Author/s:
Allison, F;Carter, M;GIBBS, M

Title:
A History of Voice Interaction in Games

Date:
2016

Citation:
Allison, F., Carter, M. & GIBBS, M. (2016). A History of Voice Interaction in Games. Digital Games Research Association.

Persistent Link:
https://hdl.handle.net/11343/282435

A History of Voice Interaction in Games

Fraser Allison, Marcus Carter and Martin Gibbs
Microsoft Research Centre for Social Natural User Interfaces
The University of Melbourne
fallison@student.unimelb.edu.au, marcusc@unimelb.edu.au,
martin.gibbs@unimelb.edu.au

Keywords
Voice interaction, speech recognition, platform studies, accessible games

INTRODUCTION
For more than forty years, voice interaction in gaming has been a mirage: a vision for how games of the future will play that has never quite arrived. Voice has been heralded as a "natural" mode of interaction (Aylett et al. 2014, 749), unbounded by the limitations of screen space and controller ergonomics. Yet in practice, it is more often described as "unnatural" and "uncomfortable" (Carter et al. 2015). Although the majority of contemporary game platforms have the technological capability to handle speech recognition, voice interaction remains a marginal and under-studied modality in games.

In this paper, we present a periodised history of voice interaction in digital games, drawing on a platform studies approach (Montfort and Bogost 2009, p. 145) to highlight the ways voice interaction has been both enabled and constrained by game platforms, and how this has contributed to the design patterns for voice interface in games. We identify six distinct, overlapping phases in the development of voice interaction for games.

The origins of voice interaction games lie in computer science research in the 1960s. Simple games were (and continue to be) convenient test cases for researchers working on speech recognition technology, as they could be directed with a limited and standardised vocabulary. The earliest voice interaction game of which we are aware is *Voice-Chess*, described by Raj Reddy and colleagues in a 1973 paper. It operated on the Hearsay-I speech understanding system, which could recognise and correctly respond to chess instructional phrases such as "Bishop to Queen Three" (Reddy et al. 1973).

By 1983, voice interaction games were beginning to reach the English-speaking market. Speech recognition add-on units were developed for the Texas Instruments TI-99/4A, Atari 2600 and Commodore 64 platforms. Perhaps the most ambitious project was the Halcyon game console by RDI Video Systems, which was to be controllable entirely through speech commands, able to recognise the player's specific voice, and able to add new words to its vocabulary. But in the wake of the North American video game crash of 1983, all of these projects failed; the Halcyon and Atari Voice Commander never reached mass production.

However, a more limited version of voice interaction found some success in Japan with the Nintendo Family Computer (Famicom) console, which launched in 1983. The Famicom came with two controllers, one of which included a built-in microphone. The microphone was not widely used by game developers, and was left out of the version of the console released internationally (as the Nintendo Entertainment System) (Altice 2015, 26).

Proceedings of 1st International Joint Conference of DiGRA and FDG

Nevertheless, voice input was a feature of several popular games on the system, notably *The Legend of Zelda* (Nintendo 1986a) and *Kid Icarus* (Nintendo 1986b).

In the two decades that followed, there was a divergence in style for voice interaction games between developers in Japan and their counterparts in North America. Japanese developers showed an earlier willingness to experiment with voice interaction, and their games typically focused on two-way, *conversational* voice interaction between the player and game characters. Voice interaction took longer to regain a foothold in North American game development, and when it arrived it was consistently used for voice *command*, as a one-way mode of sending instructions from the player to characters under their authority.

In parallel and somewhat apart from these developments, a genre of digital games was steadily emerging that gave voice interaction its broadest success to date: karaoke. The factors that set karaoke apart from other voice interaction games include the symbolism of the dedicated stage-style microphone, and the game's social role as a "glue technology that assists in crafting and strengthening social linkages" (Fletcher and Light 2011).

The introduction of the Microsoft Kinect had a substantial impact in bringing voice interaction technology closer to ubiquity in game interfaces. This has led to one of the largest waves of games to feature voice interaction in some capacity. This can likely be attributed to the devices' built-in support of voice interaction, removing the requirement for game developers to invest in their own speech recognition software.

Most recently, there has been a swell of voice interaction projects by independent game developers, in the form of both complete games and mods to add voice interaction to existing games. This is particularly notable as the first wave of voice interaction games to have arisen primarily on computers and mobile phones rather than consoles, enabled by a series of developments in hardware and software availability. Comparing these recent projects to games of the past, we find that independent developers have displayed a particularly focused engagement with the design challenges and opportunities that are inherent to voice interaction in a gaming context.

BIBLIOGRAPHY

Altice, N. 2015. *I Am Error: The Nintendo Family Computer / Entertainment System Platform*. MIT Press.

Aylett, M.P., Kristensson, P.O., Whittaker, S. and Vazquez-Alvarez, Y. 2014. "None of a CHInd: Relationship Counselling for HCI and Speech Technology." In *Proc. CHI'14 EA*, 749-760. ACM Press.

Montfort, N. and Bogost, I. 2009. *Racing the Beam*. MIT Press.

Carter, M., Allison, F., Downs, J. and Gibbs, M. 2015. "Player Identity Dissonance and Voice Interaction in Games." In *Proc. CHI Play '15*, 265-269. ACM Press.

Dragon's Lair Project. (n.d.) "Halcyon Interactive Laserdisc System." http://www.dragons-lair-project.com/community/related/homesystems/halcyon/ (accessed Jan. 2016).

Fletcher, G. and Light, B. 2011. "Interpreting Digital Gaming Practices: *SingStar* as a Technology of Work." In *Proc. ECIS 2011*: 154.

Nintendo. 1986a. *The Legend of Zelda* [Family Computer]. Nintendo, Kyoto, Japan.

Nintendo. 1986b. *Kid Icarus* [Family Computer]. Nintendo, Kyoto, Japan.

Reddy, D.R., Erman, L.D. and Neely, R.B. 1973. "A Model and a System for Machine Recognition of Speech." *IEEE Trans. on Audio & Electroacoustics*, 21(3): 229-238. IEEE.