Consolidation and retention of auditory categories acquired incidentally in performing a visuomotor task

Yafit Gabay (ygabay@edu.haifa.a.c.il)
Department of Special Education, University of Haifa
Abba Khoushy Ave 199, Haifa, 3498838 Israel

Avi Karni (avi.karni@yahoo.com)
Departments of Neurobiology and Learning Disabilities, University of Haifa
Abba Khoushy Ave 199, Haifa, 3498838 Israel

Lori L. Holt (loriholt@cmu.edu)
Department of Psychology, Carnegie Mellon University
Pittsburgh, PA 15232 USA

Abstract

A wealth of evidence indicates the existence of a consolidation phase, triggered by and following a practice session, wherein new memory traces relevant to task performance are transformed and honed to represent new knowledge. But, the role of consolidation is not well-understood in category learning and has not been studied at all under *incidental* category learning conditions. Here, we examined the acquisition, consolidation and retention phases in a visuomotor task wherein auditory category information was available, but not required, to guide detection of an above-threshold visual target across one of four spatial locations. We compared two training conditions: (1) Constant, whereby repeated instances of one exemplar from an auditory category preceded a visual target, predicting its upcoming location; (2) Variable, whereby five distinct category exemplars predicted the visual target. Visual detection speed and accuracy, as well as the performance cost of randomizing the association of auditory category to visual target location, were assessed during online performance, again after a 24-hour delay to assess the expression of delayed gains, and after 10 days to assess retention. Results revealed delayed gains associated with incidental auditory category learning and retention effects for both training conditions. Offline processes can be triggered even for incidental auditory input and lead to category learning; variability of input can enhance the generation of incidental auditory category learning.

Keywords: Category learning, auditory, incidental learning, memory consolidation, speech, statistical learning

Introduction

Although a rich literature documents early phonetic category acquisition (Werker, Yeung, & Yoshida, 2012) and there is increasing evidence for continued phonetic development in later childhood (Zevin, 2012) quite little is understood about the learning mechanisms involved. Distributional learning, by which listeners are sensitive to the statistical regularities across speech categories, is widely believed to be significant (Maye, Werker, & Gerken, 2002; Thiessen, 2007). However, there are also concerns about whether laboratory demonstrations of distributional learning may 'scale' to real speech input (Lim, Lacerda, & Holt, 2015; Pierrehumbert, 2003). Moreover, distributional learning does not itself implicate a specific learning mechanism (Lim, Fiez, & Holt, 2014).

One reason it has been challenging to establish the mechanism(s) of phonetic category acquisition is that it is difficult, if not impossible, to control the distributional detail of listeners' speech experience. Even neonates have had prenatal speech experience that shapes perception (DeCasper & Spence, 1986). Over the last decade, research has circumvented this difficulty by examining acquisition of novel non-linguistic auditory categories composed of artificial nonspeech sounds to understand the general mechanisms available to phonetic acquisition (e.g., Goudbeek, Swingley, & Smits, 2009; Holt & Lotto, 2006; Holt, Lotto, & Diehl, 2004; Mirman, Holt, & McClelland, 2004). A benefit of this approach is that experience can be tightly controlled to investigate specific mechanistic hypotheses, as has been the case in the long-standing and productive research literature on visual perceptual category learning (e.g., Maddox & Ashby, 2004).

Incidental Auditory Category Learning

As in visual category learning, most non-linguistic auditory category learning studies have used explicit tasks in which listeners are aware of the existence of categories and explicitly search for category-diagnostic dimensions by making overt decisions to maximize experimenter-provided feedback. This work has yielded insights that have translated directly to a better understanding of the mechanisms available to phonetic category acquisition (Lim & Holt, 2011). Yet, this work does not model learning conditions in which phonetic categories are acquired that are neither wholly passive, nor explicit and dependent upon overt feedback (Lim et al., 2014). Category learning often occurs under more *incidental* conditions in which listeners are *actively engaged* in environments in which auditory categories are associated with *rich multimodal cue*s and *behaviorally-relevant outcomes*.

In an attempt to model these learning contexts in the laboratory, researchers have developed several incidental learning paradigms that, while computer-based and consistent with tight experimental control, better capture task demands involved in building complex perceptual categories without awareness of the categorization task, overt category decisions, or experimenter-provided feedback about categorization (Gabay, Dick, Zevin, & Holt, 2015; Wade & Holt, 2005). Results from this research indicate that listeners can acquire complex auditory (Gabay et al., 2015; Leech, Holt, Devlin, & Dick, 2009; Liu & Holt, 2011; Roark & Holt, 2015; Wade & Holt, 2005) and phonetic (Lim et al., 2015; Lim & Holt, 2011) categories via incidental learning. Learning generalizes to novel category instances. Moreover, adult listeners who incidentally acquire complex non-native phonetic categories show transfer of the learning that scaffolds word learning in the non-native language (Liu & Holt, 2015a). Additionally, incidental learning of *non-linguistic sound categories* designed to model some of the perceptual dimensions defining difficult non-native phonetic categories generalizes to support subsequent non-native *speech* categorization (Liu & Holt, 2015b).

Altogether, these data indicate that the processes underlying incidental learning of non-linguistic sound categories inform those available to phonetic category acquisition. The distinction of incidental training versus passive or explicit training is important because there is growing evidence that these learning paradigms draw upon neural substrates with distinctive computational specialties (Doya, 1999; Maddox & Ashby, 2004; Seger & Miller, 2010). Emerging evidence suggests that incidental auditory category learning engages the procedural learning system (striatum of basal ganglia, Lim et al., 2014; Lim, Fiez, Wheeler, & Holt, 2013) and recruits putatively speech-selective cortex for processing newly-acquired *non-linguistic* auditory categories (left posterior superior temporal sulcus, pSTS; Leech et al., 2009). Significantly, striatal activation is correlated with behavioral incidental learning performance and exhibits functional connectivity with the left pSTS region sensitive to category learning mentioned above (Lim et al., 2013). In all, these results demonstrate that both speech and nonspeech signals may draw on cortical networks once thought to be speech-selective as a function of category expertise. This further substantiates the use of non-linguistic auditory categories as a test-bed for mechanisms available to phonetic acquisition, points to procedural auditory category learning (Yi, Maddox, Mumford, & Chandrasekaran, 2016), and establishes incidental learning as a valuable approach to understanding mechanisms available to phonetic acquisition.

Learning Stages in Procedural Skill Acquisition

In a parallel literature, a growing body of research indicates that skill learning is a multi-stage, dynamic process of performance and knowledge changes across time (see Karni & Korman, 2011). In addition to performance gains that occur concurrently with a learning task (*online, fast learning*), delayed performance gains may also evolve in the absence of additional practice (*offline, slow learning*). These latter changes involve consolidation processes whereby new memory traces become less susceptible to interference, but also are transformed and honed to represent new knowledge (Dudai, Karni, & Born, 2015), may require sleep (Karni, Tanne, Rubenstein, Askenasy, & Sagi, 1994), and are accompanied by measurable neural signatures (Ungerleider, Doyon, & Karni, 2002). Consolidation is considered a key feature of effective skill learning and the attainment of fluency (automaticity), central for the establishment of procedural memory (Atienza, Cantero, & Stickgold, 2004; Dudai et al., 2015). There is considerable evidence for slow learning phases reflecting memory consolidation in the motor domain (Dudai et al., 2015) and research demonstrates slow learning changes associated with consolidation in language domain (Davis, Di Betta, Macdonald, & Gaskell, 2009; Earle & Myers, 2015; Fenn, Nusbaum, & Margoliash, 2003) Although these studies examine consolidation of learning across speech signals, the learning was evoked by explicit, rather than incidental, training.

In other domains the behavioral expression of memory consolidation is considered a key signature indicating establishment of robust, automatic and efficient representations (Karni & Bertini, 1997). Therefore, investigating incidental auditory category learning across several time points will be critical in revealing how memory consolidation processes affect procedural auditory category learning and phonetic acquisition.

The Present Study

The current study is designed to examine the expression of consolidation phase gains and retention of incidental auditory category learning. These measures afford the construction of theoretical bridges to neurobehavioral evidence and mechanisms of plasticity (Dorfberger, Adi-Japha, & Karni, 2007) that putatively underlie auditory category learning.

The second issue examined in the present studies concerns variability. Research in speech category learning has emphasized the importance of experiencing high acoustic-phonetic variability in training. Experience with multiple speakers, phonetic contexts, and exemplars seems to promote non-native speech category learning and generalization among adult learners (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Iverson, Hazan, & Bannister, 2005; Jamieson & Morosan, 1989; Wang, Spence, Jongman, & Sereno, 1999). As such, the issue of variability in training has been influential in empirical and theoretical approaches to speech category learning. However, it has arisen from studies of extensive training across multiple training sessions spanning days or weeks that have examined learning via explicit, feedback-driven tasks in which listeners actively search for category-diagnostic information. In this way, it has not been investigated in a manner to assess consolidation of learning gains, or incidental learning. In a previous study of incidental auditory category learning, we observed enhanced learning when within-category acoustic variability was experienced within trials, as compared to across trials (Gabay et al. 2015) even when global variability was held constant. The present study extends this work to examine the influence of within-category acoustic variability on consolidation and retention of auditory categories.

Methods

Participants

In each experiment, young adult participants were recruited from the University of Haifa community. They received payment or course credit, had normal or corrected-to-normal vision, and reported normal hearing. 24 participants were tested in Experiment 1 and 22 were tested in Experiment 2.

Procedure

Nonspeech stimuli. Figure1a illustrates four auditory categories, drawn from prior research (e.g., Wade & Holt, 2005; Leech et al., 2009; Liu & Holt, 2011; Gabay & Holt, 2015; Gabay et al., 2015). These sounds have some of the spectrotemporal complexity of speech but are unequivocally nonspeech owing to their noise and square wave sources (Wade & Holt, 2005). Each category has 6 exemplars used in training and 5 exemplars withheld to test generalization (not shown in Figure 1). Two categories are defined by a unidimensional acoustic cue (up- or down-sweep in frequency of a higher-frequency component). The other two categories are defined in a more complex, multidimensional perceptual space (no one acoustic cue uniquely defines category membership, see Wade & Holt, 2005).

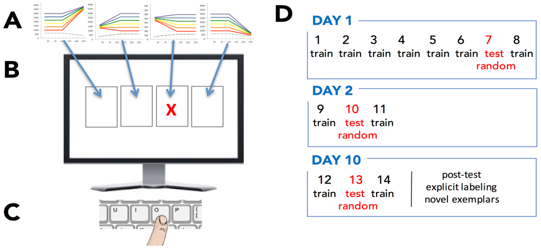

Figure 1. Overview of SMART Paradigm. (A) Four nonspeech auditory categories are defined by multiple exemplars. (B) Each category is associated with a particular visual location, thereby predicting the upcoming appearance of a visual target 'X.' (C) Participants indicate the target location with a key press. (D) Blocks include a Test Block in which the

Systematic Multi-modal Association Time (SMART) Task. In the SMART task (Figure 1), participants rapidly detect the appearance of a visual target in one of four possible screen locations and report its position by pressing a key corresponding to the visual location. The primary task is visual detection. However, a brief sequence of sounds precedes each visual target. Unknown to participants, the sounds are drawn from one of four distinct sound categories. There is a multimodal (auditory category to visual location) correspondence that relates variable sound category exemplars to a consistent visual target location and response. This mapping is many-to-one, such that multiple, acoustically-variable sound category exemplars are associated with a single visual location. Likewise, sound categories are predictive of the action required to complete the primary visual detection task; in the SMART task, auditory categories perfectly predict the location of the upcoming visual detection target and the corresponding response button to be pressed. Thus, learning to treat the acoustically variable sounds as functionally equivalent in predicting the upcoming location of a visual target may facilitate visual detection without requiring overt sound categorization decisions or even awareness of category structure. The SMART task makes it possible to investigate whether participants learn auditory categories *incidentally*, during a largely visuomotor task.

Participants completed 8 practice trials to acquaint them with the visual detection response. Sounds preceded visual targets in these practice trials, but there was no category-to-location correlation. Immediately thereafter there were 6 training blocks (96 trials, 4 sound categories x 6 exemplars x 4 repetitions) for which there was a perfect correlation between auditory sound category and visual target location (Figure 1d). In the seventh block (B7, 48 trials), any sound exemplar could precede presentation of the visual target in any position; sound category no longer predicts the position in which the visual target would appear. A final B8 training block restored the relationship between sound category and the location of the upcoming visual target.

Twenty-four hours later on Day 2, participants completed a training block (B9) and a shorter (48 trial) random-mapping block (B10) and a final training block (B11) to restore the mapping. On Day 10, participants completed B12, B13, and B14, with a structure identical to the Day 2 blocks. Subsequently on Day 10, there was an explicit labeling task in which novel sound exemplars drawn from one of the 4 auditory categories were presented on each of 96 trials and participants selected the expected visual target location; no target appeared.

Testing took place in a sound-attenuated chamber with participants seated directly in front of a computer monitor. Sounds were presented dichotically over headphones (Beyer, DT-150).

Experiments 1 and 2 were identical, except for the manner by which within-category exemplar variability was experienced. As noted above, five sound exemplars preceded the visual target on each trial. In Experiment 1, a single category exemplar was randomly chosen and presented five times such that within-category exemplar variability was experienced *across* but not *within* trials. In Experiment 2, five *unique* exemplars drawn from a category preceded the visual target. Across experiments, the categories perfectly predicted the upcoming target location and, across trials, the within-category variability experienced by participants was equivalent. However, in Experiment 2 the within-category variability was more tightly coupled with the visuomotor associations we hypothesize to promote incidental category learning. Gabay et al. (2015) hypothesized that the SMART task visumotor associations provide 'representational glue' with which to bind acoustically variable category exemplars. By this view, experiencing within-trial acoustic variability will result in more robust learning via the tighter coupling of category variability with visualmotor task demands. Here, we seek to examine how this prediction relates to consolidation and retention of incidental category learning.

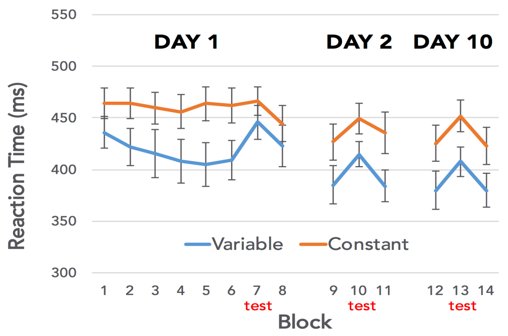

Figure 2. Reaction time (RT) to detect the visual target as a function of training block (1-14) and training type (variable vs. constant) and session (Day 1, 2 and 10).

Results

Reaction Time (RT) Cost. Following the approach of Gabay et al. (2015), we predict that if participants incidentally learn sound categories across training blocks then visual detection will be slower in the test block relative to the training block that preceded it (*RT Cost*) because the category-to-location assignment is randomized in the test block. This *covert* measure of category learning does not require overt auditory categorization decisions or explicit labeling.

Results are presented in Figure 2. Trials for which there was a visual detection error (M=2%) or response time (RT) longer than 1500 ms or shorter than 100 ms (M=2%) were excluded from analyses. A repeated-measures analysis of variance (ANOVA) was conducted with RT Cost (repeated vs. random block) and Day (1, 2, and 10) as within-subjects variables and Training Type (constant vs. variable training) as a between-subjects variable and mean reaction time to detect the visual target as the dependent variable. The main effect of training type was marginally significant, F(1, 44)=3.71, p=.07, η_p^2 = .07 suggesting that in general participants were somewhat faster in the variable condition compared with the constant training condition. There was significant main effect of session F(2, 88) = 13.324, p=.00001, η_p^2 = .23, such that participants became faster in later sessions (Days 2, 10) compared with the first session (Day 1), F(1, 44) =20.49, p=.00004. No significant difference was observed between the second and third sessions (Day 10), $F<1$, p=.595. The RT Cost main effect was also significant, F(1, 44) = 12.5, p=.00097, η_p^2 = .22 indicating that participants on average were faster to detect the visual target during the training blocks compared with the test blocks. This indicates that participants were sensitive to the relationship between sound category and visual target. The three-way interaction of session, RT Cost and training type was significant, F (2, 88) = 3.43, p=.037, η_p^2 = .07. Further analysis revealed that there was a significant increase in RT Cost magnitude in later sessions compared with the initial session for the constant training condition. Greater RT costs were observed in Day 2 compared with Day 1, F (1, 44) = 4.56, p=.038 and in Day 10 compared with Day 1, F (1, 44) = 5.09, p=.029. There were no differences in RT Cost magnitude between Day 10 and Day 2, F=.234 p=.630. For the variable training condition, RT Costs were significant in all sessions and there were no differences in RT Cost magnitude across sessions.

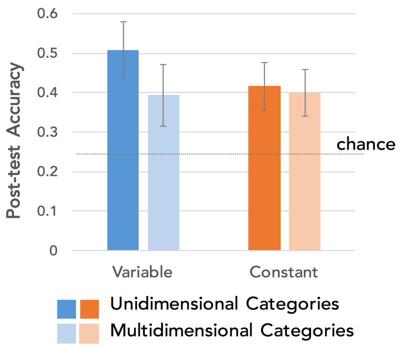

Figure 3. Overt category labeling post-test accuracy as a function of training condition and uni- versus multidimensional categories.

Posttest categorization. As an overt measure of category learning, we used participants' accuracy in explicitly matching novel sound exemplars with the visual location consistent with the category-location relationship encountered in training. The exemplars tested in the overt categorization task were not previously encountered. Thus, generalization of category learning was required for accurate matching. Results are shown in Figure 3.

Participants undergoing both constant and variable incidental training exhibited above-chance overt category labeling of novel exemplars (all p's=<.05), (Variable: Unidim, t(21)=3.77, p=.001; Multidim, t(21)=2.52, p=.0019; Constant: Unidim, t(23)=2.89, p=.008; Multidim, t(23)=3.104, p=.005). A repeated measures ANOVA with category type (Uni- vs. Multi-dimensional categories) as a within subject variable on mean accuracy and training type (constant vs. variable) as a between subject factor, showed a main effect of category type F (3, 132) = 9.63, p=.003, η_p^2 = .12 such that unidimensional categories were better learned than multidimensional categories. There was also a significant interaction between category type and training F (1, 44) =5.45, p=.024, η_p^2 = .1 such that unidimensional categories were learned better than multidimensional categories with variable training, F (1, 44) =14.17, p=.0004, whereas no such difference was observed for the constant training condition (F<1).

Discussion

The present study tested consolidation phase gains and retention in incidental auditory category learning. Building from prior research (Gabay et al., 2015), we employed the SMART task in which four novel auditory categories are consistently predict the upcoming location of a visual target to which participants respond with a keypress to indicate target location. Participants were not informed about the consistent mapping between audio and visual inputs and could potentially perform the task perfectly without relying on the auditory input. In the context of this largely visuomotor task, participants incidentally learned the auditory categories, and generalized this learning to novel category-consistent exemplars.

Here, training occurred across three sessions to assess acquisition, consolidation and retention of the incidentally learned auditory category knowledge. Furthermore, the nature of the training experience was manipulated to examine the influence of variability on each one of these processes.

Consistent with prior results examining learning within a single training session (Gabay et al., 2015), participants became reliant on auditory categories to guide visual detection, as evident in the RT Cost to visual detection reaction time upon the elimination of the consistent category-to-location mapping. Notably, however, a significant RT cost was observed only for the variable training condition by the end of the first session. This differs from the observations of Gabay et al., who observed a RT Cost in a single session of SMART training with no within-trial exemplar variability (although this effect was less robust than that observed with exemplar variability).

Nonetheless, by Day 2, learning was evident for both conditions across the delay period as a reduction in RT in Block 9 relative to Block 8 and an exaggerated RT Cost to randomization in Block 10. This pattern of results suggests that the association of the auditory categories with specific visuomotor contingencies underwent a consolidation phase and was strengthened over the delay. These results are consistent with the wealth research in motor and visual perceptual domains indicating the existence of a consolidation phase in the development of skill. Consistent with prior reports of consolidation phase 'offline' gains for category knowledge (Djonlagic et al., 2009), the present results are the first to report consolidation effects for incidental learning, and for nonspeech auditory category learning.

Although both types of training, constant or variable, elicited offline gains, a reliance on auditory categories to guide visual detection (manifested as a RT Cost) developed already in the 1st session in the variable condition. This was well-maintained over the 24-hour delay. In contrast, there were no RT Costs for the constant training condition in the initial session. Yet, category acquisition must have been underway, as the RT Costs were apparent at 24-hour post-training. This notion is consistent with a previous report (Gabay et al., 2015) of RT Costs under constant training already during the initial session, although the costs were significantly smaller than those observed under variable training. Cohort differences (the mean RT was somewhat slower in the present studies compared to the Gabay et al. study) or even biases from native language (English vs. Hebrew, in the prior and current studies, respectively) may potentially have played a role.

The current results underscore the notion that details of training influence the evolution of category acquisition over time, perhaps influencing the ultimate learned representations. As an important control, future work will need to establish the extent to which there are offline gains associated with the visual target detection task itself, independent of auditory category learning. This will help to establish the origins of the increased speed of responses observed on Day 2 for both conditions.

Despite brief, incidental training with entirely novel sound categories, the learning gains attained by Day 2 were robustly retained 10 days after initial training, as evident by robust RT Costs as well as by the above-chance overt labeling of novel category-consistent exemplars.

Taken together these results suggest that 'offline' processes resulting in performance gains can be triggered for incidental auditory experience associated with but not necessary for a visuomotor task. The present study establishes a framework for studying the evolution of category representations as they emerge over time.

Acknowledgments

This research was supported by the National Science Foundation-Binational Scientific Foundation to LLH, AK and YG under Grant (2016867).

References

Atienza, M., Cantero, J. L., & Stickgold, R. (2004). Posttraining sleep enhances automaticity in perceptual discrimination. *Journal of cognitive neuroscience, 16*(1), 53-64.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. i. (1997). Training Japanese listeners to identify English/r/and/l: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America, 101*(4), 2299-2310.

Davis, M. H., Di Betta, A. M., Macdonald, M. J., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of cognitive neuroscience, 21*(4), 803-820.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9*(2), 133-150.

Djonlagic, I., Rosenfeld, A., Shohamy, D., Myers, C., Gluck, M., & Stickgold, R. (2009). Sleep enhances category learning. *Learning & memory, 16*(12), 751-755.

Dorfberger, S., Adi-Japha, E., & Karni, A. (2007). Reduced susceptibility to interference in the consolidation of motor memory before adolescence. *PloS one, 2*(2), e240.

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural networks, 12*(7), 961-974.

Dudai, Y., Karni, A., & Born, J. (2015). The consolidation and transformation of memory. *Neuron, 88*(1), 20-32.

Earle, F. S., & Myers, E. B. (2015). Overnight consolidation promotes generalization across talkers in the identification of nonnative speech sounds. *The Journal of the Acoustical Society of America, 137*(1), EL91-EL97.

Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature, 425*(6958), 614-616.

Gabay, Y., Dick, F. K., Zevin, J. D., & Holt, L. L. (2015). Incidental Auditory Category Learning. *Journal of experimental psychology. Human perception and performance, 41*(4), 1124.

Goudbeek, M., Swingley, D., & Smits, R. (2009). Supervised and unsupervised learning of multidimensional acoustic categories. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1913.

Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisitiona). *The Journal of the Acoustical Society of America, 119*(5), 3059-3071.

Holt, L. L., Lotto, A. J., & Diehl, R. L. (2004). Auditory discontinuities interact with categorization: Implications for speech perception. *The Journal of the Acoustical Society of America, 116*(3), 1763-1773.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English/r/-/l/to Japanese adults. *The Journal of the Acoustical Society of America, 118*(5), 3267-3278.

Jamieson, D. G., & Morosan, D. E. (1989). Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology/Revue canadienne de psychologie, 43*(1), 88.

Karni, A., & Bertini, G. (1997). Learning perceptual skills: behavioral probes into adult cortical plasticity. *Current opinion in neurobiology, 7*(4), 530-535.

Karni, A., & Korman, M. (2011). *When and where in skill memory consolidation: neuro-behavioral constraints on the acquisition and generation of procedural knowledge.* Paper presented at the BIO Web of Conferences.

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J., & Sagi, D. (1994). Dependence on REM sleep of overnight improvement of a perceptual skill. *Science, 265*(5172), 679-682.

Leech, R., Holt, L. L., Devlin, J. T., & Dick, F. (2009). Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. *The Journal of neuroscience, 29*(16), 5234-5239.

Lim, S.-J., Fiez, J. A., & Holt, L. L. (2014). How may the basal ganglia contribute to auditory categorization and speech perception? *Frontiers in neuroscience, 8*, 230.

Lim, S.-J., Fiez, J. A., Wheeler, M. E., & Holt, L. L. (2013). Investigating the neural basis of video-game-based category learning. *Journal of cognitive neuroscience*.

Lim, S.-J., Lacerda, F., & Holt, L. L. (2015). Discovering functional units in continuous speech. *Journal of Experimental Psychology: Human Perception and Performance, 41*(4), 1139.

Lim, S. j., & Holt, L. L. (2011). Learning Foreign Sounds in an Alien World: Videogame Training Improves Non-Native Speech Categorization. *Cognitive science, 35*(7), 1390-1405.

Liu, R., & Holt, L. L. (2011). Neural changes associated with nonspeech auditory category learning parallel those of speech category acquisition. *Journal of cognitive neuroscience, 23*(3), 683-698.

Liu, R., & Holt, L. L. (2015a). Dimension-based statistical learning of vowels. *Journal of Experimental Psychology: Human Perception and Performance, 41*(6), 1783.

Liu, R., & Holt, L. L. (2015b). Perceptual scaffolding of non-native speech categories through videogame-based training. *The Journal of the Acoustical Society of America, 137*(4), 2386-2386.

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes, 66*(3), 309-332.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101-B111.

Mirman, D., Holt, L. L., & McClelland, J. L. (2004). Categorization and discrimination of nonspeech sounds: Differences between steady-state and rapidly-changing acoustic cues. *The Journal of the Acoustical Society of America, 116*(2), 1198-1207.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and speech, 46*(2-3), 115-154.

Roark, C. L., & Holt, L. L. (2015). Rule-based and information-integration categorization during an incidental learning task. *The Journal of the Acoustical Society of America, 137*(4), 2385-2385.

Seger, C. A., & Miller, E. K. (2010). Category learning in the brain. *Annual review of neuroscience, 33*, 203.

Thiessen, E. D. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language, 56*(1), 16-34.

Ungerleider, L. G., Doyon, J., & Karni, A. (2002). Imaging brain plasticity during motor skill learning. *Neurobiology of learning and memory, 78*(3), 553-564.

Wade, T., & Holt, L. L. (2005). Incidental categorization of spectrally complex non-invariant auditory stimuli in a computer game task. *The Journal of the Acoustical Society of America, 118*(4), 2618-2633.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America, 106*(6), 3649-3658.

Werker, J. F., Yeung, H. H., & Yoshida, K. A. (2012). How do infants become experts at native-speech perception? *Current Directions in Psychological Science, 21*(4), 221-226.

Yi, H., Maddox, W. T., Mumford, J. A., & Chandrasekaran, B. (2016). The role of corticostriatal systems in speech category learning.*Cerebral Cortex, 26*(4), 1409-1420.

Zevin, J. D. (2012). A sensitive period for shibboleths: The long tail and changing goals of speech perception over the course of development. *Developmental psychobiology, 54*(6), 632-642.