

Mode-switching in Digital Game Environments: A Multimodal Phenomenon

Abstract:

This paper investigates the phenomenon of *mode-switching*, or a quasi-synchronous shift of linguistic mode by a speaker in interaction. Video recordings of three *World of Warcraft* players are analyzed to determine the motivations for mode-switching. Ultimately, the driving force for mode-switches is a shifting participation framework; speakers use a mode which corresponds to the intended audience for their talk. Each mode has certain of topics which are appropriate for the audience present in that mode, and speakers shift between the modes seamlessly in interaction to address each of these sets of participants. The norms for the uses of modes must be learned by speakers and deployed properly in order for a speaker to be considered a competent member of the community. Mode-switching is compared to the processes of code-switching and style-shifting and is found to have numerous similarities, suggesting that the three phenomena are variations of a larger process of variety-changing.

Lauren B. Collister
Department of Linguistics
University of Pittsburgh
Second Comprehensive Paper

Advisor:
Scott F. Kiesling
Committee:
Shelome Gooden
Sara Kajder

Defended 10/19/2011
Revised and submitted: 11/22/2011

Mode-switching in Digital Game Environments: A Multimodal Phenomenon

In this paper, I examine the phenomenon of *mode-switching* in digital game environments. Before I can address the phenomenon itself, or even the concept of *mode* in language, I must first introduce the environment. Specifically, I will be discussing examples from my ethnography of the massively multiplayer online roleplaying game (MMORPG) *World of Warcraft* (Blizzard Entertainment, 2004). *World of Warcraft* (also called *WoW*) is a fantasy-style game set in a digital world, and has more than ten million active players. In this work, I will be focusing on the linguistic behavior of experienced, or “hardcore”, players of the game, who generally spend more than 20 hours per week playing the game.

I first began to observe the hardcore player community in the context of what is called a *raid*. Raids are events in which large groups of players band together to attempt the most difficult game content available, usually involving venturing into complex areas called *dungeons* to defeat particularly difficult computer-controlled enemies called *bosses*. The purpose of this is threefold. First, players raid for accomplishment, and bragging rights about who is more “progressed” in the game content. Second, players receive rare and powerful items when they defeat these bosses. These items, called *loot*, increase player performance by boosting statistics like strength or spellpower. Third, raiding is a very social event, in which a player can digitally “hang out” with up to twenty-four other players, while engaging in the same in-game activity.

My observations of raids in the early days coincided with my own personal adventure of learning to be a raider myself. The most difficult thing to learn for me, as a player, was not making my character perform in the necessary ways; rather, I had the greatest difficulty managing the intense social aspect of hardcore gaming. With more experience, I began to realize why I had this difficulty – hardcore gaming is intensely multimodal, and players are required to

manage a large amount of linguistic information all at one time.

To illustrate, I will describe typical participant setup for a typical raid night with the guild I observed in my ethnography. (In fact, this setup was in place for the data I analyze later in this paper.) I am sitting in my apartment at my desktop computer, with the game open on my screen. I control my character, Parnopaeus, using my mouse to move her around in the digital world and the keyboard to activate spells and abilities. Parnopaeus is in a raid group of twenty-five players and we are about to enter Naxxramas, a dungeon full of bosses. The raid group is chatting in a bright orange font in the chat box in lower left of my screen (see Figure 1). At the same time, members of the guild who are not currently raiding are also chatting in the same chat box, but in a bright green font. In this same chat box, I receive a whisper in bright pink font, a direct message from my friend Jahaerys, asking if I have any extra potions that he can borrow for the raid tonight. I have three textual chat channels open already.



Figure 1: A screenshot of the *World of Warcraft* interface. The chat box is in the lower left corner. Character and group information and abilities appear around the edges of the screen, and the digital world is depicted in the center.

Then, in the bright orange font that indicates the raid channel, our leader, Theon, tells us that “everybody should get on Vent”. He means Ventrilo, the voice chat program that hardcore players use during raids to facilitate communication. I put on my headset, a speaker in one ear and a microphone near my left cheek, and load Ventrilo and connect to the guild’s server. I join the RAIDING channel and now I can hear the voices of all twenty-five people in the raid group. Not everybody talks at once; some people do not want to have their voices heard, do not have a working microphone, or simply do not want to try to talk over everybody else, and use the bright orange [Raid] chat channel to talk.

With all of this, I am situated inside of an intensely social event. I must pay attention to the verbal exchanges in my one ear, as my raid leader Theon is giving instructions for the defeat of the bosses. My in-game friends might also be cracking jokes or asking me questions over Ventrilo that I should respond to. At the same time, I must pay attention to the text chat channels in the box on my screen, since communication is happening there as well. I must differentiate between the chat channels of [Raid] and [Guild] and private messages, among a host of other available channels. And, on top of this intense social situation, I must also pay attention to my character, and activate her abilities at the right times while engaging in battle with the rest of my raid group.

Hardcore gaming, and raiding in general, is an intense experience. In my first raids, I had difficulty managing all of the modes of communication. I frequently ignored private messages because they simply became lost in the wealth of textual chat. I would be slow to talk on Ventrilo because I would be too busy controlling my character. Eventually, with practice, I learned how to manage all of these modes as most experienced gamers do, but my fascination with this intense social situation remained. There must be rules, because I (as many other players before and after

me) learned how to use all of the modes to navigate the social landscape in the game. What are these rules? What strategies do players deploy when interacting in such modally dense environments? It is this that I decided to investigate in this paper, and in doing so, I observed the phenomenon of mode-switching.

Mode-switching: What is it?

Communication in general is inherently *multimodal*, meaning that language is just one part of the process of communication. A *mode* is a system of representation of meaning, with regularities and norms (Norris, 2004: 11). Modes do not have to be linguistic to carry meaning, and linguistic and nonlinguistic modes often operate in tandem to produce a meaningful utterance. For example, speakers use gesture in combination with the spoken word to add meaning to the utterance; academic presentations incorporate a range of communication sources including spoken language and gesture, but adding written words on a handout and visual images and words projected on a screen. According to scholars of multimodality like Jewitt (2009), the “representation and communication always draw on a multiplicity of modes, all of which have the potential to contribute equally to meaning” (14).

The mode-switching phenomenon is the quasi-synchronous use of two or more linguistic modes. This definition includes one speaker who switches between spoken and written modes in a single interaction. In Kress’s (2010:159) words, mode-switching is akin to “orchestrating a multimodal ensemble”, or taking advantage of the available modes. Mode-switching usually occurs in highly modally dense situations, which is, according to Norris (2004: 80), a situation in which many modes interplay and speakers are giving attention to all of the modes. That is, interactions can be more or less modally dense – if a speaker is paying attention to a large number of modes, the interaction is said to be more modally dense than if a speaker is only

paying attention to one or two modes.

In the gaming situation described in the introduction, the situation becomes progressively more modally dense. Not only is the player paying attention to the game, with its rich visual landscape and sound experience, but the player is also paying attention to the chat box in the corner of the screen. In the beginning, the situation is modally dense, but not necessarily linguistic modally dense. As the player engages in more communication besides the one bright orange chat channel (first receiving a private message in text chat, then putting on the headset and adding verbal communication with other players through Ventrilo), the linguistic landscape becomes more complex and more modally dense. It is in this kind of environment that mode-switching occurs.

Mode-switching had been addressed by other researchers. Baynham (1993) studied a mode-switching phenomenon with multilingual speakers code-switching when moving between conversation and discussion of written texts. The texts in Baynham's study were frequently written in many languages, and the language of the text affected which language the speakers used – that is, they switched languages when referring to the text or conversing amongst themselves. In her examples, the speakers were simultaneously negotiating different languages as well as different modes of communication. Sindoni (2011) described mode-switching in digital environments, specifically looking at the reasons for alternating use of voice and text chat by users of Skype. She found interlocutors who were conversing via Skype's video chat option would utilize the text chat feature (which acts somewhat like an instant message) to convey certain types of information, usually sensitive information that they did not want to be overheard.

For my study, I analyze transcripts of videorecordings of players of the massively multiplayer online roleplaying game (MMORPG) *World of Warcraft*. *World of Warcraft* is set in

a quasi-medieval digital world, and features many activities for players such as completing quests, slaying creatures, exploring the world, reading and participating in in-game stories, collecting rare items, and competing against other players in battlegrounds. More than these activities, *World of Warcraft* is an intensely social game, requiring interaction with other players to participate in most of the activities. Because of this, *World of Warcraft* players have developed their own community with its own linguistic rules and behaviors; furthermore, the interaction in the game environment is inherently multimodal. Players not only have several channels of textual chat to participate in, but voice chat is a popular addition to the gaming experience. In this particular study, I look at players who are physically co-present – meaning that they are playing together in the same room – but who are also interacting with players who are digitally present via voice chat and textual chat. This makes their experience even more modally dense and complex than the gaming experience I described in the introduction, because they add another layer of interaction. In this environment, the players have the maximum number of possible mediums regularly employed in gaming environments to use language: face-to-face, voice chat mediated by the computer, and textual chat.

When thinking about modalities in environments such as online gaming, a useful idea is that of *presence*. Jones (2004), in describing the phenomenon of “presence” online, relates the idea of *social presence* to Goffman’s (1964) *social situation*, stating that it is “an environment of *mutual monitoring possibilities*, anywhere within which an individual will find himself accessible to the naked senses of all others who are present, and similarly find them accessible to him”. Jones suggests that social presence online affords users “different sets of mutual monitoring possibilities”, which in turn have “different ways in which [technologies] allow us to be *present* to one another and to be aware of other peoples’ *presence*” (p. 23). One way that

technologies affect presence is discussed by Norris (2004: 22), who uses the notion of *anwensenheit* to capture the idea that people being co-present in a mode may not have their interlocutors as their main focus – the focus may be on a shared activity, for example. This idea of *presence*, specifically *anwensenheit*, is very important to a modally dense social space like *World of Warcraft* because players have many options for where they want to be socially present in the game world, and the management of these different modes is a method for demonstrating communicative competence (Hymes, 1986) in this particular community. Communication in this community is polyfocal, which is not exactly the same as multitasking. Polyfocality, according to Scollon et al. (1999), is when a participant’s focus skips along multiple “attentional tracks”, which sometimes intertwine and sometimes do not. That is, the components of an interaction may come from many different things at once – talk, action onscreen, gesture, and textual chat – but each of these modes may not constitute an entire different thread at once. Rather, they may combine into a multimodal ensemble comprising a multimodal interaction.

To summarize, the gaming environment is a modally dense environment, and can be very linguistically modally dense. To be considered a competent gamer, one must be able to pay an appropriate amount of attention to the many different modes, including all of the linguistic modes, and still succeed at the gaming tasks which are at the forefront. Players must attend to the task of playing the game, but at the same time navigate the complex linguistic landscape. One of the ways that players manage this polyfocality is through mode-switching, or changing linguistic mode based on the context and the knowledge of the other people who are present in that particular mode.

Mode-switching as Variety-changing

One comparison that has arisen repeatedly in the mode-switching literature is between

mode-switching and code-switching. Saville-Troike (1989:58) defines a *code* as “different language, or quite different varieties of the same language”, and *code-switching* as “change in languages within a single speech event”. Code-shifting is often discussed alongside *style-shifting*, which Saville-Troike defines as “change in language varieties which involves changing only the *code-markers*; these are variable features which are associated with such social and cultural dimensions as age, sex, social class, and relationship between speakers”. The difference between code-switching and style-shifting can be thought of in terms of how different the varieties are and how much speakers attend to the difference in the varieties. If the varieties being used are classified as two distinctly different languages or dialects (such as classical versus colloquial Arabic), then it is code-switching. If the varieties are thought of as falling within the same language, then it is style-shifting. Saville-Troike suggests that *variety-changing* would be a useful term to apply to the general process of shifting varieties of language, but notes that the term has not been suggested by others and did not wish to further complicate the picture. I believe that mode-switching can also be placed under the umbrella of variety-changing, because of its similarity to both code-switching and style-shifting.

Style-shifting is not only similar to mode-switching, but the phenomenon of style-shifting happens within linguistic modes themselves, as a body of work on computer-mediated communication has shown (e.g. Baron, 2009, on “netiquette”; Baym, 2010, on immediacy cues). People do not write e-mail to their professors or employers the same way that they write e-mail to their friends, even if it’s “all e-mail”. Styles exist within linguistic modes, just as they exist within linguistic codes. Code-switching also exists in computer-mediated communication for skilled bilinguals, such as the Swiss-German chat rooms described by Siebenhaar (2006) where dialects of German were used alongside Standard German depending on the topic of

conversation. Mode-switching, as a phenomenon, is more than just the incorporation of both style-shifting and code-switching in multimodal environments. Mode-switching itself bears similarities in its usage to the ways that code-shifting and style-shifting are deployed in spoken language. There are four major similarities that drive this belief: linguistic skill involved by speakers, modes as linguistic resources to be used in conversation, modes carrying meaning in discourse, and the usage patterns of modes.

First, code-switching is described as a strategy employed by skilled bilinguals, and by code-switching appropriately, speakers demonstrate their identity as skilled bilinguals (Winford 2003). Mode-switching functions in a similar way; it is employed in this community by skilled players who have mastered the set of rules for the modes. These rules include the practical rules like how to access each of the modes (hitting the Enter key to type in the chat box, or using a push-to-talk key in Ventrilo), but also the social rules for the modes including the type of language used in each and the appropriate situations for using each of the modes. These rules for the use of linguistic modes has arisen, and users demonstrate their competency as members of the community by using the modes correctly.

Second, Heller (1992) suggests that code-switching is a way for people to manage linguistic resources, and is part of the range of practices that people use to establish social goals. Modes are resources available to users as *channels* to carry their linguistic information. *Channel* refers to ways that language may be transmitted, usually broken down into oral, written, and signed channels (Saville-Troike 1989:23). “Channel” and “linguistic mode” may be nearly analogous terms, but in this work I use the term “linguistic mode” to avoid confusion with the concept of a “chat channel”, a term used to denote different “chat rooms” in the same general environment. Saville-Troike’s definition of “channel” indicates that a speaker’s “choice of

channel may depend on environmental conditions” (52), or, alternatively, that the environment determines the range of channels available to a speaker. Similarly, in multimodal environments, a range of modes are available to speakers depending on the nature of the situation. In virtual worlds like *World of Warcraft*, the game itself provides a platform for the use of linguistic modes like text chat, while the community has incorporated external programs like Ventrilo to facilitate the usage of other modes. With all of these modes available, speakers develop strategies for managing their interactions in all of these environments. As a linguistic resource, modes may be managed in a similar way to the management of linguistic codes, leading to the phenomenon of mode-switching.

Third, Gumperz (1982) suggests that speakers do not use language simply because of situational factors, but rather because they have many linguistic choices to exploit in order to convey intentional meaning. In this, Gumperz says, code-switching is a type of discourse strategy, which spawned a number of investigations into discourse-related codeswitching (e.g. Auer 1984). Myers-Scotton (1993) proposed a markedness model of code-switching, distinguishing between unmarked codes (which are more frequent) and marked codes (those that are less frequent); furthermore, she asserts that using a marked code conveys additional information about the utterance which is linked to the code itself. Woolard (2004: 80) suggests that the most useful application of the markedness model is as an activity of the interactants rather than as an actual property of the code – that is, the additional information conveyed by a code is negotiated by interlocutors in interaction. In this, interlocutors have some volition over their choice of code usage; some codes may be unmarked in certain situation, and the choice to use a marked code may be a strategic move on the part of the speaker to impart extra meaning onto the utterance. These ideas are relevant in informing the analysis of mode-switching in the

World of Warcraft community. Similar to code-switching, speakers have a range of modes available in multimodal environments; they can choose which mode to use to carry their linguistic information. However, the markedness of modes is not the same all the time in the game community, which affects the reasons for choosing certain modes over others. Certain modes are privileged in certain situations, such as voice chat as a tactical resource during raids (Olson, 2009), but players choose to use different modes. Why? Each mode does different work in the social environment. This question is investigated in this paper: what is the purpose of using these different, marked modes, and what social work does mode choice do in interaction?

Fourth, a thread of discussion in the code-switching literature relates to the difference between situational and metaphorical code-switching, which was first introduced by Blom and Gumperz (1972). Situational code-switching is tied to changes in the conversational situation, and in fact the code-switch is usually caused by them. The idea of situational code-switching can be applied to situational mode-switching, when speakers change modes due to a change in their environment. In the opening of this paper, I described the scenario in which my raid leader ordered our group to “get on Ventrilo”, and we loaded the program and put on our headsets. This is an example of situational mode-switching, in which an event (a raid) has begun, and a different mode is more appropriate for use (and in fact explicitly ordered by a powerful participant). While situational mode-switching is interesting and an event in the community I studied, I will instead focus on mode-switches that resemble metaphorical code-switching. Metaphorical code-switching does not need to follow a situational change in the conversation, rather speakers use the switching of codes to add meaning to their utterance by using their knowledge of the social meaning of the codes. Auer (1995: 120) proposed a list of motivations for metaphorical code-switching, including: reported speech, change of participant constellation,

side-comments, reiterations, change of activity type, topic shift, language play or shift of key, and topicalization. Many of these types of metaphorical code-switching also apply to mode-switching, especially participant constellation and topic shift, and I will present examples that illustrate each later in this paper.

Switching languages, dialects, or varieties in interaction imparts language-related meanings onto utterances; however, switching modes or mediums may not necessarily change the actual language variety being used, but instead does other social work. One of the most important uses of mode-switching, I argue, is changing the participation structure of the interaction. That is, when a player of *World of Warcraft* chooses to speak over Ventrilo, he or she is communicating only to those hearers who are currently logged in to that particular Ventrilo channel; when a player types in a certain chat channel, they are typing to be read by the occupants of the chat channel. This is one of Auer's types of metaphorical code-switching (a change in participant constellation), and is, in my analysis, the most salient motivation for mode-switching in this game environment. To demonstrate the role of participant constellation in mode-switching, I will analyze several excerpts in which mode-switching occurs. However, first I will describe the linguistic modes available to players (voice, Ventrilo voice chat, and text) and the norms of their usage, and then I will move on to the analysis of mode-switching.

The Data

My initial observations about the modes and their use come from an ethnography of *World of Warcraft* conducted from 2007-2011. However, in order to investigate linguistic mode usage more authentically, I obtained video recordings of players of *World of Warcraft*. During April 2009, two volunteers (named Owen and Matt) from the guild I was observing made recordings of themselves using two Flip Video cameras. Each camera recorded an hour of video.

The participants recorded gaming-only sessions, meaning that they were focused on gameplay for the duration of the hour. Any other participants who were playing with Owen and Matt at the time of the recording were informed of the nature of the study and consent was obtained. Owen and Matt were asked to record at least two gaming sessions during which the researcher was also present. This was done for three reasons: first, so that I could obtain recordings of the Ventrilo chat happening at the time; second, so that I could obtain informed consent of all incidental participants in the recordings aside from Owen and Matt; third, I insisted on being present so that I could take field notes during the recordings of any interesting occurrences, and simply so that I would know what was happening during the recordings. In one recording, Owen and Matt were playing with a third player, Blake, who was aware of the recordings and appeared in the background of both videos. For this particular recording, Blake's language and gestures are transcribed as well.

Owen, Matt, and Blake were chosen as participants in this study because of their expertise in the game world. They were, at the time of the recording, so-called "hardcore" players and experienced gamers who had been playing *World of Warcraft* for at least three years. They participated in many guild events and played the game more than twenty hours per week; for Owen and Matt particularly, the game served as a regular social outlet. Their expertise in playing the game also meant that they had expertise in the norms of communication in the game environment. Because of their skill in both gaming and communication, I can observe the mode-switching behavior of my participants as a type of skilled performance, which is a characterization frequently made of code-switching (Myers-Scotton 1993:47).

The Flip Video cameras were situated in such a manner as to capture the vocal utterances of the participants as well as the sound coming from their computers. The video also needed to

capture the screen of each player. For this reason, the cameras were placed behind the participants, offering an over-the-shoulder view of their hands on the keyboard and mouse as well as the screen. The faces of the participants were not recorded at all during the session.¹

Recording by any means (whether audio or video) has both its benefits and drawbacks. A benefit of recording, particularly with video, is that the recording device is not biased, especially if the camera remains stationary (Jordan and Henderson, 1995:51). Furthermore, the analyst can review the recordings many times in order to create an accurate transcript. The primary drawback to recording participants in interaction is that knowledge that they are being recorded may cause the subjects to alter their behavior because they feel self-conscious. However, many analysts (e.g. Jordan and Henderson:55) have shown that participants habituate to a camera relatively quickly, especially when it has no operator.

The videos were then transcribed using ELAN, a video transcription program that allows the user to view multiple videos at the same time and produce transcriptions with anchors in a video file. Transcriptions were made with the assistance of a research assistant and double-checked to ensure that transcriptions matched all available recordings.² The transcription style was a combination of an Interaction Analysis approach (e.g. Jordan and Henderson, 1995) and a Multimodal Discourse Analysis approach (e.g. Norris, 2004). The transcription framework is outlined in the following section.

The transcripts were analyzed for interaction hierarchy and modal density to observe how the subjects chose their mode of interaction based on the events happening both in their physical space and in the game space. Linguistic and gestural indications of speech activity were also

¹ If I were to make the recordings again, I would send extra cameras so that another view could be obtained of the player's face while playing. This would facilitate an analysis of gaze and reaction to in-game events.

² The research assistant for this project was Michelle Moser.

observed for differences in behavior depending on activity or individual differences. In this work, I focus primarily on the verbal and written modes; while gesture, layout, and environment are important contributors to communication in both the physical and the digital world, they are outside the scope of the analysis I conduct in this paper.

The Transcription Framework

The transcriptions used in this project were a synthesis of many different approaches and constitute a project of their own. The primary motivator for the transcriptions is Multimodal Discourse Analysis (Norris, 2004), which allows the analyst to discern many modes outside of verbal communication. Norris's transcription method includes notating not just the utterances of the interlocutors, but also the gestures, body positions, head movements, and gaze of the participants; the layout of the environment and the objects being used in the interaction are also accounted for in the framework. By observing all of these modes (including both "linguistic modes", in which language is employed, and "non-linguistic modes", which do not involve the use of language), an analyst may be able to make closer observations about the nature of language in interaction when combined with the other modes of communication. While I do not specifically analyze gesture, gaze, and position in this paper, these were transcribed in the process and in some cases informed the analysis.

Norris (2004, pp. 80-94) also discusses the idea of *modal density*, meaning the intensity of focus on certain modes. For instance, a telephone conversation has high intensity of one mode – that being verbal communication – because the interlocutors do not share a physical environment and cannot see each other's gestures or body movements. On the other hand, an interaction like playing a video game together may be dense through complexity because many modes are at play – the interlocutors use verbal communication, but also gestures and body

position, as well as an object (the game console or computer) in the environment. Kress's (2010:159) description of such multimodal environments likens them to orchestrating a musical ensemble for the desired outcome – he attributed agency to the speakers in their use of linguistic modes, saying that the speaker makes a conscious choice of the mode that will carry their message.

In my transcriptions, I adopt a style similar to Norris's although different in several ways to accommodate the various modes mediated by technology. First, Norris uses sequences of images with the words superimposed onto the images to depict the flow of conversation in all modes. While this gives an excellent visual representation of the interaction, in practice operating with such a transcript is cumbersome. Furthermore, not all of the categories included by Norris are relevant for the analysis done in this project, and in turn, I have created some categories which are particularly relevant to the type of interaction that I am describing in this paper but which are not always relevant to all interactions. Those new categories of transcription are the behaviors of the avatar inside of the game world – and these behaviors have an impact on the interaction of the participants in the physical world.

Because the speakers in my data are primarily interacting with each other through the mediation of an object in the environment – the computer, and with it the game – I adopted the insights of Interaction Analysis. The Interaction Analysis approach is defined by Jordan and Henderson (1995:39) as “an interdisciplinary method for the empirical investigation of the interaction of human beings with each other and with object in their environment.” The focus of Interaction Analysis is to discover how people make order of the actions of their interlocutors (41), and the method for doing this is to include an extensive documentation of these actions in the transcript. The only way to capture these actions is to record interaction with videotape for

the analyst to transcribe.

Table 1 – The Modes

Owen face-to-face	}	Linguistic Modes
Matt face-to-face		
Ventrilo		
Chat		
Matt proxemics	}	Non-linguistic Modes
Owen proxemics		
Matt avatar		
Owen avatar		

Using ELAN, I annotated the videos with the columns of transcription found in Table 1. I use the term *proxemics* to encompass physical body positions related to the computer and to each other, as well as gesture, gaze, and actions (although these do not typically fall under the term *proxemics*, it kept the transcript from becoming unwieldy). Keystrokes or mouse movements are required to make the avatar perform an action; in many cases in the transcript, these physical actions are assumed when describing the character's actions. However, as previously stated, the aural and written modes will be the focus of this analysis, although the proxemic and avatar information are coded in the original transcript. In this paper, the excerpts have been converted to a table format for ease of reading, with time in the first column and each subsequent column indicating one of the tiers that were transcribed. Some rows which do not contain important information were deleted for ease of reading.

Linguistic Modes

In this section, I will describe each of the different modes being used and the ways that the modes are typically employed. I will also discuss the participation structures – or who is present and available to speak and hear the talk being done – in each mode. The modes that I will discuss in this section are face-to-face talk, voice chat (mediated by the software Ventrilo), and

textual chat. Due to space limitations, I will not include transcripts of each individual mode in this section; however, examples of each mode will be seen in the mode-switching excerpts in the next section.

Face-to-Face

Traditional face-to-face verbal communication has been the subject of linguistic study for as long as the field has existed. In the context of this study, what I dub “face-to-face” is not actually occurring with the participants facing each other or talking directly to each other – in one video, Owen and Matt are sitting back-to-back in their shared dormitory room, and in another they are sitting next to each other at a long table with Blake in the background at another table. Regardless of their positioning, face-to-face is the label for verbal conversation that only occurs between Owen, Matt, and Blake, which may also be affected by gesture, expression, and objects in the room they are in.

When engaging in gameplay together, the participants usually use face-to-face conversation for non-game-related socialization and narrative, usually based around objects in the shared environment. In the videos, the participants discuss their pets, the food and drinks that they are currently consuming, and television shows that they had watched earlier in the evening. Since the players share a physical space, they can comment on these topics to each other; it would not make sense for Owen to use textual chat or Ventrilo to complain to Matt about having cat hair on his keyboard when Matt, the cat hair, and the keyboard are all in the immediate environment. This is an example of mode-switching due to topic shift (related to Auer’s (1995) discussion of metaphorical code-switching), specifically for topics that relate to “real-world” objects and occurrences.

Aside from socialization and discussing objects in the immediate environment, face-to-

face communication is also used to remark on what is happening in the game with a controlled audience. One of the most common things that the three players did was make disparaging remarks about the other players who could not hear them. For example, Matt questions the playing ability of other players in the raid called “tanks”, whose job it is to control the enemy creatures; Matt’s character repeatedly dies in the game, which suggests a failure of the tanks to adequately control the creatures, and thus Matt complains about this in face-to-face talk to Owen and Blake, neither of whom are tanks. This allows Matt to air his frustrations without incurring the face-threat inherent in telling another player that they are not doing their job properly. This is an example of the participation structure influencing choice of mode use.

To summarize, face-to-face talk is reserved for utterances which are appropriate only for the audience of co-present players. This is one participation structure available for the players to use.

Ventrilo (Voice Chat)

Ventrilo is a voice chat program marketed towards players of online games. Ventrilo differs from other voice chat programs like Skype in a number of ways: 1.) It has “channels” for different users to occupy, which can be set up to differentiate between different game activities such as socialization, raiding, and player-versus-player combat, 2.) It has no readily available video or text chat capabilities³, and 3.) It has a “push-to-talk” (PTT) button, which is a key on the keyboard that players must hold down while they talk in order to broadcast their speech to the chat room. The PTT button is useful for gaming because often there will be many people in a Ventrilo channel (e.g. 25 in one of my video excerpts), and having all players talking at the same time would be interactional chaos.

³There is a direct chat option like an instant message, but this is only available to owners or moderators of a server, and there is also an infrequently-used chat room option. Ventrilo has no video chat capability that I am aware of.

To be present on Ventrilo is to have the program open, be logged in, and in a channel with other players. The list of Ventrilo users is available simply by looking at the Ventrilo window, so all users know who else is occupying their channel. The Ventrilo login information is usually kept private within one guild (a sub-community of the larger *World of Warcraft* community), so access is restricted to those people who are already known to other participants. The participation structure of Ventrilo is further broken down to what users are currently occupying the same channel. If someone speaks on Ventrilo, their utterance can only be heard by those in the same channel, presuming that the others are currently at their computer.

Little work has been done on the role of voice chat in online gaming environments. Kavetsky (2008) described how Ventrilo was a tool for ascertaining the physical gender of other players, and the ways that male and female players negotiate gender both in text and in voice chat. Olson (2009) discussed the phenomenon of “tactical use of Ventrilo” in *World of Warcraft* – namely, that Ventrilo was used for tactical purposes in raiding and player-versus-player battleground environments which require much character movement and leave little time or room to use the keyboard for chat. Being able to use voice chat allowed players to communicate their needs and directions much more quickly, thus allowing better overall performance in these particular game environments. This tactical use of Ventrilo is certainly seen in my study, as it is used during the entire video of the raid, and all raiding instructions are given over Ventrilo. This privileging of the voice chat mode in these situations is an example of mode choice due to activity type.

Wadley et al. (2007) studied voice chat usage based on interviews and diaries from players of *Dungeons and Dragons Online* and *World of Warcraft*, with the basic finding that players tended to use voice chat to facilitate communication and to create a more personal feel to

the gaming environment. This finding relates to a use of Ventrilo in my study, namely for socialization with other players while playing the game. It does not need to be used strictly for tactical information, although most socialization done over Ventrilo revolves around game-related information. This is because those who are using Ventrilo are often using it while playing the game or engaging in activities together, so they have a shared environment to discuss. This is somewhat similar to the role that face-to-face talk plays, except the participants are not all physically present, and the objects in the environment that are available for reference are in a digital world (or related to it, such as one's computer, mouse, or headset).

Textual Chat

The nature of textual chat in online environments has been much-discussed in the literature on computer-mediated communication. Its role in multimodal environments, however, is less-researched. Sindoni (2011) discussed the role of text chat in video chat conversations and found that participants frequently used the text chat function to convey secrets or very personal information that they either did not want to have overheard by others in the environment or that they felt was a face threat of some sort to their interlocutors. She suggested that putting the utterance into words mitigated the face threat. Sindoni's findings support an earlier assertion by Walther (1995), who argued that computer-mediated communication (textual chat) was not less intimate than face-to-face communication in a comparative study.

In the video data from *World of Warcraft*, the uses of text chat were numerous and usually related to the participation structure of the chat channel where the talk took place. The primary use was to discuss game-related information or events with friends who were not physically present and also not present on Ventrilo. The players who were not present on Ventrilo were usually not engaging in the same in-game activity, therefore they would not be using voice chat

to coordinate. Some examples of this type of interaction are greeting a player when he or she comes online, asking questions about guild events, comparing strategies for completing in-game quests or achievements, relating stories about recent events, or discussing game content. This type of discussion was usually done in the game-created [Guild] chat channel. The participation structure for the [Guild] chat channel includes everyone in the guild who is currently online and at their computer.

The most restricted participation structure in the use of text chat is the private message, or a message sent directly from one player to another. The speaker controls the audience of the talk entirely, by restricting the message's visibility to only one other person. Another version of this is the private user-created chat channel – a channel that players can create for only their friends. One example of this which will be important later in this paper is the chat channel called [TheBackRoom], which I made for only Owen, Matt, and myself. We used this channel to coordinate the use of the video cameras for this project, but during the course of the recordings, the chat channel was used for many other things. Millions of these user-created chat channels exist, allowing select groups of players to be present in the same chat channel to facilitate in-group conversation. This is a form of controlling one's presence and availability in the textual mode -- by only allowing certain other players access to this channel, one can restrict who sees the talk.

Textual chat can also be used to coordinate game-related activities between players who do not necessarily know each other. Frequently, players in the same virtual area who are attempting to complete quests will team up to help each other out, even if they have never met before. They do this by creating a “party”, and these players are automatically put in a chat channel called [Party]. For such a short interaction with an unknown person, Ventrilo information

is typically not shared since Ventrilo is seen as the realm for friends and guildmates only. Therefore, game-related strategy and coordination between strangers will happen in the chat channel [Party] (or [Raid] for larger groups). The participation in this chat channel is limited to those in the team; once the team disbands, the [Party] chat channel is no longer available.

There are several chat channels that players are automatically subscribed to in which the participants are not known. These channels are [General], [Trade], and [LocalDefense]. Talk in these channels is considered public, and can be read by hundreds and sometimes thousands of players. The public nature of these chat channels, and the number of unknown participants, discourages many players from producing any utterances in them at all; however, some players perceive these chat channels as a stage for displaying their wit or their gaming prowess. A minority of players sees these public chat channels as a means to annoy their fellow players by spamming or saying purposefully inflammatory things (see Friedline and Collister, forthcoming, for a discussion of “spammers”). One peculiar use of textual chat in the guild I studied was the collaborative recitation of songs, chants, or popular quotes from geek culture in these public chat channels. One common recitation was “summoning Captain Planet in General Chat”, meaning using [General] to recreate the iconic chant from the cartoon *Captain Planet and the Planeteers* that summoned the superhero. This chant served no purpose in the environment of *World of Warcraft*, except to demonstrate the alignment of members of the guild with geek culture and to incite a reaction by other players by causing a disturbance. This particular use of textual chat was quite salient and important to the community (and one which happened relatively often). An example of this behavior is in Example 1.⁴ In this excerpt, players remark in both verbal modes Ventrilo (247 and 269) and face-to-face (262 and 263) about the talk happening in the chat mode.

⁴ It is important to note that this is a failed attempt to summon Captain Planet, as the order of elements is incorrect. A failure to properly recite the chat was routinely mocked by guild members, see line 271.

Example 1 -- Captain Planet

Line	Begin Time	Matt Face-to-Face	Ventrilo	Chat
243	05:25.7			[General] Jimli: FIRE!
244	05:27.0			[General] Erp: EARTH
246	05:29.1			[General] Mysero: WATER
247	05:31.0		Gregor: general was so quiet	
248	05:31.6	[laugh]		
249	05:31.8			[General] Rufus: HEART!
259	05:38.4			[General] Dora: STFU
261	05:39.8			[General] Hellias: >.<
262	05:39.8	[laugh] STFU lookit		
263	05:40.6			[General] Yubna: rofl
267	05:44.8	Gosh you guys are causing a riot on general chat		
269	05:46.4		Rufus: everyone says zoh my god SeeD's raiding today	
270	05:47.4			[General] Snoopy: HRAET
271	05:47.4			[General] Gregor: ^ FAIL

Text Chat

Finally, textual chat is used for roleplaying purposes. Roleplaying is the domain of a significant minority of *World of Warcraft* players, and involves the player acting as their character in interaction with other roleplayers. That is, players create histories and personas for their characters, and collaboratively create storylines and interactions with other players. This is a creative venture that not all players participate in, but the linguistic rules of roleplaying had an effect on the textual chat used by the players in my study. Roleplayers value “proper grammar and spelling” (see Friedline, 2008) in their textual chat, because their utterances are seen as a collaborative storytelling. Stories are usually read in novels or short stories, which are edited and use standard written English, and roleplayers have adopted the more standard style of writing for their roleplaying interactions to give them a similar feel to these written mediums. This is one example of a style-shift within a mode – the shifting to a standard written style for roleplaying

from a more casual style used in non-roleplaying interactions. Usually one specific chat channel is used for roleplaying, which is the channel [Say]. [Say] is unique because typing an utterance in this channel makes the utterance appear above the avatar's head in a speech bubble, so it seems that the talk is coming from the avatar and not from the player. Players can also use this channel to emote at each other, or cause their avatars to perform actions like cheering, dancing, or waving. This visual aspect of this channel is perhaps what contributes to roleplayers appropriating it for roleplaying utterances. Furthermore, utterances typed in [Say] can only be read by other players whose avatars are in the immediate vicinity of the speaker. This is a somewhat different type of participant structure relying on spatial organization, which is what makes its usage so unique in the community. When players in my study use [Say], they often adopt a more formal or standard style of writing. This has an effect on the use of this particular channel in mode-switching, as I will show later in the paper.

To summarize, in the video data collected, textual chat serves these main functions: discussing game-related information with players not on Ventrilo, displaying guild membership and identity in public channels, talking in controlled participant environments, roleplaying, and emoting.

Having set out the uses of each mode individually, I will now move on to documenting the ways that players engage in mode-switching, and how the players may shift modes based on the purpose of the interaction.

Mode-Switching

As stated above, mode-switching is the quasi-synchronous use of two or more linguistic modes. One simple example of mode-switching is in Example 2, in which the topic at hand causes a mode-switch from Face-to-Face to Ventrilo. In this excerpt, four players (Owen

[Jahaerys], Matt [Natholis], Theon [a non-copresent player], and Parnopaeus [the author of this paper]) are fighting a group of enemies. Parnopaeus is acting as the “tank”, or character in charge of taking damage from the enemies to protect others; Matt, as the character Natholis, is a healer, or in charge of healing damaged allies; Owen (as his character Jahaerys) and Theon are damage dealers, or DPS. What has happened is that Owen’s character died after running into a large group of enemies (which he is not supposed to do, as a DPS character). Matt informs Owen of his character’s impending death in the face-to-face mode (line 191), and then Theon and Parnopaeus comment on Owen’s character’s death over Ventrilo. The occurrence of this event in the game world makes discussion of it relevant for comment on Ventrilo because all of the Ventrilo participants are also present in the digital game space and have access to the event for reference. Matt switches to Ventrilo in line 200 to make fun of Owen by sarcastically suggesting that he “just couldn’t get that heal off fast enough”; he could have said the same thing to Owen only by using the Face-to-Face mode, but by using Ventrilo, he includes the other two participants in his joke. Because the event is game-related, Ventrilo is an appropriate medium to use here; furthermore, Matt adds meaning to his joke by holding down his PTT key long enough for the listeners to hear the snap of his fingers, indicating his sarcasm. Sarcasm is notably difficult to convey in textual chat, and the tone of Matt’s voice and the snapped fingers indicate the tone of his comment. Then, in line 202, there is another modeswitch as Parnopaeus reacts to Owen’s antics by using [Say] to emote “Parnopaeus frowns with disappointment at Jahaerys”. This emote can only be done in text; presumably she could have said over Ventrilo “I’m frowning with disappointment at you, Owen”, but such narrations of actions are infelicitous when a quick typed command can cause a character action that can be “seen” by others.

Example 2 – This Could Be a Bad Thing

Line #	Begin Time	Owen f2f [Jahaerys]	Owen Ventrilo	Matt f2f [Natholis]	Matt Ventrilo	Parn Ventrilo	Theon Ventrilo	Chat
191.	04:48.1			ooh someone's gonna die cuz I'm fea:red	Face-to-Face			
192.	04:50.3	I think it's ^me^!						
193.	04:52.7			[laugh]				
195.	04:56.9						this could be a bad thing	Ventrilo
196.	04:57.0		Owen				this is why you don't do that	
197.	04:59.6			but it's so much fun	Matt			
200.	05:30.8							
202.	05:33.0							Parnopaeus frowns with disappointment at Jahaerys. Chat

In Example 2, the speakers have agency in choosing the modes they use for communicating their ideas. Matt chooses to use Ventrilo in Line 200 to increase the audience for his sarcastic comment about Owen’s character’s death, which is a change in participant constellation. Furthermore he uses a gesture (snapped fingers) that has an aural component to indicate the tone of his remark. This sarcastic comment may be seen as a type of language play,

reminiscent of the ways that code-switching is used for puns or jokes (Auer, 1995). Owen responds in Ventrilo in line 197 to remarks made about him in the same mode, continuing the flow of conversation. Finally, Parnopaeus uses a written mode, that of the emote, in line 202 to convey extra information by having her character perform an action in the game world. For all of these turns, other mode options are available to the speaker – but the modes that are chosen by the speakers fit within the social rules for mode usage.

Different modes can be employed to do other social work besides controlling the flow of conversation. In the section describing the various modes, I mentioned that the Face-to-Face mode was employed to avoid a face threat while remarking on the playing abilities of other players. In Example 3, this concern is put into action. The excerpt comes from a raid (in a multimodal situation similar to that described in the introduction of this paper), in which 25 players are attempting to defeat an enemy “boss” called Patchwerk. This particular encounter has a counterintuitive mechanic for the players – Patchwerk will attack and instantly kill any player who has a full health bar (that is, has not taken any damage). The solution to this is for players to poison themselves with a conveniently placed river of slime, which will reduce their health, and for the healers to never heal anyone to the point that their health bar is full.

The conversation in the excerpt occurs right after the entire group has died due to a misunderstanding of this strategy. Owen (as his character Jahaerys), Matt (Zalburg), and Blake (Exelsior) are all playing DPS in this excerpt. There are two other participants in the conversation: Parnopaeus, a healer who is in charge of coordinating the healer group, and Theon, a DPS who is acting as the leader of the entire raid group. Blake notices that he was killed by Patchwerk because a healer healed him to full health, contrary to the strategy. Matt confirms this problem as it happened to him as well. They discuss it first face-to-face (lines 520-526), and then

Matt moves to textual chat, using the private chat channel [TheBackRoom] which only himself, Owen, and myself (Parnopaeus) inhabited. Matt tells Parnopaeus that one of the healers (Joannis) has misunderstood the strategy and wound up killing everybody (lines 534-592). Parnopaeus, being a healer, was then able to communicate this mishap to the rest of the healers and the leader of the group, Theon (although this text doesn't appear in this transcript, since neither Owen nor Matt had access to those channels). Following this, Theon reminds the healers (using Ventrilo) to not heal the DPS (lines 636-649). In this exchange, mode-switching is used by one player – Matt – to achieve the task of refining the strategy without publicly identifying the player who has committed the error. He switches from face-to-face verbal conversation to textual chat in order to address a different audience. Mode-switching is also done by Theon from textual chat to voice chat, although one part of the exchange is not captured in the transcript.

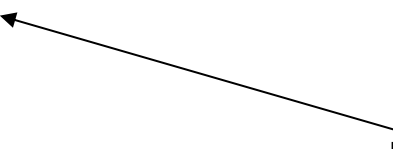
Example 3 – Melee's Getting Healed

Line #	Time	Owen f2f [Jahaerys]	Matt f2f [Zalburg]	Blake f2f [Exelsior]	Theon Ventrilo	Chat	Chat 2
519.	09:46.7			did you die too?	Face-to-Face		
520.	09:48.1			let's see here who all is up			
521.	09:48.6		the offtank died so yeah I'm all alone				
522.	09:50.1	one offtank died					
523.	09:50.8			oh it might be from the the heal from Joannis that I got			
525.	09:54.9		you got another one?				
526.	09:56.1			ah well I got a heal from Joannis as I was in fighting so I had full health when I hit mortal strike			
529.	09:59.9		Yeah				
534.	10:07.8						
552.	10:23.3		where is that name (.) I can't spell it				
560.	10:28.1		there we go				
562.	10:30.2	I just took forty					

Matt switches

Chat

[7. The-BackRoom]
Zalburg: Me-lee's getting healed



563.	10:30.5	six k to the face	[7. The-BackRoom] Zalbag: Joannis
572.	10:36.4		[7. The-BackRoom] Parnopaeus: probably a splash
575.	10:42.0		[7. The-BackRoom] Zalbag: 17.5k
587.	10:54.4		[7. The-BackRoom] Zalbag: >.>
590.	10:58.7	how did I die?	
592.	10:59.9		[7. The-BackRoom] Parnopaeus: okay that's not a splash
594.	11:00.6	I feigned death	
596.	11:03.2	oh well	
636.	11:39.2		
641.	11:43.4	oop!	
644.	11:47.0		
649.	11:52.2		

Ventrilo

um also one of the healers was healing the melee dps I don't know who was doing it but do not do that

the only people you should be healing is the tank that you are assigned to if you heal the melee dps you are going to kill them

In Example 3, mode-switching is employed to solve a problem with a particular player not following the rules of a gaming encounter. The use of many modes in Example 3 shows how speakers shift modes because of the different participation structures. Ventrilo is broadcast to all who are listening, but face-to-face talk is confined only to those in the physical area. Textual chat occupies an entirely different role, and that depends on what channel is being used. In Example 3

there is the use of the chat channel [TheBackRoom], created by the players to communicate information only to those in the channel. Again the agency of the speakers is on display here – Matt uses a mode with a restricted participation structure with very few players present in order to communicate the errors of a non-present player. He did not (as he could have) call out Joannis on Ventrilo or using [Raid] chat (both of which Joannis was present in); he chose his mode to avoid the face threat. Later in the excerpt, in line 636, Theon speaks to the entire group on Ventrilo but avoids the face threat by not naming the errant player, but rather simply saying “one of the healers” and “I don’t know who was doing it”, followed by a summary of the strategy in lines 644 and 649 to remind the errant healers of what they should have been doing. The choice of mode here is part of Theon’s agency – he uses Ventrilo, a mode normally used for discussing raid strategies, and speaks in a somewhat stern tone of voice so that the other players can know that what he is saying is important.

In Example 4 below, mode-switching is employed for face-saving between players to maintain the good relationship and negotiate an in-game benefit. Example 4 is a long excerpt and requires some explanation. Owen and Matt have teamed up with two others, Theon and Parnopaeus, to kill a boss in a dungeon called Karazhan that has a small chance to reward an exceptionally rare item – a rideable horse with flaming hooves called “The Fiery Warhorse”. This “mount run” video is a successful attempt to attain this warhorse⁵, and after some negotiation, Matt was awarded the warhorse because he won a roll of virtual dice. However, Matt was concerned about his rights to claim to this warhorse because it was his first time running with this group, and checks verbally with Owen about whether he should even attempt to roll the dice (line 357). Then, after he makes the winning roll, Matt says over Ventrilo “you guys

⁵It cannot be overstated how rare it is to actually see this item, and it is viewed as “miraculous” that I managed to capture this event on videotape.

sure? I'll pass if you want”, even though he still takes the warhorse after Theon offers to buy it from him in virtual gold (lines 382-417). During the entire process, he negotiates with Owen in face-to-face interaction about what he should do to repay the party for allowing him to win the warhorse – the fee that they decide on is that Matt will agree to go with the group on their weekly attempts to get the horse again (408-410, 492-496, 508). Finally, Matt declares to the party with a performative utterance “I hereby will aid you every week” in the textual chat mode [Say] (514-519). He even includes an emote that Natholis, his character, “signs on the dotted line”, as though he is filling out a contract, even though there was no such contract either in the virtual world or the physical world. By using this particular chat channel, which is public and can be viewed by anyone in the area, Matt finalizes his pact with the group and puts his promise into writing. He also uses a more formal style of writing, since his utterance occurs in [Say] where roleplayed utterances normally take place. What is interesting about this excerpt is that Matt uses each of the three linguistic modes to carry his promise to aid the group every week when they attempt to get the horse. These repetitions are outlined below.

Example 4 – I’ll Go With You Guys Every Week

Line	Time	Owen f2f [Jahaerys]	Owen Ventrilo	Matt f2f [Natholis]	Matt Ventrilo	Matt Proxemics	Parn Ventrilo	Theon Ventrilo	Chat
357	09:03.5			I don't know if I should roll on it or not					
359	09:04.7						kay roll it		
361	09:05.4							alright boys and girls	
362	09:08.0								Parnopaeus rolls 28 (1- 100)
363	09:09.0								Jahaerys rolls 56 (1- 100)

Line	Time	Owen f2f [Jahaerys]	Owen Ventrilo	Matt f2f [Natholis]	Matt Ventrilo	Matt Proxemics	Parn Ventrilo	Theon Ventrilo	Chat
366	09:11.0								Theon rolls 1 (1-100)
367	09:11.4						ouch!		
368	09:12.6	Ugh							
370	09:13.8			aw man					
373	09:14.7							oooh [laugh]	
374	09:16.1	you wanna roll?				types /roll			
375	09:17.0								Natholis rolls 67 (1-100)
377	09:18.6				ehhhhhh				
379	09:19.9	grats sir! [claps]							
380	09:20.6						ya:y!		
382	09:22.5				you guys sure?				
385	09:23.9	of course							
387	09:24.3				I'll pass it if you want it				
391	09:27.6							I'll pay you for it	
394	09:30.7				oh but I don't want it to go to Jahaerys [Owen] [laugh]				
399	09:34.0								
401	09:35.0							well	
406	09:38.1							I'll pay if someone wants to pass it	
408	09:42.7								
410	09:45.9	yeah! just we do this every week							

Mode mixing here – Owen affirms in face-to-face to what Matt says in Ventrilo

Face-to-Face

what if I just take it I'll come with you guys

Ventrilo

Line	Time	Owen f2f [Jahaerys]	Owen Ventrilo	Matt f2f [Natholis]	Matt Ventrilo	Matt Proxemics	Parn Ventrilo	Theon Ventrilo	Chat
414	09:48.2							I can has the money on me so:	
416	09:48.9					accepts loot			
417	09:49.0								Natholis has earned the achievement Fiery Warhorse's Reins!
424	09:54.1				I'll come with you guys every 'week'				
474	10:51.8	grats!							
475	10:52.3			I never won anything like that before [laugh]					
480	11:05.0	that's um: one of the rarest mounts in the game to get							
484	11:08.0			oh my go:d [laugh]					
485	11:09.1	so: grats							
488	11:11.3			thank you I'm glad I came [laugh]					
492	11:14.2			but yes I shall come every time					
496	11:16.8			and do uh heals for ya					

Line	Time	Owen f2f [Jahaerys]	Owen Ventrilo	Matt f2f [Natholis]	Matt Ventrilo	Matt Proxemics	Parn Ventrilo	Theon Ventrilo	Chat
499	11:23.5						let's see it!		
500	11:26.1					summons mount			
501	11:26.7			oh my god this thing's sick look- ing [laugh]					
503	11:30.7	ni:ce							
504	11:31.3						oh ho ho		
507	11:36.0								Natholis says: :O
508	11:36.8	yeah if you just want to do our mount with us from now on							
514	11:48.0								
519	11:57.0								

Matt uses [Say] for a formal roleplaying-like utterance.

Natholis signs the dotted line...
Natholis says: I hereby will aid you... every week!

In this excerpt, we can see Matt use each of the three different modes to say the same thing, that he will go with the group every week. He uses face-to-face to talk only to Owen (408 and 492), Ventrilo to speak verbally to the group of four (424), and the textual chat mode in the channel [Say] (519). This repetition shows that each of the modes does have a different meaning; if they did not, it is unlikely that Matt would bother to repeat himself. In line 408, when he uses face-to-face to talk about his plan to “come with you guys”, he is checking with Owen to see if this is a reasonable thing to do in exchange for winning the warhorse. Here he gets validation in a safe space with the most restricted audience; just like he avoided conflicts by using modes strategically in Excerpt 3, here he is checking with Owen to see if his plan is acceptable before

doing something that would cause a face-threat to the other players.

In line 424, having received validation of his plan, Matt moves to Ventrilo to tell the other party members that he was going to continue to help the rest of them get their mounts as a justification for his acceptance of the mount (line 416) despite Theon's attempt to buy it from him. Matt's use of the written mode in lines 514 and 519 serves as a way to formalize the agreement made between the four players. The performative marker "hereby" in line 519 signals this formality, referencing the type of language that would be found in a legal document (which he had just previously emoted his character signing). This evocation of another written mode – legal documents, which have no counterpart in the virtual environment of *World of Warcraft* – is a way of adding meaning to the utterance by making it permanent. This may be considered a type of language play, or key shift, and is another link to Auer's description of metaphorical code-switching. The formal "legalese" Matt uses in his utterance is a style particularly suited to the written mode, and so he mode-switches to get at the written legalese style which is available in that mode. Another factor in Matt's choice of mode is the permanence of the written language – players can take a screenshot and use it to remind Matt if he fails to adhere to his promise later.⁶

The examples above are just a few of the ways that different modes of communication may be used to create a complex event. This mode-switching is not unique to Owen and Matt – other members of my study explained that they engaged in similar behaviors when they played *World of Warcraft* (and other multiplayer games) in the presence of other gamers, and the few times that I have observed such behaviors have been very similar. The strategic use of modes appears to be a facet of more advanced gaming behaviors, and a way to demonstrate competency in the gaming environment by saying the right thing to the right group of people.

Discussion

⁶ In fact, this did happen several months later.

In the above section, I first set out the different ways that the three major linguistic modes of *World of Warcraft* are deployed in isolation, and second, I showed how the modes may be used quasi-synchronously for different purposes. Here is a summary of the uses of the three modes addressed in this paper:

Face-to-Face:	socializing with those co-present, remarking on objects in the immediate physical environment, commenting on the play of those not co-present
Ventrilo	socializing with those not co-present, tactical information, socializing during keyboard-heavy intense gameplay
Textual Chat	socialization with those not in Ventrilo, “spamming”, communicating with a controlled audience, formalizing important information

Players are aware of the many uses of each mode, and deploy them for different purposes in conversation. They are using the modes as resources for interaction, and the primary use of the resources is to access different sets of participants. The mode with the most restricted set of participants is face-to-face talk – unless a new person enters or leaves Owen and Matt's dorm room, no new participants can be added to this mode. The most free of these is textual chat – although it is restricted to those who are online at the time the talk is being done, through specific chat channels, private whispers, and modes like [Say], interlocutors can fine-tune the hearers of their talk for a number of reasons. Ventrilo occupies a unique space in the participant framework because it operates in both game space and physical space, since the utterances made over Ventrilo can be heard by those on Ventrilo as well as anyone present in the physical room.

Digital gaming, especially *World of Warcraft*, is highly socially-motivated, and in most of

the videos, the game served as the focal point for conversation. The players are all present with each other, but the conversation is not the main focus of their interaction. This is an example of Norris's (2004) notion of *anwesenheit*, or co-presence which is focused around a shared object or activity. The presence of the game in the interactional landscape has an effect on the mode usage of the players. Not only does the game serve as a platform for textual chat, it also provides a shared topic for conversation. When the topic at hand shifted to non-game entities, the conversation shifted to face-to-face verbal talk; at other times, in-game situations required communication with players who are not physically present and available for face-to-face talk. One instance is in Example 4, in which Matt has to communicate to others his decision to accept the mount that he has won. Another interlocutor, Theon, is trying to communicate with Matt through Ventrilo about "buying" the mount off of him, and Matt is required to respond to this offer or else be considered rude or indecisive. As he gains confidence in his decision to take the warhorse in face-to-face communication with Owen, Matt is able to move into another mode which has a larger audience and therefore more possible face-threats. Once he has received the warhorse, he promises the others to help them out with their ongoing attempts to get their own mounts, which he needed to do with a digitally-mediated mode or else two of his interlocutors would not have received his message. When all of this happened, the situation changed from a social gaming experience among friends to a negotiation of rewards and payment, and it is mode-switching which helps accomplish this change in interaction.

At the outset of this paper, I described some of the similarities between mode-switching and variety-changing, specifically code-switching. I have shown how different modes are resources, in that they open up different sets of participants, and that players manage their access to these sets of participants by mode-switching. When Blake comments on a healer not following

the strategy in Example 3, he tries to save face by using a mode (face-to-face) where only non-healers can hear his talk and respond to it. Then, if the participants in that mode deem it important enough, they can take it to another mode, as Matt does when he types in the extremely controlled chat channel [TheBackRoom] so that only two other interlocutors (one who is not co-present) can read his talk. Changing participation structure is one of the motivations for metaphorical code-switching proposed by Auer.

The other question I asked in the beginning of this paper is: what kind of social work is done by mode-switching? In the sequence of events occurring in Example 3, Matt and Blake achieve their goal of reporting the improper behavior of another player while minimizing the social risk to themselves; this strategic use of different modes harkens back to Migge's (2007:56) claim that speakers code-switch to “maximize interactional rewards and minimize interactional costs”. Mode-switching to control the participation structure of an interaction also bears similarities to exclusive use of code-switching, as in Genishi's (1981: 147) work showing how bilinguals code-switch to include or exclude speakers in the conversation. This is another way that the modes are employed as resources, and ways that they do social work in interaction. This social work is minimizing conflict through the use of participant structures. In Example 4, Matt mode-switches and uses three different linguistic modes to make a promise to his fellow players. At first, he uses a mode that is very controlled, that of face-to-face, to minimize conflict by asking his friend Owen if it would be acceptable for him to take the extremely rare mount in return for promising to accompany the rest of the group on their future endeavors. When he receives affirmation from Owen, Matt relays his promise to the rest of the group in an expanded participation structure. Finally, Matt mode-switches to use textual chat to document his promise in a more formal way, choosing the written mode because, in the words of Saville-Troike about

choosing between spoken and written language, “choosing oral or written channels is usually dependent on... the need for a permanent record” (52). The social work done by Matt’s mode-switching is both minimizing conflict and creating a permanent record of his words that his fellow players can use in the future.

I have shown ways in which mode-switching and code-switching are similar. There are, of course, ways that mode-switching and code-switching are not similar. One glaring example of this is that mode-switching does not occur seamlessly intra-sententially like code-switching does in some communities. I have no instances of intrasentential mode-switching in my video data from this study, and those that I have observed in other instances are usually accidental. For example, a player may accidentally press down the Ventrilo push-to-talk button while talking to another person in the room, thus inadvertently broadcasting a snippet of conversation. This is not an intentional use of mode-switching, and while it can have an interesting effect on the surrounding conversation, it does not follow my definition of mode-switching. Furthermore, this type of occurrence is somewhat rare, and I have no examples of it in the video data from this project.

Another disparity between mode-switching and code-switching is that mode-switching may occur without the changing of actual linguistic codes. When switching between face-to-face talk and Ventrilo, there are no noticeable changes in the register used by either Owen or Matt in this study, although there is a difference in topic. There is, however, a marked difference between codes when switching between casual textual chat and roleplaying textual chat, which falls under the realm of style-shifting. This, of course, is not always the case for all speakers. Throughout my ethnography, I have encountered many players who confess to attempting to sound “better” on Ventrilo than they do “in real life” – sexier in the case of one woman who was romantically

attracted to a fellow player, or tougher in the case of one male player who was attempting to exert authority over others. Whether one switches registers or codes between spoken face-to-face language and spoken language over voice chat seems to be an inherently personal style consideration.

Mode-switching exists as an entity that is somewhat similar to both code-switching and style-shifting, although it does not fall neatly into either category. The phenomenon itself looks like code-switching, especially metaphorical code-switching like that described by Auer, but the motivations for the phenomenon are most like those for style-shifting. Specifically, mode-switching seems to be a way to manage participation structure, somewhat like the Audience Design theory discussed by Bell (2001). Mode-switching appears to be related to the two phenomena, but not altogether the same. Since they share so many features, I suggest that code-switching, style-shifting, and mode-switching are all part of the same general phenomenon, that called “variety-changing” by Saville-Troike.

Conclusions

In this work, I have explored the phenomenon of mode-switching in a digital gaming environment. I have explored it through a descriptive lens, with an eye towards comparing it to a similar linguistic phenomenon (code-switching). Overall, the similarities between mode-switching, style-shifting, and code-switching are great in number; the main connection is to metaphorical code-switching in that the participation structure aspect of mode-switching that is inherent in the act of using different linguistic modes. However, there are ways that mode-switching is more like style-shifting, suggestion that the three phenomena are closely related and can be analyzed as examples of a larger category of variety-changing. With this study, I hope that the value of analyzing language switching in online communities is apparent, as new and

interesting linguistic behaviors may arise when mediated by communication technology.

It is important to understand the communicative purposes of the modes before attempting a multimodal or mode-switching analysis. For this reason, contextual information that can be provided by ethnographic observation is important to include in an analysis. In the above study, I have attempted to set out the rules for interaction as I have observed them in the community; while there may be some similarities from one online community to the next, it remains to be seen whether the uses of the linguistic modes bear any generalizable similarities across communities. Is it always about presence and participation in other communities?

When observing multimodal communication online, it is useful to remember that this is not an entirely new phenomenon. The insights gained from years of research on similar phenomena – code-switching, written language study, and multimodal communication “in real life” – can inform the study of online mode-switching. The experience of the interlocutors in a virtual environment is *inherently* multimodal; more precisely, users are accustomed to having these many modes available for communication and use them in a natural and native manner. Because the purpose of much online activity is social interaction, the multimodality itself provides social information to interlocutors, and in fact appears to be a fundamental component of interaction online.

References

- Auer, P. (1984). *Bilingual Conversation*. Amsterdam, The Netherlands: John Benjamins.
- Auer, P. (1995). "The pragmatics of code-switching: A sequential approach." In L. Milroy and P. Muysken (Eds.), *One Speaker, Two Languages: Cross-Disciplinary Perspectives on Code-Switching*, pp. 115-135. Cambridge: Cambridge University Press.
- Baron, N. (2005). Discourse structure in instant messaging: The case of utterance breaks. In S. Herring (Ed.), *Computer-Mediated Conversation*. Cresskill, NJ: Hampton Press.
- Baym, N. K. (2010). *Personal Connections in the Digital Age*. Malden, MA: Polity Press.
- Baynham, M. (1993). Code switching and mode switching: Community interpreters and mediators of literacy. In B. Street (Ed.), *Cross-cultural Approaches to Literacy*, pp. 294-314. Cambridge: Cambridge University Press.
- Bell, Allan. (2001). Back in style: Reworking audience design. In P. Eckert & J. R. Rickford (Eds.), *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- Blizzard Entertainment. (2004). *World of Warcraft*.
- Blom, J.P., and Gumperz, J.J. (1972). "Social meaning in linguistic structure: Code-switching in Norway." In J.J. Gumperz and D. Hymes (Eds.), *Directions in Sociolinguistics: The Ethnography of Communication*, pp. 407-434. New York: Holt, Rinehard, and Winston.
- Friedline, B. (2008). A linguistic profile of power and interaction in *World of Warcraft*. Unpublished master's thesis, University of Pittsburgh.
- Friedline, B., and Collister, L. (forthcoming). Power and interaction in *World of Warcraft*. In G. Voorhees et al. (eds.), *The Digital Game Studies Reader*. Continuum.
- Genishi, C. (1981). Code-switching in Chicano six-year-olds. In R. Duran (ed.), *Latino Language and Communicative Behavior*, pp. 133-152. Norwood, N.J.: Ablex.
- Goffman, E. (1964). The neglected situation. *American Anthropologist* 66: 133-35.
[doi:10.1525/aa.1964.66.suppl_3.02a00090](https://doi.org/10.1525/aa.1964.66.suppl_3.02a00090)
- Gumperz, J. J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- Heller, Monica. (1992). The politics of codeswitching and language choice. *Journal of Multilingual and Multicultural Development*, 13:2, pp. 123-142.
[doi:10.1080/01434632.1992.9994487](https://doi.org/10.1080/01434632.1992.9994487)
- Hymes, D. (1986). Models of interaction of language and social life. In J. Gumperz and D.

- Hymes (eds.), *Directions in Sociolinguistics: The Ethnography of Communication*, 35-71. Oxford: Blackwell.
- Jewitt, C. (Ed.) (2009). *The Routledge Handbook of Multimodal Analysis*. New York: Routledge.
- Jones, R. H. (2004). The problem of context in computer-mediated communication. In P. Levine and R. Scollon (eds.), *Discourse and Technology: Multimodal Discourse Analysis*, pp. 20-35. Washington, D.C.: Georgetown University Press.
- Jordan, B., and Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4:1, 39-103. [doi:10.1207/s15327809jls0401_2](https://doi.org/10.1207/s15327809jls0401_2)
- Kavetsky, J. (2008). Men behaving (not so) badly: Interplayer communication in *World of Warcraft*. Unpublished Master's thesis, Bowling Green State University.
- Kress, G. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. New York: Routledge.
- Migge, B. (2007). Code-switching and social identities in the Eastern Maroon community of Suriname and French Guiana. *Journal of Sociolinguistics*, 11(1), 53-73. [doi:10.1111/j.1467-9841.2007.00310.x](https://doi.org/10.1111/j.1467-9841.2007.00310.x)
- Myers-Scotton, Carol. (1993). *Social Motivations for Codeswitching: Evidence from Africa*. New York: Oxford University Press.
- Norris, S. (2004). *Analyzing Multimodal Interaction: A Methodological Framework*. New York: Routledge.
- Olson, C. (2009). The tactical importance of voice communication in massively multiplayer online role-playing games. Unpublished thesis. Accessed 1 November, 2010. http://stylized-lotus.com/Stylized_Lotus/Games_files/Report.pdf
- Sankoff, D. and Poplack, S. (1981). A formal grammar for codeswitching. *Papers in Linguistics*, 14(1): 3-46. [doi:10.1080/08351818109370523](https://doi.org/10.1080/08351818109370523)
- Saville-Troike, M. (1989). *The Ethnography of Communication: An Introduction*. New York: Basil Blackwell.
- Scollon, R., Bhatia, V., Li, D., and Yung, V. (1999). Blurred genres and fuzzy identities in Hong Kong public discourse: Foundational ethnographic issues in the study of reading. *Applied Linguistics* 20(1): 22-43.
- Siebenhaar, B. (2006). Code choice and code-switching in Swiss-German Internet Relay Chat rooms. *Journal of Sociolinguistics* 10(4): 481-506. [doi:10.1111/j.1467-9841.2006.00289.x](https://doi.org/10.1111/j.1467-9841.2006.00289.x)
- Sindoni, Maria Grazia. (2011, March). "Mode-switching: Speech and writing in video chats."

Paper presented at the Georgetown University Round Table, Washington, D.C.

Wadley, Greg, Martin Gibbs, and Peter Benda. (2007, December). "Speaking in Character: Using Voice-Over-IP to Communicate Within MMORPGs." Paper presented at the Australasian Conference on Interactive Entertainment, RMIT University, Melbourne. Accessed 6 August, 2011. <<http://ieconference.org/ie2007/IE2007proceedings/>>.

Walther, J. B. (1995). Relational aspects of computer-mediated communication: Experimental observations over time." *Organizational Science*, 6(3), 186-203. [doi:10.1287/orsc.6.2.186](https://doi.org/10.1287/orsc.6.2.186)

Winford, D. (2003). *An Introduction to Contact Linguistics*. Malden, MA: Wiley-Blackwell.

Woolard, Kathryn A. (2004). Codeswitching. In A. Duranti (ed.), *A Companion to Linguistic Anthropology*, pp. 73-94. New York: Wiley-Blackwell.