

Chapter 11

Uncanny Speech

Angela Tinwell
University of Bolton, UK

Mark Grimshaw
University of Bolton, UK

Andrew Williams
University of Bolton, UK

ABSTRACT

With increasing sophistication of realism for human-like characters within computer games, this chapter investigates player perception of audio-visual speech for virtual characters in relation to the Uncanny Valley. Building on the findings from both empirical studies and a literature survey, a conceptual framework for the uncanny and speech is put forward which includes qualities of speech sound, lip-sync, human-likeness of voice, and facial expression. A cross-modal mismatch for the fidelity of speech with image can increase uncanniness and as much attention should be given to speech sound qualities as aesthetic visual qualities by game developers to control how uncanny a character is perceived to be.

INTRODUCTION

As technological advancements allow for the representation of high fidelity, realistic, human-like characters within computer games, aspects of a character's appearance and behaviour are being associated with the Uncanny Valley phenomenon. (A definition of the Uncanny Valley is provided in the first section of this chapter.) It seems that one of the main factors contributing to a character being regarded as *lifeless* as opposed to *lifelike* is the character's speech. In 2006, Quantic Dream

revealed a tech demo (*The Casting*) for the computer game *Heavy Rain* (2006), in which the main character, Mary Smith, evoked a somewhat negative response from the audience (Gouskos, 2006). Criticism was made of the uncanny nature of Mary Smith's speech in that it sounded strange and out of context with the given facial expression and emotion portrayed by this character. A closer inspection of the video showed that not only were there errors in the sound recording (disparities between the acoustics and the volume and materials of the room with excessive plosives contradicting the distant camera and microphone), but a lack of correct pitch and intonation for speech and a lack

DOI: 10.4018/978-1-61692-828-5.ch011

of synchronization of speech with lip movement were factors that reduced the overall believability for this character (Tinwell & Grimshaw, 2010). A mismatch between the conveyed emotion of Mary Smith's voice with her gestures and posture exacerbated how unnatural and odd the character was perceived to be. MacDorman (quoted in Gouskos, 2006), observed that a perceived asynchrony of lip movement with speech was one of the factors that people found disturbing about Mary Smith:

In addition, there is sometimes a lack of synchronization with her speech and lip movements, which is very disturbing to people. People 'hear' with their eyes as well as their ears. By this, I mean that if you play an identical sound while looking at a person's lips, the lip movements can cause you to hear the sound differently.

Since Mary Smith was revealed in 2006, increasing technological sophistication for computer games has allowed for heightened realism of human-like characters. Cinematic animation is achieved not only for cut scenes and trailers containing full motion video (FMV) but also for animation during in-game play. For example, the phoneme extractor and facial expression tool *Faceposer* designed by Valve for titles such as *Left 4 Dead* (2008) and *Half Life 2* (2008). However it would seem that speech, as a factor integral to the uncanny phenomenon, is often overlooked when compared to the aesthetic visual qualities of behaviour of a human-like character. So far there have been limited studies to ascertain which factors contribute to the uncanny for virtual characters. In response to the hearsay in mass media raised by characters such as Mary Smith, Tinwell and Grimshaw (2010) conducted a study to investigate how the cross-modality of image and sound might exaggerate the uncanny. The results from this study are referred to throughout all sections within this chapter as the Uncanny Modality (UM) study, unless otherwise stated from another study. Prior to this, much of the work on the uncanny had been

visually-based, excluding sound as a factor. As a way towards building a conceptual framework for the uncanny and virtual characters in immersive 3D environments, this chapter defines how characteristics for a character's speech may exaggerate the uncanny by considering aspects such as synchronization of audio and video streams, articulation, and qualities of speech.

The first section provides an exposition of the Uncanny Valley describing how the theory came about, previous investigation into the theory and potential limitations of the theory in relation to virtual characters.

Previous authors (such as Bailenson et al., 2005; Brenton, Gillies, Ballin, & Chatting, 2005; and Vinayagamoorthy, Steed, & Slater, 2005) have suggested that uncanniness is increased when the behavioural fidelity for a realistic, human-like character does not match up with that character's realistic, human-like appearance. The second section discusses how a cross-modal mismatch between a character's appearance and speech may exaggerate the uncanny. For instance whether a character's speech may be perceived as belonging to a character or not, based on that character's appearance.

The third section discusses how particular qualities of speech such as slowness of speech, intonation and pitch and how monotone the voice sounds, may influence perceived uncanniness and how such qualities might work to the advantage of those characters intended to elicit an eerie sensation.

The results from the UM study (Tinwell & Grimshaw, 2010) revealed a strong relationship between how strange a character is perceived to be and the lack of synchronization of speech and lip movement. (Characters rated as close to perfect synchronization for lip movement and speech were perceived as less strange than those with disparities in synchronization.) The fourth section reviews the findings from this study and also puts forward future experiments that may

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/uncanny-speech/46794

Related Content

Linking the Elements of Learning, Assessment, and Play Experience in a Validation Framework

Chin Ike Tan (2019). *Design, Motivation, and Frameworks in Game-Based Learning* (pp. 93-122).

www.irma-international.org/chapter/linking-the-elements-of-learning-assessment-and-play-experience-in-a-validation-framework/208023

Games for and by Teachers and Learners

Peter van Rosmalen, Amanda Wilson and Hans G.K. Hummel (2015). *Gamification: Concepts, Methodologies, Tools, and Applications* (pp. 1706-1733).

www.irma-international.org/chapter/games-for-and-by-teachers-and-learners/126139

Paratext: The In-Between of Structure and Play

Daniel J. Dunne (2016). *Contemporary Research on Intertextuality in Video Games* (pp. 274-296).

www.irma-international.org/chapter/paratext/157039

Playing with Programming: A Serious Game to Start Programming

Anabela de Jesus Gomes, Álvaro Nuno Ferreira Silva Santos, César Paulo das Dores Páris and Nuno Cid Martins (2017). *Gamification-Based E-Learning Strategies for Computer Programming Education* (pp. 261-277).

www.irma-international.org/chapter/playing-with-programming/163711

The Effects of Fully and Partially In-Game Guidance on Players' Declarative and Procedural Knowledge With a Disaster Preparedness Serious Game

Ting Zhou and Christian S. Loh (2020). *International Journal of Gaming and Computer-Mediated Simulations* (pp. 23-37).

www.irma-international.org/article/the-effects-of-fully-and-partially-in-game-guidance-on-players-declarative-and-procedural-knowledge-with-a-disaster-preparedness-serious-game/268881