

July 2001, San Francisco

Speechpainting -- An Architectural Medium

Sha Xin Wei

Recording

How did people talk together before they could record and play back sound? How many times have you said very naturally during an argument, I wish I had a tape recorder? Before we could record speech, how could people reach after and snare the dying echo of what they said, so they could take it somewhere else and re-perform it?

As so many commentators on technology have pointed out, prior to the era of mechanized recording and transmission of writing, the technology of memory of speech was the art of rhetoric, of public oratory as a persuasive art. But that was, in the classical era, quite an elaborate technology of memory, which took a good part of a person's adult life to master. Not exactly a Walkman.

But there's writing? What about writing? Sitting back in our contemplative recliners, we sometimes think of writing as a recording of speech. But is it really? If you've ever looked at a transcript of any unscripted conversation, like the chatter between people at a dinner table, or even the coded chatter between air traffic controllers and pilots, you realize that speech is a world away from writing. And anyway, in the heat of spoken conversation, who ever takes the time to write down a transcript of what they say? Imagine what that would do the dynamics of a conversation. There's quite a value to the very ephemerality and superposability and local quality of speech. Sometimes it's quite important that what you say disappears after awhile.

Free speech

From the point of view of an information systems analyst, telephone call centers' speech interface systems thicken the skin of an organization against the incursions of the client (the organization's alien). The strategy is to make the client do some of the classification work as the client burrows into the organization. By the time the client reaches a human representative, the client has been categorized or even satisfied by automated responses. Typically our civilian satisfaction with such systems is pretty low, but there are gradations in this domain of thickened organizational skin. What

makes the United Airlines' help center more appealing than many of its predecessors? Is it the displays of modest uncertainty? "Did I hear you say...?" Is it the sound of him clucking to himself in a distracted way? Underneath, the logic is just as well-specified as any "if you are dead push 911" telephone interface, but what we encounter is an example of an open dialogue, a dialogue scripted where only one speaker is known in advance. It's quite interesting as a open narrative form, but I'd like to move on draw attention to something else that's happening. What we hear here is an system that tries to mediate between the organization and whole open field of un-constrained, un-scripted speech. Free speech, so to speak.

What we can begin to see from this example is that the membrane between an organization -- a social organism -- and the world need not be defined solely by a pre-set menu but by a system that can register and respond to variable verbal impressions. But if the system has to do a lot of logical processing and respond to that speech, the system inevitably feels elephantine, because you're not speaking to an individual, you're speaking to an organizational policy.

Registration

Now, what if we take away most of this elephant and consider the membrane. What if the speech system is relieved of the burden of understanding what you said, but is asked just to >register< it, the way that a pond registers a pebble tossed onto it, or the way that a pavement registers the shadow of a child playing tag. Registration in this sense, is a world away from recording and surveillance, because we are not requiring the system to understand what we conventionally call the meaning of the speech. Surveillance rears its head when we are building a system that must faithfully record and later, perhaps with the aid of a human interpreter, understand what you said. But as you walk across a surface casting your shadow on the ground, ordinarily you're not concerned about sending a message with it, and ordinarily, the pavement doesn't record any shadows.

What if we could register speech in this same spirit on surfaces? What if we could make speech hang around long enough, tangible (say visible) enough to treat it as part of the built environment the way we've done with writing for centuries [Petrucci, Public Lettering]? In the modern city, billboards and highway signs occupy most of the reduced architecture of public lettering, but this certainly does not exhaust the domain of public writing, the obvious example being graffiti.

Socially, graffiti has the qualities of improvisation and spontaneity of shadow tag, but its physics is as persistent as billboard art -- it's a pain to erase and overwrite. Although its practitioners come from only certain classes of the polis, graffiti is certainly a ubiquitous aspect of urban public architecture.

Now let's pull it all together: the playful and violent spirit of spontaneous public writing, the physical persistence of writing, the lightness and ephemerality and locality of speech, the emerging technology that responds to speech, that recognizes speech and can make it tangible as visible marks. That technology is coming into shape as a confluence of speech recognition and display technologies ruggedized for outdoor conditions. It'd be overstating the case to call this a convergence since these technologies are still being developed along their intrinsic motivations. What we can anticipate with such confluences is to work with speech, public improvised speech, as part of the architecture of urban space.

Public Speech

Yes of course, speech is used to communicate, but what's communication? Given the heavy interpenetration of telecommunication with digital computer technology, the flip side of the question is, what are computers for? What Roy Harris [The Language Connection] and Anatol Holt [Organized Activity and Its Support By Computer] in their respective terms suggest is that we use such technology not so much to transmit coded messages-in-a-box from person A to person B, but as ways to coordinate human activity. In the life of organizations, coordination, and lack of it, occupies a lot of attention. But coordination also plays a role in private or disorganized life as well. The coordination of a gaggle of children playing tag, the coordination of friends arguing as they walk down the street, the coordination of people choosing which bench to sit on in a public park.

So there are ways to use language that may not have to do with explicitly "sending a message" but more to do with people coordinating with one another. Keeping that in mind, we may come up with some novel and powerful ways to embed speech technology in very broad social settings.

The Market is beginning to take notice of speech-enabled information systems [http://www.instat.com/abstracts/ct/2001/va0101sr_abs.htm], but the domains of application, for good information processing reasons, have been restricted to special

contexts like telephone call centers with extremely restricted vocabulary, or to a single user at a desk who has the incentive to spend a lot of time training his or her personal speech recognizer's user model. But what we can begin to see here are the opportunities that open up for design when we set aside the constraint of speech as a medium of information processing, and think of speech as an architectural medium for coordinating play in public space.

References

Frances Dyson, "A Philosophonics of Space: Sound, Futurity and the End of the World," [http://www.soundculture.org/words/dyson_philosophonics.html].

Andre Leroi-Gourhan, *Gesture and Speech*.

Armando Petrucci, *Public Lettering: Script, Power and Culture*.

Sha Xin Wei, Topological Media Lab, Hubbub Speechpainting Project, <http://titanium.lcc.gatech.edu/hubbub/>.

Speech recognition information
<http://www.tiac.net/users/rwilcox/speech.html>

"Speech Recognition to Explode into the Mainstream: Software Revenues to Exceed \$ 2 Billion in Next Two Years"
http://www.instat.com/pr/2001/va0101sr_pr.htm .

Giambattista Vico, (*La Scienza Nuova*) *New Science : Principles of the New Science Concerning the Common Nature of Nations*.