

Voice in Electronic Art: From Whistle to Speech Recognition

Martha Carrer Cruz Gabriel
Unicentro Belas Artes de São Paulo / University of São Paulo
martha@martha.com.br

Introduction

Voice interfaces, whether they are speech-only or multi-modal, are a fascinating subject. The human dream of talking to computers in a natural way is not new. Science fiction books and movies that live in our imagination present several examples of this aspiration, like: *Star Trek*, *Lost in Space*, *Star Wars*, and *Blade Runner*, among others.⁶

Until recently, talking to computers was in the realm of fiction — the web has been largely mute and deaf. However in the beginning of the 21st century talking to computers has become possible and easy due the enormous advances in speech synthesis and voice recognition technologies as well as the open standards adopted by the W3C (such as VoiceXML). The accuracy level reached by voice technologies now has allowed us to use them widely on the web.

However, talking to computers adds “ears” and “mouths” to the Internet organism, changing the way we interact with it, bringing new possibilities and new challenges as well. We must face the increasing complexity that voice interfaces bring to the web while we also open new channels for digital inclusion, provide more accessibility and increase mobility through voice. All these things affect the human role inside the high-tech social structure we live in, at once causing excitement and fear.

In this context, we believe that artworks help people to understand and experience the new emergent technological world that surrounds us, where convergence and hybridization have become ubiquitous and easy, and “to talk to computers” is going to become common.

Electronic artworks & voice

Several electronic artworks have used voice in many ways — from whistles and blows to the state of the art in using speech synthesis and recognition technologies. Although the first experience with voice technologies was in the 18th century, only since the beginning of the 20th century has their commercial use really started and it was not until the end of the century, in the 1980s, that the first electronic art experiments with voice technologies were developed. Since then many electronic artworks have used voice in several creative forms. From the first voice experimentations in electronic art to the use of speech synthesis and voice recognition on the web — as we have nowadays — it has been a long journey.

According to Gabriel,⁴ adding some recent works, the following artworks draw a brief panorama of interesting relationships between electronic art and voice, contextualizing the artistic development from the beginning (in the 1980s) up to now.

1986 - Synthetic Speech Theatre, by S. Wilson	2000 - Riding the Net, by C. Sommerer and L. Mignonneau
1989 - Barbie Liberation, by Ron Kuivila	2000 - Talk Nice, by Elizabeth Vander Zaag
1990 - Talking Machine, by Martin Riches	2001 - Living Room, by C. Sommerer and L. Mignonneau
1991 - Inquiry Theater, by Stephen Wilson	2002 - RE:MARK, by G. Levin and Z. Lieberman
1994 - Oh toi qui vis la-bas, by Don Ritter	2003 - Alert, by Barbara Musil
1994 - Speech Sculptures, by Bruce Cannon	2003 - Messa di Voce, by Golan Levin et al
1995 - I Have Never Read the Bible, by Jim Campbell	2003 - Summoned Voices, by Iain Mott
1996 - Le Pissenlit, by E. Couchot and M. Bret	2003 - Universal Whistling Machine, by Marc Böhlen and JT Rinker
1996 - Orpheus, by Ken Feingold	2004 - IP Poetry, by Gustavo Romano
1999 - Universal Translator, by David Rokeby	2004 - Voice Mosaic, by Martha Gabriel
2000 - Giver of Names, by David Rokeby	2005 - Organum, by Greg Niemeyer et al
2000 - Huge Harry, by A. Elsenaar, R. Scha	2005 - Tampopo, by Kentaro Yamada
2000 - Millennium Venus, by Sharon Grace	2007 - NetAura, by Giselle Beiguelman et al
2000 - n-Cha(n)t, by David Rokeby	
2000 - netsong, by Amy Alexader	

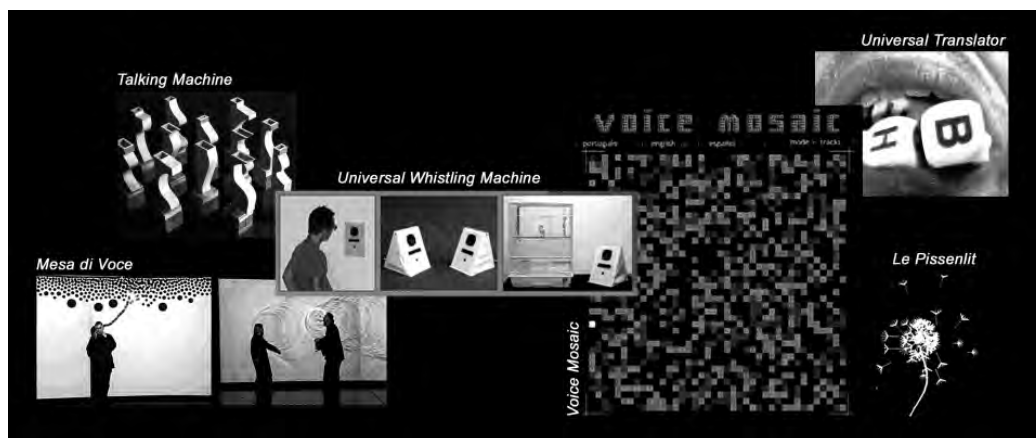


Figure 1: Pictures of artworks related to voice.

An interesting analysis of these works happens when we focus on the way they use the voice. In this sense we will point out some artworks, as follows (Figure 1).

Some artworks that implement **non-verbal voice as an input** are: *Le Pissenlit*³ — uses the blow in a digital image (a dandelion flower) that dissolves; *Universal Whistling Machine*² — deals with whistle synthesis and recognition; *Talk Nice*¹¹ — speech recognition focused on pitch; and *Universal Translator*⁸ — speech recognition for phoneme analysis.

As examples of artworks that employ the **verbal dimension of voice with speech synthesis**, we could mention: *netsong*¹ — uses speech synthesis to create a song based on the web links provided by a search on the web; and *Talking Machine*⁷ — the speech synthesis is created by a physical system that imitates the human vocal system. We can observe some other technological trends influencing the artworks, as it happens with *netsong*: besides probably being the first artwork to use speech synthesis on the web it is deeply related with search engines — one of the most important and influential interfaces of our time.

Artworks that employ the **verbal dimension of voice with speech recognition** start with *Inquire Theater*,¹⁰ in 1991. In this work, speech recognition is interpreted by a virtual navigation system in San Francisco. Another interesting artwork based on an algorithm that analyses speech transforming it into images is *Mesa di Voce*,⁵ which is concerned with the poetic implications of making the human voice visible. We can also cite *Voice Mosaic*⁹ as an artwork that uses speech recognition, however, instead of being a installation, the work happens on the web integrated in real-time with the telephone.

Conclusion

By analyzing the use of voice in electronic art we have seen that speech recognition is less frequent in the works than other voice technologies. Few artworks

have explored phoneme-less inputs such as blows and whistles so far. Although whistling is a primitive form of communication, only in 2003 has an artwork explored it as an alternative way of voice-input and human-computer interaction. As seen above, before 2004 vocal-inputs in artworks were restricted to installations, which limited their wide. Now, the web and telephone have been the latest step these works have reached by using the state of the art in voice technologies.

From now on we think that it will be possible to provide wider and deeper experimentation with voice interfaces due to the available technologies integrating the web and telephone. We expect it will probably allow us all to break frontiers and go further in artistic/human possibilities and developments.

- 1 Alexander, A. 2000. *netsong*. <http://netsong.org/>
- 2 Böhlen, M. and Rinker, J.T. 2003. *Universal Whistling Machine*. http://www.realtechsupport.org/new_works/uwm.html
- 3 Couchot, E. and Michel, B. 1996. *Le Pissenlit*. <http://www.artmag.com/techno/landowsky/projet.html>
- 4 Gabriel, M. 2006. *Interfaces de Voz em Ambientes Hipermediáticos*. Master's Degree Dissertation, University of São Paulo.
- 5 Levin, G. et al. 2003. *Mesa di Voce*. <http://tmema.org/messa/messa.html>
- 6 Perkowitz, S. 2004. *Digital People: From Bionic Humans to Androids*. Washington: Joseph Henry Press.
- 7 Riches, M. 1990. *Talking Machine*. <http://www.floraberlin.de/soundbag/index53.html>
- 8 Rockeby, D. 1999. *Universal Translator*. <http://homepage.mac.com/davidrokey/trans.html>
- 9 Gabriel, M. 2004. *Voice Mosaic*. <http://www.voicemosaic.com.br/>
- 10 Wilson, S. 1991. *Inquire Theater*. <http://userwww.sfsu.edu/~swilson/>
- 11 Zaag, E. V. 2000. *Talk Nice*. <http://www.canadacouncil.ca/news/releases/2000/un127241852363281250.htm>

Life and Art in Second Life

Martha Carrer Cruz Gabriel
Unicentro Belas Artes de São Paulo / University of São Paulo
martha@martha.com.br

Introduction

One of the many virtual worlds inspired by the cyberpunk literature movement, Second Life has attracted global attention since 2006 and counts today with millions of residents. Aligned with the Web 2.0 trends and considered by many as the best digital life at the present moment, the SL Metaverse gives flow to cybrid processes. In despite of the fact of not being a complete novelty — since 3D MUVES (Multi User Virtual Environments) and social networks have existed for more than one decade on the web — Second Life brings several new questionings and possible influences in language and personal relationships that cannot be despised.

The objective of this paper is to briefly explore the new possibilities for expression and interaction provided by Second Life and other virtual worlds on the web, especially in art, and to bring reflections about their probable influence in the future navigation interfaces on the web. Some selected artworks in Second Life will be pointed out to illustrate the paper.

Life in Second Life

Living and expressing the self in a world where one can either follow or break the laws of Physics is a power never experienced before by the flesh and bones man. However, in order to allow an avatar live that freedom in Second Life (SL), his/her physical body needs to be seated in front of a computer in the First Life (FL). Everything seems to be possible and extraordinary, but nothing really exists in SL if there will not be a physical human body behind it, feeding and supporting the brain in FL. We could say it seems that SL provides a platform that gives flow to *The Matrix* (Wachowski, 1999) paradox.

It is still too early to conclude something deeper about SL or where this road leads. However, we are already feeling its effects and influences: SL is the 10th most popular virtual world, including online games (GigaOM, 2007). At any time there is an average of 30 to 40

thousand people connected and US\$ 1 million circulates daily *inworld*.

SL is the Stephenson's Metaverse (1993) first embodiment: totally immersive 3D environment where humans interact (as avatars) with each other (socially and economically) and with software agents of the cyberspace, which uses the physical world metaphor but without the physical limitations. SL is not a complete novelty since in reality it is a mix, a co-existence of everything that already exist — games, economy, social networks, etc. However, we believe that it initiates a new engaging way of interaction.

There are some important aspects to be considered in the life in SL. First, the balance between breaking and following the laws of Physics — it makes the environment much richer than if it were just one or the other, conveying the surprise effect constantly. Second, the simulation functionalities give a new dimension to language, like: flying, teleporting and the several new view functions. Third, SL is like a playground environment — it allows social interaction and toys. The more toys a playground has, the more attractive it becomes to different audiences. Fourth, the essence of SL is social. The most popular places are those that involve the social aspect — meetings, classes, shows, parties, etc. Fifth, cybrid (Beiguelman, 2004) processes are potentialized by SL: the life of a person in SL is extension of his/her life in FL and they complete each other, widening the field of action of the being. We are increasingly living a 'Digital Life' simultaneously with a 'Life Digital' (Iskold, 2007). A very interesting art exhibition that explores the convergence — through cyberspace — of real and synthetic places is the Mixed Realities (Turbulence, 2008). Another example is the company Fabjectory (2007), which creates statues of objects and avatars that exist in SL, transforming them in things in FL.

Finally, we believe that the key-aspect of SL is that it gives a bigger power to the individual than he/she has in FL: a) power over himself/herself (appearance,