Auditory perceptual learning: It's about time

Beverly A. Wright
Northwestern University

To accurately perceive many naturally occuring sounds, including speech, listeners must be able to discern their durations, and their relative timing across frequency as well as across ears. I will describe our investigations into the human learning of four such auditory temporal-processing tasks: temporal-interval discrimination, asynchrony detection, temporal-order discrimination, and interaural-time-difference discrimination. Our primary goals were (1) to determine whether long term training helps listeners learn to improve their ability to perform each of these temporal tasks, and, if so, (2) to determine whether this learning generalizes from trained to untrained conditions. The results provide information about the plasticity and organization of temporal-processing mechanisms.

Each of our experiments included pretest, training, and posttest phases. In the pretest phase, we used an adaptive, two-interval, forced-choice procedure to determine the detection or discrimination thresholds of naive listeners on a set of conditions involving temporal processing. In the training phase, we trained a subset of those listeners for 6 to 10 hours on only one of the conditions from the pretest by requiring them to perform the task repeatedly. The remaining listeners served as untrained controls. Finally, in the posttest phase, we retested all listeners on the same conditions as in the pretest. We compared the proportional improvements from the pre- to the posttests between trained and control listeners to evaluate learning and generalization. A greater improvement by the trained than control listeners indicated training-induced learning in the trained condition, and the generalization of learning in the untrained conditions.

We first examined learning on a temporal-interval-discrimination task to gain insight into the encoding of the time interval between sounds [1,2]. Across two experiments, we trained twenty listeners to discriminate a standard sound of two, 15-ms, 1-kHz tones whose onsets were separated by a temporal interval of 100 ms, from a signal in which the tone onsets were separated by a longer interval (Fig. 1, left panel). The trained listeners improved from an average threshold of 23 ms before to 11 ms after training on the trained condition (Fig. 1, middle panel). On average, the trained listeners learned more than controls on the trained condition as well as on the trained interval (100 ms) presented at an untrained frequency (4 kHz). However, they learned no more than controls on untrained intervals (50, 200, or 500 ms) presented at the trained frequency (1 kHz; Fig. 1, right panel). Thus, training-induced learning on interval discrimination generalized across frequency, but was specific to the trained interval. We saw a similar learning pattern using somatosensory stimuli. That learning generalized across skin sites and even to auditory stimuli, but also was largely specific to the trained interval [3]. These results suggest that learning modifies a central, multisensory, timing mechanism that processes different temporal intervals separately.

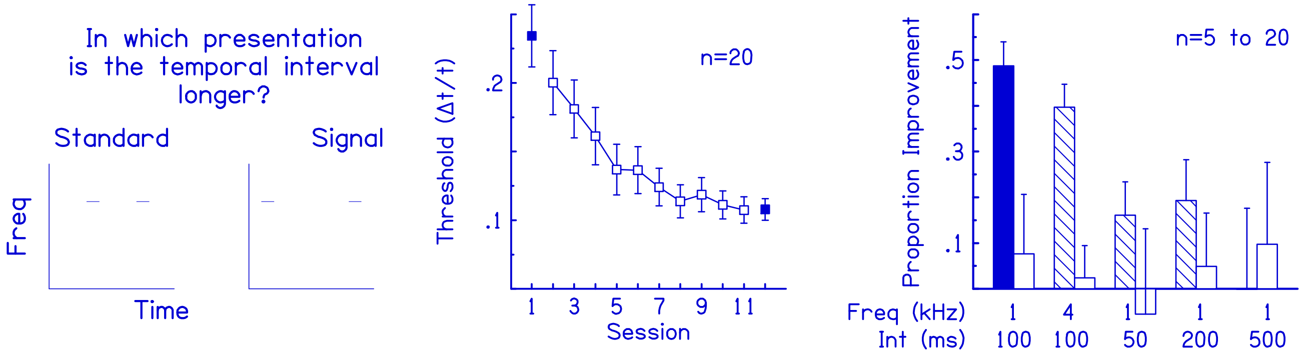

Figure 1: Human learning and generalization of auditory temporal-interval discrimination.
Left Panel: Schematic diagram of the interval discrimination task.
Middle Panel: Learning curve of the trained listeners in the trained condition (100 ms interval at 1 kHz), including mean thresholds and standard errors from the pre- and posttests (filled squares) and the training phase (open squares).
Right Panel: Proportional improvements of the trained listeners on the trained (filled bar) and untrained (hashed bars) condtions, and of the untrained controls (open bars).

We next examined learning in experiments on asynchrony-detection and temporal-order discrimination to gain insight into the encoding of the relative timing of the onsets of two sounds. In the asynchrony-detection experiment [4], we trained six listeners to discriminate a standard sound in which the onsets of a 0.25-kHz and a 4-kHz tone were simultaneous, from a signal in which the onsets of the two tones were asynchronous, with either tone leading in time. In the temporal-order-discrimination experiment [5], we trained a different group of six listeners to discriminate a standard sound in which the onset of a 4-kHz tone preceded the onset of a 0.25-kHz tone from a signal in which the onset of a 4-kHz followed the onset of a 0.25-kHz tone. In both experiments, the two tones always ended synchronously 500 ms after the onset of the 4-kHz tone. The trained listeners improved on the trained condition in both experiments, reducing their average thresholds from 79 ms to 20 ms for asynchrony detection and from 58 ms to 12 ms for temporal order discrimination. Furthermore, in both experiments, the trained listeners learned significantly more than controls on the trained condition. However, they learned only slightly or no more than controls on: (1) the trained asynchrony or order task at two sets of untrained frequencies (0.5 and 1.5 kHz or 0.75 and 1.25 kHz), (2) the trained task applied to the offsets rather than the onsets of the two tones, and (3) the opposite, untrained task, whether asynchrony or order, applied to either the onsets or the offsets of the trained frequency pair. Thus, training-induced learning on both the asynchrony and order tasks was specific to the trained frequency, temporal position (onset vs. offset), and task. Taken together, these results suggest that learning in asynchrony-detection and temporal-order tasks appears to be mediated by separate mechanisms, each of which is sensitive to the frequencies and temporal position of the input.

Finally, we examined learning on an interaural-time-difference discrimination task to gain insight into the encoding of the relative timing of two sounds on the microsecond time scale. We trained eight listeners to discriminate a standard sound of two 300-ms, 0.5-kHz tone presented in phase at 70 dB SPL to both ears over headphones (midline), from a signal presented with an interaural time difference favoring the right ear. The trained listeners improved from an average threshold of 61 μs to one of 29 μs in the trained condition. However, in contrast to the learning patterns we observed on the other temporal-processing tasks, on average, the trained and control listeners both improved by the same amount on the trained condition as well as on every untrained condition. This interaural-time-difference learning appears to be largely complete after only about 20 minutes of training [6]. It is not clear whether the lack of learning during the training phase for interaural-time-difference discrimination occurs because adult listeners are already highly practiced at this task through everyday experience, or because the encoding of interaural-time differences tapped by this discrimination task is simply not malleable.

[1] Wright, B.A., Buonomano, D.V., Mahncke, H.W., Merzenich, M.M. (1997). "Learning and generalization of auditory temporal-interval discrimination in humans," J. Neurosci., 17, 3956 3963.

[2] Wright, B.A. (1998). "Generalization of auditory-discrimination learning." Assoc. Res. Otolaryngol. Abs., Abs. 413, 104. (A)

[3] Nagarajan, S.S., Blake, D.T., Wright, B.A., Byl, N., and Merzenich, M.M. (1998). "Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality," J. Neurosci., 18, 1559-1570.

[4] Wright, B.A., Fitzgerald, M.B., and Mossbridge, J.A. (2001) "Specificity of learning in an auditory asynchrony-detection task," J. Acoust. Soc. Am., in press (A).

[5] Wright, B.A., and Mossbridge, J.A. (2000) "Specificity of learning in an auditory temporal order task," J. Acoust. Soc. Am., 107, 2882 (A).

[6] Ortiz, J.A., Wright, B.A., Fitzgerald, M.B., and Pillai, J. (2001) "Rapid improvements on interaural-time-difference discrimination: Evidence for three types of learning," J. Acoust. Soc. Am., in press (A).