

A Functional Model for Dialogic Speech in Video Game Design

J. Martin

Mediadesign University of Applied Sciences, Duesseldorf, Germany

j.martin@mediadesign.de

Abstract

This paper introduces a functional model for dialogic speech in video games. Loosely based on linguistic frameworks for real-life speech, it discusses how insights from this field can inform a more simplified model for dramatic speech. This simplified model aids two kinds of design decisions. Through a general set of speech functions (mood, feedback, information), it aids high-level decisions whether to use dialogic speech in a game or not. Through a more specialized set (inform, inquire, influence, cultivate, entertain), it aids low-level design decisions for individual beats. Moreover, this paper argues that these functions must address the player and player characters separately, and that relationships between characters are necessary for dialogic speech to work.

Keywords: *game design, speech, dialogue*

1. INTRODUCTION

Speech, like every other element in video game design, should conform to the principles of functional aesthetics as a mode of experience that creates meaning and knowledge [1]. In video game design, functional aesthetics should be manifest in terms of skill, style, and subject matter—to be and appear professional; contribute to original, recognizable patterns; and work together toward an integrated whole. With regard to a functional model of speech for video game design, the application of speech must account for skill, style, and subject matter in two different ways: its aesthetic function as a design decision for the game as a whole on the one hand, and each speech element's aesthetic function on the level of individual beats as the smallest elements of structure [2] on the other.

In this paper, a functional model of speech will be explored for speech that is both dialogic [3] and diegetic, the latter in the sense of an element that appears both for the characters in the fiction and for the audience [4], from non-player character dialogue lines to enacted dialogue-wheel decisions to diegetic speech in virtual reality games [5]. Not considered in this paper are all forms of monologic and reflective speech, e.g., soliloquies; non-diegetic speech, e.g., voice-over narration; or the use of fictional languages, e.g., Simlish [6].

2. LINGUISTIC FRAMEWORK

An important concept from the perspective of linguistic theory is the function of speech as a tool to manipulate one's environment with and through other humans, be it as speech act [7], speech event [8], power discourse [9][10], or, from the broader perspective of language, as a general tool for purposeful action (*Handlungszweck*) [11]. Particularly in the latter's view, language and speech are sometimes compared to opposable thumbs as an essential factor that enables humans not only to adapt to their environments but to adapt these environments to their needs, both individually and over evolutionary times. Against this backdrop, this paper argues that both speech as such and every individual speech event should represent a purposeful action, both within the game world from the perspective of the characters and outside the game world from the perspective of the player.

3. FUNCTIONS OF SPEECH

The first design decision with regard to speech is the decision to use speech in the game or not. Certainly, every game needs to communicate something to the player. But to communicate, speech is not essential—animals, e.g., can communicate just fine without being able to use language and speech in the very specific and qualitatively different ways human animals can [12]. Then, there is written text, which is an excellent alternative to speech when it comes to communication. Written text can *imitate* speech in many different ways as well. Japanese games in particular, like *Ōkami* [13] and—with the exception of cutscenes—even 2017's *The Legend of Zelda: Breath of the Wild* [14] deliver narrative and non-player character dialogue as written text that is accompanied by emotive sounds created by Foley artists or scrambled from voice actor samples.

Thus, the design decision to use speech in a video game is not as obvious as it appears. The same caveat applies to design decisions with regard to speech for individual beats. In both *Uncharted 4: A Thief's End* [15] and *Uncharted: The Lost Legacy* [16], e.g., the disposable enemy-mercenaries have the same limited set of speech events ("Got anything?" "Nothing." "Check Over There. I'll go this way." "Sure thing, mate."), repeated over and over. These particular speech events fulfill no discernible function, to be discussed in detail below. But, more importantly in this context, the decision to use speech was not necessary at all. Anything worthwhile communicating while securing a perimeter, from observations to orders, could and should have been communicated by these professionals through tactical signals instead of speech.

3.1 General functions

The design decisions to use speech in a game or not can be aided by a set of general functions that speech has in common with music and Foley. This general set of functions consists of *mood*, *feedback*, and *information*. Mood affects and, to a certain extent, controls the player's emotions. Feedback establishes the link between player ac-

tions and the results of these actions. Information includes meaning and cues for interpreting any given element in the game world.

As mentioned, this set is not particular to speech. Music and Foley can set the mood as well, from a lonely blues-harp to the sound of raindrops hitting the roof. Music and Foley can also link actions to their results, from the change to a major key after a successful attack to the sound of jingling change when picking up a gold coin. Finally, both music and Foley are excellent in conveying information too, from leitmotifs to characteristic sound markers that announce and differentiate approaching enemies.

That way, in the context of these general functions, speech might not be necessary because mood, feedback, and a considerable palette of information can be conveyed through music or Foley instead.

3.2 Specialized functions

To aid design decisions with regard to the use of speech in an individual beat, a more granular set of specialized speech functions is needed. Numerous models have been proposed to categorize all possible occurrences of speech into a defined number of functions, prominently among them Jakobson (referential, emotive, conative, phatic, metalingual, and poetic, each with their various subsets) [17], Leech (conceptual, connotative, social, affective, reflected, collocative, and thematic, with their various group commonalities and dependencies) [18], or Halliday (ideational, interpersonal, and textual from his systemic linguistic theory's metafunctional dimension) [19][20]. As these and similar models attempt to integrate high-level matters like semantics, poetic language, or language acquisition, they are, by and large, not well-suited to serve as templates or guidelines to designing speech events in video games. Thus, a systematically cut-down and condensed set of speech functions will be proposed, for the expressed purpose of facilitating design decisions for dialogic speech events in video games.

This set includes not more than five simple functions: *inform*, *inquire*, *influence*, *cultivate*, and *entertain*. The first two are complementary. Speech that informs expresses external or internal states, including intent. It comprises a wide range of informational activities from teaching and educating to warning or disagreeing. Speech that inquires does the opposite: it asks to be informed, be educated, be taught, and similar. Speech that influences can request, persuade, order, command, convince, scare, or inspire. Speech that cultivates can socialize, chat, assure, ascertain, introduce, and similar. Speech that entertains comprises a wide range of forms of aesthetic and poetic speech.

Importantly, these five functions define the *primary* function of a speech event, contingent on context and intent. In most cases, elements or connotations from one or more different functions will be attached, whether on purpose or not. For example, a warning can inform ("it's dangerous"), influence ("don't go there"), or cultivate ("my advice, dude"). In dramatic speech in works of art, be that tragedies, novels, movies, or video games, the primary function should always be clear, by design. (Or, alternatively, purposefully designed toward ambiguity.) Speech events in real-life are less de-

signed, more messy, and more sensitive to interference in principle; but the primary function can nevertheless be retrieved in most cases.

Applying this functional set to the above-mentioned example from *Uncharted 4: A Thief's End* [15] and *Uncharted: The Lost Legacy* [16], the mercenary dialogues in these games seem to express no primary function whatsoever. They do not inform—nothing that is exchanged exceeds anyone's previous knowledge and the dialogues give the player no tactical cues. They do not inquire—any fresh information in these situations would have been obvious and not in need of being communicated through speech. They do not influence—"this way" or "that way" is utterly random and devoid of directions in the game world. They do not cultivate relationships and, finally, they also do not entertain anyone. This can be contrasted with the game *Oxenfree* [21]. In this game, speech events inform, inquire, influence, cultivate, and entertain in exceptionally natural ways, supported by time limit and interruption mechanics.

3.3 Character-specific vs. player-specific functions

When a speech event does serve one or more functions, it needs to be designed for two distinctly different addressees. One addressee is a character or several characters within the game world, which can be, or include, the player character. The other addressee is the player who resides outside the game world.

This is necessary for dramatic reasons. If the function of a speech event addresses only the player outside the game world, these dialogues will sound as unnatural as bad movie exposition dialogue, or like the unnatural exchanges between talkative guards in action-adventures. If the function of a speech event addresses only the characters within the game world, these dialogues serve no dramatic function with respect to the player (i.e., advance the plot, portray a character, communicate an insight into the game world, advance player proficiency, and similar).

The primary functions of a speech event for the characters and for the player do not have to be the same. Indeed, speech events often work better and appear more natural if they serve different functions for characters and players. Speech event with the primary function of entertaining or socializing within the game world can have the primary function of informing or influencing the player, for example. There are many possible permutations to employ the five functions inform, inquire, influence, cultivate, or entertain for dramatic purposes in interesting ways.

3.4 Speech functions and relationships

Besides the dramatic requirements already discussed, a dialogue is also a conversational exchange, a mutual exploration of observations, ideas, and intents. Linguistic, philosophical, and political theories on dialogue often stress the fact that dialogues build and transform relationships, and that dialogues are critical for social cohesion [22]. Less focus is placed on the fact that dialogue, especially dramatic dialogue, is *built* around relationships which must already exist, however intimate or temporary. Where such relationships do not exist, the dialogues run the risk of sounding like a se-

ries of monologues. Also, speech that merely lectures or informs is not dialogic, and even trading information does not necessarily constitute a dialogue. Dialogues sound like dialogues when the characters, including the player character, have shared interests or opposing interests, and it works even better when these interests are charged with emotional values, positively or negatively. Without such a relationship, i.e., without an *interesting* relationship, dialogues often fall flat even when they are well-written.

4. CONCLUSION

Regardless whether a given speech event in a video game is highly artificial, perfectly natural, or anything in between, it should at least cover one general dramatic function, i.e., mood, feedback, or information, and ideally at least one specialized function, i.e., inform, inquire, influence, cultivate, or entertain. For dramatic purposes, moreover, every speech event must have a character-specific function and a player-specific function, drawn from the first or the second set. These functions need not be the same. Finally, speech events need an established relationship between characters with common or conflicting interests to work as dialogues.

Besides aiding design decisions, this model can be used to prevent the overuse of the general information function and specifically the inquire function as a typical pitfall for dialogue design in video games. Clicking systematically through predictable questions on a dialogue wheel does not constitute dialogue. Combining the inquire function with at least one other speech function will make dialogues both more natural and palatable. A tit-for-tat with the inform function is a good start, but less beaten paths like combining inquiry and inform with cultivate or entertain should be more interesting and captivating.

Design decisions regarding speech, finally, should not be made as stand-alone decisions. They should seamlessly integrate with design decisions regarding music and Foley, and with *their* respective general and specialized functions, for a holistic auditory experience—from converging or diverging intensity curves over the course of the game including effects of temporary silence to the soundscape of the game as a whole that is original, recognizable, and exciting.

REFERENCES

- [1] Knight, Aimee. "Reclaiming Experience: Aesthetics & Multimodal Composition." *Computers and Composition* 30 (2013). 146–155.
- [2] McKee, Robert. *Story: Substance, Structure, Style, and the Principles of Screenwriting*. New York: ReganBooks, 1997. 37.
- [3] Yakubinsky, Lev Petrovich. "On Dialogic Speech." Intr. & transl. Michael Eskin. *PMLA* Vol. 112, No.2 (March 1997). 243–56.
- [4] Ignacio, Dino. "Crafting Destruction: The Evolution of the Dead Space User Interface." GDC 2013. <https://www.gdcvault.com/play/1017723/Crafting-Destruction-The-Evolution-of>

- [5] Çamcı, Anıl. "Exploring the Effects of Diegetic and Non-diegetic Audiovisual Cues on Decision-making in Virtual Reality." In: *SMC 2019. Proceedings of the 16th Sound and Music Computing Conference*, May 28–31, 2019. Ed. Isabel Barbancho, Lorenzo J. Tardón, Alberto Peinado, & Ana M. Barbancho. Universidad de Málaga: ATIC Research Group, 2019. 195–201.
- [6] Boland, Eric. *The Sims: A Complete Guide*. Vancouver: WTYW7, 2010. 17–21.
- [7] Austin, J. L. *How to Do Things With Words. The William James Lectures delivered at Harvard University in 1955*. Oxford: Clarendon, 1962.
- [8] Jakobson, Roman. "Linguistics and Poetics." In: *Style in Language*. Ed. Thomas A. Sebeok. Cambridge: MIT P, 1960.
- [9] Foucault, Michel. *The Archaeology of Knowledge*. Transl. A. M. Sheridan Smith. New York: Routledge, 1989.
- [10] Foucault, Michel. "The Order of Discourse." Transl. Ian McLeod. In: *Untying the Text: A Post-Structuralist Reader*. Ed. & intr. Robert Young. New York: Routledge, 1981. 51–78.
- [11] Keller, Rudi. *Sprachwandel*. Tübingen: Franke, 1990.
- [12] Snowdon, Charles T. "Language Capacities of Nonhuman Animals." *American Journal of Physical Anthropology* Vol.33 S11 (1990). 215–43.
- [13] *Ōkami*. Clover Studio. Capcom, 2006.
- [14] *The Legend of Zelda: Breath of the Wild*. Nintendo EPD. Nintendo. 2017.
- [15] *Uncharted 4: A Thief's End*. Naughty Dog. Sony CE. 2016.
- [16] *Uncharted: The Lost Legacy*. Naughty Dog. Sony CE. 2017.
- [17] Jakobson, Roman. "The Speech Event and the Functions of Language." Conference lecture, 1959. In: *Roman Jakobson. On Language*. Ed. Linda R. Waugh & Monique Monville-Burston. Harvard: Harvard U, 1995. 69–79.
- [18] Leech, Geoffrey. *Semantics: The Study of Meaning*. Rev. 2nd Ed. Hammondsworth: Penguin, 1981.
- [19] Halliday, M.A.K. "Language Structure and Language Function." 173–95. In: *On Grammar. Vol.1 in the Collected Works of M. A. K. Halliday*. Ed. Jonathan Webster. New York: Continuum, 2002.
- [20] Halliday, M.A.K. *Explorations in the Functions of Language*. London: Edward Arnold, 1973.
- [21] *Oxenfree*. Night School Studio. 2016.
- [22] Allwood, Jens. "On Dialogue Cohesion." *Gothenburg Papers In Theoretical Linguistics* 65. Gothenburg U: Dept. of Linguistics, 1992.