

# Speech Invaders & Yak-man: Retrogames for Speech Therapy

Chek Tien Tan\*  
Games Studio  
University of Technology, Sydney

Andrew Johnston, Andrew Bluff, Sam Ferguson  
Creativity & Cognition Studios  
University of Technology, Sydney

Kirrie J. Ballard  
Faculty of Health Sciences  
University of Sydney

## 1 Introduction

Speech therapy is used for the treatment of speech disorders and commonly involves a patient attending clinical sessions with a speech pathologist, as well as performing prescribed practice exercises at home [Ruggero et al. 2012]. Clinical sessions are very effective – the speech pathologist can carefully guide and monitor the patient’s speech exercises – but they are also costly and time-consuming. However, the more inexpensive and convenient home practice component is often not as effective, as it is hard to maintain sufficient motivation to perform the rigid repetitive exercises.

We present two mobile retro games, Speech Invaders and Yak-man (Figure 1), that serve to enable intrinsic motivation for performing these repetitive word vocalization exercises, as well as providing continuous visual feedback of speech through gameplay. To allow realtime speech analysis, we developed a formant-based speech analysis engine that focuses on recognizing vowels.

## 2 Game Design

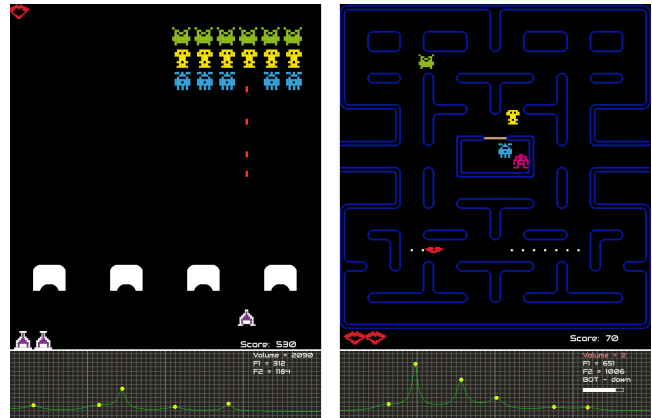
Speech Invaders and Yak-man are based on the popular classics Space Invaders and Pac-man. These games have withstood being remade and republished in various forms for nearly 40 years. Therefore, we have borrowed their core gameplay mechanics to provide intrinsic motivation for speech exercises. The game art was also re-designed to present a premise closer to speech therapy (e.g., by replacing pac-man with a mouth avatar). The games are currently freely available on the iOS platform.

## 3 Speech Engine

In both games, we added a novel input modality from the player, i.e., speech. For gameplay, high-performance continuous realtime recognition of player vocalizations is required. Current speech recognition systems mostly recognize complete words using machine learning methods. To the best of our knowledge, robust realtime performance remains a challenge today, and we did not find an algorithm that suits our realtime gameplay requirements. We also learned from our prior findings where children exhibited much frustration in the split-second lags experienced with a previous speech game we made using the Microsoft Speech SDK with the Kinect [Tan et al. 2013]. Hence we have developed our own engine that focuses only on recognizing vowels, and this method has been shown to exhibit superior realtime performance [Misra et al. 2005].

Our speech engine is centered around a Linear Predictive Coding (LPC) [Misra et al. 2005] stage that detects formant frequencies in speech utterances. Please see our main SIGGRAPH Asia MGIA

\*email: chek@gamesstudio.org



**Figure 1:** Screenshots of the Speech Invaders and Yak-man iPad games. The visualization of the activated speech formants at the bottom is designed to be viewed by the speech pathologist.

paper for the technical details [Tan et al. 2014]. By using a simple training step within the game, where the user sounds vowels into the microphone to set particular input controls for the game actions, the engine can accurately detect the specified vowels in English speech.

## 4 Future Work

Speech Invaders and Yak-man are mobile games designed to motivate patients to perform speech therapy exercises at home. We are currently running a preliminary survey-based study with speech pathologists to obtain feedback about the games. The results of this survey will be used to guide the next iteration of the game, which will be released for a large scale user study with the general public.

## Acknowledgments

This work was supported by a UTS Early Career Researcher Grant, as well as by the University of Sydney.

## References

- MISRA, A., WANG, G., AND COOK, P. R. 2005. SndTools: Realtime Audio DSP and 3D Visualization. In *Proc. 2005 International Computer Music Conference*.
- RUGGERO, L., MCCABE, P., BALLARD, K. J., AND MUNRO, N. 2012. Paediatric speech-language pathology service delivery: an exploratory survey of Australian parents. *International journal of speech-language pathology* 14, 4 (Aug.), 338–50.
- TAN, C. T., JOHNSTON, A., BALLARD, K., FERGUSON, S., AND PERERA-SCHULZ, D. 2013. sPeAK-MAN: towards popular gameplay for speech therapy. In *Proc. IE 2013*, ACM.
- TAN, C. T., JOHNSTON, A., BLUFF, A., FERGUSON, S., AND BALLARD, K. 2014. Retrogaming as Visual Feedback for Speech Therapy. In *Proc. SIGGRAPH Asia 2014 MGIA*, ACM.