

# Verbal Coordination in First Person Shooter Games

Anthony Tang<sup>1,2</sup>

<sup>1</sup>University of Calgary  
Computer Science  
tonyt@ucalgary.ca

Jonathan Massey<sup>2</sup>

<sup>2</sup>Georgia Institute of Technology  
GVU Center  
{ragingpotato | keith}@cc.gatech.edu

Nelson Wong<sup>3</sup>

<sup>3</sup>University of Saskatchewan  
Computer Science  
nelson.wong@usask.ca

Derek Reilly<sup>4,2</sup>

W. Keith Edwards<sup>2</sup>

<sup>4</sup>Dalhousie University  
Computer Science  
reilly@cs.dal.ca

## ABSTRACT

We explore how expert First Person Shooter (FPS) players coordinate actions using a shared voice channel. Our findings emphasize the importance of the temporality and spatiality of these tactical verbal communications (“call-outs”). From here, we outline potential designs to mitigate problems in the production/interpretation of call-outs to better support coordination.

## Author Keywords

First person shooter, FPS, CVE, coordination

## ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

First-person shooter (FPS) games are fast-paced games that rely on quick, precise targeting and movement actions. Players each have their own view into a 3-dimensional world (Figure 1), navigating and scanning through the virtual world to quickly locate and disable avatars of the opposing team by shooting at them. FPSs are made more complex with team-based “objective” game variants where 4- to 8-player teams work together to achieve shared goals. For example, in “Sabotage” game variants, the offensive team moves to one or more well-known locations to “plant a bomb,” while the other team works to defend those locations from the offensive team. In “Territory” game variants, teams try to capture control of certain well-known locations while defending locations from opponents.

These game variants require substantive team coordination under time constraints against an opposition. But because each player has his/her own viewport into the world, teams cannot make the shared, deictic references to objects or locations that we do in everyday life [1,5]: team members are rarely within sight, and have different visual perspectives on the situation. As a consequence, the shared voice channel becomes a principal point for contact, coordination and interaction [6]. Understanding how this coordination occurs may provide insight for building

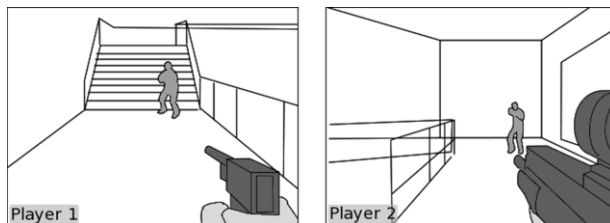


Figure 1. Players have different views in a FPS game.

technologies to help real-life mobile teams such as firefighters who also rely on a shared verbal channel.

In this paper, we explore verbal communication strategies that expert FPS players employ to coordinate their actions. Our study focuses on audio/videotaped gameplay from local competitive FPS tournaments, informed by a survey of FPS expert players. An analysis of transcribed tactical verbal communications—termed *call-outs*—reveals several characteristics of verbal communication in FPS games: call-outs are brief, highly contextual (both spatially and temporally), encoded with game-state information, and can be difficult to produce when engaged in combat.

The main contribution of our work is a detailed characterization of voice-mediated coordination in FPSs. These insights detail how experts overcome coordination problems using a shared voice channel. These ideas lay the groundwork for novel tools to support coordination in FPS games, CVEs, and potentially real-life distributed mobile expert teams (e.g. firefighters) who rely on timely communication on a shared voice channel for coordination.

## BACKGROUND

The principal challenge for FPS players is to build and maintain a shared understanding of the environment—the locations/intentions of fellow teammates, opposing players, and the state of the game. Early research exploring 3D collaborative virtual environments (predating most FPSs) foreshadowed many of the factors that impair coordination in 3D environments [1,5]: (1) narrow field of view; (2) inability to easily see others, and inability to understand what they can/cannot see; and (3) the inability to easily construct gestures to refer to objects in the environment. Moreover, the poor expressiveness of avatars, and the slow speed of text-based communication result in coordination problems [7]. Because FPS games necessitate rapid coordination, FPS players commonly use a voice channel [6]. In casual games, this voice channel is a medium for rich, social experiences for players [2,10], used for performative expression (taunting, support, accusations of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CSCW 2012, February 11–15, 2012, Seattle, Washington.

Copyright 2011 ACM XXX-X-XXXXX-XXX-X/XX/XX...\$5.00.

		Halo	TF2	CS	
Minutes collected		68:49	82:08	212:23	
Call-out rate (phrases/min)		25.5	10.5	16.2	
Call-out Codes	Intention (%)	Directive	28.4	23.6	23.7
		About self	27.5	20.4	38.7
		About enemy	12.7	32.5	13.7
		Question	5.9	5.1	3.2
		Confirmation	9.8	1.9	1.1
	Attributes (%)	Location: General	2.9	19.1	3.0
		Location: Specific	39.2	11.5	43.5
		Urgent: Temporal	13.7	7.6	11.3
		Urgent: Enemy	16.7	15.9	20.4
		Repeated	13.7	26.1	15.1

**Table 1.** Summary of our audio coding analysis.

cheating, etc.), and learning (teaching players new strategies, etc.).

Yet, what is the role of voice in tactical (i.e. moment-to-moment, low-level) coordination in FPS games? In competitive games, where players are focused specifically on winning, the character of the voice channel changes: rather than being used for socializing [10], experts use it for tactical teamwork [8] and implicit coordination [9]. Our research builds on this work, focusing on the subtle mechanics employed by expert players to overcome the problems identified by early CVE researchers.

## STUDY

Our research goal was to understand how FPS players employ the verbal channel to coordinate activity. We first conducted a small formative online survey to sensitize us to issues pertinent to FPS coordination. Using these results, we coded transcribed audio and video recordings of team FPS sessions from two local competitive tournaments.

Because tactical communication is primarily characteristic of expert players [8], we chose three games (Halo 3, Team Fortress 2, and Counter-Strike) based on their use in professional video game leagues. While each game employs unique game mechanics and gameplay, they are all FPSs, and tournament play focuses on only a handful of maps, so experts have an intimate knowledge of the environments.

*Participants.* For the online survey, we collected 24 responses solicited from relevant online forums. These respondents aged 19-32, and all but one were male. At two independent competitive tournaments, we recruited 49 different players. These ranged in age from 18-40, and were almost all males (47 males, 2 females). We targeted competitive tournaments, and focused our analysis on the final stages of tournaments (the most successful teams).

*Method.* At the tournaments, we captured voice communications by attaching an audio recording client to voice servers being used by teams; in other cases, we used a physical audio splitter at one teammate’s terminal, or put the audio recorder in front of the team. We videotaped gameplay where possible. In total, we collected

Line	Time	Call-out	Convo	codes				
				1 Intention	2 Location	3 Urgent	4 Repeat	5 By Dead
1	22:02	Where's that at?	•	Q	G			
2	22:03	I'm just prefling	•	S				
3	22:08	Squiggly door, bomb is squiggly, get lower ramp go lower now	•	ED	S	T	*	*
4	22:12	I'm going	•	S				
5	22:15	Don't peek	•	D				*
6	22:16	<inaudible> red	•					
7	22:20	You got my <inaudible>	•					
8	22:20	Get lower	•	D	S			
9	22:21	Red, go up	•	D	S			
10	22:22	He's in headshot box	•	E		E		
11	22:23	I'm lower ramp	•	S	S			
12	22:23	Upper upper upper upper	•	E	S	T	*	*
13	22:28	'Nade the site, nade the site, nade the site right now	•	D	S	T	*	*
14	22:34	I'm going to go through their site	•	S	G			
15	22:39	There's one squiggly	•	E	S	E		
16	22:41	Up in three in outer -uhm- upper	•	E	S	E		*
17	22:47	There's four upper	•	E	S	E		
18	22:52	One in smoke	•	E	S	E		
19	22:58	Alright you gotta make a move, make a move	•	D			*	*

Intention: Directive/about Self/Enemy/Question/Confirm  
Location: Specific/General  
Urgency: Temporal/Enemy

**Listing 1.** One minute coded transcript from one Counter-Strike session (some codes omitted for space). Location phrases highlighted.

approximately six hours of audio data (Table 1), along with some field notes and interview data.

*Analysis.* We inductively analyzed our transcribed audio data, coding each speech act using an iteratively developed coding scheme. This coding scheme was originally based on our field notes, and refined through the analysis process.

## FINDINGS

Players are acutely aware of how timeliness and spatial context affect verbal communication in FPS. We briefly discuss these ideas with data from our coded transcripts. We then describe strategies that require coordination, showing how they are mediated by verbal communication.

Listing 1 is an excerpt that typifies much of the verbal communication in competitive FPS teams. Table 1 shows a summary of the data we collected, and some of our codes. Call-outs were generally coded with one of the “Intention” labels: *directives* were commands; *about self* were instances of verbal shadowing, where players reported on their own status [4]; *about enemy* were labels indicating when players reported on enemy status. We also coded call-outs with “Attribute” labels: *location* if the call-out contained a general (e.g. “left” or “right”) or specific label (e.g. “sewer”); *urgent* if it was urgent regarding time or an enemy, and also *repeated* if it was repeating information.

Table 1 is suggestive of some patterns, namely that directives and awareness (about self/opposition) heavily dominate tactical communication. Of these classes, location plays a big role: 43.5% of directive call-outs have a location; 69.5% of call-outs about oneself or the opposition contain a location.

## Verbal Communication in FPSs

As Listing 1 illustrates, “conversations” in the traditional sense (with explicit back-and-forth) are rare, or are brief. Instead, the vast majority of statements are declarative: referring to *specific locations* (code 2), and often addressed to the entire team. Players tend to listen to all call-outs to maintain awareness: “[I am] always talking to [the] entire team, even when referring to a single player since his/her actions might have important change in gameplay. [P1]”

*Timing.* Immediately striking is the *paucity of speech*—most call-outs are 8 words or fewer. This reflects the importance of timely communication, and the importance of crafting useful and meaningful call-outs under duress: “[I am] Very specific when I'm not under pressure; otherwise I just blurt out the general area. [P2]” *Temporal urgency* (code 3) becomes a prevailing theme, referring to important game states. Note how rapidly this information becomes stale: because the opposing team is also moving quickly, most information is out of date within a few short seconds.

*Location.* As indicated earlier, spatial location information is also important. In the vast majority of cases, experts used *named locations* rather than relative/general references to locations. For instance, in Line 3, the player not only indicates the current location (“squiggy”), but anticipates its location in the very near future (“lower ramp”). These named locations are how experts employ common references, as they provide all teammates information regardless of where they are in the environment. While some locations had well-known names, either from unusual/distinctive landmarks in the game or from “*hearing other people refer to them during games [P3]*”, many players indicated, “*You just make them up, unless they have a label. [P4]*” Some players noted that external sources, such as screenshots and maps on forum posts, played a role in making names concrete for the entire community.

*Production.* It seemed generally difficult to produce interpretable/correct utterances for many players while under duress (navigating or engaged in combat): Listing 1 omits many such unintelligible call-outs, and Line 16 shows an example where the player hesitates and changes the information due to an incorrect call-out at the outset. Some players overcome this problem by performing *verbal shadowing* [4], where they call out their own actions as they are undertaken. In many games, “dead” players can continue to observe the game from the viewports of teammates who are still “alive” in the game. As illustrated in Listing 1, these eliminated players, freed from the responsibility of “playing,” and become more involved in providing call-outs: Lines 3,5,12,13,16,19 (code 5) are examples of “dead” players contributing.

*Understanding/Inference.* Experts operate with a deep, shared understanding of the game environments (i.e. the maps) and game mechanics: this allows them to use short phrases, knowing that others will interpret those phrases in a rich, appropriate way. Line 13, for example, is a

command (to “[throw grenades at] the site”); to an expert player, it reveals far more: first, the opponent is “planting the bomb”; second, that at least one fellow teammate is close enough to throw grenades onto the site; third, it is not clear whether there are other opponents nearby, hence the call for grenades. Accordingly, experts also understand what is to be perceived as out of the ordinary. In Line 1, the player is reacting to a nearby teammate shooting his gun, and asks, “Where are the enemies?” The other player responds (Line 2) that it is a pre-emptive shot.

*Repetition.* Repetition was fairly common (code 4)—in our sample, we found that many utterances were repeated (Table 1). Players provided several reasons for this redundancy: first, voice communication software might drop the first milliseconds, resulting in unintelligible speech; second, repetition is used to convey urgency, allowing players to emphasize important information, and third, repetition gave listeners “a second chance” to hear a call-out if they “missed it” the first time around.

*Anticipation Ratio.* As outlined by prior work, implicit coordination can be assessed by calculating the ratio of utterances with information “pushed” out *vs.* utterances with information requested (“pulled”) (e.g. [9]). As evidenced by Listing 1, the vast majority of utterances are “pushes”—our expert players are anticipating necessary and/or useful information with the call-outs.

## Coordination in FPSs

Within the context of the game, call-outs help coordinate team tactics. Although an exhaustive taxonomy of FPS tactics is beyond the scope of this paper, it is worth noting how verbal coordination supports some of these tactics.

*Pairing* involves two players working closely side-by-side. In TF2, this strategy was most often employed by “Heavies” (slower characters with powerful weapons) and “Medics” (weaker characters that can heal other characters). In these cases, there is a dual responsibility between the two characters to look out for one another, and it is understood that they are paired: relative location call-outs (“behind you”) are extremely common, and are made with the partner’s frame of reference.

*Scouting* is most common in TF2 with the “Scout” character, who is capable of moving very quickly through the map. The Scout is capable of “sighting” the opponent before quickly escaping. Players in this role will often run ahead of their team, scouting and calling out the locations and anticipated actions of opposing players. Players on the Scout’s team monitor the verbal channel and ask questions if the call-outs do not make sense, or are in conflict with expectation.

*Pushing/Falling back* are strategies that require timed, coordinated movement of the entire team to be successful. Pushing means to simultaneously attack in tight formation as a large group to overwhelm the opponent; falling back is a strategy used when the team’s forces are weak.

These strategies are generally very difficult to coordinate temporally and spatially without verbal call-outs.

### DISCUSSION AND OPPORTUNITIES FOR DESIGN

As foreshadowed by the research in CVEs, FPS games present challenges to team coordination [1,5,6]: it is difficult to see where a teammate is, what they see, and to know what a teammate is referring to. Yet, survey participants reported having to evolve and learn strategies to overcome these problems: for instance, how to maintain an awareness of their surroundings (expert players frequently glance around themselves even as they move in a given direction), the names of specific locations, the contextual implications of certain utterances (e.g. Line 13 from Listing 1), and how to produce meaningful call-outs.

While this learning takes place in a specific game context, and well-known environments, these lessons open an interesting design space for supporting verbal coordination. We see three types of potential tools: supporting the *interpretation* of call-outs by others, supporting the *production* of useful call-outs; and enhancing *current* behaviours. While these tools may not be desirable from a game design perspective (i.e. fun), we take the view that exploring these designs (even as thought experiments) provides valuable insight into verbal coordination in other contexts.

*Interpretation* is hampered by an inability to identify the speaker, the location of the speaker (particularly if they are out of view), and to understand the context of the speaker (i.e. what s/he sees). Providing visualizations of this information in a quickly-understood manner would be useful: for example, a simple arrow to cue the player to the relative location of the speaker, or perhaps highlighting the speaker's avatar *through* the walls of the map altogether (i.e. so that the player would be able to judge the location/distance and orientation of the speaker). Alternately, the ability to quickly send an image of one's view to a teammate, or to quickly toggle between one's own view and a teammate's would help support interpretability of call-outs..

Similarly, players suggested enhancements that would help the *production* of call-outs. For example, in most Counter-Strike variations, the player's named location in the map is displayed prominently on the UI (many games do not have this): players can use these labels for a call-out. Further, an enhanced UI could also generate these call-outs, akin to verbal shadowing, on behalf of the player, indicate location, direction, enemy location, etc. These call-outs might include screen shots and text messages.

It may also be interesting to enhance *current* behaviours. For example, call-outs are often repeated multiple times so that if players miss the call-out the first time, they can hear it again. This is a crude way of approximating temporal persistence of information, something that can be emphasized by providing immediate transcriptions of call-

outs in the UI. These transcriptions can fade over time, but provide the temporal permanence that repetition creates.

We also see potential for their application in ubiquitous technologies for real-life distributed mobile teams (e.g. firefighters), as members of these teams also have unique perspectives on the scene

### CONCLUSION

Coordination is generally a tricky problem, as it involves building and acting on a shared mental model. This problem is exacerbated when people do not share the same view of the world (as in FPS games). In this paper, we demonstrated that expert players overcome these problems over the voice channel using nuanced, de-contextualized utterances (call-outs), communicated rapidly in short bursts, which are repeated for emphasis. These insights have implications for tools to support both the interpretation and production of call-outs in FPSs, and more generally for CVEs.

### ACKNOWLEDGEMENTS

Thanks to GeorgiaTech GameFest and Atlanta LANFest. Huge thanks to reviewers for excellent feedback.

### REFERENCES

1. Fraser, M., Benford, S., Hindmarsh, J., and Heath, C. Supporting awareness and interaction through collaborative virtual interfaces. *UIST '99*, 27-36.
2. Halloran, J. It's talk but not as we know it: using VoIP to communicate in war games. *IEEE Games & Virtual Worlds for Serious Applications '09*, 133-140.
3. Halloran, J., Rogers, Y., and Fitzpatrick, G. From text to talk: multiplayer games and voiceover IP. *Level Up '04*, 130-142.
4. Heath, C. & Luff, P. Collaborative activity and technological design: task coordination in London underground control rooms. *ECSCW '91*, 65-80.
5. Hindmarsh, J., Fraser, M., Heath, C., Benford, S., and Greenhalgh, C. Fragmented interaction: establishing mutual orientation in virtual environments. *CSCW '98*, 217-226.
6. Manninen, T. Virtual team interactions in networked multimedia games—case: “counter-strike” – multiplayer 3D action game. *Presence 2001*.
7. Moore, R. J., Gathman, E. C., Ducheneaut, N., and Nickell, E. Coordinating joint activity in avatar-mediated interaction. *CHI '07*, 21-30.
8. Rambusch, J., Jakobsson, P., and Pargman, D. Exploring e-sports: a case study of gameplay in counter-strike. *Proc DiGRA 2007*, 157-164.
9. Toups, Z. O., Kerne, A., Hamilton, W. A., and Shahzad, N. Zero-fidelity simulation of fire emergency response: improving team coordination learning. *CHI '11*, 1959-1968.
10. Wright, T., Breidenbach, P., and Boria, E. Creative player actions in FPS online video games - playing counter-strike. *Intl J of Comp. Game Res.* 2 (2), 2002.