messa di voce
tmema * blonk * la barbara

A performance & installation for voice and interactive media
in which the speech, shouts and songs produced by two
vocalists are radically augmented in real-time by custom
interactive visualization and sonification software

Interactive Software: **Tmema (Golan Levin + Zachary Lieberman)**
Composition/Performance: **Jaap Blonk + Joan La Barbara**
Web site: **http://www.tmema.org/messa**

Contents.

Overview.

Messa di Voce (Ital., "placing the voice") is a concert performance in which the speech, shouts and songs produced by two experimental vocalists are radically augmented in real-time by custom interactive visualization software. The 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. In addition to the performance itself, which features the composer/vocalists Joan La Barbara and Jaap Blonk, a separate installation version of *Messa di Voce* makes select software modules available for public play and exploration.

Our software transforms every vocal nuance into correspondingly complex, subtly differentiated and highly expressive graphics. These visuals not only depict the singers' voices, but also serve as controls for their acoustic playback. While the voice-generated graphics thus become an instrument which the singers can perform, body-based manipulations of these graphics additionally replay the sounds of the singers' voices — thus creating a cycle of interaction that fully integrates the singers into an ambience consisting of sound, virtual objects and real-time processing.

Messa di Voce lies at an intersection of human and technological performance extremes, melding the unpredictable spontaneity and extended vocal techniques of two master composer-improvisers with the latest in computer vision and speech analysis 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 language.
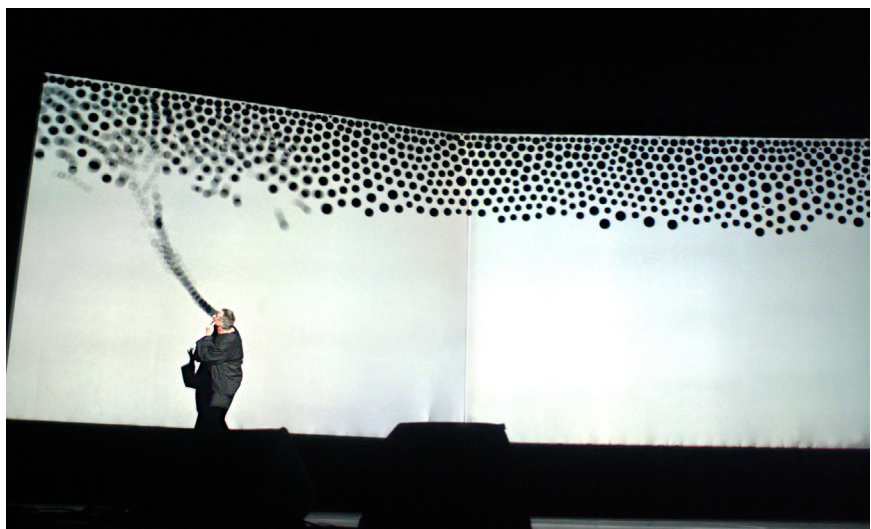
Background.

Messa di Voce presents the fiction that the voice can be seen. Like our previous augmented-reality artworks *RE:MARK* and *Hidden Worlds of Noise and Voice*, it is concerned with the poetic implications of this proposition, in an attempt to address the question: "if we could see our speech, what might it look like?"
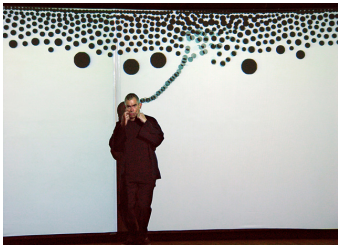
In *Messa di Voce*, a head-tracking system permits in-situ visualizations of speech. Graphic forms are positioned to be coincident with the sound's point of origin: the performer's mouth.

In *Messa di Voce*, the core technology which makes this possible is a custom software system which integrates real-time computer vision and speech analysis algorithms. Specifically, a computer uses a video camera in order to track the locations of the performers' heads. This computer also analyses the audio signals coming from the performers' microphones. In response, the computer displays various kinds of visualizations on a projection screen behind the performers; these visualizations are synthesized in ways which are tightly coupled to the sounds spoken and sung by the performers. Owing to the head-tracking system, moreover, these visualizations can be projected such that they appear to emerge directly from the performers' mouths. In some of the visualizations, the projected graphical elements not only represent vocal sounds visually, but also serve (bidirectionally) as a *playable interactive interface* by which the sounds they depict can be re-triggered by the performers.

Our group's interest in phonesthesia, or phonetic symbolism, is at the heart of the *Messa di Voce* project. According to this idea, the sounds of words tend to reflect, to some extent, associated connotations from other perceptual domains such as shape or texture. A classic illustration of the phonesthetic principle can be found in Wolfgang Köhler's pioneering psychology experiment from 1927, in which he asked subjects, "which of the figures below represents the sound *maluma*, and which one represents the sound *takete*?" Nearly all viewers respond with the same answer—suggesting rich research opportunities for both artmaking and cognitive psychology alike.
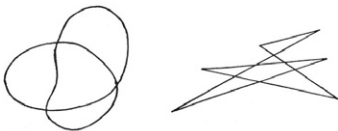
Köhler's phonesthesia experiment, 1927

Messa di Voce brings together findings from phonesthesia research with our group's shared interests in abstract language, extended vocal techniques, interactive systems, and live audiovisual performance. In our concert, these ideas take shape in a series of twelve brief vignettes which explore different symbolic, tactile and audiovisual aspects of phonesthetic relationships. The entire *Messa di Voce* performance generally runs 30 to 40 minutes in length.

Exhibition History.

Messa di Voce has been exhibited both as a performance, featuring the collaboration of vocalist-composers Jaap Blonk and Joan La Barbara, and as a standalone interactive installation, in which select software modules are made available for public play. The information below is current as of April 2004.

Performance, 29 November 2003:
Ultrasound Festival, Huddersfield, England.
Huddersfield Art Gallery, Princess Alexandra Way.
Projection on two screens: 8m*3m.

Performance, 7-8 November 2003:
London Institute for Contemporary Art (ICA), England.
ICA Main Theater, The Mall, London.
Projection on two screens: 8m*3m.

Installation, 12 October - 13 November 2003:
Beta Launch '03: Eyebeam Artists-in-Residence.
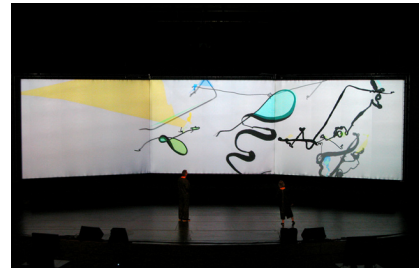Eyebeam Atelier Gallery Space, Chelsea, NYC.
Projection on one screen: 4m*3m.

Performance, 7 September 2003 [Premiere]:
Ars Electronica Festival, Linz, Austria.
Grosser Saal, Brucknerhaus Linz.
Stereo-projection on three screens: 18m*4m.
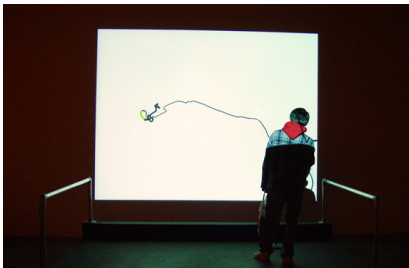
Messa di Voce, installation on one projection at Eyebeam.
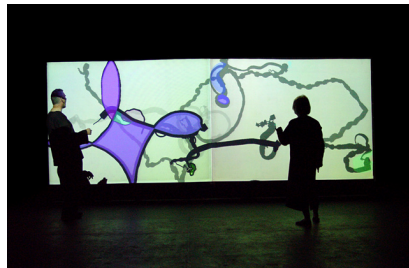
Messa di Voce, concert on two projections at ICA London.

Messa di Voce, concert on three projections at Ars Electronica.

Messa di Voce has also been tentatively scheduled to appear at:
London International Poetry Festival, England, 26 October 2004, and Stockholm New Music Festival, Sweden, 26 February 2005.

Press and Reviews.

"A Logiciel Ouvert: En Autriche, Ars Electronica, baromètre annuel du numérique, marque un retour aux sources." *Libération*, p.31. Paris, 11 September 2003.

Bersten, Rosanne. "Singing a stream of floating images." *AustralianIT*, 28 October, 2003.

Bliss, Abi. "Ultrasound." *Leeds Guide*, 19 November 2003.

Feilmayr, Ingrid. "Faszinierende 'Zaubereien' mit Stimme und Technik." *Neues Volksblatt*, Vienna, 9 September 2003.

Garulli, Lavinia. "Ultrasound." *FlashArt News*, Fall 2003.

Judmayer, Irene. "Magere Stimm-Bilder: Blonk/La Barbara im Bruckner-haus." *OÖ Nachrichten*, Graz, 9 September 2003.

Levin, Golan and Lieberman, Zachary. "*In-Situ* Speech Visualization in Real-Time Interactive Installation and Performance." *Proceedings of The 3rd International Symposium on Non-Photorealistic Animation and Rendering*, June 7-9 2004, Annecy, France.

"Messa di Voce." Press release, *Ars Electronica Festival*, Linz, Austria, 20 June 2003.

Noonan, Mickey. "Q&A with Golan Levin: Messa di Voce." *Huddersfield Metro*, 28 November 2003, p. 18.

Richards, Kate. "Ars Electronica: Interface Futures." *Realtime Arts*, #58, September 2003.

Somajni, Chiara. "Dipingere con la voce." *Il Sole 24 Ore*, 14 September 2003.

Turco, Marina. "Tutti nella rete digitale Ars Electronica a Linz, tante performance e un solo 'code'." *Il Manifesto,* 13 September 2003.

Software Modules and Score.

The *Messa di Voce* performance is divided into 12 audiovisual vignettes, which create a narrative progression from pure sound, to speech, to song.

1. Brightness/Balance. *Messa di Voce* begins with a brief, *Genesis*-like exposition of its own fundamental materials: light and breath. The performers, offstage, breathe illumination and animation into the rectangle of the electric canvas. As Joan breathes, the canvas becomes brighter; when Jaap breathes, it begins to rotate.
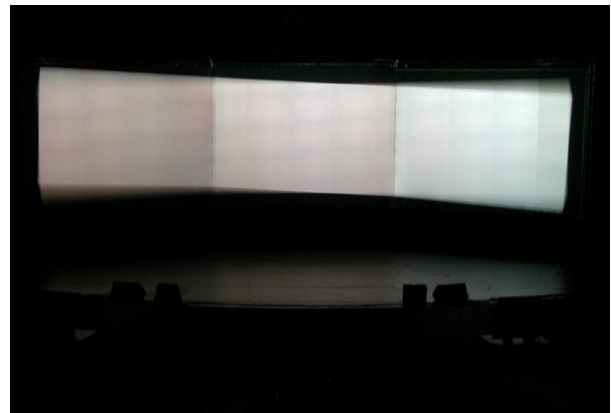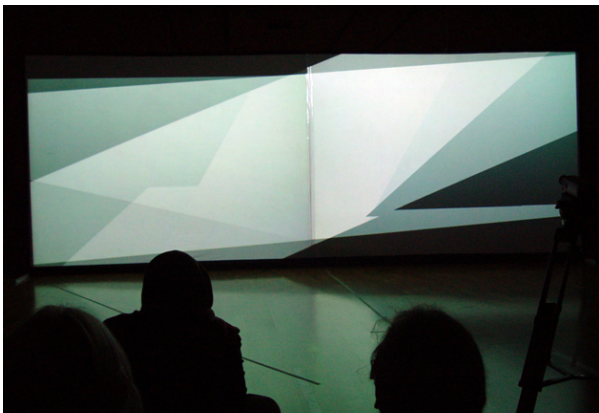
"hhhhhh" cloud


"sssssss" cloud

2. Clouds. The performers enter the stage and continue their elaboration of elementary wind, air and breath sounds. Each respiration is recorded, represented, and replayed by an animated cloud whose visual texture evolves with the timbre of the sound it portrays. Bright frequencies in the breathing sounds are mapped to higher spatial frequencies in the clouds, while duller sounds produce smoother clouds. After a time, the hall and screen are filled with a swirling soup of clouds and breath sounds.
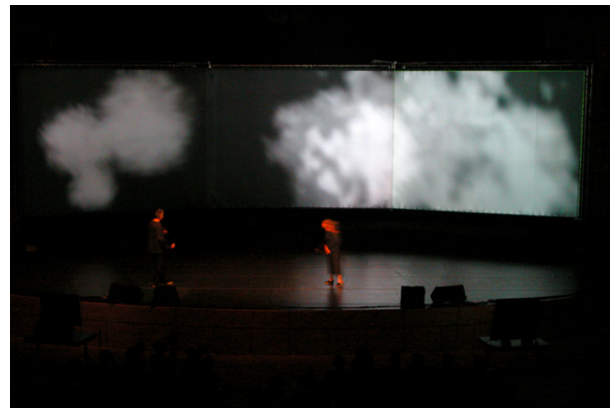
3. Bodystamp. The performers obtain concrete forms for the first time, drawing our awareness to the relationship between a sound and the body of its maker. Each time one of the performers vocalizes, their speech sound—as well as the moving silhouette of their body—are simultaneously recorded. These sounds and their corresponding body-animations are then synchronously replayed in periodic cycles. The result is an animated chorus of previous selves.

In this software module, the performers' vocal sounds and body movements are simultaneously recorded whenever they speak. Here, Jaap is running and shouting at the same moment.
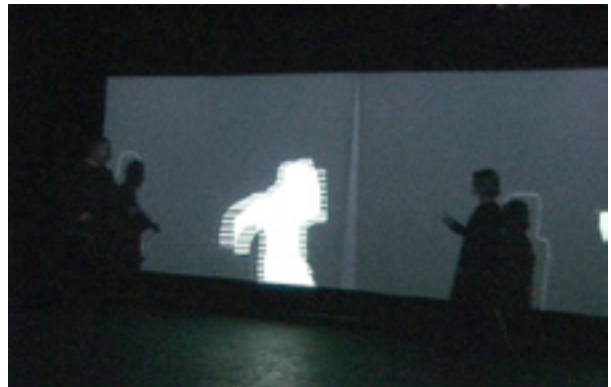
Here, Jaap's recorded animation plays back, accompanied by its corresponding sound. The animation and sound loop periodically according to a period set by the length of his audiovisual gesture.

4. Bounce (Jaap's Solo). A man enters an empty white void. He emits a stream of bubbles by making a special cheek-flapping sound. As his sounds grow more vigorous, his bubbles fill up the screen. But the resulting cloud of jostling sound-bubbles is unstable. Turning to admire his work, his cloud bursts—raining bubbles that, when they fall onto him or crash to the ground below, replay recordings of his cheeky sounds. He tries to contain the noisy torrent, but, failing this, storms off in distress.

Jaap makes a cheek-flapping sound, producing a bubble for each small mouth-sound. Bright sounds make larger bubbles.

He gradually fills the upper part of the screen with bubbles. Meanwhile, each small mouth-sound is being recorded.


He turns to 'admire' the scene he has created.

The bubbles begin to fall, bouncing when they collide with Jaap or the floor. When they bounce, they trigger the replay of the individual mouth-sound that created them.


Jaap tries to catch or collect the bubbles. The bubbles cascade in a noisy torrent, triggering hundreds of simultaneous mouth-sounds when they bounce.


Jaap aggravatedly clears the screen by kicking the bubbles.
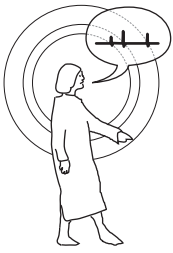
5. Ripple. Tension is dispelled as our attention turns to a pair of chirping, clicking wetland creatures, whose curious chatter perturbs their watery environs. Every little sound causes a unique ripple to emerge from its creator's mouth. Wakes emerge in the liquid surface when a performer moves and sings simultaneously.
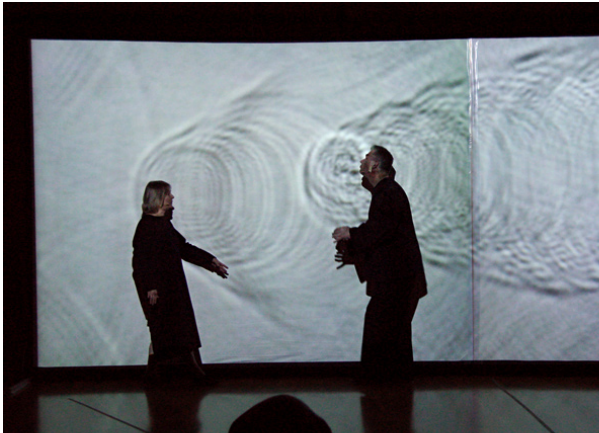
6. Insect Nature Show. The performers adopt different roles: Joan, as a peculiar kind of animal; and Jaap, as narrator or "nature show host", who discourses (in an abstract language) about the remarkable qualities of this creature. This section of the performance marks a transition in focus from pure vocal sounds, to more speechlike forms.
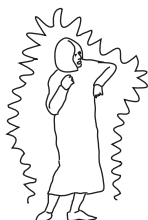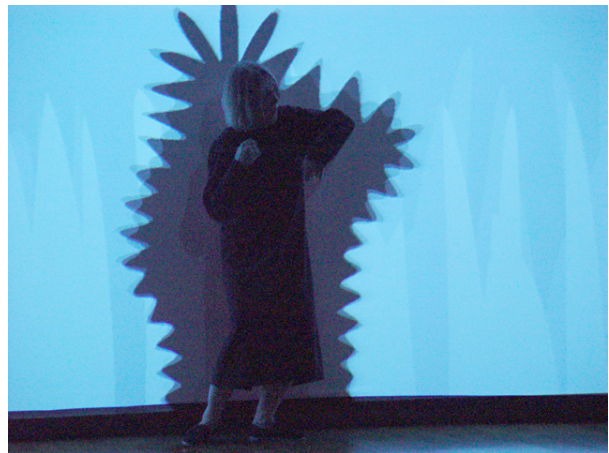

In this module, Joan's silhouette is augmented by an interactive projected shape. This black shape is a larger, bouncier version of her own outline. The edge of this shape, moreover, also changes in response to her speech: it develops wiggles which reflect the timbre and pitch of her voice. At a key point in this section of the concert, Joan makes a special hissing sound which causes her shape to release a stream of bubbles.
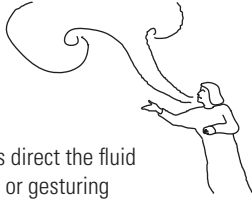
Performers direct the fluid by leaning or gesturing while speaking or singing.

7. Fluid. The players return to a primordial liquid world. In this hushed dialogue of soft, lulling sounds, a glowing fluid or plasma appears to emerge from the performers' mouths when they speak and sing. The performers direct the flow and movement of the fluid with their bodies, passing it back and forth. (The color of the fluid is related to the vowel character of their sounds; for example, brighter vowels like '*ee*' produce greenish and yellow chromas.)
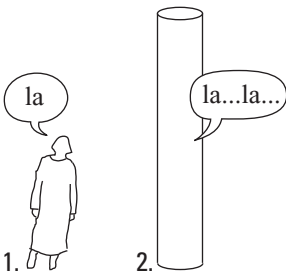
8. Rothko (Joan's Solo). Joan builds a layered atmosphere of soaring vocal tones. She sings long melodies, which are recorded in real-time, and then loop according to their own natural periods. Each melody is represented by a colored column, which marks the locations where Joan created it. Subtle changes in the color and position of the columns reflect the timbre, pitch and stereolocation of Joan's melodies. If Joan walks or moves while recording a column, the column will replay the positions of her movements.

9. Stripe. Jaap joins Joan, and the two sing a slowly-evolving duet built from simple pure tones and subtle dissonances. Their pitches and timbres are visualized in the softly-changing stripes behind them. The separation of and colors of the stripes are linked to the singer's pitches, making their small dissonances especially noticeable.
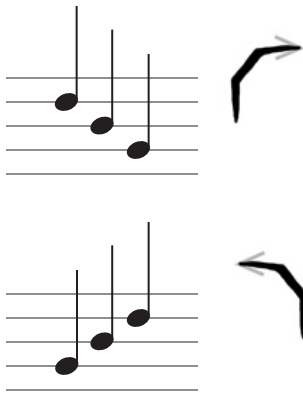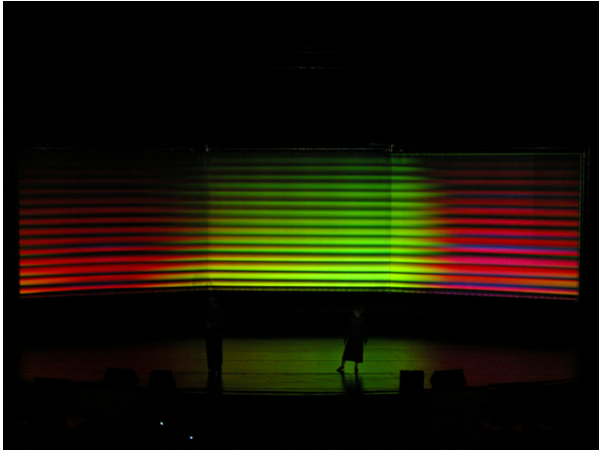
10. Pitchpaint. The performers take their most direct control of the canvas, painting bold gestures by singing. In this section, descending pitches create lines which curl clockwise, while rising notes produce lines which curl counter-clockwise, and unchanging tones produce straight lines. Because an octave is precisely mapped to 180° degrees of arc, the singers can create attractive regular polygons by singing common scales and arpeggios. The performers erase their marks by making the sound, "Ssh!"

Marks originate at the position of the singer's head. Stroke thickness is governed by the singer's loudness; the color of closed regions is linked to vowel quality. From the hushed, simple tones of the previous section, the performers develop an expanded vocabulary of quickly-changing glissandos, melodic fragments, and abstract speech-songs.
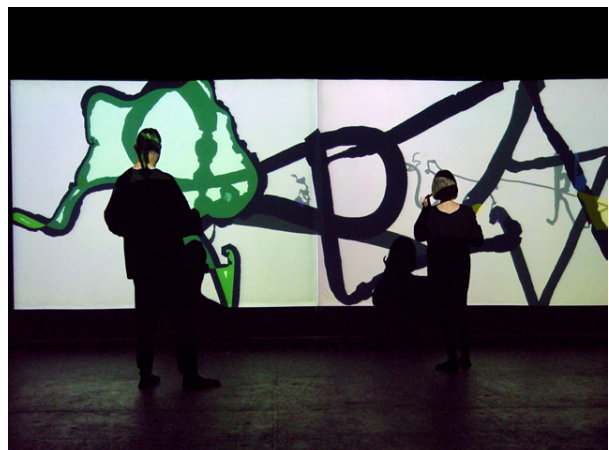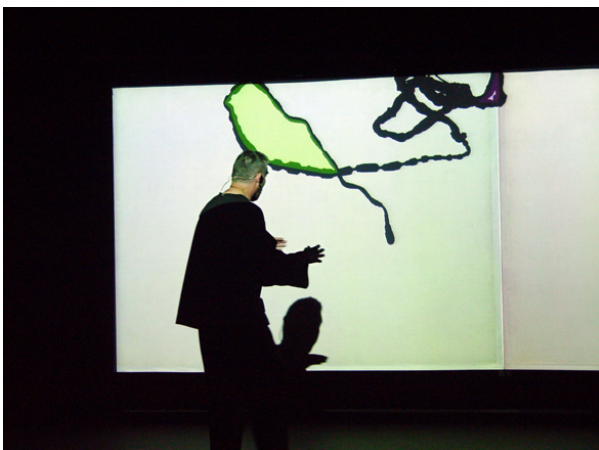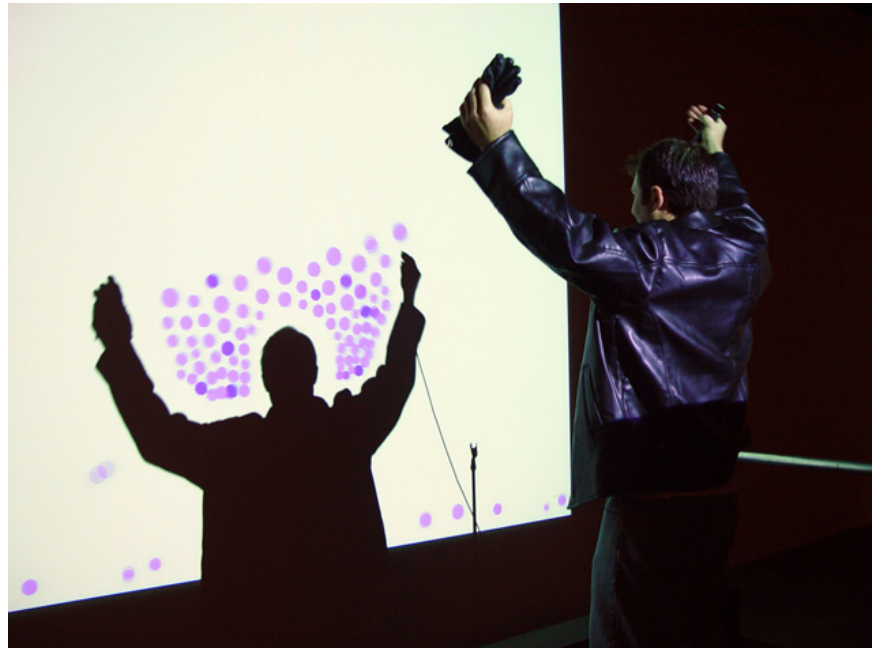
11. Shush-Fade. The performers clear the screen, in the previous section, by slowly saying "ssh". As they continue to shush the screen, it gradually dithers to black, and the performance is finished.
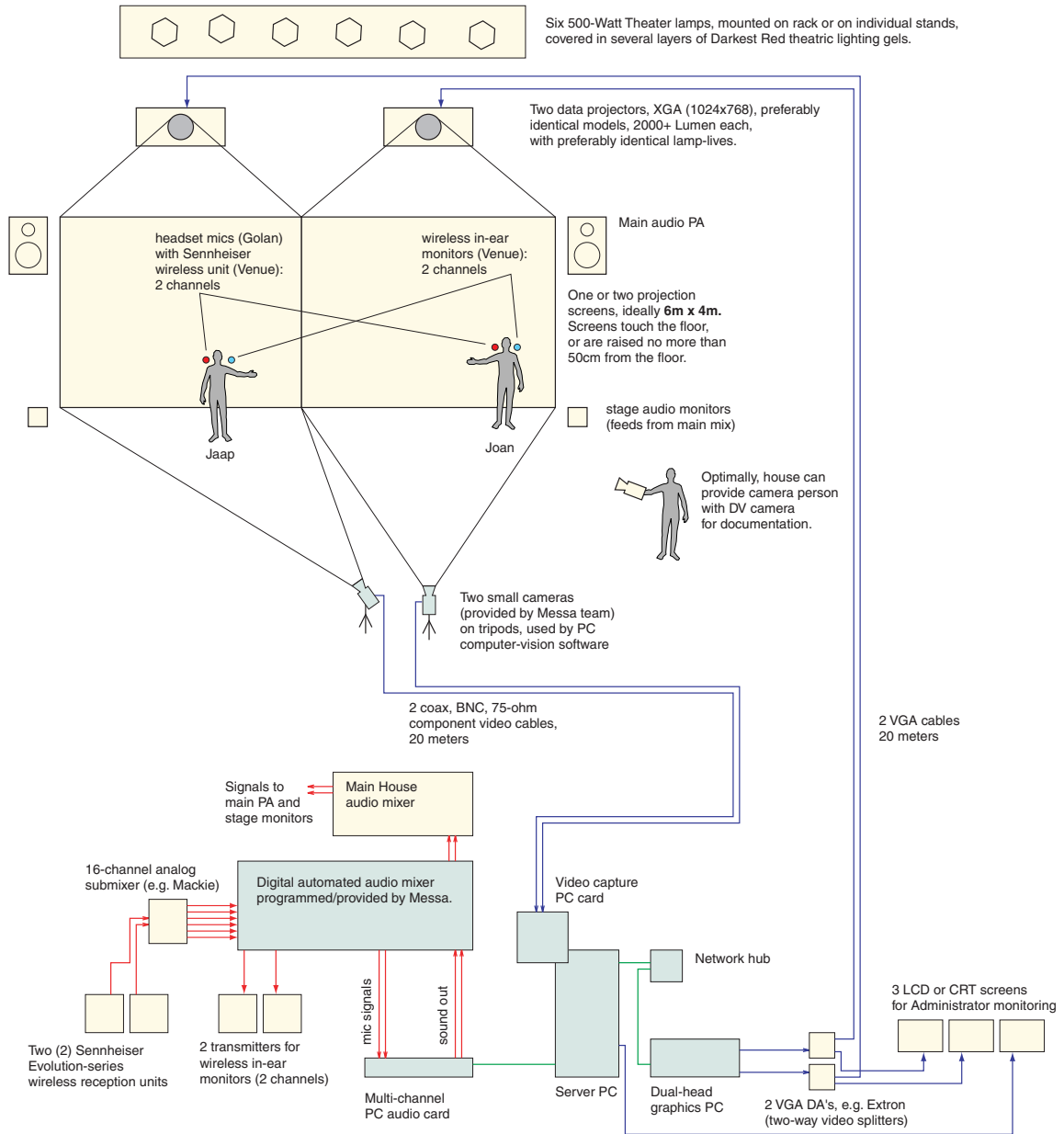

12. Credits. The creators of *Messa di Voce* announce each other's names. As they do, cartoon word-balloons containing their names appear above their heads.

Installation version. *Messa di Voce* has also been presented in the form of an interactive installation for two participants. Five or six of the software modules from the performance, such as *Pitchpaint, Fluid, Insect,* or *Bounce* (pictured below, at the Eyebeam Gallery in New York City) are presented for public play. Visitors easily find and understand how to use the microphones with the projected graphics.
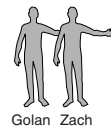
Performance Diagram.

Six 500-Watt Theater lamps, mounted on rack or on individual stands, covered in several layers of Darkest Red theatric lighting gels.

Two data projectors, XGA (1024x768), preferably identical models, 2000+ Lumen each, with preferably identical lamp-lives.

Main audio PA

headset mics (Golan) with Sennheiser wireless unit (Venue): 2 channels

wireless in-ear monitors (Venue): 2 channels

One or two projection screens, ideally **6m x 4m.** Screens touch the floor, or are raised no more than 50cm from the floor.

stage audio monitors (feeds from main mix)

Jaap

Joan

Optimally, house can provide camera person with DV camera for documentation.

Two small cameras (provided by Messa team) on tripods, used by PC computer-vision software

2 coax, BNC, 75-ohm component video cables, 20 meters

2 VGA cables 20 meters

Signals to main PA and stage monitors

Main House audio mixer

16-channel analog submixer (e.g. Mackie)

Digital automated audio mixer programmed/provided by Messa.

Video capture PC card

Network hub

3 LCD or CRT screens for Administrator monitoring

mic signals

sound out

Two (2) Sennheiser Evolution-series wireless reception units

2 transmitters for wireless in-ear monitors (2 channels)

Multi-channel PC audio card

Server PC

Dual-head graphics PC

2 VGA DA's, e.g. Extron (two-way video splitters)

Golan   Zach

NOTES:
Equipment in PALE GREEN is provided by Messa Team
Equipment in PALE YELLOW is provided by Venue:
House sound system, main mixer, 2 stage monitors
2 XGA Data projectors + projection screens
2 Sennheiser Evolution Wireless sets (RX+TX) (- we have mics already fitted for Sennh.)
2 channels of wireless in-ear monitoring (RX+TX), including earsets
1 sixteen-channel analog submixer
2 VGA Distribution Amplifiers (two-way video splitters)
4 LCD or CRT screens
3 VGA cables, 20+meters
2 coax BNC 75-Ohm video cables, 20+meters

Acknowledgements.

Sponsors.

Artist Biographies.

Golan Levin is an artist, engineer, and composer interested in developing artifacts and events which explore supple new modes of interactive expression. His work focuses on the design of systems for the creation, manipulation and performance of simultaneous image and sound, as part of a more general inquiry into non-verbal communications protocols in cybernetic systems. Levin is Assistant Professor of Electronic Time-Based Art at Carnegie Mellon University, Pittsburgh.

Zachary Lieberman is an artist, engineer, and educator whose work is focused on exploring the creative and human uses of technology. He produces installations, on-line works and concerts concerned with the themes of kinetic and gestural performance, interactive imaging and sound synthesis. Lieberman lives and works in New York City, where he teaches courses in audiovisual synthesis and creative image processing at Parsons School of Design.
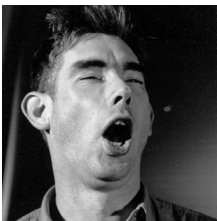
Jaap Blonk is a composer, voice performer and sound poet. In reciting poetry, especially the works of Antonin Artaud and Kurt Schwitters, he discovered the directness and flexibility of vocal utterance, whether using meaningful words or not. At present, he has developed into a prolific writer/composer and a specialist in the performance of sound poetry, supported by a powerful stage presence and an almost childlike freedom in improvisation.

Joan La Barbara has been called one of the great vocal virtuosas of our time. A pioneer in the field of contemporary and sound art, she has developed a unique vocabulary of experimental and extended vocal techniques, including multiphonics, circular singing, ululation and glottal clicks. In addition to her own compositions, she has premiered landmark compositions written for her by noted American composers, including Robert Ashley, John Cage, Philip Glass, Steve Reich, Morton Subotnick and James Tenney.

Related Work.

We believe it is important to recognize prior art and contemporary context in our documentations. Several related projects form important conceptual and aesthetic precursors to this one. These include:

Pavlátová, "Reci Reci Reci"

Michaela Pavlátová, "Reci Reci Reci" ("Words Words Words") [1991]. A delightful hand-drawn animation in which the conversations of café patrons are abstracted and visualized. *http://www.kratkyfilm.com/catalogue/html/358.htm*

Grady Klein, "Afterbabble" [2002]. An animation and children's book which speculates about what happens to our words after we speak them. *http://silvertone.princeton.edu/~grady/pages/illustrationPages/afterbabble/*
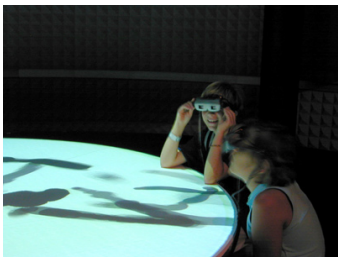
Levin and Lieberman, "Hidden Worlds"

Golan Levin and Zachary Lieberman with the Ars Electronica Futurelab, "RE:MARK" and "Hidden Worlds of Noise and Voice" [2002]. Our own prior work, a pair of interactive installations in which participants are able to "see" their voices. *http://www.flong.com/remark*

Toshio Iwai and Ryuichi Sakamoto, "Music Plays Images X Images Play Music" [1997]. A masterful audiovisual piano/graphics performance, which explores a wide variety of possible relationships between physical action and virtual representation. *http://www.aec.at/en/archives/prix_archive/prix_projekt.asp?iProjectID=2494*

Iwai and Miyata, "Sho of Light"

Toshio Iwai and Mayumi Miyata, "Sho of Light" [2003]. An audiovisual performance in which real-time computer graphic projections are used to augment Ms. Miyata, a virtuosa performer of Japan's traditional wind instrument. "Sho of Light" premiered just a few days before *Messa di Voce*. *http://www.daf-tokyo.jp/eng/program/01.html*