynchrony controls, $F_{(1,36)} = 12.78$, $p = 0.001$; order controls, $F_{(1,80)} = 51.45$, $p < 0.001$) but no time-by-condition interaction (in both cases, $p \geq 0.312$). Thus, in both experiments, pretest exposure without multihour training had an overall positive impact.

To determine whether multihour training benefited listeners, on each condition, we compared the posttest thresholds of trained listeners and controls, taking into account differences in pretest thresholds by using an analysis of covariance (ANCOVA) with pretest threshold as the covariate. We also analyzed the thresholds using a two-group (trained vs control) $\times$ two-time (pretest vs posttest) ANOVA with repeated measures on time; an interaction between group and time indicated that trained listeners learned more than controls on that condition. The results of both analyses always agreed.

For each trained group, multihour training enhanced learning on the trained conditions, but there was no generalization of this learning to untrained conditions. Trained listeners learned significantly more between the pretests and posttests on their respective trained conditions than controls [asynchrony trained vs controls, ANCOVA, $F_{(1,10)} = 9.70$, $p = 0.011$; order trained vs controls, ANOVA, $F_{(1,30)} = 10.98$, $p = 0.002$; an ANCOVA was precluded because a test of the homogeneity of regression was significant (Hays, 1994)]. However, although they did improve, neither trained group learned significantly more than controls on any of the following: (1) the untrained task at the trained temporal position with the trained frequency pair, (2) the untrained temporal position with the trained frequency pair, regardless of whether the task was the trained or untrained task, or (3) the trained task at the trained temporal position with any untrained frequency pairs (all ANCOVA p values ≥ 0.157). Thus, training-

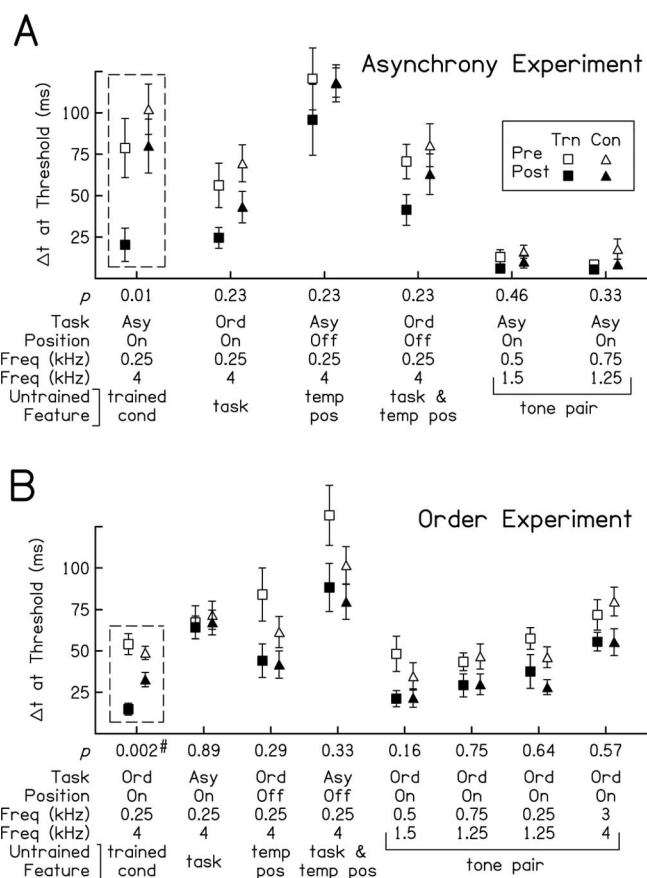


Figure 3. Improvement on pretest and posttest conditions. Mean thresholds on the pretests (Pre; open symbols) and posttests (Post; filled symbols) for trained listeners (Trn; squares) and controls (Con; triangles) on each condition tested. Results are shown for six trained listeners and seven controls in the asynchrony experiment (*A*) and for 6–14 trained listeners and 8–18 controls (see Materials and Methods, Listeners) in the order experiment (*B*). Error bars indicate ± 1 SEM across listeners. Condition parameters and the untrained features are listed along the abscissa. p values are given for ANCOVAs or, in one case, an ANOVA (marked with a # next to the p value). Boxes indicate conditions on which trained listeners learned significantly more than controls. Freq, Frequency; Asy, asynchrony; Ord, order.

induced learning did not generalize across any of the tested parameters.

Pretest thresholds versus improvement

In both experiments, poor pretest performance was associated with greater improvement between the pretest and posttest. To examine the relationship between pretest thresholds and the magnitude of improvement, we determined the linear regression of the amount learned (pretest minus posttest threshold, y -axis) on the pretest threshold (x -axis) for trained listeners and controls in each condition (Fig. 4*A,B*). Separate regression lines were fitted for trained listeners and controls only on the trained conditions because these were the only conditions for which the two groups were statistically separable. For all other conditions, one regression line was fitted to all of the data. All regression lines had positive slopes, and 12 of the 16 lines were significantly so, indicating that listeners who started poorly tended to learn more between the pretest and posttest than those with lower starting thresholds. Note, however, that even trained listeners with very low pretest thresholds on their respective trained conditions (asynchrony, <10 ms; order, <25 ms) still showed significant learning on these conditions during the training phase.

Retention

Trained listeners and controls retained their posttest performance for at least 1 month after training. All six listeners who were trained on asynchrony-onset detection and 3 of the 18 listeners trained on order-onset discrimination returned for a second posttest 4–8 weeks later. Paired *t* tests revealed no differences on any condition between the first and second posttest thresholds of these listeners (asynchrony trained, all *p* values ≥ 0.286 ; order trained, all *p* values ≥ 0.133). In addition, four of the seven controls in the asynchrony experiment returned for a second posttest. These listeners actually showed a trend toward improvement on one condition (order onset at 0.25 and 4.0 kHz, $t_{(3)} = 3.13$; *p* = 0.052) and no difference between the two posttests on the others (all other *p* values > 0.215). It is not known to what extent the exposure to the first posttest contributed to the retention of learning in trained listeners and controls. Regardless, the results suggest that the training and/or testing regimen we used produced benefits that were retained over the long term.

Discussion

Multihour training on asynchrony and order judgments at sound onset resulted in significant improvements in performance that were specific to the trained task, temporal position, and frequency pair. These results appear to be the first formal demonstration of learning on auditory asynchrony detection and echo previous reports of learning on auditory order discrimination (see Introduction). These data also apparently provide the first information about the generalization of learning, or as it happened, the lack thereof, on these relative-timing tasks. We conclude that the specificity of this learning resulted from the constraints of the neural circuitry modified during training and discuss other behavioral and physiological evidence consistent with these constraints below. However, first we briefly discuss the learning that was induced by the pretest, address the possibility that listeners may have used an alternative cue to perform the trained tasks, and then consider and reject a strategy-based explanation for what is perhaps the most surprising outcome of these experiments, the lack of generalization between the asynchrony and order tasks.

Pretest-induced learning

Although the present experiments were designed to examine the effects of multi-hour training, it is worth noting that pretest exposure in itself resulted in signifi-

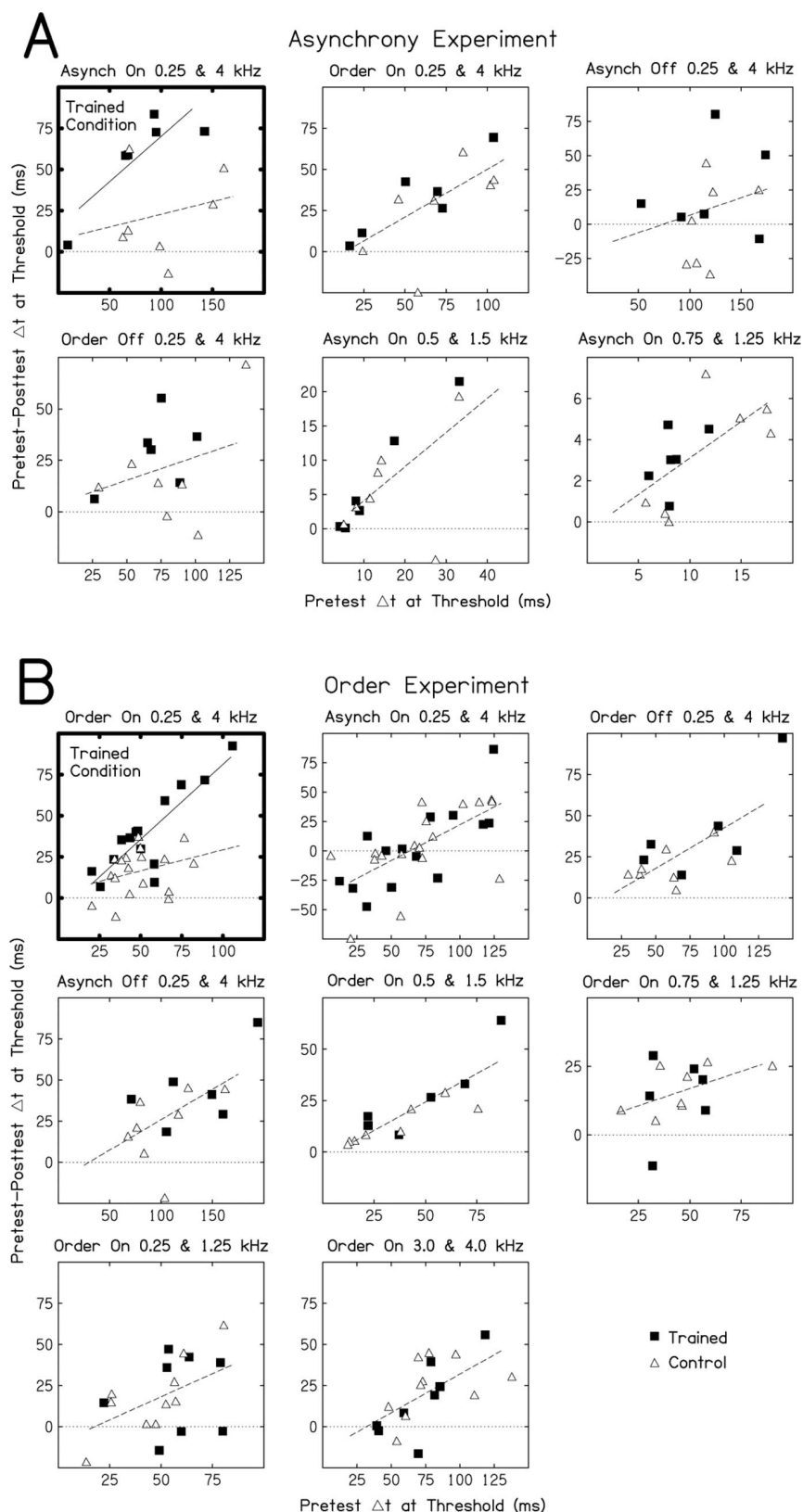


Figure 4. Pretest thresholds versus improvement. The magnitude of improvement [pretest threshold subtracted from posttest threshold (*y*-axes)] plotted as a function of pretest threshold (*x*-axes) for trained listeners (filled squares) and controls (open triangles) in the asynchrony (*A*) and order (*B*) experiments. The linear regression of the magnitude of improvement on the pretest thresholds was determined for each dataset. Separate regression lines were estimated for trained listeners (solid) and controls (dashed) only on the trained conditions (darker frames), because those were the only conditions on which analyses of pretest and posttest thresholds indicated a significant difference between the two groups. On each remaining condition, the regression line for the combined data are shown (dotted). The dotted horizontal line at *y* = 0 indicates no learning; values above this line indicate that performance improved between the pretests and posttests.

cant learning. Controls, who were only exposed to a pretest and posttest, showed a general improvement across all conditions. Improvement resulting from single-session training has been variably attributed to learning of the procedure(s), the task(s), or the stimuli used during that session (Fiorentini and Berardi, 1980; Poggio et al., 1992; Shiu and Pashler, 1992; Karni and Sagi, 1993; Zohary et al., 1994; Rubin et al., 1997; Recanzone, 1998; Ahissar, 2001; Wright and Fitzgerald, 2001; Maye et al., 2002; Hawkey et al., 2004). Here, we cannot make any firm conclusions about which type of learning was induced by the pretest because there was more than one condition on each pretest and the condition order was randomized. Thus, improvement on any single condition could have resulted from practice on that condition itself, or the generalization of learning from another condition. The key conclusions drawn here do not depend on this pretest-induced learning, because they are based on comparisons between trained listeners and controls.

Interestingly, it appears that pretest exposure induced learning that did not decay for at least 8–10 d. We assumed that, had the thresholds of the controls been measured the day after the pretest, the controls would have shown as much improvement as the trained listeners did at this point in time. Given this assumption, if the controls improved from the pretest to posttest as much as the trained listeners did from the pretest to the first day of training, it would indicate that the controls did not lose any of the learning induced by the pretest exposure. Suggesting such retention, the second five threshold estimates on the trained condition (obtained on the first day of training for the trained listeners or on the posttest for the controls), adjusted for differences in pretest threshold, did not differ significantly between the groups in either experiment (both ANCOVA p values ≥ 0.404). In addition, the improvement of controls who returned for a second posttest was retained for 4–8 weeks after the initial posttest. The retention of single-session perceptual learning for at least 7 d is consistent with several previous reports (Rubin et al., 1997; Wright and Fitzgerald, 2001; Fitzgerald and Wright, 2005). Thus, the current results provide additional support for the long-lasting retention of improvement resulting from single-session exposure.

A possible alternative cue

It is also worth noting that listeners likely used relative-timing cues to perform the trained asynchrony and order tasks, although an alternative duration cue was available. Similar to the stimuli used in other investigations of relative-timing performance (Hirsh, 1959; Pisoni, 1980; Parker, 1988; Zera and Green, 1993a), in all the present conditions, the higher-frequency tone remained the same duration, whereas the lower-frequency tone consistently differed in duration between the signal and standard stimuli. Thus, both the asynchrony- and order-trained listeners could have learned to distinguish between the different durations of the lower-frequency tone rather than learning to judge the relative timing of tone onsets. However, this does not appear to be the case, for at least two reasons. First, trained listeners had posttest thresholds of 15–20 ms on the relative-timing conditions on which they were trained. These thresholds were considerably below the ~ 50 ms duration-discrimination thresholds that have been reported previously for listeners highly trained with a 500 ms standard duration (the average duration of the lower-frequency tone used here, $\pm \Delta t$) (Abel, 1972a,b). Second, the learning resulting from training did not generalize to untrained frequencies of the lower tone, whereas that of listeners explicitly trained on duration discrimination generalizes to untrained frequencies (Wright et al., 1997; Karmarkar and Buonomano,

2003). Here, training-induced learning did not even generalize to the untrained offset conditions, in which the lower-frequency tone had exactly the same frequency as in the trained condition (0.25 kHz). Thus, it appears that trained listeners did not learn to discriminate between the durations of individual tones but instead learned to make the intended comparison of the timing of tone onsets.

Strategy-based explanation

Furthermore, it appears that the lack of generalization between the asynchrony and order tasks cannot be explained by an inability of trained listeners to recategorize stimuli when they performed the untrained task on the posttest. Asynchrony-trained listeners learned to categorize as a single group two different signal stimuli (higher tone first and lower tone first). In contrast, when performing the order-onset conditions on the posttest, these listeners had to categorize into two groups (signal vs standard stimuli) the two stimuli they had learned previously to include in a single group (signal stimuli). Conversely, order-trained listeners were taught to discriminate between two stimuli (higher tone first in the signal vs lower tone first in the standard) that they had to recategorize into a single group (signal stimuli) when they performed the asynchrony-onset condition. Thus, confusion caused by this need to recategorize stimuli on the posttest could have been responsible for the lack of generalization between the two tasks. Arguing against this interpretation, we have preliminary data indicating that training on order-onset discrimination does not generalize to asynchrony-onset detection even when the asynchrony condition is performed with only one signal stimulus, and even when that signal is the same as in the trained order condition (higher frequency onset first) (E. S. O'Connor and B. A. Wright, unpublished data).

Neural circuitry constraints

We are therefore led to the conclusion that the present training modified neural circuitry involved in the early stage(s) of relative-timing processing and that the specificity of the observed learning reflects the selectivity of this circuitry. The most counter-intuitive aspect of the present results is the specificity of the learning to the trained tasks. It would seem that, to determine the order of two tones, they must be perceived as asynchronous. However, the lack of generalization between the trained tasks suggests that learning did not modify a serial pathway, in which order judgments depend on a primary judgment of asynchrony (Hirsh, 1959; Hirsh and Sherrick, 1961) because, in that case, learning on the order task should have resulted in improved asynchrony performance. In addition, it is not likely that training affected a global perceptual-timing decision function (Sternberg and Knoll, 1973), because learning should then have generalized between both tasks. Rather, the mutual lack of generalization between trained tasks indicates, surprisingly, that the neural circuitry modified by training on asynchrony judgments is independent of that modified by training on order judgments. The lack of generalization across temporal positions or frequency pairs suggests that the task-specific circuits engaged during training at sound onset govern performance only at sound onset and are subdivided to separately process different frequency combinations.

Previous behavioral and physiological observations, although not previously considered as a whole, are consistent with these conclusions. First, supporting the task specificity of relative-timing circuitry, performance on visual asynchrony and order tasks varies differentially as stimulus parameters change (Mitrani et al., 1986). Such parameter-dependent variations in perfor-

mance have also been reported for asynchrony and order tasks performed with audiovisual stimuli (Fujisaki et al., 2004). Unfortunately, similar comparisons have not been made using auditory stimuli alone. In addition, in the present experiments, pretest thresholds for the trained frequency pair were significantly higher on the asynchrony than the order task ($t_{(43)} = 4.39$; $p = 0.005$), further supporting the idea that asynchrony does not need to be detected before order is discriminated. Physiological support for this implication arises from investigations in the mustached bat revealing neurons that respond to either synchronous sounds or ordered sounds, but not both (Portfors and Wenstrup, 1999; Leroy and Wenstrup, 2000).

Second, the specificity of auditory relative-timing circuitry to the temporal position of the disparity (sound onset or offset) is supported by studies that reveal a distinct difference in relative-timing performance and in physiological responses to sounds at these two temporal positions [behavior (Raphael, 1972; Pastore et al., 1982; Pastore, 1983; Zera and Green, 1993a; Phillips et al., 2002) and physiology (Brugge and Merzenich, 1973; Pflingst and O'Connor, 1981; He et al., 1997; Recanzone, 2000; He, 2002; Takahashi et al., 2004)]. For one example, asynchrony detection in sounds with 20 frequency components is ~ 10 times better at onset than at offset (Zera and Green, 1993a). There is also evidence that the neural substrate for separate processing of these two temporal positions exists. Neurons that are specifically activated at either sound onset or offset, but not both, have been observed in monkey (Brugge and Merzenich, 1973; Pflingst and O'Connor, 1981; Recanzone, 2000) and cat (He et al., 1997) auditory cortex as well as in guinea pig auditory thalamus (He, 2002), and offset population responses recorded by evoked potentials have been observed to be spatially segregated from onset responses in rats (Takahashi et al., 2004).

Third, the specificity of relative-timing circuitry to particular frequency pairs is also supported by previous results. Psychophysical data show that the specific combination of frequencies being compared affects performance on relative-timing tasks. Asynchrony detection at sound onset has been reported to become more difficult as the component tones become more removed from one another in frequency (Parker, 1988; Zera and Green, 1993b), and a similar result has been obtained in some (Divenyi and Hirsh, 1974; Wier and Green, 1975; Kelly and Watson, 1986; Barsz, 1988, 1996), although not all (Hirsh, 1959; Pastore et al., 1982; Pastore, 1983), auditory order and related sequence-identification experiments. In addition, neurons that are activated only by specific frequency combinations have been observed in the forebrains/cortices of several species (Suga et al., 1979; Margoliash, 1983; Margoliash and Fortune, 1992; Tian and Rauschecker, 1994; Lewicki and Konishi, 1995; Ohlemiller et al., 1996; Esser et al., 1997; Brosch et al., 1999; Leroy and Wenstrup, 2000; Kilgard and Merzenich, 2002). It is not clear from present or previous results whether frequency-combination sensitivity is attributable to the specific frequencies themselves or to a more abstract relationship between the frequencies, such as the range between them. Overall, both behavioral and physiological evidence is consistent with the hypothesis that at least a portion of auditory relative-timing circuitry is subdivided to independently process task, temporal position, and frequency-pair parameters.

Potential neural correlates of learning

Suggesting that the pretest- and multihour-induced improvements observed here resulted from changes in different neural substrates, there is considerable neurophysiological evidence from human subjects that different neural circuitry is engaged at

different time points in the learning process. For example, within a single 15 min training period, functional imaging revealed a shift in activation for both a verbal production task as well as a spatial/motor (maze tracing) task (Petersen et al., 1998). For each task, the activation shifted from regions the authors referred to as “scaffolding” areas (left frontal cortex, anterior cingulate, and right cerebellum for the verbal task; right premotor, right parietal, and left cerebellum for the motor task) to “storage” areas (insular cortex for the verbal task; supplementary motor area for the motor task). In addition, from before to after a 6-min training session on auditory sequencing, a task related to order discrimination, the implied sources for event-related potentials (ERPs) appeared to have shifted from primarily unilateral (left) to bilateral superior temporal cortex (Gottselig et al., 2004). Similarly, the activity pattern in motor cortex changed during 30 min of training on a finger-opposition task (Karni et al., 1998). Yet another pattern emerged, an increase in the area of activation in the motor cortex for the trained sequence, after 3 weeks of training on this same task (Karni et al., 1998). Furthermore, on an auditory pattern-discrimination task, the implied sources of ERP components differed depending on whether recordings were made 24 h (primary auditory cortex) or 36 h (nonprimary auditory thalamus and cortex) after a single 6–22 block practice session (Atienza et al., 2002). Finally, in monkeys, behavioral data obtained after unilateral lesions in somatosensory cortex indicated that initial motor skill learning is dependent on the projection from somatosensory to motor cortex, but that, once acquired, performance of the same task does not rely on this connection (Pavlidis et al., 1993).

Among the potential modifications occurring in the circuits affected by training (for review, see Buonomano and Merzenich, 1998; Weinberger, 2004), one potential synaptic mechanism for the present improvements on asynchrony and order performance is spike-timing dependent plasticity (STDP). In STDP learning models, synapse strengths are affected by the temporal association between presynaptic and postsynaptic activation that occurs with repeated exposure to stimuli. These alterations may lead to reductions in synaptic latency (Song et al., 2000), which could conceivably sharpen the timing of the activated neural circuitry. Suggesting that STDP is a good candidate mechanism for the synaptic changes underlying the learning reported here, learning models based on STDP have been proposed for the ordering of inputs (Rao and Sejnowski, 2001; Karmarkar and Buonomano, 2002; Drew and Abbott, 2003; Legenstein et al., 2005), and physiological responses consistent with these order-learning models have been observed in auditory cortex in animals trained for either single (Fu et al., 2002) or multiple (Kilgard and Merzenich, 2002) sessions.

Summary

Overall, the current results show that auditory training induces learning on asynchrony and order tasks at sound onset, implying that at least some portion of the neural circuitry governing performance on these relative-timing tasks is plastic. Improvements resulting from pretest exposure and those resulting from multihour training possibly reflect changes in different underlying neural circuitry at different time points in the learning process. Most importantly, the lack of generalization of learning on either task to the untrained task (asynchrony or order), temporal position (offset), or frequency pairs indicates that the circuitry affected by training is highly specialized and thus is not likely to process global auditory-timing cues. Other behavioral and physiological evidence supports this conclusion. These results could

guide the development of training regimens for both clinical and nonclinical populations who would benefit from improved auditory relative-timing skills and also lead to the refinement of neural, cognitive, computational, and neural models of relative-timing performance.

References

- Abel SM (1972a) Duration discrimination of noise and tone bursts. *J Acoust Soc Am* 51:1219–1223.
- Abel SM (1972b) Discrimination of temporal gaps. *J Acoust Soc Am* 52:519–524.
- Ahissar M (2001) Perceptual training: a tool for both modifying the brain and exploring it. *Proc Natl Acad Sci USA* 98:11842–11843.
- Atienza M, Cantero JL, Dominguez-Marín E (2002) The time course of neural changes underlying auditory perceptual learning. *Learn Mem* 9:138–150.
- Barsz K (1988) Auditory pattern perception: the effect of tonal frequency range on the perception of temporal order. *Percept Psychophys* 48:293–303.
- Barsz K (1996) Accuracy of same/different judgments of sequences of complex tones differing in tonal order under various levels of fundamental frequency range, listener training, and type of standard sequence. *J Acoust Soc Am* 99:1660–1669.
- Bregman AS, Ahad PA, Kim J (1994) Resetting the pitch-analysis system. 2. Role of sudden onsets and offsets in the perception of individual components in a cluster of overlapping tones. *J Acoust Soc Am* 96:2694–2703.
- Broadbent DE, Ladefoged P (1959) Auditory perception of temporal order. *J Acoust Soc Am* 31:1539.
- Brosch M, Schulz A, Scheich H (1999) Processing of sound sequences in macaque auditory cortex: response enhancement. *J Neurophysiol* 82:1542–1559.
- Brugge JF, Merzenich MM (1973) Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation. *J Neurophysiol* 36:1138–1158.
- Buonomano DV, Merzenich MM (1998) Cortical plasticity: from synapses to maps. *Annu Rev Neurosci* 21:149–186.
- Carlyon RP (2004) How the brain separates sounds. *Trends Cogn Sci* 8:465–471.
- Darwin CJ (1997) Auditory grouping. *Trends Cogn Sci* 1:327–333.
- Divenyi PL, Hirsh IJ (1974) Identification of temporal order in three-tone sequences. *J Acoust Soc Am* 56:144–151.
- Drew PJ, Abbott LF (2003) Model of song selectivity and sequence generation in area HVC of the songbird. *J Neurophysiol* 89:2697–2706.
- Esser KH, Condon CJ, Suga N, Kanwal JS (1997) Syntax processing by auditory cortical neurons in the FM-FM area of the mustached bat *Pteronotus parnellii*. *Proc Natl Acad Sci USA* 94:14019–14024.
- Fiorientini A, Berardi N (1980) Perceptual learning specific for orientation and spatial frequency. *Nature* 287:43–44.
- Fitzgerald MB, Wright BA (2005) A perceptual learning investigation of the pitch elicited by amplitude-modulated noise. *J Acoust Soc Am* 118:3794–3803.
- Fu YX, Djupsund K, Gao H, Hayden B, Shen K, Dan Y (2002) Temporal specificity in the cortical plasticity of visual space representation. *Science* 296:1999–2003.
- Fujisaki W, Shimojo S, Kashino M, Nishida S (2004) Recalibration of audiovisual simultaneity. *Nat Neurosci* 7:773–778.
- Gottselig JM, Brandeis D, Hofer-Tinguey G, Borbely AA, Achermann P (2004) Human central auditory plasticity associated with tone sequence learning. *Learn Mem* 11:162–171.
- Hawkey DJ, Amitay S, Moore DR (2004) Early and rapid perceptual learning. *Nat Neurosci* 7:1055–1056.
- Hays WL (1994) Statistics, Ed 5. Orlando, FL: Harcourt Brace.
- He J (2002) OFF responses in the auditory thalamus of the guinea pig. *J Neurophysiol* 88:2377–2386.
- He J, Hashikawa T, Ojima H, Kinouchi Y (1997) Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. *J Neurosci* 17:2615–2625.
- Hirsh IJ (1959) Auditory perception of temporal order. *J Acoust Soc Am* 31:759–767.
- Hirsh IJ, Sherrick Jr CE (1961) Perceived order in different sense modalities. *J Exp Psychol* 62:423–432.
- Karmarkar UR, Buonomano DV (2002) A model of spike-timing dependent plasticity: one or two coincidence detectors? *J Neurophysiol* 88:507–513.
- Karmarkar UR, Buonomano DV (2003) Temporal specificity of perceptual learning in an auditory discrimination task. *Learn Mem* 10:141–147.
- Karni A (1996) The acquisition of perceptual and motor skills: a memory system in the adult human cortex. *Brain Res Cogn Brain Res* 5:39–48.
- Karni A, Bertini G (1997) Learning perceptual skills: behavioral probes into adult cortical plasticity. *Curr Opin Neurobiol* 7:530–535.
- Karni A, Sagi D (1993) The time course of learning a visual skill. *Nature* 365:250–252.
- Karni A, Meyer G, Rey-Hipolito C, Jezard P, Adams MM, Turner R, Ungerleider LG (1998) The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc Natl Acad Sci USA* 95:861–868.
- Kelly WJ, Watson CS (1986) Stimulus-based limitations on the discrimination between different temporal orders of tones. *J Acoust Soc Am* 79:1934–1938.
- Kewley-Port D, Watson CS, Foyle DC (1988) Auditory temporal acuity in relation to category boundaries; speech and nonspeech stimuli. *J Acoust Soc Am* 83:1133–1145.
- Kilgard MP, Merzenich MM (2002) Order-sensitive plasticity in adult primary auditory cortex. *Proc Natl Acad Sci USA* 99:3205–3209.
- Legenstein R, Naeger C, Maass W (2005) What can a neuron learn with spike-timing-dependent plasticity? *Neural Comput* 17:2337–2382.
- Leroy SA, Wenstrup JJ (2000) Spectral integration in the inferior colliculus of the mustached bat. *J Neurosci* 20:8533–8541.
- Levitt H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49 [Suppl 2]:467.
- Lewicki MS, Konishi M (1995) Mechanisms underlying the sensitivity of songbird forebrain neurons to temporal order. *Proc Natl Acad Sci USA* 92:5582–5586.
- 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* 3:1039–1057.
- Margoliash D, Fortune ES (1992) Temporal and harmonic combination-sensitive neurons in the zebra finch's HVC. *J Neurosci* 12:4309–4326.
- Maye J, Werker JF, Gerken L (2002) Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82:B101–B111.
- McFarland DJ, Cacace AT, Setzen G (1998) Temporal-order discrimination for selected auditory and visual stimulus dimensions. *J Speech Lang Hear Res* 41:300–314.
- Merzenich MM, Jenkins WM, Johnston P, Schreiner C, Miller SL, Tallal P (1996) Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* 271:77–81.
- Miller JD, Wier CC, Pastore RE, Kelly WJ, Dooling RJ (1976) Discrimination and labeling of noise-buzz sequences with varying noise-lead times: an example of categorical perception. *J Acoust Soc Am* 60:410–417.
- Mitrani L, Shekerdjiski S, Yakimoff N (1986) Mechanisms and asymmetries in visual perception of simultaneity and temporal order. *Biol Cybern* 54:159–165.
- Nickerson RS, Freeman B (1974) Discrimination of the order of the components of repeating tone sequences: effects of frequency separation and extensive practice. *Percept Psychophys* 16:471–477.
- Ohlemiller KK, Kanwal JS, Suga N (1996) Facilitative responses to species-specific calls in cortical FM-FM neurons of the mustached bat. *NeuroReport* 7:1749–1755.
- Parker EM (1988) Auditory constraints on the perception of voice-onset time: the influence of lower tone frequency on judgments of tone-onset simultaneity. *J Acoust Soc Am* 83:1597–1607.
- Pastore RE (1983) Temporal order judgment of auditory stimulus offset. *Percept Psychophys* 33:54–62.
- Pastore RE, Farrington SM (1996) Measuring the difference limen for identification of order of onset for complex auditory stimuli. *Percept Psychophys* 58:510–526.
- Pastore RE, Harris LB, Kaplan JK (1982) Temporal order identification: some parameter dependencies. *J Acoust Soc Am* 71:430–436.
- Pastore RE, Layer JK, Morris CB, Logan RJ (1988) Temporal order identification for tone/noise stimuli with onset transitions. *Percept Psychophys* 44:257–271.
- Pavlidis C, Miyashita E, Asanuma H (1993) Projection from the sensory to

- the motor cortex is important in learning motor skills in the monkey. *J Neurophysiol* 70:733–741.
- Petersen SE, van Mier H, Fiez JA, Raichle ME (1998) The effects of practice on the functional anatomy of task performance. *Proc Natl Acad Sci USA* 95:853–860.
- Pfingst BE, O'Connor TA (1981) Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. *J Neurophysiol* 45:16–34.
- Phillips DP, Hall SE, Boehnke SE (2002) Central auditory onset responses, and temporal asymmetries in auditory perception. *Hear Res* 167:192–205.
- Pisoni DB (1980) Adaptation of the relative onset time of two-component tones. *Percept Psychophys* 28:337–346.
- Poggio T, Fahle M, Edelman S (1992) Fast perceptual learning in visual hyperacuity. *Science* 256:1018–1021.
- Portfors CV, Wenstrup JJ (1999) Delay-tuned neurons in the inferior colliculus of the mustached bat: implications for analyses of target distance. *J Neurophysiol* 82:1326–1338.
- Rao RP, Sejnowski TJ (2001) Spike-timing-dependent Hebbian plasticity as temporal difference learning. *Neural Comput* 13:2221–2237.
- Raphael LJ (1972) Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in Am English. *J Acoust Soc Am* 51:1296–1303.
- Recanzone GH (1998) Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc Natl Acad Sci USA* 95:869–875.
- Recanzone GH (2000) Response profiles of auditory cortical neurons to tones and noise in behaving macaque monkeys. *Hear Res* 150:104–118.
- Rubin N, Nakayama K, Shapley R (1997) Abrupt learning and retinal size specificity in illusory-contour perception. *Curr Biol* 7:461–467.
- Shiu LP, Pashler H (1992) Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Percept Psychophys* 52:582–588.
- Song S, Miller KD, Abbott LF (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat Neurosci* 3:919–926.
- Sternberg S, Knoll RL (1973) The perception of temporal order: fundamental issues and a general model. In: *Attention and performance IV* (Kornblum S, ed), pp 629–685. New York: Academic.
- Suga N, O'Neill WE, Manabe T (1979) Harmonic-sensitive neurons in the auditory cortex of the mustache bat. *Science* 203:270–274.
- Summerfield Q (1982) Differences between spectral dependencies in auditory and phonetic temporal processing: relevance to the perception of voicing in initial stops. *J Acoust Soc Am* 72:51–61.
- Swisher L, Hirsh IJ (1972) Brain damage and the ordering of two temporally successive stimuli. *Neuropsychologia* 10:137–152.
- Takahashi H, Nakao M, Kaga K (2004) Cortical mapping of auditory-evoked offset responses in rats. *NeuroReport* 15:1565–1569.
- Tian B, Rauschecker JP (1994) Processing of frequency-modulated sounds in the cat's anterior auditory field. *J Neurophysiol* 71:1959–1975.
- Tillmann B, Bharucha JJ (2002) Effect of harmonic relatedness on the detection of temporal asynchronies. *Percept Psychophys* 64:640–649.
- Warren RM (1974) Temporal discrimination by trained listeners. *Cognit Psychol* 6:237–256.
- Weinberger NM (2004) Specific long-term memory traces in primary auditory cortex. *Nat Rev Neurosci* 5:279–290.
- Wier CC, Green DM (1975) Temporal acuity as a function of frequency difference. *J Acoust Soc Am* 57:1512–1515.
- Wright BA, Fitzgerald MB (2001) Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proc Natl Acad Sci USA* 98:12307–12312.
- Wright BA, Buonomano DV, Mahncke HW, Merzenich MM (1997) Learning and generalization of auditory temporal-interval discrimination in humans. *J Neurosci* 17:3956–3963.
- Zera J, Green DM (1993a) Detecting temporal onset and offset asynchrony in multicomponent complexes. *J Acoust Soc Am* 93:1038–1052.
- Zera J, Green DM (1993b) Detecting temporal asynchrony with asynchronous standards. *J Acoust Soc Am* 93:1571–1579.
- Zohary E, Celebrini S, Britten KH, Newsome WT (1994) Neuronal plasticity that underlies improvement in perceptual performance. *Science* 263:1289–1292.