Irvine, D. R. F. 2018. Auditory perceptual learning and changes in the conceptualization of auditory cortex. *Hearing Research*. **366**: 3-16.

doi: 10.1016/j.heares.2018.03.011

Auditory Perceptual Learning and Changes in the Conceptualization of Auditory Cortex

Dexter R. F. Irvine

Bionics Institute, East Melbourne, Victoria 3002, and School of Psychological Sciences, Monash University, Victoria 3800, Australia

Correspondence:

Prof. D. R. F. Irvine
Bionics Institute
East Melbourne
VIC 3002
Australia

Phone: 61 3 93874123
Email: dirvine@bionicsinstitute.org

Abstract

Perceptual learning, improvement in discriminative ability as a consequence of training, is one of the forms of sensory system plasticity that has driven profound changes in our conceptualization of sensory cortical function. Psychophysical and neurophysiological studies of auditory perceptual learning have indicated that the characteristics of the learning, and by implication the nature of the underlying neural changes, are highly task specific. Some studies in animals have indicated that recruitment of neurons to the population responding to the training stimuli, and hence an increase in the so-called cortical "area of representation" of those stimuli, is the substrate of improved performance, but such changes have not been observed in other studies. A possible reconciliation of these conflicting results is provided by evidence that changes in area of representation constitute a transient stage in the processes underlying perceptual learning. This expansion – renormalization hypothesis is supported by evidence from studies of the learning of motor skills, another form of procedural learning, but leaves open the nature of the permanent neural substrate of improved performance. Other studies have suggested that the substrate might be reduced response variability - a decrease in internal noise. Neuroimaging studies in humans have also provided compelling evidence that training results in long-term changes in auditory cortical function and in the auditory brainstem frequency-following response. Musical training provides a valuable model, but the evidence it provides is qualified by the fact that most such training is multimodal and sensorimotor, and that few of the studies are experimental and allow control over confounding variables. More generally, the overwhelming majority of experimental studies of the various forms of auditory perceptual learning have established the co-occurrence of neural and perceptual changes, but have not established that the former are causally related to the latter. Important forms of perceptual learning in humans are those involved in language acquisition and in the improvement in speech perception performance of post-lingually deaf cochlear implantees over the months following implantation. The development of a range of auditory training programs has focused interest on the factors determining the extent to which perceptual learning is specific or generalises to tasks other than those used in training. The context specificity demonstrated in a number of studies of perceptual learning suggests a multiplexing model, in which learning relating to a particular stimulus attribute depends on a subset of the diverse inputs to a given cortical neuron being strengthened, and different subsets being gated by top-down influences. This hypothesis avoids the difficulty of balancing system stability with plasticity, which is a problem for recruitment hypotheses. The characteristics of auditory perceptual learning reflect the fact that auditory cortex forms part of distributed networks that integrate the representation of auditory stimuli with attention, decision, and reward processes.

Keywords

Frequency discrimination
Generalization
Musical training
Plasticity
Stability
Synaptic weights

1. Introduction

As discussed elsewhere (Irvine, 2018), evidence for various forms of plasticity in the primary auditory cortex (AI) has led to major changes in the way in which auditory cortex (AC) is conceptualized. In contrast to the traditional view that AI, as the final processing stage of the "lemniscal line" system, serves simply as a high level auditory analyser, it is now seen as a component of complex distributed networks that integrate the representation of auditory stimuli with attention, decision and reward processes.

One of the forms of experience that has been shown to influence processing in auditory and other sensory cortices is practice on tasks involving the detection, discrimination, or identification of sensory stimuli. Improvement in perceptual performance as a result of such practice or training is referred to as perceptual learning. It should be emphasized that the term does not apply to improved performance on associative learning tasks that involve discrimination between two readily discriminable stimuli (as is required in various forms of differential conditioning). Perceptual learning involves an improvement in sensory performance itself, typically measured by an increase in d′ or in accuracy, or a reduction in threshold or in the duration of the observation interval needed to make the discrimination.

The fact that perceptual judgements can be improved by practice/training has, of course, been recognized for many years, and a variety of professions (e.g., wine-tasters, radiologists) have been cited as exemplars. It was also acknowledged in the traditional procedure in psychophysical experiments of using a small number of highly trained participants, often the experimenters themselves. This procedure was adopted to ensure optimal performance of the task used to measure the performance of the sensory system under investigation, and the possibility that sensory processing mechanisms themselves might be modified by such training was not considered. For example, in a classic review of the topic, Gibson (1953) suggested that improvements in perceptual judgements involved reduction of error in the mapping of response categories onto the stimulus continuum.

The need to train human and non-human participants on the particular task used to measure perceptual performance gives rise to an important general issue in perceptual learning research: distinguishing between improved performance attributable to learning the task (commonly termed “procedural” learning) and improvements in perceptual performance itself (see Ortiz and Wright (2009) for a detailed discussion of these issues). Although it has commonly been assumed that the early fast phase of learning on a perceptual learning task is procedural rather than perceptual, this assumption is unwarranted in some cases (e.g., Hawkey et al., 2004). One solution to this problem has been to introduce a preliminary task-practice phase (e.g., Irvine et al., 2000) in which participants are trained on a different sensory discrimination using the same task.

The recent upsurge of interest in perceptual learning was prompted by findings in the 1970s and 80s suggesting that some forms of visual perceptual learning (VPL) in fact involve changes in early stages of visual processing (see Sagi (2011) and Watanabe and Sasaki (2015) for reviews). Because of the importance of these studies, and the fact that theoretical developments in the field have been largely driven by VPL research, this review commences with a brief account of the visual literature.

2. Visual perceptual learning and models

The surprising finding that prompted renewed interest in VPL was that many forms of such learning were highly specific to the particular attributes of the training stimuli (e.g., orientation, spatial frequency) and/or to the region of the retina to which the stimuli were presented (see Sagi (2011) and Watanabe and Sasaki (2015) for reviews). This specificity led to the view that the most likely site of the underlying changes was at relatively early stages of cortical processing characterized by retinotopy and neuronal selectivity for those attributes (viz., in primary visual cortex (V1)). Such “early-stage” models have been supported by subsequent electrophysiological and imaging studies, although the changes described in V1 have been relatively modest (see Watanabe and Sasaki (2015) for review). For example, Schoups et al. (2001) reported that in monkeys that were trained and showed marked improvement on an orientation identification task, there was no increase in the proportion of neurons tuned to the training orientation, but a change in the tuning of the neurons that conveyed most information about the training orientation (viz., those tuned to nearby orientations). In a study by Crist et al. (2001), there was no change in cortical magnification, receptive field size or orientation tuning in V1 of monkeys trained on a bisection discrimination task, but the effect of a contextual stimulus placed outside the receptive field changed with training. An important feature of this study is that the cortical recordings were made while the animals were performing one of two tasks – the bisection discrimination task or a simple task requiring them to detect dimming of the fixation dot. The reported changes in neuronal response characteristics were seen only when the animals were actually performing the bisection task, an observation that will be considered further below (see Section 9.2). In a number of other studies of the neural correlates of VPL, much larger changes were seen in higher visual areas than in V1, in some cases in the absence of any change in V1 itself (e.g., Raiguel et al., 2006; Yang & Maunsell, 2004), giving rise to what Watanabe and Sasaki (2015) call "mid-stage" models.

Although the argument that the specificity of many forms of VPL suggests that the underlying neural changes occur in sensory cortical representation is compelling, alternative explanations are possible. The first to point this out were Mollon and Danilova (1996), who argued that what the subject in a perceptual learning study has to learn is what subset of neural channels provides the information to support optimal performance. On this argument, specificity arises because, when the stimulus is changed, the subject has once again to go through the process of determining the most-informative channels. Neurophysiological evidence in support of such a model is provided by a study in which monkeys were trained on a task requiring determination of the direction of visual motion (Law and Gold, 2008). The responses of neurons in the motion-sensitive medial temporal area did not change with training, but those of neurons in the lateral intraparietal area did change with training, and the changes reflected the improvements in behavioral sensitivity. In this situation, learning did not involve a change in the representation of sensory information in visual cortex, but a change in the way in which the sensory representation was interpreted, a decision process that determined the behavioral response. Psychophysical and modelling studies have supported the suggestion that VPL reflects changes in the weighting or readout of inputs to decision (e.g., Dosher and Lu, 1999), which in turn might involve basing decisions on the activity of a small set of the most informative neurons (e.g., Jacobs, 2009). These proposals, suggesting that perceptual learning involves changes in higher–order cognitive processes associated with attention and decision making, constitute what Watanabe and Sasaki (2015) have termed "late-stage" models.

It is important to note that although this classification of models suggests that changes take place at one level or another, it is entirely possible that changes take place at multiple levels. For example, Yan et al. (2014) reported that improvement by monkeys on a task requiring detection of camouflaged visual contours involved changes in both early visual processing and readout and decision processes. One influential theory of the processes underlying perceptual learning – reverse hierarchy theory (Ahissar et al., 2009) – postulates that the critical process is a shift in the level of processing on which performance is based.

3. Auditory perceptual learning

Auditory perceptual learning by humans has been investigated for a wide range of stimulus attributes (viz., frequency, level, temporal interval, interaural level and time disparities) (see Wright and Zhang (2009a) for review). A detailed account of the psychophysical research in the area is beyond the scope of this paper, which will focus on studies that directly investigate the nature of the underlying neural, and particularly cortical, mechanisms. However, some psychophysical studies have important implications for the nature of these mechanisms, and will therefore be considered in the context of the issues on which they bear. One important general point emphasised by Wright and Zhang (2009a) is that the characteristics of auditory perceptual learning (e.g., its time course, the shape of the learning curve, patterns of across-task and across-stimulus generalization) are highly attribute dependent. As they point out, these differences suggest differences in the nature of the neural processes contributing to learning on these tasks.

3.1. Neurophysiological studies in animals

The neurophysiological evidence from animal studies relates mainly to changes at early levels of processing. The most detailed evidence concerns pure-tone frequency discrimination, and because this evidence raises some important general and theoretical issues it will be considered in some detail.

3.1.1. Frequency discrimination

The first evidence on this issue was presented by Recanzone et al. (1993). They trained monkeys on a frequency discrimination task, and the monkeys' performance improved with training to generate just noticeable differences that were similar to those reported previously for non-human primates. Subsequent neurophysiological recordings revealed a massive increase (by a factor of ~7 to ~9, depending on frequency) in the area of AI containing neurons with characteristic frequencies (CF) in the frequency range used in training (i.e., in the "area of representation" of those frequencies) compared to control animals or animals trained at other frequencies (Fig. 1A).

As noted elsewhere, although the term "area of representation" as defined above has been widely used as a convenient shorthand term in the plasticity literature, the term carries some surplus unjustified meaning. It is well established that a given frequency is not represented in AI by activity restricted to neurons with that CF or best frequency (BF); rather, it is represented in a complex, intensity-dependent fashion across a population of neurons that respond to that frequency (e.g., Phillips et al., 1994). Despite this caveat, the term will be used here, partly because it is widely used in the literature under review, and partly because it is difficult to come up with an alternative term that isn't excessively unwieldy.

FIG 1 ABOUT HERE

Cortical recruitment resulting in an enlarged representation of a given frequency range must presumably occur at the expense of the representation of adjacent frequency ranges. If the increased representation is the substrate of improved performance, then a decreased area would presumably result in a decrement in performance. Recanzone et al. (1993) presented evidence in support of this. Most of their animals were trained to detect increments in frequency (i.e., +ΔF), but in one monkey they interpolated trials with a frequency decrement (- ΔF). The psychophysical data for this animal indicated that improvement on the + ΔF task was associated with a decrease in performance on the -ΔF task, and subsequent recordings from the AI indicated that there was no area with CFs in the - ΔF frequency testing range.

Enlargement of the representational area in the AI was also reported by Polley *et al*. (2006), who trained rats in an operant task to identify a target tonal stimulus, defined by either a particular frequency or a particular sound pressure level (SPL), from a set of distracter stimuli varying in frequency and SPL. An ingenious feature of this study was that the groups of rats trained on the frequency and level discrimination tasks received the same distracter stimuli and therefore had almost identical stimulus exposure, but were required to learn discriminations based on different stimulus characteristics. For both groups, discrimination thresholds decreased over the course of training. In rats trained on the frequency discrimination task, the area of representation of frequencies around the target frequency (viz. 5 kHz) in AI was found to be enlarged (by a factor of ~2) (Fig. 1B), but there was no change in responses to SPL, whereas in rats trained on the level discrimination task the relative area containing neurons with best level (viz. the level eliciting the maximum response) close to the target level (viz. 35 dB) was enlarged, but there was no change in frequency responses. The different patterns of results in groups that received the same stimuli but learned different detection tasks demonstrate the importance of top-down influences on perceptual learning.

In contrast to the results of these two studies, Brown *et al*. (2004) found no change in the frequency organization of AI in cats that were trained and showed perceptual learning on a frequency discrimination task. The area of representation of the training frequencies, defined in terms of either CF, BF, or the frequency-intensity combinations used in training, did not differ between trained and control animals (Fig. 1C). However, there was a tendency for neurons with CF at and above the training frequency to have slightly broader tuning and shorter latency in the trained cats than in controls. The cats in this study were trained on a + ΔF discrimination task, but small blocks of -ΔF trials were interpolated during training, and -ΔF thresholds decreased in parallel with +ΔF thresholds.

Yet another pattern of results was reported by Witte and Kipke (2005), who recorded chronically from AI of cats in the course of frequency discrimination training, using implanted microwires. They reported *decreases* in the summed response of auditory cortical neurons in AI (and in the anterior auditory field) to the training frequencies, and in the number of neurons with CF in that range. The summed response at frequencies below the training frequencies increased, such that there was a local minimum in the response at the training frequencies, a pattern of change reminiscent of that observed by Ohl and Scheich (1996) in a differential conditioning study.

Talwar and Gerstein's (2001) finding that the enlarged representation of a particular frequency range in rat AI produced by intracortical microstimulation was not associated with improved frequency discrimination at those frequencies has sometimes been interpreted as indicating that an enlarged cortical representation is not the substrate of improved frequency discrimination. However, this argument is weakened by the consideration that enlarged auditory cortical representations produced by different procedures can reflect different mechanisms. For example, the enlarged representation of lesion edge-frequencies associated with restricted cochlear lesions is not dependent on basal forebrain cholinergic input (Kamke et al., 2005) and is apparently the result of homeostatic plasticity mechanisms (see Irvine (2018) for review). In contrast pairing stimulation of the cholinergic basal forebrain with a tonal stimulus results in an enlarged representation of that frequency (see Metherate (2011) for review). The change in cortical circuitry associated with the increase in representational area produced by intracortical microstimulation could therefore be of a type that would not support improved frequency discrimination performance. This argument is supported by Han et al.'s (2007) report that the enlarged representation of exposure frequencies in rats reared in a single-frequency tonal environment was associated with impaired, rather than improved, discrimination of the over-represented frequencies.

The very different patterns of results in the four major studies reviewed above might be explicable in terms of the numerous procedural differences between them. They differ in the species studied, the nature of the frequency discrimination tasks and conditioning procedures, and the measurement and analytic procedures (both psychophysical and neurophysiological) employed. It is now well established that the changes in the spectral sensitivity of AI neurons produced by classical and instrumental conditioning with acoustic stimuli depend critically on a range of task characteristics (e.g., the nature of the task itself, the reinforcement contingencies, the strategy employed by the animal) (see Irvine (2018) for review). It would therefore not be surprising if the changes associated with perceptual learning were to vary in a similar fashion with task variables, as suggested by Wright and Zhang (2009a).

Nevertheless, there are some other lines of evidence that bear on the nature of the changes that are likely to underlie improvements in frequency discrimination. One is provided by human psychophysical studies of perceptual learning on such tasks. In the overwhelming majority of these studies, substantial generalization of learning from trained to untrained frequencies has been reported (see Wright and Zhang (2009b) for review). The consistent finding of across-frequency generalization suggests that it is unlikely that the substrate of the learning is changes restricted to neurons tuned to the training frequency or frequency range. Furthermore, as noted above, an increase in the representation of the training frequencies must occur at the expense of the representation of other frequencies, and thus implies the loss of discriminative ability at those other frequencies. The evidence on generalization indicates that this is not the case. Even when immediately adjacent frequency ranges are considered, humans show generalization from a +ΔF to a -ΔF discrimination task (Irvine et al., 2005), as did the cats in the Brown et al. (2004) study.

There is also psychophysical and neurophysiological evidence suggesting that the mechanisms of auditory perceptual learning might involve changes other than those in in sensory representation *per se.* Jones et al. (2013) presented evidence that improvement by humans trained on a frequency discrimination task was best modelled as a decrease in the internal noise, and

Micheyl et al. (2009) have argued that a reduction in internal noise can explain the surprising finding of improvement in frequency discrimination thresholds after training with physically identical tones (Amitay et al., 2006). In a neurophysiological study of perceptual learning by gerbils on an amplitude modulation (AM) detection task, von Trapp et al. (2016) recorded from core AC (specific fields not specified) after AM detection thresholds had reached asymptotic levels. Neural sensitivity to AM was improved when animals were performing the task, compared to when they were presented with the same stimuli when sitting quietly, and the neural correlate of improved performance was a decrease in response variability (equivalent to a reduction in internal noise) rather than in neural tuning. It is not clear whether this change was the substrate of improved performance or simply a correlate of task performance, but the data nevertheless point to the importance of changes in response variability. Further support for the importance of internal noise is provided by the finding that performance of a tone-detection task by gerbils was associated with a reduction in auditory cortical spontaneous activity (internal noise) (Buran et al., 2014). Reduction in internal noise has been identified by Dosher and Lu (1999, 2017) as a major factor in VPL.

Two other issues relating to the issue of whether perceptual learning involves an increase in the area of cortical representation remain to be considered. One is the suggestion by Reed et al. (2011) that such expansion occurs in the course of training and enables the changes in circuitry that underlie improved performance, but that once these changes have occurred the representation renormalizes. This "expansion – renormalization" model was based on their finding that improved frequency discrimination was associated with an increase in AI representational area in rats that were trained on a frequency discrimination task or received paired tone – nucleus basalis (NB) stimulation, but the increase in representational area subsequently reversed without any decrement in discrimination performance. The notion that changes in sensory cortex during perceptual learning constitute a temporary phase that is necessary for learning but are not the substrate of improved performance has some support from other systems. Yotsumoto et al. (2008) trained human subjects on a visual texture discrimination task and scanned participants at various stages during training. In two different experiments, relative BOLD activation in V1 increased in the region corresponding to the trained location as performance improved, but then decreased to pre-test levels while performance remained at a high level. It is also important to note, however, that the activated region size in V1 did not expand with training (i.e., there was no change in representational area). There is also evidence from motor learning for transient increases in activity and/or activated area (see Section 6).

The expansion – renormalization model suggests that the differences in the results of the four studies of frequency discrimination discussed above might be attributable to differences in the time during training at which the cortical measures were obtained. However, the fact that Brown et al. (2004) observed progressive improvements in discrimination at non-trained frequencies (including the -ΔF testing range) when these were tested intermittently in the course of training would argue against the occurrence of an expanded representation of training frequencies (and concomitant loss of representation of non-trained frequencies) in the course of training. Furthermore, the fact that Witte and Kipke (2011), who recorded cortical activity during, rather than after, training, failed to see an increase in area indicates that learning on their task did not involve such a change. It should also be noted that if expansion-renormalization occurs on some tasks, the nature of the permanent circuit changes that constitute the substrate of subsequent improved performance on those tasks remains to be determined (see Section 9.2).

A final and more general issue concerning the proposal that an increased area of representation is the substrate of perceptual learning relates to the “stability-plasticity dilemma” which arises for all (biological and artificial) learning systems (Mermillod et al., 2013). It concerns the problem of preventing the learning of new tasks from distorting the neural representation of previously learned tasks. In neural networks, in which learning involves changes in synaptic weights, learning a new task results in “catastrophic forgetting” when connection weights in the network are changed (Kirkpatrick et al, 2017). In contrast, human and non-human animals are able to learn multiple tasks sequentially, without such forgetting. Given that recruitment, and the consequent enlargement of cortical representation, implies that improvement on one task necessarily involves a loss of ability on other tasks, it is unlikely to be a general mechanism of perceptual (or other) learning. This issue is considered further in Section 9.2.

3.1.2. Other stimulus dimensions

Changes in auditory cortical activity have also been described in association with perceptual learning on other stimulus attributes, and some of these studies have important implications for the nature of the changes underlying perceptual learning.

FIG 2 ABOUT HERE

In two recent studies of perceptual learning by gerbils on an AM detection task, evidence has been found that the neural changes associated with perceptual learning are specific to the behavioural context. As described above, von Trapp et al. (2016) reported that neural sensitivity to AM was improved during task performance. A similar context specificity was reported by Caras and Sanes (2017): improvements in core AC neural sensitivity to AM (as measured by neurometric functions) were large during task performance and were correlated with improvements in behavioral sensitivity (Fig. 2 A, B). In contrast, neural changes during “disengaged” listening (when the animals were not performing the task) were much smaller, though statistically significant (Fig. 2C). The authors argue that the small changes in the disengaged state reflect bottom-up changes in neural processing of AM, while the larger changes in the engaged state reflect top-down modulation of AC activity. These demonstrations of context specificity are reminiscent of Crist et al.’s (2001) finding (discussed in Section 2) that neural changes in V1 of monkeys trained on a bisection discrimination task were seen only when the animals were actually performing the task. Together these studies suggest that the substrate of perceptual learning is not a permanent change in cortical circuitry, but a change that is gated in a task- or context-dependent fashion (see Section 9.2 for further discussion).

Evidence suggesting task-specificity of the changes underlying perceptual learning is also provided by psychophysical studies of learning on sound localization tasks (see Wright and Zhang (2006) for review). Hofman et al. (1998) reported that vertical localization by adult human subjects was massively disrupted immediately after insertion of bilateral moulds that modified the spectral shape cues to sound source elevation, but improved greatly over 3–6 weeks experience with the moulds (Fig. 3). Remarkably, however, participants were immediately able to perform at the levels exhibited prior to the experiment when the moulds were removed (points marked C in Fig. 3). This finding suggests that experience with the moulds resulted in a new representation of auditory space (or the mapping of the head-related transfer function onto auditory space), but that the old and the new representation coexisted and the subject could switch rapidly between them.

FIG 3 ABOUT HERE

A similar rapid switching between auditory spatial representations or cues is suggested by studies in which animals and humans learned to localize sounds in the azimuth with one ear occluded. Kumpik et al. (2010) reported that insertion of an ear-plug in human participants produced a marked decrement in localization performance, but 7-8 days of training with the earplug in place resulted in marked improvement as participants learned to assign greater weight to monaural spectral cues. Immediately after removal of the earplug, when localization again depended primarily on interaural disparity cues, participants performed at normal levels. In a similar study of the effects of ear-plugging on azimuthal localization by ferrets, learning to localize with the earplug in place was associated with a brief but transient bias when the plug was removed, and re-plugging had a much less disruptive effect on performance than the initial plugging (Kacelnik et al., 2006). These authors discuss their results in terms of reweighting of different localization cues, a process akin to that postulated by Mollon and Danilova (1996), except that the immediacy of the shifts when the cues are restored suggests, as did Hofman et al.'s results, that both weighting systems are retained. All three studies indicate that the learning associated with changed localization cues does not involve a permanent change in sensory processing circuitry.

Changes in AI response characteristics have been reported as correlates of perceptual learning on a number of other stimulus dimensions. As noted above, Polley et al. (2006) reported that in rats trained on a level discrimination task the relative area containing neurons with best level close to the target level was enlarged. Keeling et al. (2008) trained cats to discriminate differences between vowel-like spectra; in cats that learned the task, improvement in discrimination thresholds was associated with narrowing of the transfer functions for spectral envelope frequencies and narrowing of pure-tone tuning curves in AI. Improved temporal responses in AI of rats trained in a "sound maze" in which approach to reward was signalled by an increase in the repetition rate of noise pulses has also been described in terms of perceptual learning (Bao et al., 2004). However, although the rats' performance on the maze improved, no evidence was presented for improvements in their ability to detect or discriminate between different repetition rates. Thus, although this study provides compelling evidence of learning-related plasticity in AC temporal processing mechanisms, it is not clear that this plasticity was associated with perceptual learning as defined in Section 1.

3.2. Imaging studies in humans

Changes in auditory cortical (and subcortical) activity in humans associated with perceptual learning have been investigated using a range of electrophysiological and imaging techniques. In addition to experimental studies of learning on a range of psychophysical tasks, the effects of musical training have emerged as a valuable model of auditory learning (see Herholz and Zatorre (2012) and Strait and Kraus (2014) for reviews).

3.2.1. Standard psychoacoustic tasks

Changes in AC event-related potentials (ERPs) measured by either electro-or magneto-encephalography have been reported in association with perceptual learning on a range of tasks, including pure-tone frequency and more complex pitch and vowel discrimination tasks. The most commonly reported effect of training is an increase in the amplitude and a decrease in the latency of early response components (those with latencies in the 100- to 200-ms range) (e.g., Bosnyak et al., 2005; Carcagno and Plack, 2011b; Reinke et al., 2003; Tong et al., 2009; van Wassenhove and Nagarajan, 2007). In most of these studies, training took place over a number of days (or weeks), but Alain et al. (2007) reported increases in early evoked responses (and in performance) in the first hour of training on a vowel identification task. Most of these studies did not use an untrained control group however, and Sheehan et al. (2005), who also saw enhancement of the P2 component of the ERP in an untrained control group, have suggested that this particular enhancement might reflect exposure to the stimulus rather than learning *per se*.

There have also been reports of decreases in the amplitude of ERP response components (Ben-David et al., 2011; Cansino and Williamson, 1997) and in hemodynamic responses in the AC (Jäncke et al., 2001) in association with auditory perceptual learning. As with the animal perceptual learning (and behavioral conditioning) data, it is likely that the specific pattern of response changes observed depends on details of the task and its attentional demands.

It should also be noted that the precise nature of the neural changes underlying changes in ERPs cannot be determined with certainty – the changes reported in these studies could reflect changes in synchrony and/or in the number and/or the response strength of activated neurons. Furthermore, the particular cortical field or fields in which the changes occur cannot be specified with certainty, because the response components that exhibit changes have multiple generators, including (but not restricted to) primary and belt areas of AC.

Although most studies of the neural correlates of auditory perceptual learning have concentrated on the cortex, there is also substantial evidence for changes in the brainstem. Increases in the strength and the accuracy of pitch encoding in the brainstem frequency-following response (FFR) have been reported in association with perceptual learning on pitch discrimination and speech-in-noise tasks (e.g., Carcagno & Plack, 2011a; Song et al., 2008, 2012). The observed changes in the FFR could reflect changes intrinsic to brainstem and midbrain structures, cortical contributions to the FFR (Coffey et al., 2017; Tichko and Skoe, 2017), and/or corticofugal modulation of brainstem and midbrain generators. Compelling evidence for the involvement of efferent systems in auditory perceptual learning is provided by the fact that increases in medial olivocochlear bundle activity have been reported in association with improved performance on a speech-in-noise discrimination task (de Boer and Thornton, 2008).

3.2.2. Musical training.

Musical training has emerged as a valuable model for the study of learning-related plasticity, but the information it provides on auditory perceptual learning is qualified by two considerations. One is that it almost always involves the acquisition of a sensori-motor rather than a purely perceptual skill, and the sensory components are typically multimodal (auditory, visual and somatosensory). The second is that, other than in (rare) purely experimental studies, there is no way of carrying out the control procedures necessary to separate the effects of learning from the effects of exposure to the acoustic stimuli involved in the training.

The early studies in this area were almost all cross-sectional, comparing groups of trained musicians or musicians at various stages of training with non-musicians. These studies demonstrated both structural and functional differences between the ACs of musicians and non-musicians, viz., greater grey matter volume (right-lateralized) and cortical thickness, greater amplitude of auditory evoked responses to musical tones, and stronger functional connectivity at rest between auditory and premotor cortex, in the former (see Herholz and Zatorre (2012), Palomar-García et al. (2017), Pantev and Herholz (2011), and Zatorre et al. (2012) for reviews). Musical training is also associated with stronger FFRs to spectrally complex sounds (e.g., speech sound harmonics; see Strait and Kraus (2014) for review) (Fig. 4). These data are compelling, but their interpretation is qualified by the fact that they are correlational, and that the direction of causality cannot be determined. It might be that people with these characteristics are attracted to musical careers, rather than that the musical training produces the changes. A recent study of monozygotic twins discordant for musical training (piano practice) revealed greater cortical thickness in the auditory-motor network of the left hemisphere in the musically active twins (de Manzano and Ullén, 2018), indicating that at least some structural differences are attributable to the training rather than to genetic predisposition The fact that the magnitude of the difference between musicians and non-musicians in the FFR increases with age (Fig. 4) also suggests that duration of training is the critical factor. Finally, the enhancement of evoked responses can be specific to the timbre of the musician's specific instrument (e.g., Pantev et al., 2001), an observation that makes an interpretation in terms of predisposition less likely, but does not rule it out.

FIG 4 ABOUT HERE

Many of the apparent effects of musical training on brain structure and on cortical and brainstem responses reported in cross-sectional studies have been confirmed in longitudinal studies, the majority of which have used children participating in various musical training programs, and therefore describe changes that reflect developmental plasticity (e.g., Fujioka et al., 2006; Hyde et al., 2009; Habibi et al., 2016, 2017; Tierney et al., 2013, 2015) Although such longitudinal studies provide stronger evidence for the dependence of the observed changes on training, they do not allow control for the effects of exposure to the stimuli encountered in the course of training.

Experimental studies of musical training make it possible to include such controls, but there appears to have been only one study in which this has been done. Lappe et al. (2008) trained a group of non-musicians over two weeks to play a musical sequence on the piano; a control group listened to and made judgements about the music that had been played by participants in the experimental group. Both groups showed improved discrimination of errors in the chord progression used for training, but the training group showed larger changes in the magnetoencephalographically-recorded mismatch negativity

Two other experimental studies of musical training have generated results of interest with respect to both the neural consequences of such training and the neural correlates of a predisposition to benefit from training. Zatorre et al. (2012a) trained adult non-musicians on a micromelody discrimination task (i.e., on a purely sensory musical task); the participants' performance improved with training and the improvement generalized to untrained frequencies. Functional magnetic resonance imaging (fMRI) revealed that training resulted in a *decrease* in the extent to which AC activity was modulated by the size of pitch intervals, and in an overall decrease in superior temporal gyrus responses to all stimuli and increased activity in a region within the right dorsolateral frontal cortex. The authors interpreted the decreased AC activity in terms of more efficient encoding within regions sensitive to pitch patterns, such that fewer neural resources were required to process the same information as a consequence of training. This proposed mechanism is in marked contrast to the recruitment models considered in previous sections. Zatorre et al. (2012a) also reported that AC activity reflected a predisposition to benefit from training: individuals with greater pre-training cortical sensitivity to the size of pitch intervals learned more rapidly. Further fMRI evidence for activity reflecting a predisposition for learning is provided by Herholz et al.'s (2016) report that pre-training activity in right AC and right hippocampus predicted the subsequent learning rate of non-musicians who received six weeks of piano training. That training resulted in increases in melody-evoked activity in premotor and fronto-parietal cortex and in the cerebellum, which were interpreted in terms of storage of auditory-motor associations, but no changes in auditory regions. In providing evidence of AC neural correlates of a predisposition to benefit from musical training, these studies also support the concern that predisposition is a confounding factor in longitudinal studies of the effects of such training.

Given the critical role of pitch in most forms of music, it is not surprising that trained musicians exhibit superior pitch discrimination (e.g., Micheyl et al., 2006) and more robust encoding of linguistic pitch patterns in the FFR (e.g., Wong et al., 2007) than non-musicians. It is less clear whether the advantages of musical training generalize to tasks that are less closely related to music. Parbery-Clark et al. (2011) reported better speech-in-noise perception in musically-trained listeners, a finding confirmed in a longitudinal study by Slater et al. (2015). However, this finding was not confirmed by Ruggles et al. (2014) or Boebinger et al. (2015). Zendel and Alain (2012) saw a difference only in older listeners, suggesting that musical training might reduce age-related decline in the central processing mechanisms involved in this important aspect of speech perception.

In summary, the musical training literature has provided a wealth of evidence on plasticity in auditory cortical and subcortical structures and in auditory-motor networks. However, the information this research has provided on auditory perceptual learning *per se* and its mechanisms is qualified by the fact that most training is multimodal and sensori-motor in nature, and by the relative paucity of experimental studies allowing the control of confounding variables. The wide range of changes observed in AC as a result of musical training, like that seen in experimental studies of auditory perceptual learning in animals, suggests that the nature of the underlying changes depends critically on a number of task-related variables.

4. Perceptual learning and language acquisition

Perhaps the most remarkable form of developmental auditory perceptual learning is the process by which human infants learn to perceive the phonetic units of speech. At birth, infants have the capacity to detect the phonetic contrasts in all languages, but in the first year of life their ability to discriminate native language contrasts increases, while the ability to perceive contrasts in non-native language decreases (see Kuhl and Rivera-Gaxiola (2008) for review) This aspect of what has been termed native language neural commitment (Kuhl, 2004) is reflected in differences in ERPs in response to native and non-native contrasts in infants at 7.5 months of age (Kuhl et al., 2008). Early native-language phonetic learning is also reflected in the FFR in adults (Intartaglia et al., 2016).

Learning phonetic contrasts in a non-native language as an adult is also, of course, an instance of auditory perceptual learning. Such learning is associated with changes in activity in auditory and language cortical areas (e.g., Golestani, and Zatorre, 2004), and features of AC anatomy predict the ability to learn non-native language contrasts (Golestani et al. 2007).

5. Perceptual learning and auditory prostheses

A clinically important example of brain plasticity is provided by the improvements in the speech perception performance of post-lingually deaf cochlear implantees over the months and years following implantation (e.g., Tyler et al. 1997; Fu and Galvin, 2007). Plasticity in a number of hearing- and language-related networks undoubtedly contributes to this improvement, including the effects of reactivation of the auditory pathway by intracochlear electrical stimulation (see Fallon et al. (2008) for review). In both adult (e.g., Pantev et al., 2006; Sandmann et al., 2015) and child (see Kral and Sharma (2012) for review) implantees, auditory cortical evoked responses increase in amplitude, decrease in latency, and achieve normal or near-normal configurations over months following the implant being switched on. The implantee's semantic and lexical knowledge is also critical to interpretation of the degraded input provided by the implant. In a clinical context of this sort, it is difficult to separate the contributions of the various factors, but the likely contribution of perceptual learning was recognized early (Watson, 1991), and is supported by the success of post-implantation training regimes (e.g., Fu and Galvin, 2007). However, the extent to which these benefits reflect improved auditory perception or improved language-related cognitive skills is unclear. Moore and Shannon (2009) have argued that training to use the information provided by an implant may be as important as further improvements in implant technology.

There have been numerous studies of the effects of computer-based auditory training on speech perception and related auditory abilities in adults with age-related hearing loss (many of them hearing-aid users). Although the quality of many of these studies was questioned in a systematic review of this literature (Henshaw and Ferguson, 2013), benefits have been documented in a number of well-controlled studies (e.g., Anderson et al., 2013; Ferguson and Henshaw; 2015; Whitton et al., 2017). As with cochlear implant training regimes, the relative roles of auditory perceptual and broader cognitive changes is unclear, although an association of improvement with enhanced FFR responses (Anderson et al., 2013) indicates that auditory system plasticity is involved in at least some cases.

6. Similarities between perceptual and motor learning

Perceptual and motor learning are both forms of procedural learning, and it has been argued that they exhibit analogous properties and share common general mechanisms (e.g., Censor, 2013; Censor et al., 2012). Although motor performance *per se* can obviously improve with training, the fact that most skills have both sensory and motor components suggests that skill learning involves a complex interaction between three types of learning: perceptual, sensorimotor integration, and motor (see Makino et al. (2016) for discussion). Nevertheless, the nature of the changes in motor cortex associated with motor learning are of interest with respect to the mechanisms of perceptual learning. In rodents trained on skilled forelimb motor tasks, the area of the forelimb representation (or the population of excited neurons) in motor cortex is initially increased, but subsequently reverts to normal while the motor skill is retained, (e.g., Molina-Luna et al., 2008; Peters et al., 2014; Pruitt et al., 2016). This finding is in accord with the expansion-normalization hypothesis of Reed et al. (2011). In a recent review, Makino et al. (2016) state that an expansion of the neural ensembles in motor cortex is characteristic of the early stages of motor learning, and suggest that it allows the sampling of a number of new circuit options. In a study by Tennant et al. (2012), expansion and renormalization of the motor cortex forelimb representation was seen in young mice who learned the reaching task, but aged mice learned the task equally well without any change in area of representation or in movement threshold. This study therefore introduces an additional variable (the effects of age) that does not seem to have been studied in the perceptual learning literature.

Wenger et al. (2017) have argued that the expansion – renormalization hypothesis extends to structural brain changes (i.e., increases in grey matter) associated with motor learning. As they point out, expansion of brain volume with each skill acquired is implausible, as it implies a perpetual process of increase in brain volume.

Perhaps the most important finding with respect to the mechanisms of motor learning is that such learning is associated with an increase in spine density on the dendrites of cortico-spinal motor neurones, and that different motor skills are associated with changes in different sets of synapses (e.g., Wang et al., 2011; Xu et al., 2009). The implications of these and related results for the mechanisms of perceptual learning are considered further in Section 9.2.

7. Specificity and generalization of perceptual learning

As noted in Section 1, the upsurge of interest in perceptual learning in the last 30 years or so was prompted by evidence for the remarkable specificity of some forms of VPL. More recently there has been an emphasis on the conditions on which such learning can generalize (e.g. Censor, 2013; Fahle, 2005; Sagi, 2011; Wright & Zhang, 2009b). This emphasis has been driven in part by the development of cognitive and sensory rehabilitation training regimes, which itself has been driven by developments in the study of brain plasticity. Clearly, the value of any training regime depends on the extent to which improvements generalize to stimuli and tasks other than those used in training (e.g., Jacoby and Ahissar, 2015; Simons et al., 2016). In the case of auditory training, the clinical implications of generalization relate to training of cochlear implantees and hearing-aid users (Section 5), and to the treatment of communication disorders (e.g. Moore et al. 2009). In general, the conditions under which perceptual learning generalizes are not well understood; in their review of the generalization of auditory perceptual learning, Wright and Zhang (2009b) concluded that the available data did not allow the identification of any simple rule that could be used to predict the pattern of generalization on a given task. With respect to mechanisms, however, it seems clear that generalization would not be expected if the underlying neural changes were restricted to those neurons activated by the training stimuli.

8. Consolidation of perceptual learning

A remarkable early finding in VPL was that improvements with training on a texture discrimination task were strongly dependent on rapid eye movement (REM) sleep (Karni et al. 1994). Although some forms of auditory perceptual learning occur rapidly (within the first hour of training) (e.g., Alain et al., 2007; Hawkey et al., 2004), there is also evidence that some forms require a between-sessions (or overnight) consolidation phase. Wright and Sabin (2007) reported that in subjects trained on either a frequency discrimination or temporal-interval task, improvements occurred *between* the daily training sessions, rather than within sessions. A similar need for consolidation between sessions is indicated by Kumpik et al.'s (2010) finding that the improvement in azimuthal sound localization seen over a number of days of training in participants with an ear blocked was not seen when the same number of trials were given as a single block on one day. Sleep was not manipulated in those studies, so it is not clear whether the between-sessions delay or sleep during that delay was the critical factor. Roth et al. (2005) reported that improvements on a task requiring identification of consonant-vowel stimuli in noise were delayed for some hours after training but that sleep during the delay was not necessary. Why performance on some tasks improves during the initial training session while improvement on others requires post- or between-session consolidation is not clear. Fenn et al. (2003) reported an interesting pattern of results in a study of perceptual learning on a synthetic speech recognition task: performance improved immediately after training, was degraded over a day's retention interval, but completely recovered following sleep. Given that consolidation (and its dependence on sleep) is a characteristic of all or most forms of learning (e.g., Stickgold, 2005), it would be remarkable if it were not important for auditory perceptual learning.

9. Mechanisms of perceptual learning

9.1. The need for causal analysis

The overwhelming majority of studies that have attempted to identify the neural substrates of auditory perceptual learning (and of perceptual learning in other modalities) have employed between-subject experimental designs, comparing neural response characteristics in groups of trained and untrained participants. Although such studies can establish the existence of associations between brain changes and learning, they do not allow the detailed nature of the relationship between the two changes to be determined. Comparisons between the neural and perceptual changes within individual subjects, such as those made by Caras and Sanes (2017) and illustrated in Fig. 2A and B, provide valuable evidence on the quantitative relationship between neural changes and perceptual performance.

It should also be emphasised that the evidence from both between- and within-subject comparisons of these sorts is correlational, and does not establish that the neural changes described are responsible for the perceptual improvement. Although the neural changes described in most studies are plausible candidates for the cause of the improved perceptual performance, it is also possible that they are simply epiphenomenal, or that they merely have a permissive role. Direct evidence that a particular neural change is necessary and/or sufficient for the perceptual improvement is required to establish that they have a causal role. As noted in an earlier discussion of the changes in auditory cortical response characteristics associated with auditory association learning (Irvine, 2007), establishing the causal linkage presents a daunting experimental challenge.

As noted in that earlier paper, lesion studies do not adequately address this issue. In the context of perceptual learning, they do not allow the roles of the lesioned auditory (cortical or subcortical) structure in the processing of the auditory information and in the perceptual improvement to be separated. An elegant method of separating these effects is provided by Caras and Sanes's (2017) study of perceptual learning on an AM detection task by gerbils. They found that bilateral AC infusion of a carefully titrated low dose of muscimol blocked improvement on the task without (in most animals) interfering with performance of the task (i.e. with AM detection). This result establishes that an enhanced level of activity in ACX is necessary for perceptual learning, but it does not establish the necessity of a particular form of changed activity. In general, it is difficult to see how loss-of-function studies of this sort could establish the necessity of particular changes in neural activity.

It is perhaps more likely that gain-of function studies might provide such evidence. In this context, the ability of optogenetic techniques to selectively increase or decrease the activity of particular classes of neurons in a given structure (Deisseroth, 2015) offers particular promise. Evidence that photo-activation or -suppression of excitatory or inhibitory neurons in AC can influence auditory discrimination acuity (Aizenberg et al., 2015; Briguglio et al., 2018) suggests that these techniques could provide valuable information on causal relationships underlying auditory perceptual learning.

9.2. The stability-plasticity dilemma and a multiplexing model of perceptual learning

The most commonly reported neural change associated with auditory perceptual learning is an increase in the amplitude of responses to the training stimuli, in some cases associated with recruitment and an increase in cortical representational area. As discussed in Section 3.1.1, issues related to the stability – plasticity dilemma suggest that it is unlikely that recruitment and expansion of the area of representation constitute the permanent substrate of perceptual learning. A possible solution to this issue is provided by the expansion – renormalization hypothesis (Reed et al., 2011), and a number of studies of motor learning (Section 6) and of somatosensory perceptual learning (e.g., Albieri et al., 2015) provide support for the view that an expansion in representational area is a transient component of the neural changes underlying perceptual learning. As described in section 3.1.1, however, there was no evidence for transient increases in representational area (Witte and Kipke, 2005) or of the perceptual consequences of such increases (Brown et al., 2004) in two studies of auditory learning by cats on frequency discrimination tasks. It should also be noted that the expansion – renormalization hypothesis leaves open the nature of the permanent changes in circuitry that remain after the renormalization process.

There have also been a number of reports of *decreased* activity or activation in sensory cortex associated with perceptual learning, interpreted as reflecting more efficient processing of the relevant stimuli. Although these studies point to changes in sensory cortex, they also presuppose selection of the most efficient circuitry by some higher-level process, and are thus reminiscent of late-stage models as described by Watanabe and Sasaki (2015).

An alternative account of the substrates of auditory perceptual learning is suggested by evidence for the task- or context- dependency of the changes associated with such learning (e.g., von Trapp et al., 2016; Caras and Sanes, 2017) and for the ability of listeners to switch rapidly between different cues or representations of auditory space (Section 3.1.2). The context specificity of the changes associated with some forms of VPL prompted Li and Gilbert (2008) to propose that the critical process underlying learning relating to a particular stimulus attribute is that a particular subset of the diverse inputs to a given cortical neuron related to that attribute is gated by top-down influences. A similar dynamic multiplexing by individual cortical neurons has been suggested by Fritz et al. (2007) to underlie the task dependency of changes in AC spectro-temporal receptive fields produced by associative learning and attention. A critical feature of this hypothesis is that it does not involve recruitment of neurons into the activated population; learning related to multiple attributes can be represented in the same neurons, and the organism can switch rapidly between different sets of inputs depending on the task being performed. This postulated mechanism also avoids the stability – plasticity dilemma, because the changes in synaptic weights are restricted to a limited set of task-specific synapses.

This multiplexing hypothesis involves two major assumptions: that a given perceptual task involves only a subset of the synapses on a given cortical neuron, and that there are (feedback) mechanisms that allow rapid switching between different subsets of inputs. Neither of these assumptions has been examined experimentally in the context of auditory perceptual learning, but evidence from other areas establishes their plausibility.

Learning is likely to involve both strengthening of existing synapses and the formation of new synapses, and evidence for selectivity in both processes is provided by studies of motor leaning in mice. Cichon and Gao (2015) reported that different motor learning tasks caused long-lasting potentiation of post-synaptic spines on different apical tuft branches of pyramidal neurons in mouse motor cortex. Two-photon imaging studies of mice trained on motor tasks indicate that learning is associated with the formation of new dendritic spines on motor cortex pyramidal neurons and that different skills are encoded by different sets of synapses in the same neurons (e.g., Xu et al., 2009; Yang et al., 2009, 2014). With respect to the consolidation issues discussed in Section 8, it is of interest that branch-specific spine formation was promoted by post-training (non-REM) sleep (Yang et al., 2014). In a remarkable recent study, the causal role of potentiation of specific subsets of spines was demonstrated by disruption of motor learning on one task by optical shrinking of the spines potentiated by that task but not of spines potentiated by another task (Hayashi-Takagi et al. 2015).

Although synapse potentiation or formation in association with auditory perceptual learning has not yet been investigated, auditory fear conditioning has been shown to result in an increase in spine formation in mouse AC (Moczulska et al., 2013). The fact that inputs to spines on individual dendrites of AI neurons cover a remarkably diverse range of frequencies (Chen et al., 2011) indicates that, at least in the frequency domain, the substrate to allow selection of subsets of inputs is present.

With respect to switching between different sets of inputs, a modelling study by Vogels & Abbott (2009) suggests ways in which individual neurons can switch their responsiveness between various input signals by adjustment of excitatory – inhibitory balance. Such adjustment was demonstrated in a recent study in which mice switched between a task requiring recognition of a target tone of a particular frequency and passive listening to the same (target and foil) tones (Kuchibhotla et al., 2017). During the recognition task, multiple inhibitory interneuron types were activated, adjusting inhibitory synaptic inputs and thereby modulating neuronal output. Although this type of context switching is rather different from switching between different perceptual tasks, it is reasonable to assume that similar mechanisms would be involved in the two forms of context dependency. Vogels and Abbott (2009) suggested that the cholinergic modulatory system might be the mechanism by which excitatory – inhibitory balance was adjusted, and Froemke (2015) has reviewed evidence that this and other neuromodulatory systems play a critical role in regulating changes in excitatory – inhibitory balance that are associated with many forms of plasticity. Kuchibhotla et al. (2017) also presented evidence that supports this proposal: cholinergic axons increased activity and directly depolarized inhibitory neurons during performance of the recognition task. Although the role of the cholinergic and other neuromodulatory systems in the neural changes associated with auditory perceptual learning has received relatively little attention, these systems are known to be critically involved in most of other forms of auditory cortical plasticity (see Metherate (2011) and Edeline (2012) for reviews). Froemke et al.'s (2013) demonstration that pairing acoustic stimuli with stimulation of the cholinergic fibers originating in the NB in awake rats led to improved detection and recognition of the paired acoustic stimuli strongly implies a role of the cholinergic system in auditory perceptual learning.

At a more general level, and in accordance with these suggestions, Roelfsema and Holtmaat (2018) have recently proposed that neuromodulatory influences on sensory cortices and feedback processes that tag particular subsets of synapses and gate their plasticity combine to improve the functioning of sensory cortical networks and optimize behavioral performance.

10. Conclusions

As summarised in this review, the last 25 years or so have generated a wealth of psychophysical and neurophysiological information on auditory perceptual learning. This evidence indicates that both the characteristics of the learning and the nature of the associated neural changes are highly attribute- and task-dependent, and thus far we have only a limited grasp of the general principles underlying such learning. As noted above, there is a pressing need for studies that establish the causal relationship between neural changes and improved perceptual performance. Nevertheless, perceptual learning and other forms of auditory system plasticity have led to a major reconceptualization of the AC, and led to recognition of its place as a component of complex distributed networks. This reconceptualization applies to all sensory cortices, and is elegantly and succinctly expressed by Dosher and Li's (2017; p. 343) statement concerning VPL: "Originally seen as a manifestation of plasticity in the primary visual cortex, perceptual learning is more readily understood as improvements in the function of brain networks that integrate processes including sensory representations, decision, attention, and reward and balance plasticity with system stability".

Acknowledgements

I am grateful to an anonymous reviewer for very helpful comments on an earlier version of this manuscript, and to James Fallon and the authors from whose papers figures have been modified for assistance in producing the figures. The author's research in this area has been funded by the National Health and Medical Research Council of Australia. The Bionics Institute acknowledges the support it receives from the Victorian Government through its Operational Infrastructure Support Program.

Abbreviations

AC	auditory cortex
AI	primary auditory cortex
AM	amplitude-modulated
BF	best frequency
CF	characteristic frequency
ERP	event-related potential
FFR	frequency-following response
NB	nucleus basalis
V1	primary visual cortex
VPL	visual perceptual learning

References

Ahissar, M., Nahum, M., Nelken, I. & Hochstein, S. (2009) Reverse hierarchies and sensory learning. *Philos. Trans. R. Soc. B - Biol. Sci.*, **364**, 285-299.

Alain, C., Snyder, J.S., He, Y. & Reinke, K.S. (2007) Changes in auditory cortex parallel rapid perceptual learning. *Cer. Cortex*, **17**, 1074-1084.

Albieri, G., Barnes, S.J., de Celis Alonso, B., Cheetham, C.E.J., Edwards, C.E., Lowe, A.S., Karunaratne, H., Dear, J.P., Lee, K.C. & Finnerty, G.T. (2015) Rapid bidirectional reorganization of cortical microcircuits. *Cereb. Cortex*, **25**, 3025-3035.

Aizenberg, M., Mwilambwe-Tshilobo, L., Briguglio, J.J., Natan, R.G. & Geffen, M.N. (2015) Bidirectional regulation of innate and learned behaviors that rely on frequency discrimination by cortical inhibitory neurons. *PLoS Biology*, **13**, e1002308.

Amitay, S., Irwin, A. & Moore, D.R. (2006) Discrimination learning induced by training with identical stimuli. *Nat. Neurosci.*, **9**, 1446-1448.

Anderson, S., White-Schwoch, T., Parbery-Clark, A. & Kraus, N. (2013) Reversal of age-related neural timing delays with training. *Proc. Nat. Acad. Sci. USA*, **110**, 4357-4362.

Bao, S., Chang, E.F., Woods, J. & Merzenich, M.M. (2004) Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. *Nat. Neurosci.*, **7**, 974-981.

Ben-David, B.M., Campeanu, S., Tremblay, K.L. & Alain, C. (2011) Auditory evoked potentials dissociate rapid perceptual learning from task repetition without learning. *Psychophysiol.*, **48**, 797-807.

Boebinger, D., Evans, S., Rosen, S., Lima, C.F., Manly, T. & Scott, S.K. (2015) Musicians and non-musicians are equally adept at perceiving masked speech. *J. Acoust. Soc. Am.*, **137**, 378-387.

Bosnyak, D.J., Eaton, R.A. & Roberts, L.E. (2004) Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cer. Cortex*, **14**, 1088-1099.

Briguglio, J.J., Aizenberg, M., Balasubramanian, V. & Geffen, M.N. (2018) Cortical neural activity predicts sensory acuity under optogenetic manipulation. *J. Neurosci.*, **38**, 2094.

Brown, M., Irvine, D.R.F. & Park, V.N. (2004) Perceptual learning on an auditory frequency discrimination task by cats: association with changes in primary auditory cortex. *Cer. Cortex*, **14**, 952-965.

Buran, B.N., von Trapp, G. & Sanes, D.H. (2014) Behaviorally gated reduction of spontaneous discharge can improve detection thresholds in auditory cortex. *J. Neurosci.*, **34**, 4076-4081.

Cansino, S. & Williamson, S.J. (1997) Neuromagnetic fields reveal cortical plasticity when learning an auditory discrimination task. *Brain Res.*, **764**, 53-66.

Caras, M.L. & Sanes, D.H. (2017) Top-down modulation of sensory cortex gates perceptual learning. *Proc. Nat. Acad. Sci. USA*, **114**, 9972-9977.

Carcagno, S. & Plack, C. (2011a) Subcortical plasticity following perceptual learning in a pitch discrimination task. *J. Assoc. Res. Otolaryngol.*, **12**, 89-100.

Carcagno, S. & Plack, C. (2011b) Pitch discrimination learning: specificity for pitch and harmonic resolvability, and electrophysiological correlates. *J. Assoc. Res. Otolaryngol.*, **12**, 503-517.

Censor, N. (2013) Generalization of perceptual and motor learning: A causal link with memory encoding and consolidation? *Neuroscience*, **250**, 201-207.

Censor, N., Sagi, D. & Cohen, L.G. (2012) Common mechanisms of human perceptual and motor learning. *Nat. Rev. Neurosci.*, **13**, 658-664.

Chen, X., Leischner, U., Rochefort, N.L., Nelken, I. & Konnerth, A. (2011) Functional mapping of single spines in cortical neurons in vivo. *Nature*, **475**, 501-505.

Cichon, J. & Gan, W.-B. (2015) Branch-specific dendritic Ca2+ spikes cause persistent synaptic plasticity. *Nature*, **520**, 180-185.

Coffey, E.B.J., Musacchia, G. & Zatorre, R.J. (2017) Cortical correlates of the auditory frequency-following and onset responses: EEG and fMRI evidence. *J. Neurosci.*, **37**, 830.

Crist, R.E., Li, W. & Gilbert, C.D. (2001) Learning to see: experience and attention in primary visual cortex. *Nat. Neurosci.*, **4**, 519 - 525

de Boer, J. & Thornton, A.R.D. (2008) Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *J. Neurosci.*, **28**, 4929-4937.

Deisseroth, K. (2015) Optogenetics: 10 years of microbial opsins in neuroscience. *Nat. Neurosci*, **18**, 1213-1225.

de Manzano, Ö. & Ullén, F. (2018) Same genes, different brains: Neuroanatomical differences between monozygotic twins discordant for musical training. *Cereb. Cortex*, **28**, 387-394.

Dosher, B.A. & Lu, Z.-L. (1999) Mechanisms of perceptual learning. *Vis. Res.* **39**, 3197-3221.

Dosher, B. & Lu, Z.-L. (2017) Visual perceptual learning and models. *Annu. Rev. Vis. Sci.*, **3**, 343-363.

Edeline, J.-M. (2012) Beyond traditional approaches to understand the functional role of neuromodulators in sensory cortices. *Front. Behav. Neurosci.* **6:45.**

Fahle, M. (2005) Perceptual learning: specificity versus generalization. *Curr. Opin. Neurobiol.*, **15**, 154-160.

Fallon, J.B., Irvine, D.R.F. & Shepherd, R.K. (2008) Cochlear implants and brain plasticity. *Hear. Res.*, **238**, 110-117.

Fenn, K.M., Nusbaum, H.C. & Margoliash, D. (2003) Consolidation during sleep of perceptual learning of spoken language. *Nature*, **425**, 614-616.

Ferguson, M.A. & Henshaw, H. (2015) Auditory training can improve working memory, attention and communication in adverse conditions for adults with hearing loss. *Front. Psychol.*, **6**.

Fritz, J.B., Elhilali, M., David, S.V. & Shamma, S.A. (2007) Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hear. Res.*, **229**, 186-203.

Froemke, R.C. (2015) Plasticity of cortical excitatory-inhibitory balance. *Ann. Rev. Neurosci.*, **38**, 195-219.

Froemke, R.C., Carcea, I., Barker, A.J., Yuan, K., Seybold, B.A., Martins, A.R.O., Zaika, N., Bernstein, H., Wachs, M., Levis, P.A., Polley, D.B., Merzenich, M.M. & Schreiner, C.E. (2013) Long-term modification of cortical synapses improves sensory perception. *Nat. Neurosci*, **16**, 79-88.

Fu, Q.-J. & Galvin, J.J. (2007) Perceptual learning and auditory training in cochlear implant recipients. *Trends Amplif.*, **11**, 193-205.

Fujioka, T., Ross, B., Kakigi, R., Pantev, C. & Trainor, L.J. (2006) One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, **129**, 2593-2608.

Gibson, E.J. (1953) Improvement in perceptual judgments as a function of controlled practice or training. *Psychol. Bull.*, **50**, 401-431.

Golestani, N., Molko, N., Dehaene, S., LeBihan, D. & Pallier, C. (2007) Brain structure predicts the learning of foreign speech sounds. *Cereb. Cortex*, **17**, 575-582.

Golestani, N. & Zatorre, R.J. (2004) Learning new sounds of speech: reallocation of neural substrates. *NeuroImage*, **21**, 494-506.

Habibi, A., Cahn, B.R., Damasio, A. & Damasio, H. (2016) Neural correlates of accelerated auditory processing in children engaged in music training. *Dev. Cog. Neurosci.*, **21**, 1-14.

Habibi, A., Damasio, A., Ilari, B., Veiga, R., Joshi, A.A., Leahy, R.M., Haldar, J.P., Varadarajan, D., Bhushan, C. & Damasio, H. (2017) Childhood music training induces change in micro and macroscopic brain structure: results from a longitudinal study. *Cereb. Cortex*, 1-12.

Han, Y.K., Kover, H., Insanally, M.N., Semerdjian, J.H. & Bao, S. (2007) Early experience impairs perceptual discrimination. *Nat. Neurosci.*, **10**, 1191-1197.

Hawkey, D.J.C., Amitay, S. & Moore, D.R. (2004) Early and rapid perceptual learning. *Nat. Neurosci.*, **7**, 1055-1056.

Hayashi-Takagi, A., Yagishita, S., Nakamura, M., Shirai, F., Wu, Y.I., Loshbaugh, A.L., Kuhlman, B., Hahn, K.M. & Kasai, H. (2015) Labelling and optical erasure of synaptic memory traces in the motor cortex. *Nature*, **525**, 333-338.

Henshaw, H. & Ferguson, M.A. (2013) Efficacy of individual computer-based auditory training for people with hearing loss: A systematic review of the evidence. *PLoS One*, **8**, e62836.

Herholz, S.C., Coffey, E.B.J., Pantev, C. & Zatorre, R.J. (2016) Dissociation of neural networks for predisposition and for training-related plasticity in auditory-motor learning. *Cereb. Cortex*, **26**, 3125-3134.

Herholz, S. C. & Zatorre, R.J. (2012) Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron*, **76**, 486-502.

Hofman, P.M., Van Riswick, J.G.A. & Van Opstal, A.J. (1998) Relearning sound localization with new ears. *Nat. Neurosci.*, **1**, 417-421.

Hyde, K.L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A.C. & Schlaug, G. (2009) Musical training shapes structural brain development. *J. Neurosci.*, **29**, 3019-3025.

Intartaglia, B., White-Schwoch, T., Meunier, C., Roman, S., Kraus, N. & Schön, D. (2016) Native language shapes automatic neural processing of speech. *Neuropsychologia*, **89**, 57-65.

Irvine, D.R.F. (2007) Auditory cortical plasticity: Does it provide evidence for cognitive processing in the auditory cortex? *Hear. Res.*, **229**, 158-170.

Irvine, D.R.F. (2018) Auditory system plasticity. *Hear. Res.,* In press.

Irvine, D.R.F., Martin, R.L., Klimkeit, E. & Smith, R. (2000) Specificity of perceptual learning in a frequency discrimination task. *J. Acoust. Soc. Am.*, **108**, 2964-2968.

Irvine, D., Brown, M., Martin, R. & Park, V. (2005) Auditory perceptual learning and cortical plasticity. In Koenig, R., Heil, P., Budinger, E., Scheich, H. (eds.) *The Auditory Cortex: A Synthesis of Human and Animal Research*. Lawrence Erlbaum, Mahwah, N.J., pp. 409-428.

Jacobs, R.A. (2009) Adaptive precision pooling of model neuron activities predicts the efficiency of human visual learning. *J. Vis.* **9(4):22**, 1-15.

Jacoby, N. & Ahissar, M. (2015) Assessing the applied benefits of perceptual training: Lessons from studies of training working-memory. *J. Vis.*, **15(10):6,** 1-13.

Jäncke, L., Gaab, N., Wustenberg, T., Scheich, H. & Heinze, H.-J. (2001) Short-term functional plasticity in the human auditory cortex: an fMRI study. *Cog. Brain Res.*, **12**, 479-485.

Jones, P.R., Moore, D.R., Amitay, S. & Shub, D.E. (2013) Reduction of internal noise in auditory perceptual learning. *J. Acoust. Soc. Am.*, **133**, 970-981.

Kacelnik, O., Nodal, F.R., Parsons, C.H. & King, A.J. (2006) Training-induced plasticity of auditory localization in adult mammals. *PLoS Biology*, **4**, 627-638.

Kamke, M.R., Brown, M. & Irvine, D.R.F. (2005) Basal forebrain cholinergic input is not essential for lesion-induced plasticity in mature auditory cortex. *Neuron*, **48**, 675-686.

Karni, A., Tanne, D., Rubenstein, B.S., Askenasy, J.J.M. & Sagi, D. (1994) Dependence on REM sleep of overnight improvement of a perceptual skill. *Science*, **265**, 679-682.

Keeling, M.D., Calhoun, B.N., Krüger, K., Polley, D.B. & Schreiner, C.E. (2008) Spectral integration plasticity in cat auditory cortex induced by perceptual training. *Exp Brain Res.*, **184**, 493-509.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A.A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D. & Hadsell, R. (2017) Overcoming catastrophic forgetting in neural networks. *Proc. Nat. Acad. Sci. USA,* **114**, 3521-3526.

Kral, A. & Sharma, A. (2012) Developmental neuroplasticity after cochlear implantation. *Trends Neurosci.*, **35**, 111-122.

Kuchibhotla, K.V., Gill, J.V., Lindsay, G.W., Papadoyannis, E.S., Field, R.E., Sten, T.A.H., Miller, K.D. & Froemke, R.C. (2016) Parallel processing by cortical inhibition enables context-dependent behavior. *Nat. Neurosci.*, **20**, 62-71.

Kuhl, P.K. (2004) Early language acquisition: cracking the speech code. *Nat Rev Neurosci*, **5**, 831-843.

Kuhl, P.K., Conboy, B.T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. & Nelson, T. (2008) Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Phil. Trans. Roy. Soc. B: Biol. Sci.*, **363**, 979-1000.

Kuhl, P. & Rivera-Gaxiola, M. (2008) Neural substrates of language acquisition. *Ann. Rev. Neurosci.*, **31**, 511-534.

Kumpik, D.P., Kacelnik, O. & King, A.J. (2010) Adaptive reweighting of auditory localization cues in response to chronic unilateral earplugging in humans. *J. Neurosci.*, **30**, 4883-4894.

Lappe, C., Herholz, S.C., Trainor, L.J. & Pantev, C. (2008) Cortical plasticity induced by short-term unimodal and multimodal musical training. *J. Neurosci.*, **28**, 9632.

Law, C.-T. & Gold, J.I. (2008) Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat Neurosci*, **11**, 505-513.

Li, W. & Gilbert, C.D. (2008) Perceptual learning. In Masland, R.H., Albright, T.D. (eds) *The Senses: A Comprehensive Reference, volume 2, Vision II.* Elsevier, pp. 303-328.

Makino, H., Hwang, E.J., Hedrick, N.G. & Komiyama, T. (2016) Circuit mechanisms of sensorimotor learning. *Neuron*, **92**, 705-721

Mermillod, M., Bugaiska, A. & Bonin, P. (2013) The stability-plasticity dilemma: investigating the continuum from catastrophic forgetting to age-limited learning effects. *Front. Psychol.*, **4**.

Metherate, R. (2011) Modulatory mechanisms controlling auditory processing. In Trussell, L.O., Popper, A.N., Fay, R. (eds) *Synaptic Mechanisms in the Auditory System* Springer, NY, pp. 187-202.

Micheyl, C., Delhommeau, K., Perrot, X. & Oxenham, A.J. (2006) Influence of musical and psychoacoustical training on pitch discrimination. *Hear. Res.*, **219**, 36-47.

Micheyl, C., McDermott, J.H. & Oxenham, A.J. (2009) Sensory noise explains auditory frequency discrimination learning induced by training with identical stimuli. *Atten. Percept. Psychophys.*, **71**, 5-7.

Moczulska, K.E., Tinter-Thiede, J., Peter, M., Ushakova, L., Wernle, T., Bathellier, B. & Rumpel, S. (2013) Dynamics of dendritic spines in the mouse auditory cortex during memory formation and memory recall. *Proc. Nat. Acad. Sci.*, **110**, 18315-18320.

Molina-Luna, K., Hertler, B., Buitrago, M.M. & Luft, A.R. (2008) Motor learning transiently changes cortical somatotopy. *NeuroImage*, **40**, 1748-1754.

Mollon, J.D. & Danilova, M.V. (1996) Three remarks on perceptual learning. *Spat. Vis.*, **10**, 51-58.

Moore, D.R. & Shannon, R.V. (2009) Beyond cochlear implants: awakening the deafened brain. *Nat. Neurosci*, **12**, 686-691.

Moore, D.R., Halliday, L.F. & Amitay, S. (2009) Use of auditory learning to manage listening problems in children. *Phil. Trans.Roy. Soc.B: Biol. Sci.*, **364**, 409.

Ohl, F.W. & Scheich, H. (1996) Differential frequency conditioning enhances spectral contrast sensitivity of units in auditory cortex (field Al) of the alert Mongolian gerbil. *Eur. J. Neurosci.*, **8**, 1001-1017.

Ortiz, J.A. & Wright, B.A. (2009) Contributions of procedure and stimulus learning to early, rapid perceptual improvements. *J. Exp. Psychol.-Hum. Percept. Perform.*, **35**, 188-194.

Palomar-García, M.-Á., Zatorre, R.J., Ventura-Campos, N., Bueichekú, E. & Ávila, C. (2017) Modulation of functional connectivity in auditory–motor networks in musicians compared with nonmusicians. *Cer. Cortex*, **27**, 2768-2778.

Pantev, C., Dinnesen, A., Ross, B., Wollbrink, A. & Knief, A. (2006) Dynamics of auditory plasticity after cochlear implantation: A longitudinal study. *Cer. Cortex*, **16**, 31-36.

Pantev, C. & Herholz, S.C. (2011) Plasticity of the human auditory cortex related to musical training. *Neurosci. Biobehav. Rev.*, **35**, 2140-2154.

Pantev, C., Roberts, L.E., Schulz, M., Engelien, A. & Ross, B. (2001) Timbre-specific enhancement of auditory cortical representations in musicians. *NeuroReport*, **12**, 169-174.

Parbery-Clark, A., Strait, D.L. & Kraus, N. (2011) Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, **49**, 3338-3345.

Peters, A.J., Chen, S.X. & Komiyama, T. (2014) Emergence of reproducible spatiotemporal activity during motor learning. *Nature*, **510**, 263.

Phillips, D.P., Semple, M.N., Calford, M.B. & Kitzes, L.M. (1994) Level-dependent representation of stimulus frequency in cat primary auditory cortex. *Exp. Brain Res.*, **102**, 210-226.

Polley, D.B., Steinberg, E.E. & Merzenich, M.M. (2006) Perceptual learning directs auditory cortical map reorganization through top-down influences. *J. Neurosci.*, **26**, 4970-4982.

Pruitt, D.T., Schmid, A.N., Danaphongse, T.T., Flanagan, K.E., Morrison, R.A., Kilgard, M.P., Rennaker, R.L. & Hays, S.A. (2016) Forelimb training drives transient map reorganization in ipsilateral motor cortex. *Behav. Brain Res.*, **313**, 10-16.

Raiguel, S., Vogels, R., Mysore, S.G. & Orban, G.A. (2006) Learning to see the difference specifically alters the most informative V 4 neurons. *J. Neurosci.*, **26**, 6589-6602.

Recanzone, G.H., Schreiner, C.E. & Merzenich, M.M. (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.*, **13**, 87-103

Reed, A., Riley, J., Carraway, R., Carrasco, A., Perez, C., Jakkamsetti, V. & Kilgard, Michael P. (2011) Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron*, **70**, 121-131.

Reinke, K.S., He, Y., Wang, C. & Alain, C. (2003) Perceptual learning modulates sensory evoked response during vowel segregation. *Cog. Brain Res.*, **17**, 781-791.

Roelfsema, P.R. & Holtmaat, A. (2018) Control of synaptic plasticity in deep cortical networks. *Nat. Rev. Neurosci.*, **19**, 166.

Roth, D.A.-E., Kishon-Rabin, L., Hildesheimer, M. & Karni, A. (2005) A latent consolidation phase in auditory identification learning: Time in the awake state is sufficient. *Learn. & Mem.*, **12**, 159-164.

Ruggles, D.R., Freyman, R.L. & Oxenham, A.J. (2014) Influence of musical training on understanding voiced and whispered speech in noise. *PLoS One*, **9**, e86980.

Sagi, D. (2011) Perceptual learning in *Vision Research*. *Vis. Res.*, **51**, 1552-1566.

Sandmann, P., Plotz, K., Hauthal, N., de Vos, M., Schönfeld, R. & Debener, S. (2015) Rapid bilateral improvement in auditory cortex activity in postlingually deafened adults following cochlear implantation. *Clin. Neurophysiol.*, **126**, 594-607.

Schoups, A., Vogels, R., N., Q. & Orban, G. (2001) Practising orientation identification improves orientation coding in V1 neurons. *Nature*, **412**, 549-553.

Sheehan, K.A., McArthur, G.M. & Bishop, D.V.M. (2005) Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Cog. Brain Res.*, **25**, 547-553.

Simons, D.J., Boot, W.R., Charness, N., Gathercole, S.E., Chabris, C.F., Hambrick, D.Z. & Stine-Morrow, E.A.L. (2016) Do “brain-training” programs work? *Psychol. Sci. Pub. Int.* **17**, 103-186.

Slater, J., Skoe, E., Strait, D.L., O’Connell, S., Thompson, E. & Kraus, N. (2015) Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behav. Brain Res.*, **291**, 244-252.

Song, J.H., Skoe, E., Banai, K. & Kraus, N. (2012) Training to improve hearing speech in noise: biological mechanisms. *Cer. Cortex*, **22**, 1180-1190.

Song, J.H., Skoe, E., Wong, P.C.M. & Kraus, N. (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. *J. Cog. Neurosci.*, **20**, 1892-1902.

Stickgold, R. (2005) Sleep-dependent memory consolidation. *Nature*, **437**, 1272-1278

Strait, D.L. & Kraus, N. (2014) Biological impact of auditory expertise across the life span: Musicians as a model of auditory learning. *Hear. Res.*, **308**, 109-121.

Talwar, S.K. & Gerstein, G.L. (2001) Reorganization in awake rat auditory cortex by local microstimulation and its effect on frequency-discrimination behavior. *J. Neurophysiol.*, **86**, 1555-1572.

Tennant, K.A., Adkins, D.L., Scalco, M.D., Donlan, N.A., Asay, A.L., Thomas, N., Kleim, J.A. & Jones, T.A. (2012) Skill learning induced plasticity of motor cortical representations is time and age-dependent. *Neurobiol. Learn. Mem.* **98**, 291-302.

Tichko, P. & Skoe, E. (2017) Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators. *Hear. Res.*, **348**, 1-15.

Tierney, A., T., Krizman, J. & Kraus, A. (2015) Music training alters the course of adolescent auditory development. *Proc. Nat. Acad. Sci. USA*, **112**, 1062-1067.

Tierney, A., Krizman, J., Skoe, E., Johnston, K. & Kraus, N. (2013) High school music classes enhance the neural processing of speech. *Front. Psychol.* **4**.

Tong, Y., Melara, R.D. & Rao, A. (2009) P2 enhancement from auditory discrimination training is associated with improved reaction times. *Brain Res.*, **1297**, 80-88.

Tyler, R.S., Parkinson, A.J., Woodworth, G.G., Lowder, M.W. & Gantz, B.J. (1997) Performance over time of adult patients using the Ineraid or Nucleus cochlear implant. *J. Acoust. Soc. Am*, **102**, 508-522

van Wassenhove, V. & Nagarajan, S.S. (2007) Auditory cortical plasticity in learning to discriminate modulation rate. *J. Neurosci.*, **27**, 2663-2672.

Vogels, T.P. & Abbott, L.F. (2009) Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. *Nat. Neurosci.*, **12**, 483-491.

von Trapp, G., Buran, B.N., Sen, K., Semple, M.N. & Sanes, D.H. (2016) A decline in response variability improves neural signal detection during auditory task performance. *J. Neurosci.*, **36**, 11097-11106.

Wang, L., Conner, J.M., Rickert, J. & Tuszynskia, m.H. (2011) Structural plasticity within highly specific neuronal populations identifies a unique parcellation of motor learning in the adult brain. *Proc. Nat. Acad. Sci. USA*, **108**, 2545-2550.

Watanabe, T. & Sasaki, Y. (2015) Perceptual learning: toward a comprehensive theory. *Annu. Rev. Psychol.*, **66**, 197-221.

Watson, C.S. (1991) Auditory perceptual learning and the cochlear implant. *Am. J. Audiol.*, **12, Supplement**, 73-79.

Wenger, E., Brozzoli, C., Lindenberger, U. & Lövdén, M. (2017) Expansion and renormalization of human brain structure during skill acquisition. *Trends Cog. Sci.,* **21**, 930-939.

Whitton, J.P., Hancock, K.E., Shannon, J.M. & Polley, D.B. (2017) Audiomotor perceptual training enhances speech intelligibility in background noise. *Curr. Biol.*, **27**, 3237-3247.e3236.

Witte, R.S. & Kipke, D.R. (2005) Enhanced contrast sensitivity in auditory cortex as cats learn to discriminate sound frequencies. *Cog. Brain Res.*, **23**, 171-184.

Wong, P.C.M., Skoe, E., Russo, N.M., Dees, T. & Kraus, N. (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci*, **10**, 420-422.

Wright, B.A. & Sabin, A.T. (2007) Perceptual learning: how much daily training is enough? . *Exp. Brain Res*, **180**, 727-736.

Wright, B.A. & Zhang, Y. (2006) A review of learning with normal and altered sound-localization cues in human adults. *Int. J. Audiol.*, 45 (Supplement 1), S92-S98.

Wright, B.A. & Zhang, Y. (2009a) Insights into human auditory processing gained from perceptual learning. In Gazzaniga, M.S. (ed) *The Cognitive Neurosciences*. MIT Press, pp. 353-365.

Wright, B.A. & Zhang, Y.X. (2009b) A review of the generalization of auditory learning. *Philos. Trans. R. Soc. B-Biol. Sci.*, **364**, 301-311.

Xu, T., Yu, X., Perlik, A.J., Tobin, W.F., Zweig, J.A., Tennant, K., Jones, T. & Zuo, Y. (2009) Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature*, **462**, 915-919.

Yan, Y., Rasch, M.J., Chen, M., Xiang, X., Huang, M., Wu, S. & Li, W. (2014) Perceptual training continuously refines neuronal population codes in primary visual cortex. *Nat. Neurosci.*, **17**, 1380-1387.

Yang, G., Pan, F. & Gan, W.-B. (2009) Stably maintained dendritic spines are associated with lifelong memories. *Nature*, **462**, 920-924.

Yang, G., Lai, C.S.W., Cichon, J., Ma, L., Li, W. & Gan, W.-B. (2014) Sleep promotes branch-specific formation of dendritic spines after learning. *Science*, **344**, 1173.

Yang, T. & Maunsell, J.H.R. (2004) The effect of perceptual learning on neuronal responses in monkey visual area V4. *J. Neurosci.*, **24**, 1617-1626.

Yotsumoto, Y., Watanabe, T. & Sasaki, Y. (2008) Different dynamics of performance and brain activation in the time course of perceptual learning. *Neuron*, **57**, 827-833.

Zatorre, R.J., Delhommeau, K. & Zarate, J.M. (2012a) Modulation of auditory cortex response to pitch variation following training with microtonal melodies. *Front. Psychol.*, **3**.

Zatorre, R.J., Fields, R.D. & Johansen-Berg, H. (2012b) Plasticity in gray and white: neuroimaging changes in brain structure during learning. *Nat. Neurosci.,* **15**, 528-536

Zendel, B.R. & Alain, C. (2012) Musicians experience less age-related decline in central auditory processing. *Psychol. & Aging.,* **27**, 410-417.

Figure Legends

Fig. 1. Results of three studies of changes in AI associated with perceptual learning on frequency discrimination tasks. **A**. Increase in area of representation of training frequency in AI of monkeys in Recanzone et al. (1993) study. Plots show cortical area of representation (defined by CF) of the frequency ranges used in training monkeys on 2.5 kHz, 5 kHz, and 8 kHz frequency discrimination tasks (upper, middle, and lower panels, respectively). Bars indicate the area of representation of the specified frequency for individual monkeys (identified in abscissa label on lower panel), and for the left (L) and right (R) hemispheres in monkey OM1. Monkeys OM1, 3 and 4 were trained on the auditory frequency discrimination task at the frequency indicated in the baseline label, and solid black bars therefore indicate the area of representation at the training frequency for those monkeys. CM monkeys were passive control animals that received the auditory stimuli presented in one of the frequency training tasks (5kHz for CM2; 8Khz for CM1; hatched bars) while learning a tactile discrimination task. Monkeys N1, N2, and N3 were behaviourally naïve control monkeys. Crosses represent hemispheres in which no cortical locations were recorded with a CF within the frequency range used in training at the specified frequency. The “Mean Untrained” bars indicate the mean area and standard deviation for that frequency in all hemispheres other than those in the monkey(s) trained at that frequency. Reproduced (in modified form) with permission from Recanzone et al. (1993). **B.** Task-specific increase in area of representation of training frequency in AI of rats in Polley et al. (2006) study. Histograms show percentages of recording sites in AI with CFs in 3⁄4-octave-wide bins centered on the frequency value shown, for all recordings obtained in control rats, and in rats trained on frequency (FR; black) and level (LR; grey) tasks. Dashed lines indicate the trained frequency range. Reproduced (in modified form) with permission from Polley et al. (2006). **C.** Lack of change in area of representation of training frequency in AI of cats in Brown et al. (2004) study. Histograms show the mean area of 0.5-kHz characteristic-frequency (CF) (upper panel) and best- frequency (BF) (lower panel) iso-frequency bands for four untrained cats and five cats

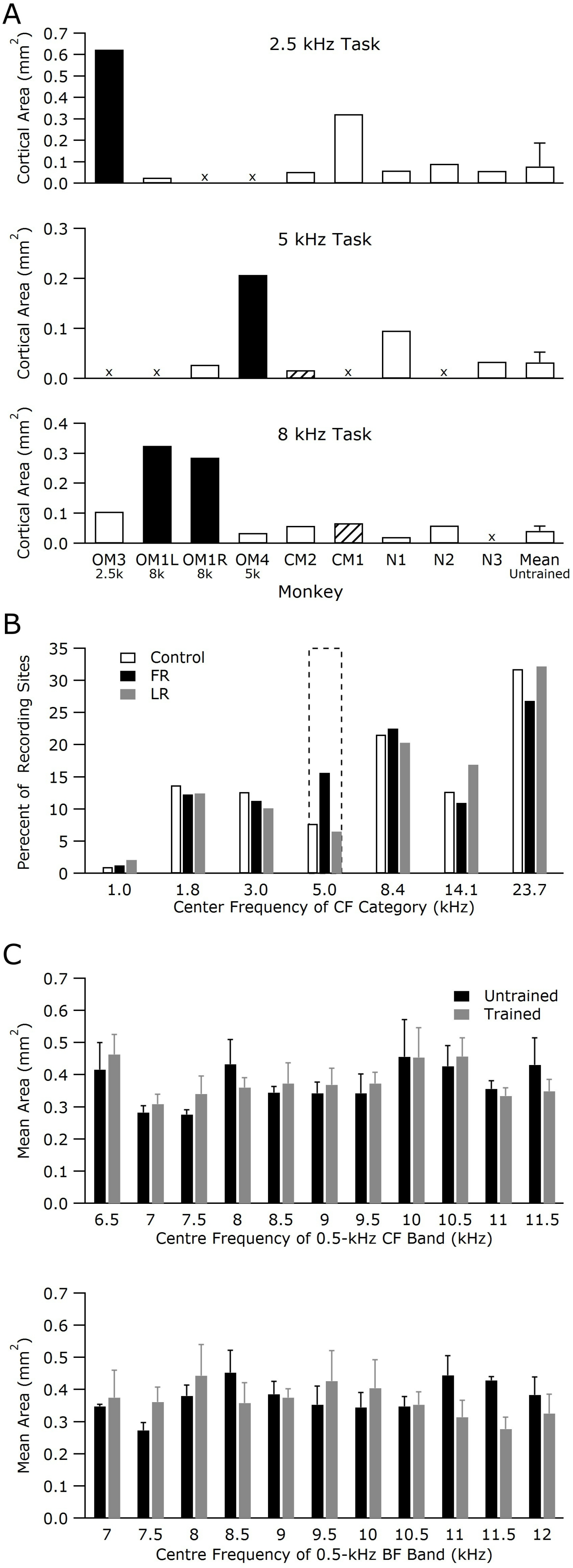
trained on an 8-khz frequency discrimination task. Iso-frequency contours were fitted objectively using an inverse distance method, and histogram bars are centred on a frequency midway between the isofrequency contours between which the areas were measured. Error bars show standard error of the mean. Reproduced (in modified form) with permission from Brown et al. (2004).

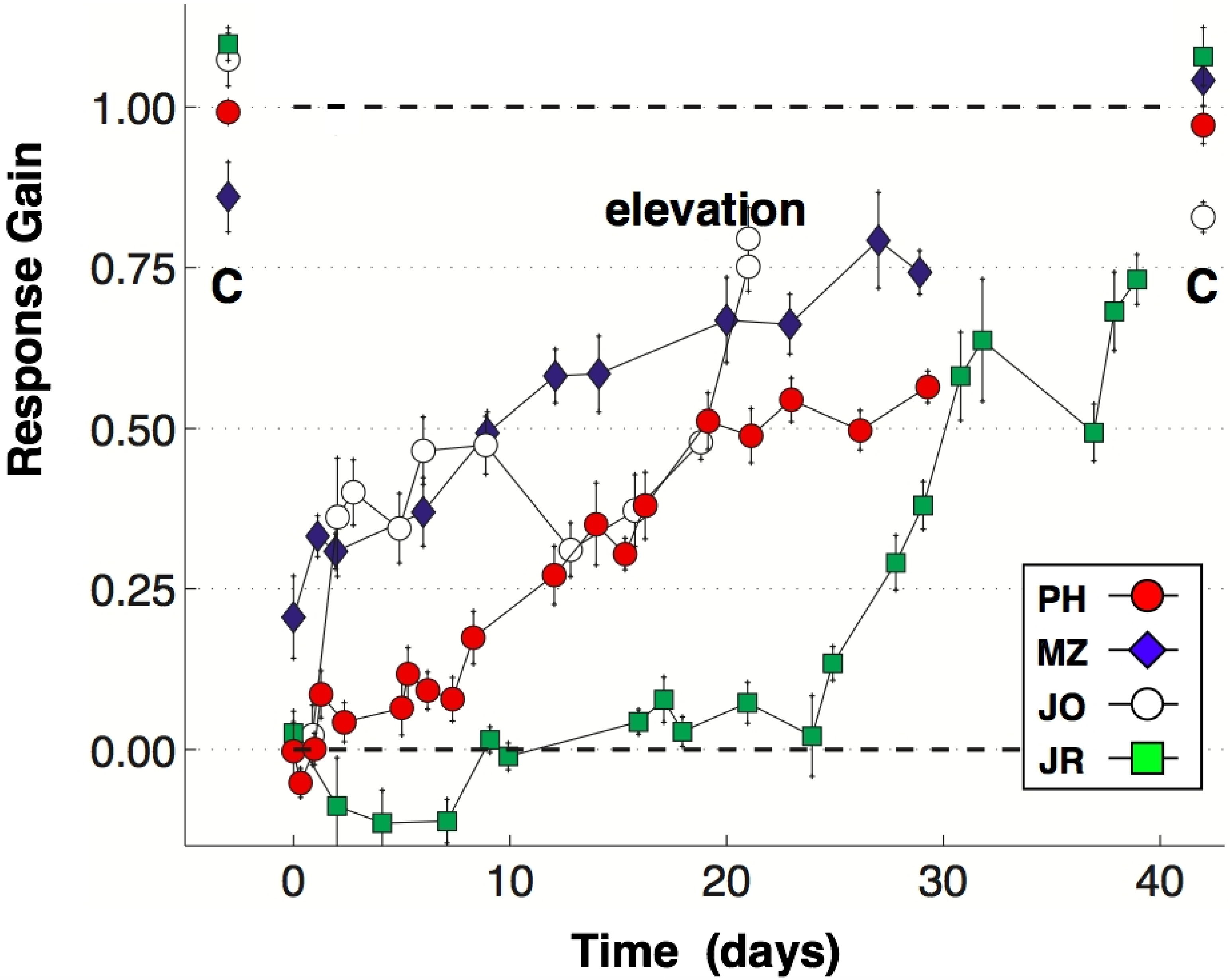
Fig. 2. Comparisons of behavioral sensitivity and sensitivity of neurons in core auditory cortex of gerbils trained on an amplitude modulation detection task. **A.** Behavioral and neural thresholds (mean ± standard error of mean (SEM)) improve in parallel over training days in an individual animal (based on data from 30 recording sites; 4 to 7 sites per day). **B**. Near-perfect correlation between behavioral and neural thresholds in an individual animal. **C.** Changes in behavioral (black) and neural thresholds (mean ± SEM) in engaged (blue) and disengaged (green) conditions (group data). n = number of (multi- or single-unit) recording sites (number of sites per day ranged from 29 to 39 for engaged condition and 5-14 for non-engaged condition). Increasing difference between engaged and disengaged thresholds is indicated by difference between orange brackets at day1 and day 7. Reproduced (in modified form) with permission from Caras and Sanes (2017).

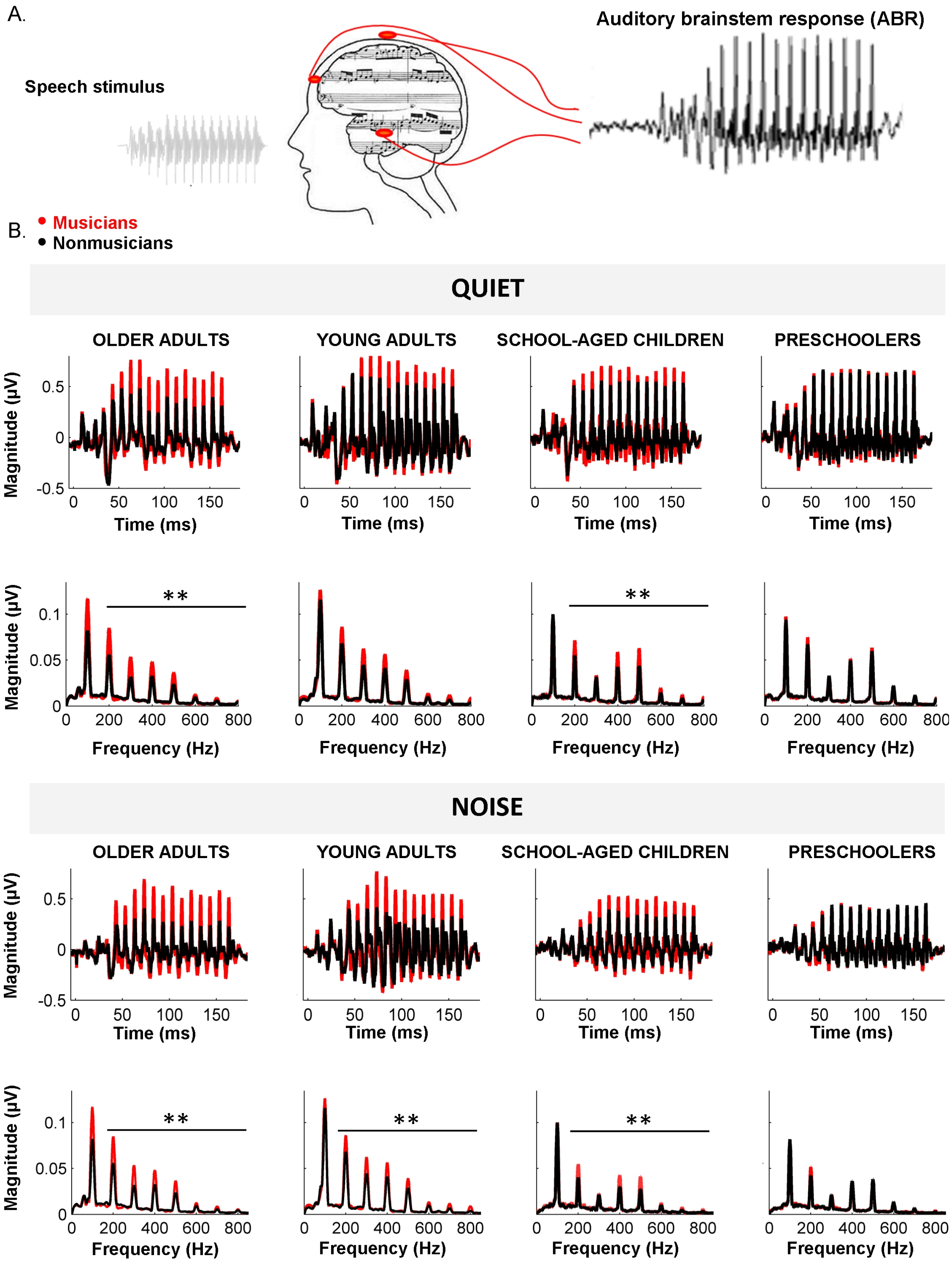
Fig. 3. Changes in vertical localization accuracy of four participants over period of experience with pinna moulds that modified the spectral cues to sound elevation. Accuracy measure (response gain) is the slope of the best-fit regression line between target and response coordinates on each recording day. Responses were saccades to broad-band white-noise sound bursts presented in darkness in an echo-free room at randomly chosen locations. Standard deviations of the gains were obtained by bootstrapping the data 100 times. Results for the pre- and post-adaptation performance without moulds are indicated as control (C) values; for each

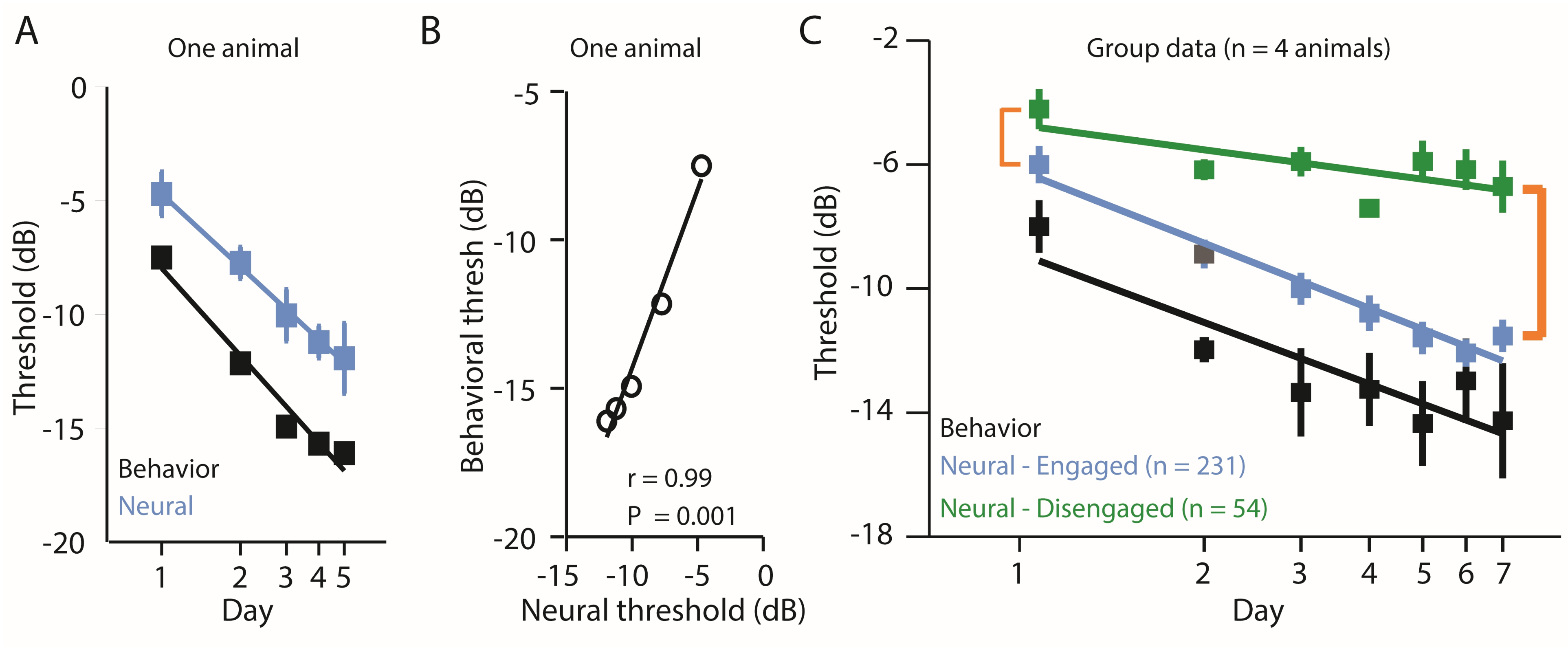
participant, post-adaptation control measurements were made immediately after removal of the moulds. Reproduced (in modified form) with permission from Hofman et al. (1998).

Fig. 4. Effects of musical training on frequency following response (FFR) evoked by speech stimulus. **A.** Averaged auditory brainstem response (the FFR) evoked by speech sound stimulus /da/, as recorded using three scalp electrodes. **B.** Responses of musicians and non-musicians (red and black, respectively) in four different age groups, and for stimuli presented in quiet and in noise. For each stimulation environment, responses are presented in both temporal and spectral domains. $^{**}p < 0.01$. Reproduced with permission from Strait and Kraus (2014).


A
2.5 kHz Task
5 kHz Task
8 kHz Task
Cortical Area (mm^2)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
x
OM3
2.5k
OM1L
8k
OM1R
8k
OM4
5k
CM2
CM1
N1
N2
N3
Mean
Untrained
Monkey
B
Control
FR
LR
Perecent of Recording Sites
0
5
10
15
20
25
30
35
1.0
1.8
3.0
5.0
8.4
14.1
23.7
Center Frequency of CF Category (kHz)
C
Untrained
Trained
Mean Area (mm^2)
6.5
7
7.5
8
8.5
9
9.5
10
10.5
11
11.5
Centre Frequency of 0.5-kHz CF Band (kHz)
7
7.5
8
8.5
9
9.5
10
10.5
11
11.5
12
Centre Frequency of 0.5-kHz BF Band (kHz)


Response Gain
1.00
0.75
0.50
0.25
0.00
C
elevation
C
PH
MZ
JO
JR
0
10
20
30
40
Time (days)


A.
Speech stimulus
Auditory brainstem response (ABR)
B.
● Musicians
● Nonmusicians
QUIET
OLDER ADULTS
YOUNG ADULTS
SCHOOL-AGED CHILDREN
PRESCHOOLERS
Magnitude (µV)
0.5
0
-0.5
0
50
100
150
Time (ms)
Magnitude (µV)
0.1
0.05
0
0
200
400
600
800
Frequency (Hz)
**
NOISE
OLDER ADULTS
YOUNG ADULTS
SCHOOL-AGED CHILDREN
PRESCHOOLERS
Magnitude (µV)
0.5
0
-0.5
Time (ms)
Magnitude (µV)
0.1
0.05
0
Frequency (Hz)
**


A
One animal
0
-5
-10
-15
-20
Threshold (dB)
1
2
3
4
5
Day
Behavior
Neural
B
One animal
-5
-10
-15
-20
Behavioral thresh (dB)
-15
-10
-5
0
Neural threshold (dB)
r = 0.99
P = 0.001
C
Group data (n = 4 animals)
-2
-6
-10
-14
-18
Threshold (dB)
1
2
3
4
5
6
7
Day
Behavior
Neural - Engaged (n = 231)
Neural - Disengaged (n = 54)