

Messa di Voce

A Performance of Visualized Speech and Song

Tmema / Jaap Blonk / Joan La Barbara

Messa di Voce is a new concert performance in which the speech, shouts, barks and songs produced by a duet of experimental vocalists are augmented in real-time by custom interactive visualization software. The forty-minute performance touches on themes of abstract communication, synaesthetic relationships, cartoon language and writing and scoring systems, within the context of a sophisticated, playful and virtuosic audiovisual narrative.

Messa di Voce lies at an intersection of human and technological performance extremes, melding the unpredictable spontaneity and extended vocal techniques of two masterful composers/improvisers with the latest in computer vision and speech recognition technologies. Utterly wordless, yet profoundly verbal, *Messa di Voce* is designed to provoke questions about the meaning and effects of speech sounds, speech acts and the immersive environment of our language.

The core technologies which scaffold the performance are computer vision and speech recognition. A computer incorporates a video camera in order to locate and track the positions of the vocalists; it also analyses the speech and song coming from their wireless microphones. In response, the computer projects various kinds of visualizations on a projection screen behind the performers. These visualizations are synthesized in direct response to the sounds spoken and sung by the performers. With the help of the computer-vision tracking system, these visualizations are projected in such a way that they appear to emerge directly from the performers' heads or mouths. Where and when circumstances permit, stereographic projections are used to enhance the illusion that the vocal figurations are hovering in the same space as the performers. The end result presents the fiction that speech can be truly visible.

Messa di Voce emerges from a pair of commissions created by Golan Levin and Zachary Lieberman as Artists-in-Residence at the Ars Electronica Futurelab. During July and August 2002, we produced two new installations for the Ars Electronica Museum, called *Re:mark* and *The Hidden Worlds of Noise and Voice*. These artworks combined state-of-the-art speech analysis technologies (such as vocal stress estimation, pitch tracking, and realtime phoneme recognition) with camera-based motion tracking, augmented reality displays, and other sensing techniques, in order to allow museum-goers to engage vocally and viscerally with different sorts of highly reactive speech visualizations. Participants in the Hidden Worlds exhibit, for example, are able



LETTERS: The singer speaks short phonemes. The computer recognizes these phonemes and places large letters in the screen. The letters crowd around the singer, but never touch him.



to “see” each others’ voices, made visible (through stereographic VR glasses) in the form of animated graphic figurations that appear to emerge from the participants’ mouths. For the *Messa di Voce* performance, we have developed these core technologies into a “professional” version, customised for a duet of vocalists well known for their experimental repertoires and extended vocal techniques: Joan La Barbara (USA) and Jaap Blonk (Holland). In *Messa di Voce*, Joan and Jaap engage in abstract dialogues with themselves and each other. Visual representations of their speech and song are computed in real-time, appropriately positioned around their heads and bodies by a computer-vision algorithm, and projected onto a series of spatially contiguous video screens. The visual representations are “aware” of the locations of the performers, and also aware of the other visual representations. In this way, it is possible for the performers/interactants to conduct a “game” with their speech-gestures. Over the course of the performance, increasingly sophisticated forms of speech recognition are employed, and used to develop an overall narrative from purely formal abstraction to symbolic representations and transcriptions.

Tmema is: Golan Levin and Zachary Lieberman.

Commissioned by Ars Electronica with the support of SAP, la Fondation Daniel Langlois pour l’art, la science et la technologie, Eyebeam Atelier, The Rockefeller MAP Fund, and the Arts Council England.



FLOWER: This idea is a really silly one-liner. After struggling and groaning and grunting for a while, a flower is abruptly released from the singer’s head when he creates a loud “POP” sound.