PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Reinforced statistical learning of auditory categories: A preliminary report of cognitive, cortical and computational mechanisms

Bob MCMURRAY[1]; Samantha CHIU[1]; John FREEMAN[1]

[1] University of Iowa, USA

ABSTRACT

Learning to categorize speech sounds is crucial for language acquisition and a struggle for second language learners. Most research has examined supervised learning in which learners receive feedback. This offers a useful model for second language acquisition. However, work in first language acquisition suggests people may acquire categories by tracking the statistics of the input without feedback. Few studies examine how these forms of learning work together and little cognitive neuroscience examines unsupervised learning. To study this interaction, we developed the reinforced statistical learning paradigm in which learners are first exposed to statistically structured stimuli in an unsupervised paradigm and then receive supervised training. We present preliminary data from two studies using this paradigm. It suggests that the unsupervised phase benefits later supervised learning, but only in limited conditions that may vary according to whether items are blocked or interleaved. We examined the cortical consequence of learning with electro-corticographic recordings from the surface of the brains of two epilepsy patients. A machine learning analysis suggests some surprising ways in which auditory cortical representations change with learning. This suggests the need for theories that explain the varied effects of unsupervised learning on auditory processing.

Keywords: Hearing, Categorization, Learning, Unsupervised Learning

1. INTRODUCTION

Speech perception is built on categorization. Phonemes are defined by many variable acoustic cues, and listeners must generalize across talkers and contexts. Understanding how auditory categories are learned is crucial for real-world problems. Deficits in phonological processing (based in part on auditory categorization) have been linked to developmental language disability and dyslexia (1). Additionally, in second language (L2) acquisition, many adults never learn to discriminate non-native sounds (2). Finally, hearing remediations like cochlear implants distort the input, requiring users to relearn auditory categories. Thus, auditory category learning has broad impacts.

Substantial work has investigated the acquisition of auditory categories (3). Most examines supervised learning (via reinforcement) (4-7). This is a useful model for second language acquisition, where sound categories can be explicitly taught, and learners have access to vocabulary, orthography and articulation to serve as implicit supervisory signals. However, in first language acquisition, infants may not have access to such signals, and may learn via unsupervised statistical learning (8).

Supervised and unsupervised learning may need to work together. Adult L2 learners can use unsupervised statistics (9). Moreover, L1 acquisition extends through adolescence (10), a time when top-down knowledge (e.g. spelling, vocabulary) is available to guide learning. However, there is little work on how supervised and unsupervised learning relate. This manuscript describes preliminary results from ongoing investigations of this interaction. We develop a new paradigm, *reinforced statistical learning* to investigate the ways in which these modes of learning interact. We start by offering a brief background on both forms of learning before presenting results from three experiments.

1.1 Mechanisms of Learning

Supervised learning (using feedback or reinforcement) has been studied as a model of L2 learning

[1] bob-mcmurray@uiowa.edu

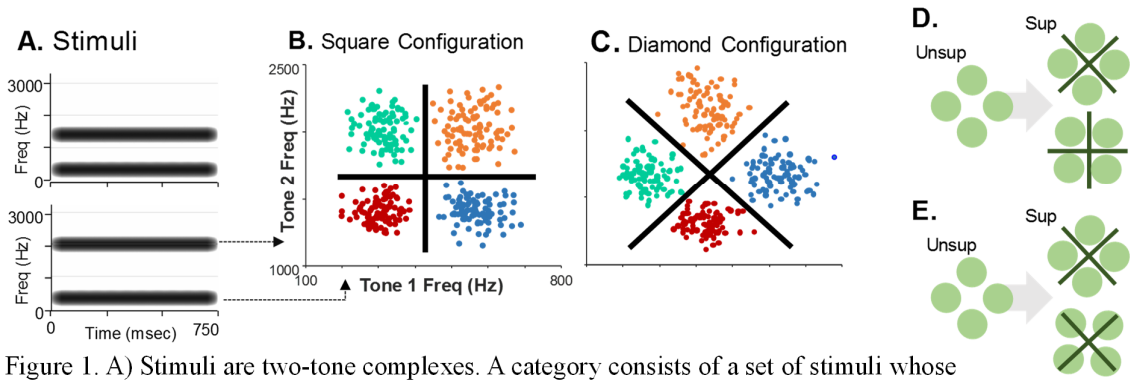

Figure 1. A) Stimuli are two-tone complexes. A category consists of a set of stimuli whose frequencies come from a Gaussian distribution. B) In square configurations, boundaries are drawn in one dimension. C) In diamond configurations, boundaries require integrating both. D) Reinforced statistical learning starts with passive exposure to stimuli from one configuration, followed by supervised training on a configuration that either matches (top) or mismatches (bottom) the unsupervised phase. E) As a control, we hold the boundaries constant during supervised learning and change the prototypes.

in adults. It has been implemented in operant paradigms in which learners hear a stimulus, choose a category and receive feedback (4-7), and in incidental paradigms where learners associate sounds with unambiguous visual cues (11). The type of feedback matters (12), and operant training is more robust than incidental (11). One theoretical model that has been recently been applied to auditory categories is COVIS (13), which posits separate learning systems for categories defined by a boundary along a single feature (rule-based or RB categories), and for categories whose boundaries integrate several dimensions (information integration or II categories). II categories are more difficult to acquire (7, 11), and unlike RB categories, they may require supervision.

In contrast, work in first language acquisition has focused on ***unsupervised statistical learning*** (8, 14). Infants rapidly tune into the categories of their language before they can produce speech or know many words that could serve as a source of feedback. Thus, most accounts suggest infants learn categories by tracking the statistics of the individual acoustic cues. For example, Voice Onset Time (VOT) is the most important cue for distinguishing /b/ and /p/. Across utterances the distribution of VOTs in English displays a bimodal distribution with one cluster corresponding to /b/ and the other to /p/. Infants could thus track the frequency of individual VOTs and identify such clusters to learn the categories of their language. Experiments show that passive exposure to stimuli that form such statistical clusters can alter discrimination in infants and adults (8, 15, 16).

1.2 Reinforced Statistical Learning

This manuscript presents preliminary results from work on the *Reinforced Statistical Learning* paradigm, which combines unsupervised and supervised learning. Our initial implementation captures a sequence characteristic of L1 acquisition, in which supervised learning builds on prior unsupervised learning. Natural speech categories are difficult to fully control and hard for adults to learn due to interference from existing categories. Thus, we use non-speech categories (11, 16) which offer a useful model. Stimuli are two-tone complexes (Figure 1A). For any exemplar, the frequency of each tone is randomly generated from a Gaussian distribution reflecting the prototype structure of the category. We taught subjects four-categories in a two dimensional space. Categories can be arranged in multiple ways (c.f., 11). For example, categories can be arranged in a "square" with boundaries along individual dimensions (RB categories; Figure 1B), or in a diamond with diagonal boundaries (II, Figure 1C).

Learners start with passive exposure to a series of stimuli from one configuration. They next learn the labels for each category in a supervised paradigm. The boundaries during this phase can match or mismatch the unsupervised configuration (Figure 1D). We infer unsupervised learning as a difference in performance as a function of whether the boundaries match across phases. Of course, a mismatch effect could also reflect the fact that the statistical distribution of the stimuli has shifted (not just the mismatch between the boundaries and the unsupervised distributions). We thus manipulate a third factor, shifting the locations of the prototypes across phases but not the boundaries (Figure 1E).

2. Experiment 1

2.1 Methods

Subjects. 159 normal hearing adults were recruited from the University of Iowa community. The intended sample size is 240 (30 / condition), so this should be treated as a preliminary report.

Design. Subjects were randomly assigned to one of 8 conditions that crossed three factors: 1) the configuration of the prototypes during the unsupervised phase (square or diamond); 2) the nature of the boundaries during supervised training (horizontal/vertical vs. diagonal); 3) the distribution of tokens during supervised training (which was included to control for fact that a mismatch between unsupervised and supervised learning could reflect either the boundaries or the prototypes).

During the unsupervised training, participants heard 700 tokens (175 / category) over about 30 minutes. Stimuli were blocked by category. This was followed by 336 trials of supervised training in which categories were fully interleaved (156 / category). Every 72 trials, subjects had 12 test trials (termed periodic test trials) on the prototypes of each category (defined by the boundaries).

Stimuli. Stimuli consisted of two tone complexes. Each was 750 msec long with a 50 msec ramp. Stimuli were generated from a multi-modal Gaussian distribution which specified the mean of each of the two tones. Stimuli were unique on each trial for each subject and generated in bark space. In the square configuration, these means were [3, 9.5], [3, 11.5], [5, 9.5], [5, 11.5]. In the diamond configuration [4, 9.09], [2.59, 10.5], [4, 11.91], [5.41, 10.5]. The standard deviation in both conditions was set to such that the means in each dimension were 5.5 standard deviations apart.

Procedure. After providing informed consent, participants underwent unsupervised training. This was conducted in a sound attenuated room with stimuli delivered at a comfortable listening level over high quality headphones. They were given no particular instructions, and colored in a book during this time. After passive training, the supervised portion began. On each trial, participants heard a single stimulus and saw two novel objects on the screen – one corresponding to the target category, and the other randomly selected from the other three. They then clicked on the matching object, and heard a buzz or a soft beep to indicate if their response was incorrect or correct, respectively.

2.2 Results

Figure 2A shows accuracy over the course of supervised training with trials binned into blocks. Subjects learned this task very rapidly reaching asymptotic performance by around 100 trials. Contrary to the predictions of COVIS (13) subjects learning diagonal (II) boundaries learned more robustly than those learning vertical/horizontal (RB) boundaries.

Figure 2B shows performance on the periodic testing trials (which were equated on difficulty) as a function of condition. There was little change in accuracy over time (since the first periodic testing trial occurred at trial 72, at which point subjects were near asymptote). Thus we collapse across block. Figure 2B again shows stronger performance for the diagonal than vert/horz boundaries: the two sets of bars on the left exceed those for the right. There is also an effect of the supervised configuration – when learning the diagonal boundaries, the diamond configuration (which matched the prototypes for those boundaries) exceeded the square condition (where the training exemplars overlapped the boundaries). The reverse was seen for the vert/horz boundaries. Finally, when learning the diagonal boundaries, there was a small benefit when the unsupervised exposure used the matching, diamond configuration. However, there appeared to be no effect of unsupervised learning for square boundaries.

This was evaluated statistically in a logistic mixed effects model that evaluated the effect of test block, boundary, unsupervised configuration and supervised configuration (all centered) on performance. This model included random intercepts of subject, and a random slope of block (the maximal model). This model should be interpreted cautiously as planned sample sizes have not been achieved. Thus, we only report the most important results. We discuss results pertaining to the supervised and unsupervised portion separately.

Effects on Supervised Learning. We found a highly significant main effect of boundary (B=-.55, SE=.092, Z=5.93, p<.0001): diagonal boundaries were learned better than square. We also found an effect of the location of the supervised prototypes relative to the boundary (e.g., Figure 1E), the supervised configuration × boundary interaction (B=.77, SE=0.19, Z=4.18, p<0.0001). This was due to the fact that diagonal boundaries showed better learning when the supervised configuration was diamond than square, and while vert/horz boundaries showed better learning with square. Learning was better when the frequently trained regions were in the center of the category defined by the boundary (Figure 1E, top), than the when they were along the boundary (Figure 1E, bottom).

Unsupervised learning. Contrary to our predictions, unsupervised configuration did not interact with the boundaries (B=0.146, SE=0.19, Z=0.78, p=0.43). While the means suggested a benefit when

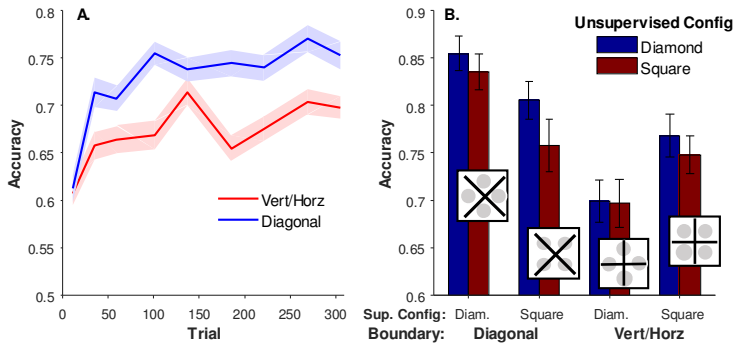

Figure 2. A) Performance on all trials during supervised learning as a function of boundary and trial block. B) Performance on periodic testing trials as a function of unsupervised configuration (red vs. blue bars), supervised configuration, and boundary. Inset boxes show the supervised configuration and boundary for each pair of bars. Error bars are standard error of the mean.

the unsupervised prototypes match the boundary, there was little statistical support for this.

Given that the unsupervised training was blocked by category, one possible explanation for these weak effects is a recency effect whereby only the last category shows any evidence for unsupervised learning (the earlier exposed ones either decayed or were overwritten by the later one). Indeed, Figure 3A suggests this might be the case. In panel A, the final category shows a large benefit of unsupervised exposure: when learning diagonal boundaries, performance showed a 10% gain for unsupervised learning in the diamond configuration; whereas no such effect was seen with square. In contrast, for the earlier three categories (Figure 3B), there is no effect of unsupervised training.

To assess this we conducted an additional mixed effects model which was structured similarly to the prior one, but with an additional term indicating whether the tested category was the final category or not. This was an exploratory (unplanned) analysis so results should be interpreted cautiously. Here we observed a significant three way interaction between boundary, the unsupervised configuration and whether or not we were testing the final category in the unsupervised phase (B=0.58, SE=0.27, Z=2.17, p=0.030). Thus, when unsupervised learning is blocked, learning may be limited to the last category presented, and may further only appear in the diamond configuration.

2.3 Discussion

People quickly and easily learned the categories during the supervised training. Further, when we consider the factors affecting supervised learning only we saw two key patterns. First, contrary to COVIS, learning was better when the boundaries were diagonal (II). Second, we found an unexpected effect of the supervised configuration (e.g., Figure 1E). In particular, participants learned better when supervised training focused on stimuli in the middle of the categories, rather than when they split the boundary. Recall that the periodic testing trials were the same in all conditions (and equally easy – they were in the center of the categories defined by the boundary). Thus, this effect represents a genuine effect on learning, not a confound of difficulty. This pattern of results is consistent with the idea that learners may be acquiring prototypes and not boundaries (regardless of the type of boundary).

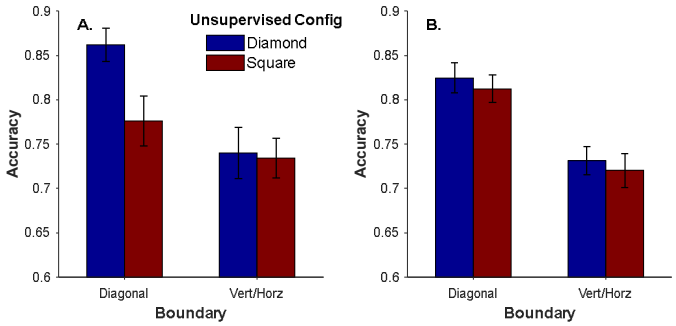

Figure 3. A) Performance on periodic testing trials as a function of unsupervised and supervised configurations, and boundary (diagonal vs. vert/horz), for the final category in the unsupervised phase. B) For all other categories. Error bars represent standard error of the mean.

That is, if people were learning boundaries or rules, one would expect that training which focused on the region of the space near the boundary would yield better performance (e.g. the bottom configuration in Figure 1E). However, instead, training that focused on the region of the space at the center of the category seemed to yield better performance.

When we turn to the unsupervised learning, we were surprised that there were few effects overall. However, when we considered the order of presentation, we found strong effects of the most recently presented category. (Figure 3A). This should be interpreted with caution as data collection is not complete, and these analyses were exploratory. However, these took the form of a consistency effect whereby unsupervised exposure to what would be center of the category (defined by the boundaries) improved performance. The fact that this was only observed in the diagonal categories is inconsistent with COVIS which posits that unsupervised learning is only possible with RB (here, Vert/Horz) boundaries. We return to this in the general discussion.

3. Experiment 2: Interleaved unsupervised training

Given these analyses, we initiated a second protocol in which all four categories were randomly interleaved during unsupervised training. We present preliminary results here, though the sample size is limited so these results should be treated cautiously.

3.1 Methods

Subjects. 74 normal hearing young adults were subjects. These were randomly assigned to one of the same 8 conditions as in Experiment 1 (~9.25 / condition). Our intended sample size is 30 / condition, so these results are preliminary.

Design, Stimuli and Procedures. All methods are identical to Experiment 1 with the exception that the order of stimuli in the unsupervised portion was completely random.

3.2 Results

Figure 4A shows performance during supervised training. Again, subjects learned rapidly and showed a clear benefit for diagonal boundaries. Figure 4B shows performance on the periodic testing trial. Contrary to Experiment 1, there was no effect of unsupervised training on diagonal boundaries, but a large (6%) gain with vert/horz boundaries. Again this is a consistency benefit, but now supervised learning was better if unsupervised learning was in the matching, square configuration.

This was examined with a mixed model using a similar structure to Experiment 1. We again found a significant effect of boundary (B=-0.51, SE=.15, Z=3.50, p=.00047). There was also a three-way interaction between boundary, unsupervised and supervised configuration (B=1.71, SE=0.58, Z=2.95, p=.0032). This supports the benefit of square unsupervised learning for vert/horz boundaries; however, the three way interaction suggests it may have been moderated by the supervised configuration (it was much stronger when the supervised configuration was square [matched the boundaries]). Given the small sample size, we are cautious about interpreting this moderation.

3.3 Discussion

Contrary to COVIS and to (11), Experiment 2 again showed stronger learning for diagonal boundaries than vertical/horizontal ones. More importantly, we also observe an influence of

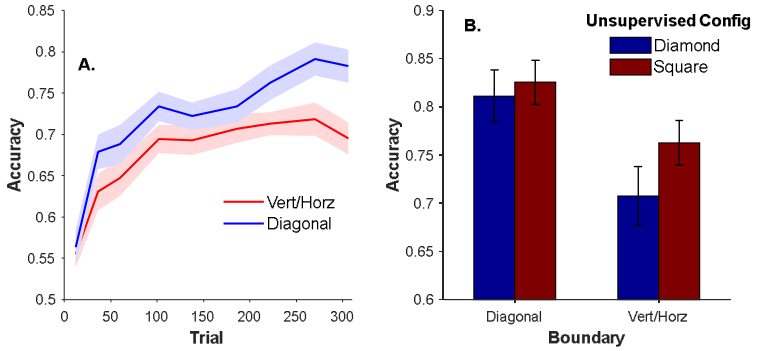

Figure 4. A) Performance on all trials in the supervised learning task as a function of trial block and boundary. B) Performance on periodic testing trials as a function of unsupervised configuration (red vs. blue bars), and boundary (diagonal vs. vert/horz). Error bars represent standard error of the mean.

unsupervised training, but one that differed from Experiment 1. While this influence again took the form of a consistency effect (a benefit to learning when the unsupervised prototypes matched the prototypes defined by the boundaries), it only appeared in the square configuration.

So why would the unsupervised benefit be limited to vert/horz boundaries in this experiment, but appear only with diagonal boundaries in Experiment 1 (although in a limited way)? Theoretical work emphasizes the utility of unsupervised processes for learning *categories* (e.g., the location and extent of the categories in cue space). However, during unsupervised exposure, learners have no goal, and no task; thus it is possible that unsupervised learning is shaping other levels of the system. In particular, unsupervised learning could serve to scaffold the *pre-categorical* representation, utilization and weighting of the auditory cues. Here, blocked exposure (Experiment 1) leads listeners to identify features that are characteristic or defining of a category. This would be enhanced in the diamond configuration, where individual frequencies of one tone or the other can be used to identify individual categories. For example, in Figure 1C, the green category is uniquely associated with the lowest values of Tone 1 (the X-axis), while the red category is associated with the lowest values of Tone 2 (the Y axis). In contrast, unsupervised learning in the square configuration may promote attention to potentially contrastive features. This is supported by Carvahlo and Goldstone (17) who showed in a supervised learning paradigm that interleaving exemplars from different categories promotes attention to contrastive features, while blocked training promotes attention to defining features.

4. Experiment 3: Cortical processes

Experiments 1 and 2 suggest unsupervised learning may alter the pre-categorical representational space. However, this is difficult to assess with behavioral measures. In contrast, neuroscientific techniques may be able to isolate them. Thus, we have been conducting experiments using electro-corticography. In this technique, electrodes are implanted directly on the surface of the brain of patients awaiting treatment for epilepsy. During this time, patients are awake and can complete experiments. The temporal lobes is a common locus of seizures, thus many of these patients receive electrodes on auditory and related areas offering an excellent opportunity to address these questions by targeting the representation of the stimuli in auditory space.

4.1 Methods

Subjects. Subjects were two adult patients (pt. 405 and 409) undergoing treatment for epilepsy. Both were monolingual English speakers with normal hearing, and neurologically normal other than epilepsy. Both were implanted on the left hemisphere.

Procedures. Both patients first conducted a pre-test in which they heard an 8×8 grid spanning the entire Tone1×Tone2 space in a passive listening procedure. Here the goal was to simply collect neural data for the entire range of stimuli. On three subsequent days they then underwent three days of unsupervised learning in the square configuration (768 stimuli/day). On day 5 they underwent an identical post-test. On day 6, patient 409 underwent supervised training (in the square configuration) with post-test on day 7.

Neural Recordings. Recordings were made from subdural electrodes in a 4x8 grid (5 mm spacing) centered on the temporal lobes, and from a depth electrode centered along Heschl's gyrus. See (18) for complete details on recording, and patient care issues.

4.2 Results and Discussion

Our preliminary analyses focused on the degree of activity at targeted electrodes along the Superior Temporal Gyrus (STG) and Heschl's Gyrus (HG). To quantify localized activity, we used a time-frequency approach to quantify activity in the high gamma frequency band (75-150 Hz) (19). Figure 5 shows the results for characteristic electrodes. For all recording sites that were responsive to the stimuli (in both patients), we saw a reduction in activity after unsupervised learning. This suggests that one of the consequences of unsupervised learning may be adaptation. Such a mechanism could play out quite differently in the diamond configuration, where some frequencies (e.g., the middle frequencies of Tone 1 and 2 in Figure 1E), occur more often than others. In contrast, in the square configuration, each frequency will be equally likely (and adapt equally).

In order to capture the representational space, we are conducting machine learning analyses (20) to recover the tone that was heard from the neural recordings. In this analysis, a support vector regression was trained to predict the frequencies of the each tones of the stimulus from the distributed pattern of activity on that trial. The predicted frequencies serve as a metric representation of the

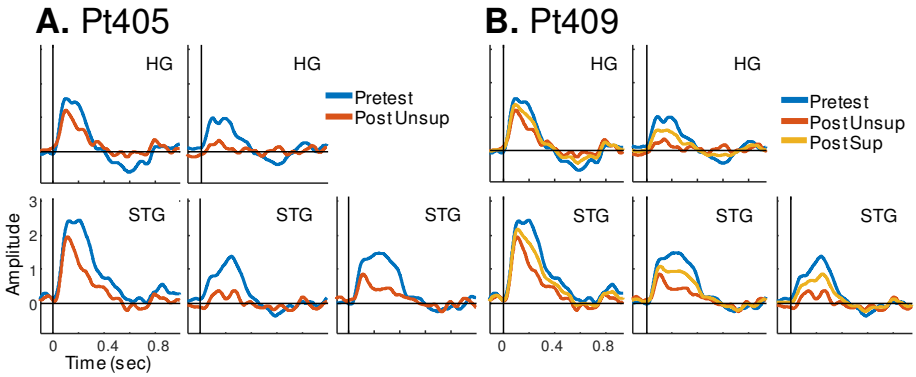

Figure 5. High gamma band activity (in dB) for selected recording sites for each patient.

stimulus space. Figure 4 shows predicted tone frequencies of each point in the grid. There were high correlations between the predicted and actual tones for both patients (Pt405: r=0.89; Pt409: r=0.88). When we compared this space before and after unsupervised learning we found that this space expanded with training (in Pt405; Figure 6), but no such changes were observed in Pt 409. This suggests that one consequence of learning may be a non-specific enhancement of the space, though this awaits confirmation in other subjects.

5. Conclusions

The results of these studies are preliminary and await confirmation. However, they support several conclusions. First, Experiments 1 and 2 suggest that listeners acquire categories with diagonal boundaries (II) better than those with square boundaries (RB). This conflicts with predictions from COVIS (11), and it remains to be seen why the benefit for RB categories was not observed with these stimuli. Second, learners benefit from training focused on the center of the space rather than the boundary regions. This suggests that for auditory categories, people primarily learn prototypes which define the central tendencies of the category rather than boundaries. Third, the effects of unsupervised learning (if present) are weak relative to the much stronger effects of supervised training – unsupervised exposure offered only a 5-10% gain over supervised training and only in some sub-conditions. Fourth, the small effects of unsupervised learning appeared to be moderated by whether stimuli were blocked or interleaved: when unsupervised learning was blocked by category, a small benefit was seen for diagonal boundaries, but when unsupervised learning interleaved all categories, the benefit was seen for vert/horz boundaries. This is consistent with the idea that interleaved training promotes the extraction of discriminative features that contrast categories, while blocked training promotes identification of characteristic features that identify a category (17). Here we extend this to unsupervised learning. Finally, our preliminary ECOG results suggest that unsupervised exposure may offer alter the system in several ways, including local adaptation of the frequency detectors, and potentially an expansion of the representational space.

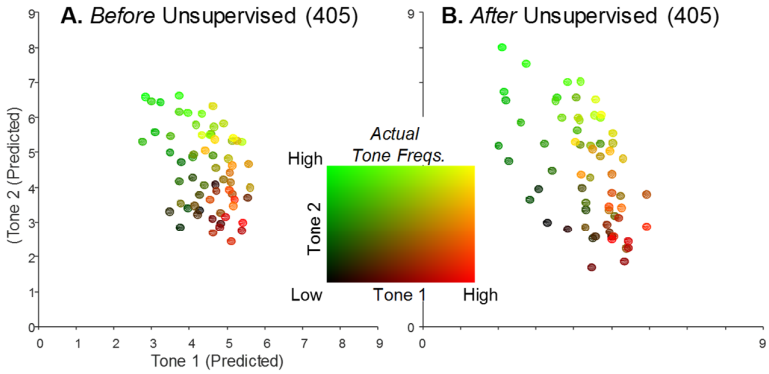

Figure 6. Reconstructed perceptual space for one patient before and after unsupervised learning. X and Y axes represent the predicted tone value; the actual tone is indicated by the color (green = high Tone 2, low tone 1, red = high tone 1, low tone 2).

ACKNOWLEDGEMENTS

The authors would like to acknowledge Matthew Howard, Kirill Nourski and McCall Sarrett for assistance with the ECOG experiments, and Ted Abel for helpful discussions on auditory category learning and Rodica Curtu for suggesting adaptation as a mecahnism. This project was supported by P50DC000242 awarded to B. Gantz, R21DC017596 awarded to B. McMurray and by the Iowa Neuroscience Institute.

REFERENCES

1. Bishop DVM, Snowling MJ. Developmental dyslexia and specific language impairment: same or different? Psychological Bulletin. 2004;130:858-86.
2. Strange W, Shafer VL. Speech perception in second langauge learners: The re-education of selective perception. In: Hansen JGH, Zampini M, editors. Phonology and Second Language Acquisition. Philadelphia: John Benjamins; 2008. p. 153-92.
3. Samuel AG, Kraljic T. Perceptual learning for speech. Attention, Perception, & Psychophysics. 2009;71(6):1207-18.
4. Perrachione TK, Lee J, Ha L, Wong PCM. Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. . Journal of the Acoustical Society of America. 2011;130 (1):461-72.
5. Logan JS, Lively SE, Pisoni DB. Training Japanese listeners to identify English/r/and/l: A first report. The Journal of the Acoustical Society of America. 1991;89(2):874-86.
6. McCandliss BD, Fiez JA, Protopapas A, Conway M, McClelland JL. Success and failure in teaching the [r]-[l] contrast to Japanese adults: tests of a Hebbian model of plasticity and stabilization in spoken language perception. Cognitive, Affective, & Behavioral Neuroscience. 2002;2:89-108.
7. Chandrasekaran B, Yi H-G, Maddox WT. Dual-learning systems during speech category learning. Psychonomic Bulletin & Review. 2014;21(2):488-95.
8. Maye J, Werker JF, Gerken L. Infant sensitivity to distributional information can affect phonetic discrimination. Cognition. 2003;82:101-11.
9. Escudero P, Williams D. Distributional learning has immediate and long-lasting effects. Cognition. 2014;133(2):408-13.
10. McMurray B, Danelz A, Rigler H, Seedorff M. Speech categorization develops slowly through adolescence. Developmental Psychology. 2018;54(8):1472-91.
11. Roark CL, Holt LL. Task and distribution sampling affect auditory category learning. Attention, Perception, & Psychophysics. 2018;80(7):1804-22.
12. Yi H-G, Chandrasekaran B. Auditory categories with separable decision boundaries are learned faster with full feedback than with minimal feedback. The Journal of the Acoustical Society of America. 2016;140(2):1332-5.
13. Ashby FG, Maddox WT. Human category learning 2.0. Annals of the New York Academy of Sciences. 2011;1224(1):147-61.
14. McMurray B, Aslin RN, Toscano JC. Statistical learning of phonetic categories: Insights from a computational approach. Developmental Science. 2009;12(3):369-79.
15. Escudero P, Benders T, Wanrooij K. Enhanced bimodal distributions facilitate the learning of second language vowels. The Journal of the Acoustical Society of America. 2011;130(4):EL206-EL12.
16. Goudbeek M, Swingley D, Smits R. Supervised and unsupervised learning of multidimensional acoustic categories. Journal of Experimental Psychology: Human Perception and Performance. 2009;35(6):1913-33.
17. Carvalho PF, Goldstone RL. Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. Memory & cognition. 2014;42(3):481-95.
18. Nourski KV, Howard MA, 3rd. Invasive recordings in the human auditory cortex. In: Celesia GG, Hickok G, editors. Handbook of clinical neurology. 1292015. p. 225-44.
19. Kaiser J, Lutzenberger W. Human gamma-band activity: A window to cognitive processing. NeuroReport. 2005;16(3):207-11.
20. Nourski KV, Steinschneider M, Rhone AE, Oya H, Kawasaki H, Howard MA, 3rd, et al. Sound identification in human auditory cortex: Differential contribution of local field potentials and high gamma power as revealed by direct intracranial recordings. Brain and Language. 2015;148:37-50.