

THE OF FUTURE SPEECH IN GAMES



BY JEFF FOLEY, A FORMER BRAND MANAGER AT ATARI, NOW WORKS AT NUANCE COMMUNICATIONS, INC.

Perhaps you've noticed a growing trend in video gaming – extending the environment beyond the traditional interface to better bring gamers into the world. The push for immersion first appeared in coin-op arcade games (Spin the turret! Punch the pads! Pedal faster!) and has been steadily moving its way to home console systems (Steel Battalion, anyone?). Just witness the recent excitement surrounding the prototype controller for the Nintendo Revolution. From force feedback controllers to bongos, from the EyeToy to the dance pad, we're seeing more and more innovative offerings that encourage players to actually participate in the game rather than merely mashing buttons. What's the logical next step for immersive gaming? My prediction: speech recognition.

Automatic Speech Recognition (ASR) has been around on consoles for years, ever since its debut in the wisecracking Seaman game for the Sega Dreamcast. Yet for the most part, ASR has been a novelty, not a must-have. Its most acknowledged successes have been with games like SOCOM, Ghost Recon, and Rainbow Six, allowing you to bark orders at your computer-controlled teammates instead of navigating menu trees or remembering button sequences. The most innovative use of ASR was in Lifeline – an underappreciated PlayStation 2 game that gave you only a bank of monitors and your voice to guide an anime waitress-turned-heroine around an alien-infested space station. Though not without its design flaws, Lifeline demonstrated that innovations in speech recognition shouldn't be pigeonholed into the squad shooter genre.

Innovation is a good thing for an industry suffering from imitation. With well over a 1,000 console and PC releases every year, many of them well-established brands or sequels, it's hard for developers and publishers to get space on crowded store shelves, let alone pique mass interest in a title. Speech recognition can spice up an existing franchise (such as the GameCube's Mario Party 6) or differentiate a new title (such as the PSP's Talkman). Let's face it – unless a video game adds something unique and innovative to the market, it's usually dismissed as just another clone of today's genre-defin-

ing shooters, RPGs, and platformers.

With any discontinuous innovation, there are inevitable barriers to acceptance. Any Trekkies who have tried earlier versions of speech recognition were probably underwhelmed with the results. It's true – even the best implementations have limitations on what you can say as well as the occasional misrecognitions. One way to deal with that limitation has been to turn it into part of the game – such as Nintendogs has done. Who would blame a dog for misunderstanding your commands? But fortunately, the computationally intensive algorithms central to speech recognition have improved considerably in the last several years – just in time to take advantage of a new generation of console systems with enough juice to feature larger vocabulary ASR without dropping frame rates.

Mind you, speech recognition doesn't make sense for every game any more than steering wheels or light guns do. It's much easier to move a thumbstick than it is to say "go right – no, go left", and it's certainly faster to fire a weapon by button than by voice. (Though c'mon, admit it, wouldn't yelling "PEW PEW PEW!" at your console be kinda fun?) Nevertheless, ASR is a great fit for many games, especially when it can do the following:

Enhance the fantasy of the game. In many games, a player or his avatar needs to speak to a computer-controlled character – perhaps a teammate, a shopkeeper, or an arch-rival. Sometimes it feels more natural to order characters around than it does to remember button mappings. Imagine having a back-and-forth conversation with a key game character instead of navigating a list of responses or hitting controller buttons. That's a new style of interaction that brings the player further into the game's environment.

Create a more usable interface. Similarly, a complicated series of drop menus or an obscure sequence of buttons can interfere with enjoyment of the game. You'd rather say "Wield the purple lightsaber," than call up an inventory screen, scroll, scroll, (where is that damn thing anyways), scroll, right, right, select (which button was equip again?), left, left, select, exit...Or consider the painful roster management screens in most sports games – it'd

“

...speech recognition doesn't make sense for every game any more than steering wheels or light guns do.

”

be much easier to say "Put Modano on the second line" or "Trade Ramirez for A-Rod." Players want to be in the game, not diving into menus looking for an option, browsing through long lists, or cycling through a dizzying array of choices.

It's easy to see why squad shooters have embraced speech recognition – ordering "Bravo: Move To Checkpoint Zulu" is easy to use and fits the fantasy. But as the new generation of consoles raise the bar on what ASR can accomplish, we should expect more innovative uses. Role-playing games are all about immersion in a new world – so why not actually say conversation choices instead of selecting menu responses? And "Show me your best armor" or "I'll pay 50 for that, no more" make for more interesting haggling with in-game merchants than scroll bars. Shooters rely more on quick reflexes than quick tongues, but today's shooters typically feature a storyline – mysterious characters to interact with, computer terminals to access, and spoken passwords to get past locked doors. Besides, "Gimme the BFG" is more satisfying than 5 button presses while squinting at weapon icons. Sports games are ripe for ASR – what if you actually called out your plays from the huddle, or gave a verbal audible at the line of scrimmage? Told your infield to move in? Warned the pit crew to get ready? Not to mention the possibilities for embedding cheats and Easter eggs in a game that are unlocked with the proper voice command.

So what barrier remains? Simply put: demand. Consumers and reviewers alike evaluate games based on today's most popular features – the best graphics, the best audio, the best storylines – not the lack of a discontinuous innovation like speech. (Who would criticize Halo for not supporting weapon-swapping by voice?) In the end, it's up to us gamers to encourage technology such as speech recognition to take root. Reward the studios that innovate by talking up their titles within your gamer communities, and then vote with your wallet. The alternative? To be doomed to a gaming life of unimaginative sequels. ■ ■ ■

The views and opinions expressed on this page are strictly those of the author and not necessarily those of Game Informer Magazine or its staff.

If you work in the industry and would like to share your opinion, contact senior editor Matt Helgeson at matt@gameinformer.com.