A Reinforcement Learning Approach to Speech Category Acquisition

Craig Thorburn, Ellen Lau, and Naomi Feldman

1. Introduction

Language learners need to map a continuous, multidimensional acoustic signal to discrete abstract speech categories. The complexity of this mapping poses a difficult learning problem, particularly for second language learners who struggle to acquire the speech sounds of a non-native language, and almost never reach native-like ability. A common example used to illustrate this phenomenon is the distinction between /r/ and /l/ (Goto, 1971). While these sounds are distinct in English and native English speakers easily distinguish the two sounds, native Japanese speakers find this difficult, as the sounds are not contrastive in their language. Even with much explicit training, Japanese speakers do not seem to be able to reach native-like ability. (Logan et al., 1991; Lively et al., 1993)

A particularly effective strategy for learning speech categories as an adult, however, is through implicit learning within a video game paradigm (Wade and Holt, 2005; Gabay et al., 2015; Lim and Holt, 2011). In these experiments, participants control a spaceship at the center of a screen with aliens entering from various locations. Participants must shoot or capture these aliens and increase their score every time they do so. Auditory tokens are presented before each alien enters and are sampled from a category that corresponds to the aliens' location and color, providing an early cue for identifying which direction the participant should face. As the experiment progresses, aliens begin to move faster and it becomes increasingly difficult to turn and shoot or capture without attending to the auditory information. Improved performance for the recognition of categories between the start and end of gameplay is observed for various types of stimuli — both synthesized noise (Lim et al., 2019) and speech tokens (Lim and Holt, 2011). In the latter experiment, adults show as much improvement in the discrimination of non-native speech sounds after 2 hours of video game training, as they do during 10 hours of explicit training over a period of 2-4 weeks (Logan et al., 1991). A potential reason that this paradigm is so effective could lie in its engagement of a

* Craig Thorburn, Department Of Linguistics and UMIACS, University Of Maryland, craigtho@umd.edu; Ellen Lau, Department Of Linguistics, University Of Maryland; Naomi Feldman, Department Of Linguistics and UMIACS, University Of Maryland. This research was supported by NSF grant BCS-2120834. We thank William Idsardi, the UMD Language Acquisition Lab, and the UMD Phonology Circle group for helpful comments and discussion.

 Proceedings of the 46th annual Boston University Conference on Language Development, ed. Ying Gong and Felix Kpogo, 797-811. Somerville, MA: Cascadilla Press.

reward-based learning algorithm. When undergoing learning in this environment, neural activity is observed in the striatum (Lim et al., 2014, 2019) - a subcortical area of the brain associated with reinforcement and reward mechanisms. This region is considered to play an important role in reward through the activation of dopaminergic circuits. While studies on the striatum have typically focused on reward prediction, recent studies have suggested that the basal ganglia could implement reinforcement learning specifically (Kawagoe et al., 1998; Joel et al., 2002; Cohen and Frank, 2009; Dabney et al., 2020)

In reinforcement learning (Sutton and Barto, 1998), an agent takes actions within an environment, receiving a reward upon reaching a favorable outcome. At each timestep the agent receives information about the state of the environment and predicts the value of each action available to it. The agent then takes an action, observes the reward received, and updates its parameters according to the mismatch between the predicted and observed reward value. This process is iterative, enabling the agent to take actions that lead to reward in the future. Even if the agent does not receive a reward immediately, it will still take actions that put it in a better position to receive a reward at another timestep.

The construction of the video game paradigm has components which lend itself clearly to the implementation of a reinforcement algorithm: a participant acts within an environment and receives rewards. This paradigm has been used in many experiments and results in robust learning across different variations of stimuli. At the same time, reinforcement mechanisms have been proposed to play a role in other speech learning paradigms (Nixon, 2020; Harmon et al., 2019). However, these two ideas have not yet been explicitly unified.

In this paper, we formalize the video game training paradigm explicitly as a reinforcement learning problem. We present two simulations. The first simulates findings of Lim et al. (2019), providing proof in principle that a reinforcement learning algorithm can successfully capture human results in a video game where adults learn novel categories of noise tokens. Our second simulation extends this to speech sounds and demonstrates that our algorithm mimics second language learners' improvement on discrimination of a non-native speech contrast. Together these two simulations show that reinforcement learning provides an accurate model of human learning in this paradigm and provide evidence supporting the hypothesis that this mechanism could play a key role in effective speech category learning in adults. In the long term, being able to identify the algorithms employed in this paradigm could provide many avenues for pedagogical changes in second language learning and let teachers harness the processes that allow for efficient learning and improvement of non-native perceptual ability.

2. Simulation 1: Video Game With Noise Categories

Our first simulation models the results of Lim et al. (2019) and demonstrates that a reinforcement algorithm is able to correctly discriminate between auditory noise categories with simple boundaries. We simulate human experiments by training a neural network to map between states and actions within a video game environment. We then evaluate the model by observing how much reward the agent receives for various categories of stimuli. We compare these results across various token types with the behavioral results observed.

2.1. Computational Model

We define the reinforcement framework formally as a Markov Decision Process where the environment is given by a set of states $S = \{s_1, s_2, s_3, ...\}$, one of which is presented to the agent at each time step t. The agent then has a set of actions $A = \{a_1, a_2, a_3, ...\}$ that it can take. At each timestep, the agent receives information about the state, chooses an action, and receives a reward r_t before the environment transitions to the next state. The transition between states can depend on both the actions of the agent and external factors of the environment. The goal of the agent is to find a policy mapping between states and actions (s, a) that will maximize the total reward for this and future steps, by updating its parameters depending on the reward it receives.

In this paper, we use model-free reinforcement learning, where the policy of the agent is determined by the Q-function, defined as follows.

$$Q(s_t, a_t) = r_t + \gamma * \text{max} Q(s_{t+1}, a_{t+1}) \tag{1}$$

This function outputs the maximum possible reward for a specific action, by summing the reward at the current timestep r_t plus the maximum possible reward from the next timestep max $Q(s_{t+1}, a_{t+1})$ multiplied by a discount parameter γ. This function is iterative since the maximum value at the next timestep will apply the same function, allowing the model to account for future reward. This mapping is not directly observable for the agent, and so it must estimate the value of each action at each state by approximating this function. While there are various methods for performing this approximation, we use a specific implementation of deep Q learning (Mnih et al., 2015) to build a deep neural network that represents the Q-function (DQN). This network takes state information as input and outputs a distribution of values over all actions the agent can take.

The agent's decisions in this model follow an ϵ-greedy algorithm to allow for a tradeoff between exploring the environment to discover patterns and choosing the most optimal action to receive a reward. Throughout the simulation, the agent will choose a random action with probability ϵ and choose the most optimal outcome with probability $1 - \epsilon$, where the optimal outcome is the action with the highest value, as determined by the network. The parameter decays during training to ensure that the agent has enough time to explore the full spectrum of state-action pairings at the start of the simulation, before starting to behave more optimally.

In the experimental paradigm (Lim and Holt, 2011; Lim et al., 2019), the player receives both auditory and visual cues as they can see the location/color of the alien and hear the corresponding sound. Over time, the game speed increases and in order to have a chance at shooting the alien, the participant must rely increasingly on the auditory cues as these are presented before the alien enters the screen. In our simulations, we present the agent with one repetition of an auditory token on each timestep, as well as the current direction the spaceship is facing. We do not present visual information identifying where the alien is situated, meaning that the agent does not have grounding for the correspondence between alien locations and the auditory category. This would be like a human playing a version of the game where they can see the direction that the spaceship is facing, but cannot see the location of any aliens. In theory, this should make the task more difficult as the agent cannot bootstrap the learning of auditory categories from any visual information present. This ensures that any success from our model is due to the exploration of the reinforcement algorithm and the reward received, rather than simple correlation between visual and auditory signals.

We discretize the continuous nature of the game, where the participant can take actions at any time, into timesteps where at each step the participant receives information about their location and environment and selects one action out of those available to them. Actions consist of turning left or right, shooting an alien, or capturing an alien. The reward is derived from the increase in score by correctly performing these actions.

2.2. Stimuli

We use stimuli constructed following the parameters from Lim et al. (2019), with offset and onset categories created across an experimental and control condition (Figure 1). Offset categories are consistent across the two conditions and can be identified by just one dimension - the trajectory of change after 150ms. Onset categories are constructed differently between the two conditions. In the experimental condition, one must attend to both the trajectory of change during the first 150ms of the stimulus and the starting frequency in order to identify the correct token. In the control condition, onset tokens are randomized and have no identifiable category boundary. The construction of stimuli in this way yields three types of category boundaries, which we refer to as SIMPLE, COMPLEX, and INCOHERENT. Offset categories in both conditions have a SIMPLE boundary, as these can be distinguished along only one dimension. Onset categories have a COMPLEX boundary in the experimental condition due to the two-dimensional nature of the category boundary, whereas onset categories in the control condition are INCOHERENT as there is no category boundary available.

We simulate the synthesized sounds in this experiment by using a Gaussian distribution centered around each frequency peak at each timestep. This means that the simulation does not differentiate the offset and onset stimuli by wave type, requiring it to discover specific patterns in the acoustic signal. Each Gaussian distribution has a peak of 10 and a variance of 150Hz. For each of the 4 categories, 11 tokens are synthesized with 4 presented during training and the remaining 7 withheld for testing.

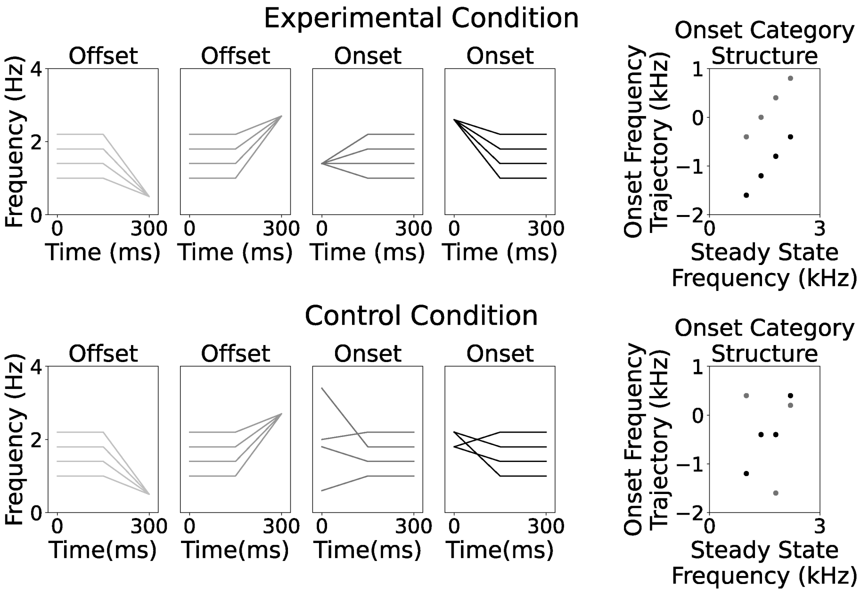

Figure 1: Frequency profiles (left) and onset category structure (right) for simulation 1 noise categories in the experimental (top) and control (bottom) condition. Profiles from Lim et al. (2019).

2.3. Game Environment

The environment for the video game in our simulations is constructed with 8 different directions in which the agent can face (N, NE, E, SE, S, SW, W, NW), where a left or right action turns the agent by 45°. Aliens can enter from 4 directions (N, E, S, and W), and each of these is associated with a different category. Each timestep consists of the presentation of one auditory token and one action by the agent. If the agent is facing the opposite direction to the location of a newly presented alien, it will need a minimum of 5 actions to turn toward and successfully shoot it. To allow enough time for this to happen, each stimulus is presented 8 times. If the agent does not shoot correctly within this time, then the alien disappears without the agent receiving any reward and the next stimulus is presented. During training, there are 4 tokens within each category presented during training and the agent has 3 actions available to it: TURN LEFT, TURN RIGHT or SHOOT. The agent receives one point of reward for every alien it shoots (unlike the original study, our simulations do not have a 'capture' mechanic). Location information is presented to the model as a one-hot vector of length 8 — ie., the value for the corresponding direction where the agent is facing is equal to 1 with all other values set to 0. This is not visual information and simply instills knowledge of the direction that the agent is facing into the model. The model does not receive any information about the location of the alien. The network combines the auditory representation with the location information and outputs a value assigned to each action using a deep Q network, trained with the algorithm outlined above and a network representation illustrated in Figure 2. The network is trained over a total of 2000 episodes, each consisting of 128 tokens and tested on an additional 500 episodes of 176 tokens with weights frozen. We defined performance by looking at the proportion of total possible reward the agent receives for each category type during this test phase.

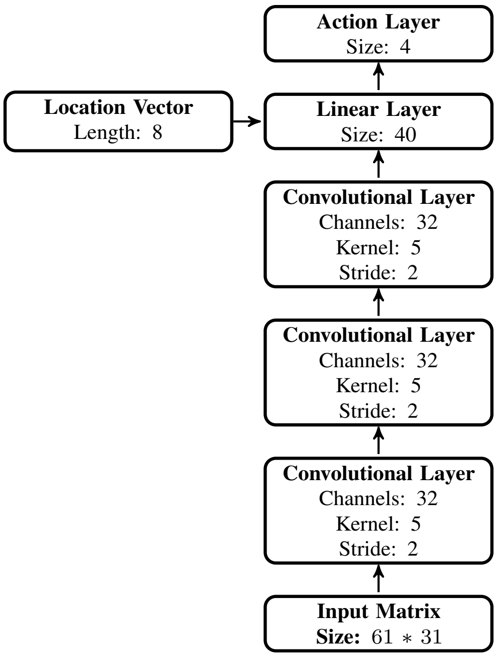

Figure 2: Network diagram for simulation 1.

2.4. Results

Results during the testing period align very closely with the post-test categorization responses from the original experiment (Figure 3). As expected, the model performs successfully in the paradigm and exhibits better performance for tokens presented during training than novel tokens. Neural networks are good at memorizing specific tokens during training, so novel tokens provide a better test of generalization. Generally, higher performance is observed for offset tokens than onset tokens, replicating human behavior. This is because the boundary between categories has lower dimensionality in the offset condition than in the onset condition. Onset tokens, however, have a more complex boundary, being defined by both the frequency trajectory and the steady-state frequency. To simplify the presentation of tokens for the incoherent boundary condition in this simulation, 4 specific tokens are presented, rather than a range of randomized tokens as in the original experiment, resulting in our model having a higher success rate for these tokens than the original experiments. As there is no boundary to generalize, however, we see that performance for novel tokens is low - mirroring the human results. The model also shows worse generalization for simple stimuli in the control condition than the experimental condition for novel tokens, despite the fact that these tokens are the same. These results are also seen in the original study and suggest that the noise provided by the incoherent tokens is enough to decrease generalization across the board, and not just for the onset tokens themselves. The above results are all consistent across various parameter settings of the network.

This simulation suggests that it is possible to gather implicit knowledge about auditory categories through a reinforcement signal, giving rise to human-like behavior. While the model is never shown directly which tokens are members of each category, it generalizes to novel stimuli within the same category. Having shown that a reinforcement algorithm is able to use perceptual information within this task, the next step is to show that the same algorithm can overcome native language knowledge and improve specifically in discrimination of non-native speech categories.

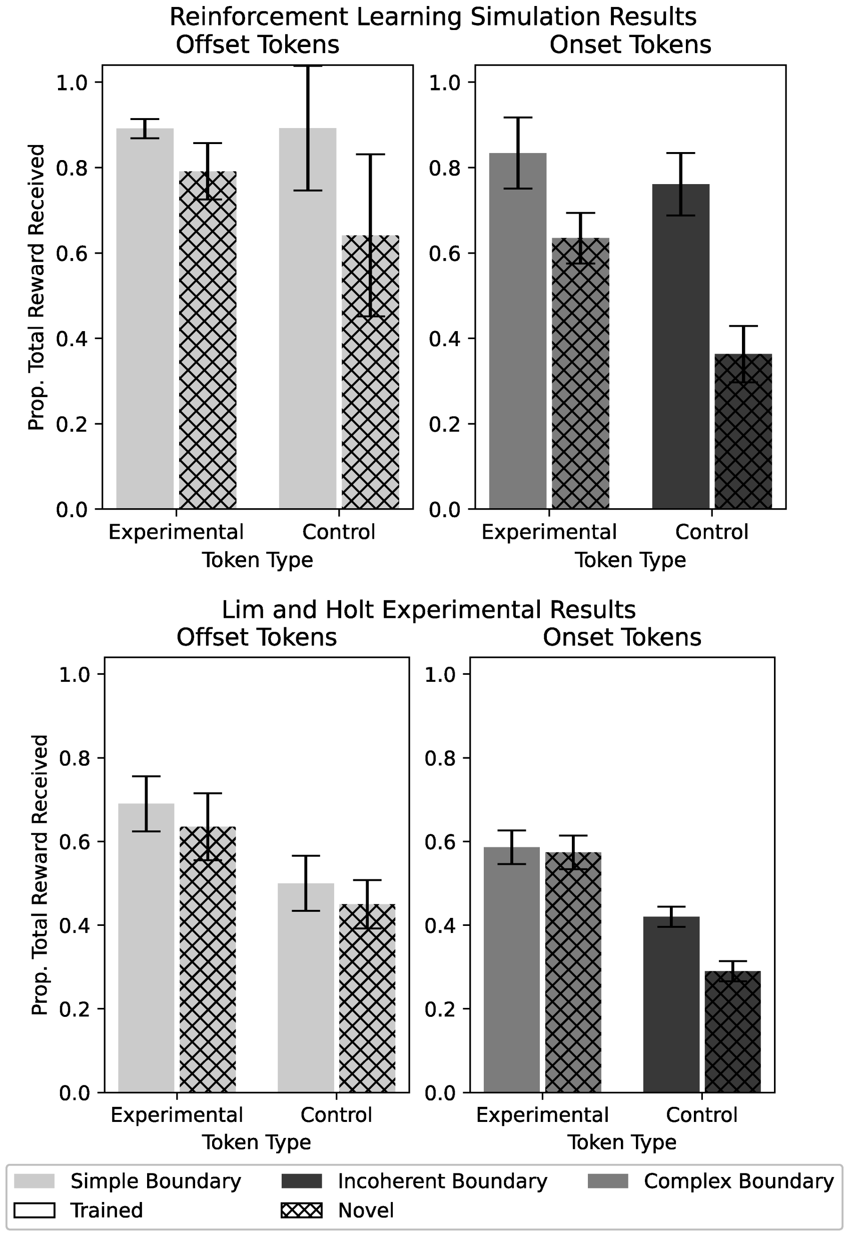

Figure 3: Simulation 1 results (top) and experimental results from Lim et al. (2019) (bottom) with 95% confidence intervals.

3. Simulation 2: Video Game With Speech Categories

Our second simulation models a learning experiment where Japanese and English speakers participate in the video game with English speech tokens presented as categories (Lim and Holt, 2011). Participants hear tokens of [ra], [la], [da], and [ga] before each alien enters and Japanese speakers show improvement on [r]/[l] discrimination after 5 days of training without reaching native-level performance, as measured by performance on an identification task presented before and after the training period. By modeling this experiment with reinforcement learning, we aim to show that our reinforcement algorithm is able to overcome native knowledge to improve on discrimination between the auditory categories presented - mirroring the human results. This simulation consists of two parts: native language training where we model the knowledge of a native Japanese speaker, and video game training, where we expose this native model to the video game environment.

3.1. Methods

We model native language knowledge by training a supervised phoneme classifier on Japanese speech. This is a simple yet efficient way to instill native language knowledge into the model. The network takes acoustic parameters as input and outputs a distribution over phonemes, indicating the identity of the phone at the center of the input window. Training data consists of auditory tokens sampled from labeled corpora as described below. The model is presented with a window of speech and a corresponding vector indicating the phone at the center of the window and is trained using stochastic gradient descent over the network's parameters. We do not mean this supervised training to be an accurate model of first language acquisition, but rather, a way to quickly create networks that have knowledge of a "native" language. We similarly train an English model to use as a native language control.

The native network (shaded in Figure 4) consists of 3 convolutional and batch normalization layers followed by 1 linear layer. Data for native language training is derived from the Wall Street Journal Corpus for English (Paul and Baker, 1992) and the GlobalPhone Japanese Corpus (Schultz et al., 2013) for Japanese. The model is presented with 19.5 hours of speech and training occurs for 25 epochs.

We subsequently expose our native language model to training in the video game paradigm, where phonetic information is combined with the location vector into a linear layer as in simulation 1 (unshaded in Figure 4) and the model once again outputs its expected value of each action. The training signal is allowed to backpropagate through all layers of the network and can change the initial parameters relating to the phone classification native language training. The environment and presentation of stimuli in the video game occurs as described in experiment 1, over 750 episodes.

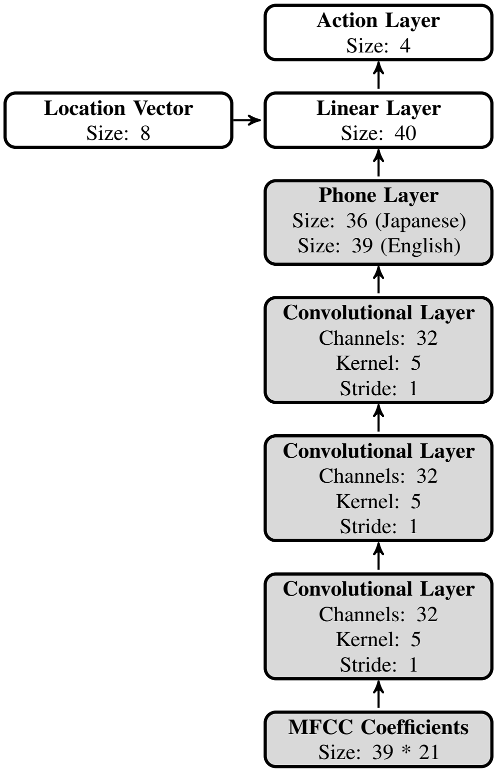

Figure 4: Network diagram for simulation 2

We sample four sounds in each category ([r], [l], [d], and [g]) from the Wall Street Journal Corpus for use in video game training. These stimuli are those that the English native language training successfully categorizes and are kept constant throughout the experiment. Once again, there are 4 tokens for each category, where each category reflects a different location of the alien. Acoustic input during native language training and video game training is given as Mel Frequency Cepstral Coefficients (MFCCs) (Mermelstein, 1976), which are designed to describe the overall shape of the acoustic at any one point in time. We take 13 coefficients and first- and second-order derivatives, giving 39 total dimensions. A 200ms speech window is segmented into frames that are 25ms wide at 10ms intervals and zero-padded, yielding a total of 21 frames. Real speech data sampled from corpora inherently are noisier than controlled lab-prepared stimuli. Using this more naturalistic data for both training and testing demonstrates that the model can deal with variability and is not sensitive only to stimuli created specifically for this experiment.

We simulate the [r]/[l] pre- and post-test discrimination tasks from the original experiment using a machine ABX test, which is a parameter-free method of measuring the distance between model representations (Schatz et al., 2013; Schatz, 2016). We take a vector of activations for a presented token at the ‘phone’ layer of the model. Two tokens are taken from one category - A and X - and a third token from a different category - B. We take the Euclidean distance between vectors A and X, and B and X and determine which distance is shorter. If X is determined to be close to A, then the trial is a success, if X is closer to B, then the trial has failed. The ABX success rate is defined as the probability of success for two tokens from these categories selected at random from the corpus. An ABX success of 1 indicates perfect discrimination, with 0.5 being chance performance. We run the ABX task over 12,000 samples of [r] and [l] drawn from a portion of the WSJ corpus withheld for testing.

3.2. Results

Pre- and post-training ABX success rates are shown in Figure 5. The Japanese model improves in [r]/[l] discrimination ability after video game training. This improvement is small but significant ($t = 2.77$, $p = 0.01$), and the performance does not reach that of the native English model. This is consistent with results from Japanese native speakers who participate in the experiment who show increased [r]/[l] discrimination performance but do not reach native-level performance.

These results show that in principle, a reinforcement signal is enough to alter speech representations and enable an agent to improve their discrimination of ‘non-native’ speech sounds, strengthening the case the the video game paradigm results could come from the use of reinforcement learning. It is once again important to note that in this experiment, at no point is the model given precise information about the identity of these tokens. It discovers any structure in the information through the reward signal alone.

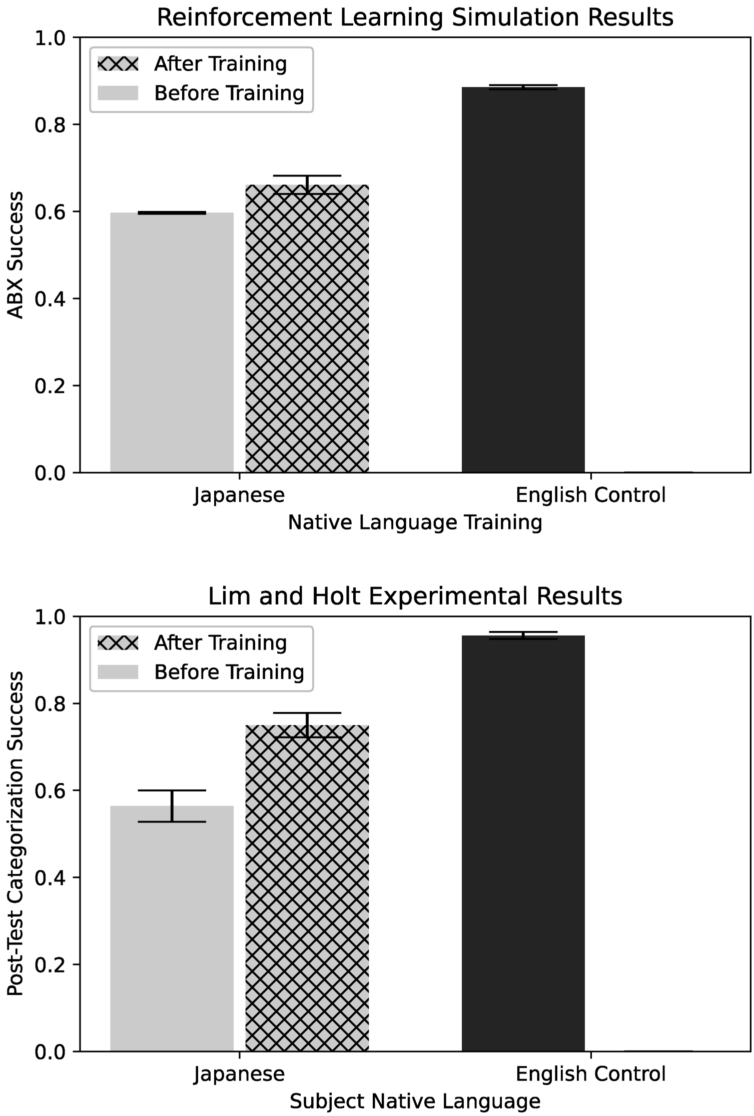

Figure 5: Simulation 1 results (top) and experimental results from Lim et al. (2019) (bottom) with 95% confidence intervals.

4. Discussion

Simulations in this paper show that a reinforcement learning signal is sufficient for a model to exhibit human behavior in two auditory category learning experiments. By developing a deep Q network that processes acoustic and location information, we model experimental paradigms with outcomes that closely match human behavior. These results provide additional evidence towards the theory that reinforcement learning plays a critical role in adults' speech category acquisition. Taken alongside neural evidence showing the activation of reward centers of the brain during participation, we suggest that reward mechanisms could be key to updating perceptual representations.

Second language learners struggle to learn non-native speech sounds and appear to perform poorly in natural environments, with just passive exposure. By framing the video game paradigm as reinforcement learning we explicitly conceptualize a learning signal as a reward, perhaps explaining why this paradigm can be effective without explicit feedback. One possible account of this process neurally is that a reward signal from the striatum passes to areas which deal with auditory processing directly. Neuroimaging of the video game not only shows activation of the basal ganglia during category learning, but participants who exhibit learning show strong functional connectivity between the striatum and Superior Temporal Sulcus—the primary centers of speech processing (Liebenthal et al., 2005). This pattern of top-down feedback to perceptual regions is also observed in other speech category literature. Myers and Swan (2012) indicate that top-down category information from frontal regions of the brain can direct information to temporal regions to reshape the perceptual space.

It is possible that reinforcement learning could play a role in non-native speech sound learning more generally - an area which warrants further investigation. Our work is consistent with other behavioral work pointing to the utilization of reinforcement in speech category learning (Nixon, 2020). Harmon et al. (2019) demonstrate that reinforcement learning models human behavior in a different paradigm where the down weighting of a phonetic cue will only occur if there is a more informative cue present. Additionally, the neural signatures observed in initial video game experiments are observed in other category learning paradigms (Golestani and Zatorre, 2004), which show activation of the caudate nucleus in the striatum during explicit category learning. Future work could also expand this algorithm to native language learning in infants, who appear to acquire speech sound categories in a passive environment.

Our framework opens avenues to investigating the flexibility of representations of speech and future work can manipulate which layers of the network can be influenced by the reinforcement learning signal, and how those weight updates occur. This will allow one to test whether there are particular parts of the processing system that are less flexibly updated, or not updated at all, during second language learning.

Finally, the simulations in this paper have implications for L2 teaching pedagogy. Learning the sounds of a second language has always been challenging, particularly when those sounds do not match up with one's native language phonology. Our work reinforces the idea that implicit learning can provide an effective strategy for learning non-native speech sounds. Understanding why this paradigm is so effective will allow us to consider other tasks and activities that could lead to successful learning. Many people use mobile game-like applications to learn foreign languages, which often include a level system where learners receive rewards such as cosmetic upgrades, items, or points for a leaderboard system. These could provide effective incentives for efficient learning as in a reinforcement framework. With the current prevalence of these applications, reinforcement learning, where improvement is motivated through rewards, is a particularly relevant topic and further simulations on the impact of rewards could greatly benefit this kind of learning.

References

Cohen, Michael X., and Michael J. Frank. 2009. Neurocomputational models of basal ganglia function in learning, memory and choice. *Behavioural Brain Research* 199:141–156.

Dabney, Will, Zeb Kurth-Nelson, Naoshige Uchida, Clara Kwon Starkweather, Demis Hassabis, Rémi Munos, and Matthew Botvinick. 2020. A distributional code for value in dopamine-based reinforcement learning. *Nature* 577:671–675.

Gabay, Yafit, Frederic K. Dick, Jason D. Zevin, and Lori L. Holt. 2015. Incidental Auditory Category Learning. *Journal of experimental psychology. Human perception and performance* 41:1124–1138.

Golestani, Narly, and Robert J Zatorre. 2004. Learning new sounds of speech: reallocation of neural substrates. *NeuroImage* 21:494–506.

Goto, Hiromu. 1971. Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia* 9:317–323.

Harmon, Zara, Kaori Idemaru, and Vsevolod Kapatsinski. 2019. Learning mechanisms in cue reweighting. *Cognition* 189:76–88.

Joel, Daphna, Yael Niv, and Eytan Ruppin. 2002. Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Networks: The Official Journal of the International Neural Network Society* 15:535–547.

Kawagoe, Reiko, Yoriko Takikawa, and Okihide Hikosaka. 1998. Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neuroscience* 1:411–416.

Liebenthal, Einat, Jeffrey R. Binder, Stephanie M. Spitzer, Edward T. Possing, and David A. Medler. 2005. Neural substrates of phonemic perception. *Cerebral Cortex (New York, N.Y.: 1991)* 15:1621–1631.

Lim, Sung-Joo, Julie Fiez, and Lori Holt. 2019. Role of the striatum in incidental learning of sound categories. *Proceedings of the National Academy of Sciences* 116:201811992.

Lim, Sung-Joo, Julie A. Fiez, and Lori L. Holt. 2014. How may the basal ganglia contribute to auditory categorization and speech perception? *Frontiers in Neuroscience* 8.

Lim, Sung-Joo, and Lori L. Holt. 2011. Learning Foreign Sounds in an Alien World: Videogame Training Improves Non-Native Speech Categorization. *Cognitive Science* 35:1390–1405.

Lively, S. E., J. S. Logan, and D. B. Pisoni. 1993. Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America* 94:1242–1255.

Logan, John S., Scott E. Lively, and David B. Pisoni. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *The Journal of the Acoustical Society of America* 89:874–886.

Mermelstein, P. 1976. Distance measures for speech recognition, psychological and instrumental. *Pattern Recognition and Artificial Intelligence* 374–388.

Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. 2015. Human-level control through deep reinforcement learning. *Nature* 518:529–533.

Myers, Emily B., and Kristen Swan. 2012. Effects of Category Learning on Neural Sensitivity to Non-native Phonetic Categories. *Journal of Cognitive Neuroscience* 24:1695–1708.

Nixon, Jessie S. 2020. Of mice and men: Speech sound acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition* 197:104081.

Paul, Douglas B., and Janet M. Baker. 1992. The Design for the Wall Street Journal-based CSR Corpus. In *Speech and Natural Language: Proceedings of a Workshop Held at Harriman, New York, February 23-26, 1992*.

Schatz, Thomas. 2016. ABX-disciminability measures and applications. Doctoral Dissertation, Universite Paris 6 (UPMC).

Schatz, Thomas, Vijayaditya Peddinti, Francis Bach, Aren Jansen, Hynek Hermansky, and Emmanuel Dupoux. 2013. Evaluating speech features with the Minimal-Pair ABX task: Analysis of the classical MFC/PLP pipeline. In *INTERSPEECH 2013: 14th Annual Conference of the International Speech Communication Association*, 1–5. Lyon, France.

Schultz, Tanja, Thang Vu, and Tim Schlippe. 2013. GlobalPhone: A Multilingual Text & Speech Database in 20 Languages.

Sutton, Richard S., and Andrew G. Barto. 1998. *Reinforcement Learning: An Introduction*. MIT Press.

Wade, Travis, and Lori L. Holt. 2005. Perceptual effects of preceding nonspeech rate on temporal properties of speech categories. *Perception & Psychophysics* 67:939–950.

Proceedings of the 46th annual Boston University Conference on Language Development

edited by Ying Gong
and Felix Kpogo

Cascadilla Press Somerville, MA 2022

Copyright information

ISSN 1080-692X
ISBN 978-1-57473-077-7 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com