

# Supporting Research into Sound and Speech Learning through a Configurable Computer Game

Garrett Kimball, Rodrigo Cano, Jingyi Feng, Lei Feng, Erica Hampson, Evan Li, Michael G. Christel  
Entertainment Technology Center  
Carnegie Mellon University  
Pittsburgh, USA  
{gkimball, rcano, jingyif, lf1, ehampson, xiuyuanl}@andrew.cmu.edu, christel@cs.cmu.edu

Lori L. Holt, Sung-joo Lim, Ran Liu, Matthew Lehet  
Department of Psychology &  
Center for the Neural Basis of Cognition  
Carnegie Mellon University  
Pittsburgh, USA  
loriholt@cmu.edu,  
{sungjol, ranliu, mil}@andrew.cmu.edu

**Abstract**— Cognitive neuroscientists studying sound and speech learning have successfully used videogames as a research vehicle. This paper describes the development of a game built to entice participants to longer periods of play, and to enable researchers to configure presentation parameters in support of varying studies. A space-themed game polished through the use of shaders and a radial cannon shooting mechanic is detailed, along with lessons learned from iterative playtesting. A preliminary study indicates the game's effectiveness for implicit learning of sounds. The template by which this game can be tuned to explore language learning is presented.

**Keywords**—game design, sound learning, speech learning.

## I. INTRODUCTION

A recent *Nature* article offered examples and evidence that playing games can have beneficial effects on the brain [1]. The article continues with a call for neuroscientists to work with game developers in producing compelling video games that boost brain function and improve well-being. Within Carnegie Mellon University, the call was answered by researchers and game developers from the Center for the Neural Basis of Cognition (CNBC) and Entertainment Technology Center (ETC) working together to produce a configurable game for sound and speech learning.

Humans are adept at decoding speech, capable of inferring meaning despite wide variation in enunciation and pronunciation. How does this learning occur? Can the process be studied through video games? In fact, the CNBC researchers have established a track record of using video games to study language learning. In one experiment, the only sounds heard by 77 adult players working with a game were deliberately distorted speech, unintelligible in any language. With just two hours of play, the participants could reliably extract word-length sound categories from continuous alien sounds and apply that learning to advance throughout the game [2]. In another study, native Japanese speakers learning the unfamiliar English "r" and "l" sound categories played a video game for 2.5 hours across 5 days. The game emphasized association among sound categories, visual information, and players' responses to game characters rather than overt categorization or explicit feedback. The subjects demonstrated improvement in "r"-l" perception on par with two to four weeks of explicit

categorization training in previous research [3]. In a third study [4], post-game sound categorization patterns demonstrated that even without explicit feedback, game players can learn complex non-speech auditory categories within a rather short period (30 minutes) of incidental exposure when higher-dimensional acoustic cue relationships are present. In a posttest, 42 college students were able to match visual characters from the game reliably to sound category exemplars encountered during game play, as well as to novel sound exemplars drawn from the same category distributions [4]. This study showed that adult listeners can acquire categories for spectrally varying acoustic exemplars, in a task in which categorization is largely incidental.

The base game used and modified for these experiments is discussed in Section II. It has a simplistic presentation and shooting mechanic that CNBC researchers anecdotally noted as becoming less and less appealing to participants over the years as the participant pool becomes more accustomed to current game capabilities. As CNBC proposes longitudinal studies where players interact with the game across numerous hours, the game's appeal is important to keep players actively engaged throughout the experiment. The base game required a large time investment to modify cueing or audio prompts in order to change study parameters (e.g., to support studies as reported in [2], [3], and [4]). In summary, a more up-to-date, configurable, appealing game was envisioned to open up more possibilities for game-based studies into language learning, bringing together the CNBC and ETC teams in making the game reported in Section III.

## II. BASELINE CNBC GAME

The baseline game was developed nearly ten years ago in the first-person shooter style with a visual look reminiscent of the 1978 *Space Invaders* arcade game. A screenshot of a representative game scene is shown in Fig. 1. For the duration of a game, the player moves forward at a constant perceived speed by animating the star field. The player first hears, and then sees, various in-game characters like the green one shown in Fig. 1. The characters are easily distinguishable from one another by shape, motion, aural, and color patterns. The player has to identify which alien characters are friends and which are enemies, following an introductory tutorial. These characters

enter the field of view of the player from different directions and at different pacing throughout the game. The player's tasks are to adjust the visual line of sight with keyboard arrow keys, in order to shoot the enemies and capture the friends.

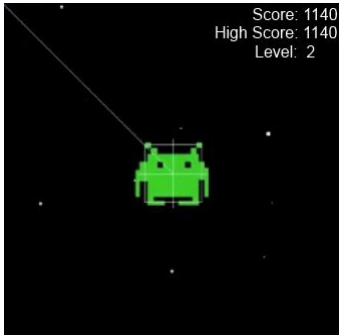


Fig. 1. Screen shot from original baseline game, showing the aiming crosshairs on an alien about to be shot through mouse and keyboard input.

The audio signaling in the game is very important. A player can hear a friend or foe before seeing them. The gap in time between when a character is heard and when it is seen on screen is a game variable that can be manipulated for both level progression (later levels can increase challenge by reducing the audio cueing time and increasing the approach speed of the characters), and for experimental design. The successful player learns sound categories to correctly predict character type and location before visual cues are used, and this implicit learning of audio cues can be tested outside of the game to see whether the learning transfers out of the game experience. The player's focus and goal is on the capture or shooting of characters, but along the way learning is happening implicitly regarding the aural communication; people also pick up language skills while embedded in richer contexts and performing other tasks. Comparisons with explicit learning, e.g., teaching native Japanese speakers the "l" and "r" sounds with traditional feedback-based training, show that the game approach with implicit learning of these categories can be done much more efficiently [3].

### III. UPDATING THE GAME: NEURALTONE

An ETC team, "neuraltone," worked with the Center for the Neural Basis of Cognition (CNBC) to update the game without losing the implicit learning of audio cues. A number of game ideas were discussed. Some, like "whack-a-mole" where sounds help predict where the moles appear, were dismissed because the learning might be too explicit without a rich enough play context. Other ideas showed promise from the research discussion, so quick playable prototypes were made to investigate three directions: a deep sea diver, a side scroller, and a radial shooter. Tens of college students ran through the prototypes and filled out surveys, revealing some disorientation with the diver prototype; concerns about the communication of controls and explicitness of sounds with the side scroller; and immediate usability and appeal of the radial shooter prototype. The decision was made to pursue the radial shooter mechanic, where the player views a cannon shooter that is always centered in the screen, clicks to set a general focus area and then moves the mouse to precisely target within that focus area. This choice left a great deal of flexibility in terms of game

theme, with suggestions ranging from fairylands to deep sea to forests to caves. The choice was made to remain with a space theme because (1) it is relatively easy to manage from the art side, allowing researchers to change character assets later if interested in varying the visual cueing; (2) space can be made more appealing with lighting and shader effects, rather than handpainted backdrops; (3) the theme affords the presence of strange audio cueing by arguing the sounds have alien origin; and (4) the CNBC research team has been working with a space theme for years.

Fig. 2 shows the radial shooting cannon as a spherical flying saucer. The saucer is always centered in the player's full view, "flying" in a slow, tight circle that animates the center "upper" portion to give an illusion of depth and communicate the saucer ship more completely. Here, the player has clicked to the right of the cannon to set the cannon turret in that direction. The turret can move side to side with 120 degrees of freedom via mouse movement in order to target approaching characters (the three dots within the radial shooter mark the limits of turret repositioning). The red sphere glowing around the cannon is the shield, which decays as ships hit the cannon to the point where the saucer explodes and player loses scoring ability. The shield immediately recovers to support the continued fighting of alien ships, while implicitly learning sound categories.



Fig. 2. Partial screen shot from neuraltone game: the turret general direction is set with a click and then precise movement of turret is tied to mouse position showed by the cross-hairs cursor, currently over a purple alien ship.

Fig. 3 shows a full game shot, hinting at the shading work that went into the rendering of stars, planets, galaxies, and other backdrops that drift slowly by to give the illusion of motion through space. These effects move the neuraltone game into a much more appealing space than its predecessor (Fig. 1). College-age playtesters commented favorably on the visual backdrops and the theme was found appealing to both genders. The aural backdrop, an original musical score reinforcing the space theme, was also appreciated and noted by many playtesters as making the setting more intriguing.

A simple parallax shader was used for the star field to create a twinkle effect when scrolled slowly. Earth-like planets use rim lighting, bump maps, and scrolling clouds. Glossiness changes depending on whether the surface is land, sea, or cloudy. Flow maps are used in order to create varied effects of lightning inside the plasma projectile. Flow maps are also used on more gaseous planets to give a more violent fluid effect on the surface of the planet. Finally, the shield is created by

mapping a screen grab to a sphere and applying rim lighting to it. Additive and emission lighting is used extensively to give many objects in space a glow-like feel.

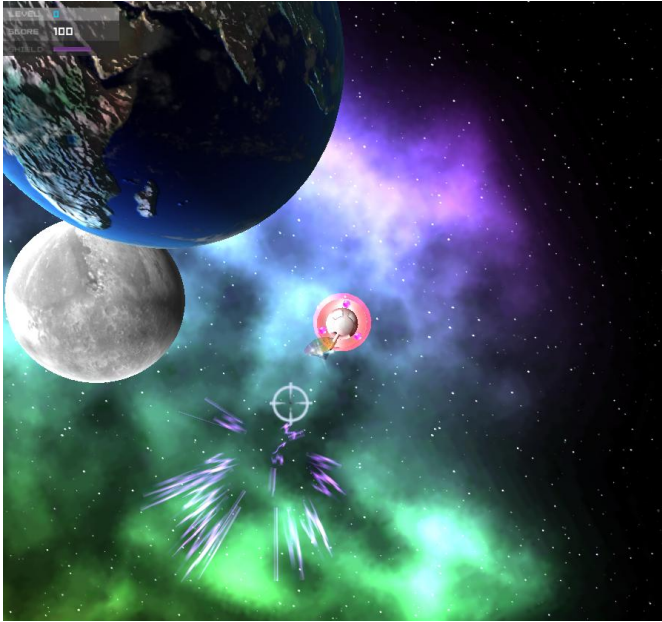


Fig. 3. Neuraltone game with particle effect remnants of a just-shot alien shift to the lower left of the cannon flying saucer.

A number of lenses are offered by Jesse Schell to guide and inspire the game designer, including the Lens of Pleasure, the Lens of Juiciness, and the Lens of Flow [5]. Neuraltone uses visual pleasure to reward the player to progress through the game; the backdrops as seen in Fig. 3 shift to other settings. "Juicy" systems reward the player many ways at once [5], and a musical score, sound effects on player actions, and accompanying visual effects on player actions all contribute to the juiciness. Fig. 3 shows the end of a particle effect animating the explosion of a destroyed ship; the aiming, shooting, and response all have aural and visual effects as well. These were missing elements in the baseline game of Fig. 1. Ideas for increased juiciness come from the CMU faculty (ETC and CNBC) during game reviews, guest game designers visiting the ETC to deliver seminars, and the playtesters themselves, who offered suggestions such as exploding particle effects in different colors for different ships.

These game design lenses in particular (pleasure, juiciness, flow) link very closely with cognitive neuroscience ideas about environments that effectively drive brain reward systems. As noted in [3], video games may be highly effective at activating the striatal reward system of the brain, providing implicit learning signals. Juicy, optimal flow tasks in which participants have incentive to perform well are most likely to robustly activate striatal reward system processing [6]. The neuraltone game was developed precisely so that such relationships between game design and cognitive neuroscience could be further explored.

As another example of a juicy interface, Fig. 4 shows a rewarding interface for advancing from one level to another: the stars warp with sound effects in an ever-increasing

explosion of white that eventually fills the screen, then recedes to show the player in a new space setting for the next level.

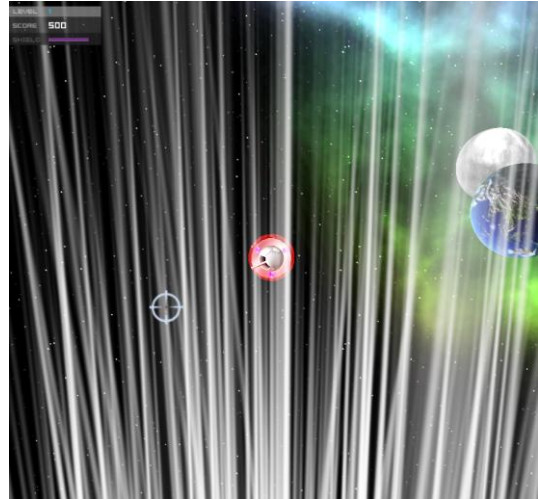


Fig. 4. Screen shot from within a few second warp effect sequence rewarding the player for completing a level and whiting out the screen in order to change space setting (e.g., to different planets, star fields, and nebulas).

As for flow, the player may be involved with the game for hours: does the game keep the player neither bored nor frustrated through a steady stream of gradually increasing challenges [5]? Initial playtests have set up a framework for how players will progress through the game, varying the pace of ship encounters and approaches and timing for aural and visual cues. The actual progression will very much be under the configuration control of researchers who may tailor the experience for particular participant pools, learning targets, or time on task objectives. For example, the advancement to a new level as shown in Fig. 4 occurs when the player score hits 500, but this could be set to a different score value or time value.

#### IV. PRELIMINARY INVESTIGATIONS AND CONFIGURABILITY

The goal of neuraltone was to update the player experience from 1970s arcade game styling to more current theming while not losing the ability to support implicit learning. A preliminary study was run with 13 college students to assess the latest game. Participants used the game for an hour, and then completed 176 sound classifications in a replication of an earlier experiment [4]. That earlier work found average accuracy values between 40% and 50% (where 25% is chance guessing of a sound from four in a multiple choice offering). In the trial with the neuraltone game, the average accuracy was 38.4%, not quite as good as the original game but much better than chance. This test was run with an early version of the cannon which required a click on the cannon center to reload it, a mechanic that players disliked and which has since been discarded in favor of an open flow of ships approaching. This test was also run with a default configuration of the game imposed by the game designers at the ETC. The plan is for subsequent tests to be run by the CNBC researchers in control of changing various game configuration parameters, to at first tweak the experience to meet or exceed the learning performances established in [4]. At that point, the researchers

are free to experiment with other sound families and questions regarding prompting, pacing, and rewards.

The configuration is accomplished through a separate GUI tool that allows non-programmer cognitive science researchers to specify various elements in a forms-based interface. The audio files and visual prompts associated with the alien ship inventory are specified.

The makeup of each level is specified in the form-based tool as well, i.e., types of ships on the level, minimum and maximum approach speeds, directions, length of sound cueing before the ship is visible on the screen, criteria necessary to pass the level. By adjusting the criteria, the experiment design moves from being hard-coded in the game to being in the hands of researchers such as the author team from CNBC, allowing them to specify when the rewards of level advancement (Fig. 4) are experienced and the next level with presumably more difficulty is introduced.

The configuration tool exports the various settings into an xml file. The neuraltone game serializes this xml file into its game classes, allowing various versions of games to be archived by saving the configuration xml file with the experiment. Along with updating the look and feel of the original game, this desire to have the game be modifiable by researchers to address different sounds, visuals, timing, and other questions is the driving motivation behind the neuraltone work.

As a concrete example, consider the dilemma of introducing a space-themed musical score into neuraltone. The presence of the score is recognized and appreciated by game players, and the absence of a score leads to a less enjoyable game based on playtesting. However, there may be research questions to explore in terms of whether the volume of that background score affects the implicit learning of the alien ship sounds. Does the makeup of that score, e.g., the genre perhaps qualified by the number of beats per minute, affect implicit learning of sounds? These questions can be studied by running the configuration tool, changing the files comprising the musical score and/or their playback volume, producing the xml file specifying that particular game setup, and then running participants through the game detailed in that particular way.

## V. FUTURE WORK

First, the game will be tested under various configurations to align it more closely to that of [4]. It may be that the visual richness and "juiciness" present cognitive overload challenges taking away from attention directed to implicitly learning different sound categories, but that is not the expectation. Based on earlier work [2][3][4], the research team anticipates that with the reworked reloading mechanic and other tweaks made to the game interface since the first study, the players will be more engaged and focused on ship approaches from various directions, and will learn sound categorization implicitly in order to score well and advance through the game. Correlation was found between game scoring/performance and category learning accuracy in [4]: the scoring algorithm may need to be tweaked in order to have such alignment in the neuraltone game.

The configuration capability will allow the game to be used in studies on foreign language learning, sound categorization, varying the audio prompting cues, and adjusting the visual composition and variance of the ships. The level progression can be tweaked as well, e.g., exploring a linear increase in difficulty versus an accelerated logarithmic one, focusing on the ideal "flow" [7] for retaining player engagement and/or achieving language learning. In conjunction with tracked game data, players can be observed with functional magnetic resonance imaging in order to view real-time brain reactions to the game [8].

The motivation for producing a game that "does you good" [1] and has learning effects in the domain of language learning stems from the challenge of running experiments with ecological validity in this domain. Experience plays an essential role in shaping auditory perception. Researchers cannot know the volume of experience shaping participants' listening and understanding skills when they enter a study on language learning [4], e.g., a participant might do well because of prior exposure to the language or familiarity with a given setting. However, a game can produce its own environment and rules to level the field: all players come in with no experience of the strange noises encountered [2]. For neuraltone, the strange noises belong to alien ships in space.

The setting is meant to draw the players in comfortably to keep working within the game, perhaps for long periods of time, with the collected data helping researchers better study the implicit learning of language. Fundamentals of good game design, including flow, pleasure, and juiciness, may be effective and appealing to players precisely because they tap into neural systems of reward: the same systems cognitive neuroscientists are trying to capitalize on for language learning.

The manner by which neuraltone is configurable can serve as a template for other brain training and learning games. These games can be tunable data collection instruments for their broader research teams.

## REFERENCES

- [1] D. Bavelier and R. Davidson, "Brain training: Games to do you good," *Nature*, vol. 494, pp. 425-426, Feb. 2013.
- [2] S.-J. Lim, F. Lacerda, and L. L. Holt, "Learning acoustically complex word-like units within a video-game training paradigm," *J. Acoustic Society of America*, vol. 129, no. 4, p. 2661, 2011.
- [3] S. Lim and L. L. Holt, "Learning foreign sounds in an alien world: Videogame training improves non-native speech categorization," *Cogn. Sci.*, vol. 35, no. 7, pp. 1390-1405, 2011.
- [4] T. Wade and L. L. Holt, "Incidental categorization of spectrally complex non-invariant auditory stimuli in a computer game task," *J. Acous. Soc. Am.*, vol. 118, no. 4, pp. 2618-2633, 2005.
- [5] J. Schell, *The Art of Game Design: A Book of Lenses*. Burlington, MA: Morgan Kaufmann, 2008.
- [6] M. R. Delgado, V. A. Stenger, and J. A. Fiez, "Motivation-dependent responses in the human caudate nucleus," *Cerebral Cortex*, vol. 14, no. 9, pp. 1022-1030, 2004.
- [7] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. New York, NY: Harper and Row, 1990.
- [8] S.-J. Lim, J. Fiez, M. E. Wheeler, and L. L. Holt, "Investigating the neural basis of video-game-based category learning," *Cognitive Neuroscience Society*, April 2013.