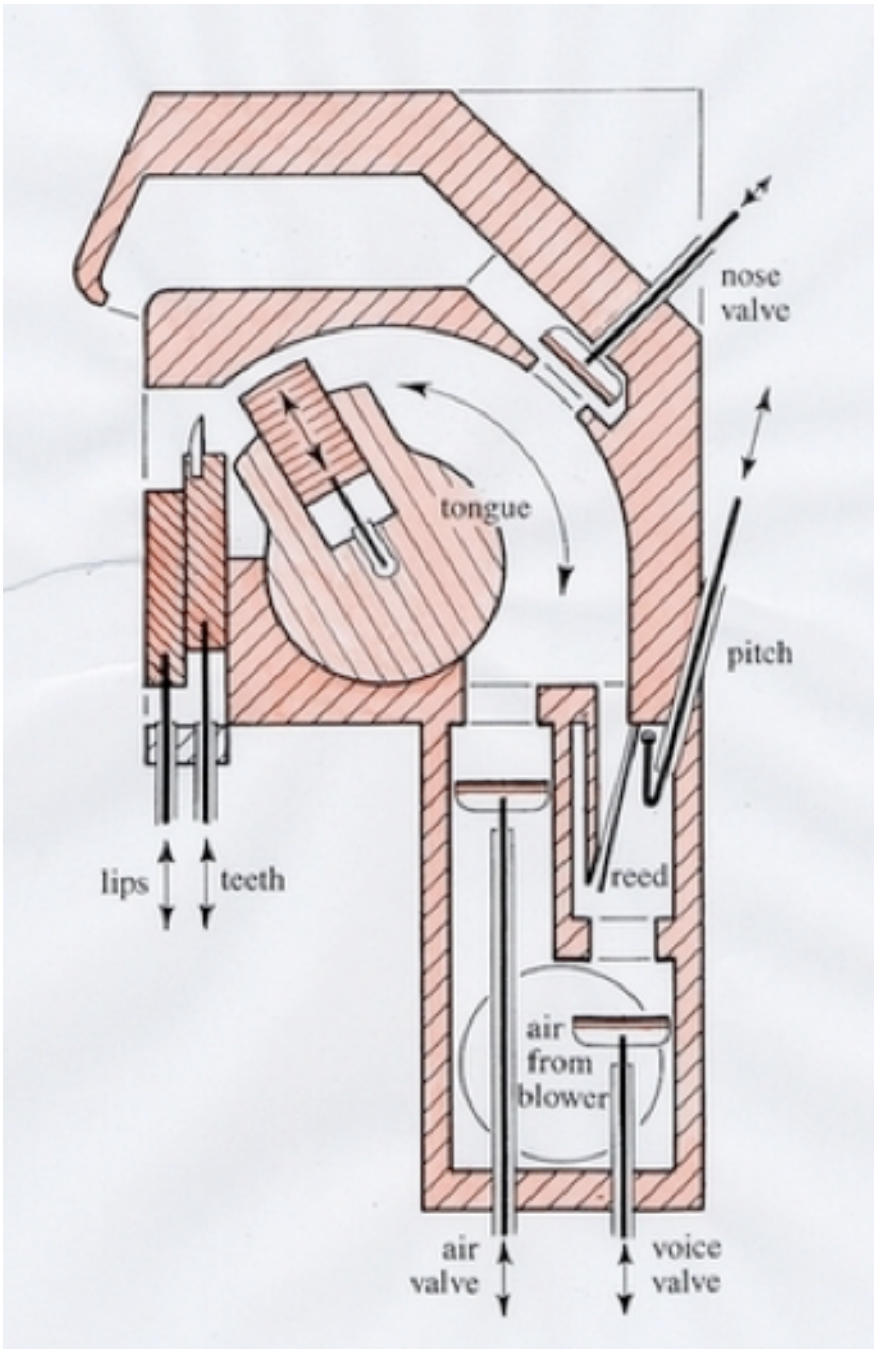
The Art of Voice Synthesis



11 – 12 – 13 May 2016

Amsterdam

Cover: design of *Motor Mouth* (1999) by Martin Riches

The Art of Voice Synthesis

symposium, expert meeting, workshop, concert

11 – 12 – 13 May 2016
University of Amsterdam, Orgelpark, STEIM
Amsterdam
www.artificialvoice.nl

Synthetic speech is part of modern everyday life. Artificial voices not only occur in multifaceted technological uses, but they also feed back into researching the natural human voice. Moreover, artists, musicians and composers find a source of inspiration in the artificial sound of such voices. The symposium inquires both the richness of the human voice and the limits and surplus of its theoretical modelling and mechanical and digital imitation. We are specifically interested in modelling and synthesizing so-called "extended vocal techniques", all sounds the human voice can produce, that exceed conventional singing and speaking. The symposium will cover the history of the artificial voice, extended vocal techniques, aspects of theoretical modelling and technical realization, and the role of the artificial voice in contemporary music. Academics, scientists and artists will come together to exchange ideas and insights in three days of presentations, meetings, workshops and a concert. With a group of international experts we will place the artificial voice in a broad perspective of historical, technical, socio-cultural, artistic and musical investigation.

Symposium: 11 & 12 May at the University of Amsterdam (Doelenzaal) and Orgelpark
Concert: 11 May at Orgelpark
Academic expert meeting: 12 May at the University of Amsterdam (Belle van Zuylenzaal)
Artistic expert meeting: 13 May at STEIM
Workshop: 13 May at STEIM

Doelenzaal, University Library, Singel 425, 1012 WP Amsterdam
Belle van Zuylenzaal, University Library, Singel 425, 1012 WP Amsterdam
Orgelpark, Gerard Brandtstraat 26, 1054 JK Amsterdam
STEIM, Achtergracht 19, 1017 WL Amsterdam

Organization:
prof. dr. ir. Remko Scha (1945-2015), dr. Hannah Bosma, prof. dr. Julia Kursell.

With financial and practical support of:
University of Amsterdam: Computational Linguistics, Musicology, Amsterdam School for Cultural Analysis ASCA; Royal Netherlands Academy of Arts and Sciences KNAW; Orgelpark, Organ Studies Vrije Universiteit Amsterdam.

Contact:
s.muziekwetenschap-fgw@uva.nl
www.artificialvoice.nl

Wednesday 11 May 2016

Doelenzaal

9:00 Registration (coffee/tea)

9:15 Welcome – Julia Kursell
9:30 Voice synthesis in perspective: introduction – Hannah Bosma
10:00 The history of voice synthesis – Fabian Brackhane

10:45 coffee/tea

11:00 The artificial voice as a scientific object – Julia Kursell
11:45 Models of the voice and the voice as model in electroacoustic music – Bruno Bossis

12:30 lunch

Orgelpark

14:15 Vox Humana – Hans Fidom
14:30 Speech synthesis and organ technology – Fabian Brackhane
15:00 Kempelen's speaking machine: demonstration of a replica – Fabian Brackhane
15:30 Research by practical mechanics of the mouth – Jaap Blonk

15:45 coffee/tea

16:00 Martin Riches: demonstration of *The Talking Machine*, *MotorMouth* and the *Singing Machine*

20:15 Concert Orgelpark

* Ute Wassermann – *Organ Pipe Chitter* (new work), for voice, small organ pipes & bird whistles
* Martin Riches – The *Singing Machine* with an extract from *hitonokiesari - people vanish* (2013) by Masahiro Miwa *&* Sadakazu Fujii
* Ute Wassermann & Martin Riches – *Pneuma* (new work), for voice and *Talking Machine*

Interval:
Martin Riches – Hands-on demonstration of the *Talking Machine* with members of the audience

* Michael Edgerton – *The Elements of Risk in Creation* (2001), an electronic composition with optional voice
* Nicolas d' Alessandro – *Lost in Layers* (new work), for soprano and tenor HandSketch instruments
* Nicolas d' Alessandro – demonstration and explanation of his synthetic voice instruments
* Ute Wassermann, Michael Edgerton, Nicolas d' Alessandro – improvisation

Intermezzi: organ improvisations by Guus Janssen

Thursday 12 May

Doelenzaal

9:00 Registration (coffee/tea)

9:15 Objectification of sound and sound source, modelling of emotions: singing with care, vocalecture by Jannie Pranger
9:30 Expanding the voice, lecture-workshop by Ute Wassermann
10:15 Controlled use of nonlinear phenomena in contemporary vocal music – Michael Edgerton

11:00 coffee/tea

11:15 Digital voice synthesis: how to model the "grain" of the voice – Peter Pabon

12:00 lunch

13.15 Concatenative Speech Synthesis: Playing the Imitation Game – Arthur Dirksen
14.15 Get Objects to Talk and Sing: Creating and controlling artificial voice – Nicolas d'Alessandro

15:00 drinks

Belle van Zuylenzaal

15:30 academic expert meeting
(on invitation only)

Friday 13 May

STEIM

10:00 – 13:00 Workshop Nicolas d'Alessandro: Creating and controlling synthetic voices
(€25, registration required via tijs@steim.nl – max. 12 participants)

14:00 – 17:00 Artistic expert meeting
(on invitation only, max. 12 participants)

In memoriam Remko Scha

The Art of Voice Synthesis is an initiative of Remko Scha, artist and Professor Emeritus of Computational Linguistics at the Institute for Logic, Language and Computation at the University of Amsterdam.

Remko Scha passed away on 9 November 2015.

We organize this conference in sad and thankful remembrance of his enthusiasm, generosity, keen interest and inspiring ideas.

The Art of Voice Synthesis: Models and Replications

Creating an artificial voice has been a preoccupation already for several centuries. The first mechanical models imitated parts of the human body that were most clearly involved in vocal production. Later models were based on theoretical principles for the speaking (and sometimes: singing) voice. Electroacoustic synthetic or re-synthesized (quasi-)vocal sounds have been used in contemporary music and art already for decades. Nowadays, artificial voices are abundant: such as in voice response systems (telephone), in navigation systems, as aid for the vocally or visually disabled, and as commercial singing synthesis computer programme (Yamaha's Vocaloid). The current success of artificial voices comes with a change of focus from synthesis by rules, based on a model of the voice, to the use of recorded voices that are analyzed, cut into tiny fragments, manipulated and put together to form new utterances, facilitated by the enormous increase of computer data capacity and computing power.

We wish to look at the full range of techniques that are used and that have been used: from mechanical replication of the synthesis process (Von Kempelen, 1791), through theory-based modelling (electro-mechanical: Helmholtz, 1863; or digital: Klatt, 1980s; synthesis by rule; physical modelling), to methods that are based on audio recordings (synthesis by analysis, analysis/re-synthesis; and concatenation, such as Vocaloid).

Our point of departure is contemporary music. Therefore, a system's capacity to produce acceptable-sounding speech is not our ultimate evaluation criterion. Also, the replication of classical opera technique, though definitely interesting, is not enough. Our reference frame includes the use of extended vocal techniques in the 20th century avant-garde (such as developed by Cathy Berberian, Trevor Wishart and others), as well as the singing styles of various popular and ethnic traditions. A new research question thus emerges: the artificial generation of the complete repertoire of human vocal possibilities.

What are the limitations of the existing voice synthesis models and techniques? And what do these limitations reveal of the complexity and diversity of real, embodied human voices? Is it possible to synthesize "the grain of the voice" (R. Barthes)?

Voice synthesis is, to varying degrees, based on a model of the voice informed by phonetics and voice acoustics. Time-to-frequency transformation (Fourier analysis), sound spectrum analysis, formants, and the source–resonance principle (larynx – vocal tract) are among its basic concepts. Musical instruments functioned as models, objects and inspiration for the science of acoustics; and with respect to the voice, the organ seems to be a privileged metaphor. What attitude(s) towards voice and sound do the existing voice synthesis models and the underlying concepts imply? How is this related to conceptions of sound, music, voice, body, gender and nature?

What alternative models have been conceived of the voice and its artificial synthesis? What alternative models could we think of? If temporal acuity is central in auditory processing (Oppenheim & Magnasco 2013), and the ear does not (only) perform spectral analysis, what are the consequences for the prevalent models of the voice in which vocal spectra (with formants) are of primary importance?

How are artificial voices used in music and other arts? Does this offer a different perspective on the models and methods of voice synthesis? And does the artistic use of voice synthesis offer different perspectives on the voice in general or on specific voices in particular?

In the realm of these questions we are organizing a conference where technical and historical experts meet to discuss the potentials, limitations, implications and contexts of the different voice synthesis techniques.

Remko Scha & Hannah Bosma, 2014

Contributors & Abstracts

Hannah Bosma

Dr. Hannah Bosma defended her doctoral dissertation *The Electronic Cry: Voice and Gender in Electroacoustic Music* at the University of Amsterdam 11-12-'13. Promotores were prof. dr. ir. Remko Scha and prof. dr. Rokus de Groot. As a follow-up, Remko Scha and Hannah Bosma started to set up the symposium The Art of Voice Synthesis.

Previously, the pleasure of the voice led to such diverse activities as singing in the boy's choir of Bach's *Matthaeus Passion*, making a weekly programme on/with all kinds of voice sounds for a local experimental radio station, singing choral works of Xenakis, Feldman and Boerman in the Koor Nieuwe Muziek, studying extended vocal techniques and phonetics during her musicology studies at the UvA, and writing about the voices of Madonna and Cathy Berberian.

Her articles are published in *The Routledge Companion to Sounding Art* (ed. by M. Cobussen, V. Meelberg and B. Truax; Routledge, 2016, forthcoming), *Contemporary Music Review* (2016), *Cathy Berberian: Pioneer of Contemporary Vocality* (ed. by P. Karantonis, F. Placanica, A. Sivuoja, P. Verstraete; Ashgate, 2014) and *Sonic Mediations: Body, Sound, Technology* (ed. by C. Birdsall and A. Enns; Cambridge Scholars Publishing, 2008), i.a. In 2006, she co-organized the conference Cathy Berberian: Pioneer of Contemporary Vocality, at the University of Amsterdam (musicology, ASCA) and programmed the concert series 1000 Volt at Muziekgebouw aan 't IJ. She was a guest editor for a thematic issue on gender and music technology of *Organised Sound* (8/1, 2003). She lectured at various universities in the Netherlands and abroad, for example the course 'Madonna: De muziek en het fenomeen' at the University of Amsterdam in 1997. In 1999 the book *Madonna: De vele gezichten van een popster,* by Hannah Bosma and Patricia Pisters, was published.

From 1998 to 2013, she worked as project manager and specialist electroacoustic music (NEAR) at the Dutch music institute Donemus / Music Center the Netherlands. In 2014–2015, she was a researcher for a project of the Culturele Coalitie Digitale Duurzaamheid and Stichting Digitaal Erfgoed Nederland on born digital heritage in film, photography, architecture and media art.

Fabian Brackhane

The history of speech synthesis
Speech synthesis and organ technology
Kempelen's speaking machine

The aim of producing speech without a human speaker is nearly as old as history. We already know about legendary beginnings in antiquity. From then on the idea of "talking heads" was determining the whole topic. But nearly all of this "sounding brass" was a fraud. It lasted until the late 18th century to get first really functioning and really anatomically based speech synthesis concepts, which then partly were multiple revived until the 20th century. Two aspects during this long development were of special interest: The epoch-making work of Wolfgang von Kempelen and the idea of the pipe organ as a prototype of speech synthesis.

The character of Wolfgang von Kempelen was very often misconstrued by his contemporaries and equally romanticized and ostracized by later generations. This is also for his respectable and fascinating work on speech synthesis.

The connection between the pipe organ and speech synthesis is not very easy to see for us now. But in the 17th and 18th century there was a widespread belief that especially one particular organ stop is the starting point for a real speech synthesis: The so called "vox humana" stop.

In my talks I will sketch the general history of (mechanical) speech synthesis in general as well as the connection between the pipe organ and speech synthesis and their foundations and will present and explain the speaking machine of Wolfgang von Kempelen based on my own replica of this fascinating piece of science history.

Dr. Fabian Brackhane is a German phonetician and musicologist. Via his interest in organs and organ building in particular, he got in contact with the work of the polymath Wolfgang von Kempelen (1734-1804). Through this, since many years his research work is mostly dedicated to the history of speech synthesis. He built several replicas of von Kempelen's speaking machine and is co-editor of a modern, German-English edition of von Kempelen's book *Mechanismus der menschlichen Sprache* (published in 2016).

Julia Kursell

The artificial voice as a scientific object

In this talk the artificial voice is discussed as an object of experimentation in 19th-century experimental physiology. More precisely, it address shift from studying the vocal tract and its use in speaking to the ear as the arbiter of the quality of the voice that enabled a far reaching new approach to encoding sound. For this, Hermann von Helmholtz's "apparatus for the artificial composition of vowels" will be introduced, which is nowadays known as the first "syntheziser." This apparatus bears no similarity with the human vocal organs. Instead, Helmholtz aimed to make audible a viable mathematical description of the vowel sounds with the help of a complicated arrangement of tuning forks, resonators and an electromagnet. In the experimentum crucis for his theory of hearing, he then tested the extent to which the ear recognized the manipulations that this description was able to grasp. This experiment was part of his research into sensory perception that was based on the assumption that perception uses smallest elements that form the actual images or sounds as do the colored dots on tapestry. Helmholtz's idea of such a "stramin" or embroidery canvas proved more lasting than the theory of the physiological process of hearing he suggested. The ear could not be seen in actu, which would have been necessary to understand its physiological function. The synthesis of vowels, in contrast, explored the ability of the sense of hearing to detect differences in sounds that could not have been controlled in a naturally produced sound. The talk will conclude on some thoughts on the accessibility or inaccessibility of the elements involved in this research.

Prof. dr. Julia Kursell is professor of Musicology at the University of Amsterdam. She studied musicology, Slavic philology, and comparative literature and completed her doctoral studies at Munich University with a thesis on music in the early Russian avant-garde. Before coming to Amsterdam, she worked as a research fellow at the Max Planck Institute for the History of Science in Berlin. Her research interests include 20th-century composition, the relation of music and science and the history of musicology and of research in hearing. Her book *Epistemologie des Hörens*, on the role of music in Hermann von Helmholtz's physiology of hearing, is forthcoming with Fink Verlag. Recent publications include a focus section (co-edited with Rens Bod) in the journal *Isis* on the history of the humanities.

Bruno Bossis

Models of the voice and the voice as model in electroacoustic music

During the xx^{e} century, computer science deeply renewed the traditional classification between different categories: the art of composing, the instrumental and vocal interpretation and the instrument making. More precisely, a programmable electronic device must be considered on the one hand as a simple tool, dealing with notes and sounds, and on the second hand as something that is able to manipulate knowledge, to behave in reaction of different solicitations, and to memorize and even learn.

In addition, the dichotomy between voice and instrument is no longer strictly and utterly effective. Sometimes, the sound of an instrument can be more or less transformed by formantization, filtering, or some hybridization, the result being a sound material related to the vocality. Moreover, ambiguity is often the result of these subtle interactions between voice and no-voice. In the history of electroacoustic, numerous examples of this work on the boundaries of our usual perception can be found. Besides, the composing process is sometimes deeply influenced by the structure of the voice spectral elements. In this case, vocality is used, not as a sound reference during the concert, but as a structural model, i.e. more as a concept than as an object. Different examples demonstrate the efficiency of this idea and its implications in terms of expression, sensibility and imagination.

The more interesting consequences of these two kinds of interaction between voice, electronics and music, i.e. “models of the voice and the voice as model” are generated by their combination, or more exactly their intimate intertwining. One of the latest and more powerful examples of this particular way of composing with vocality rather than with voice can be found in *Speakings* by Jonathan Harvey. In this piece for live electronics and large orchestra, the latest developments of Ircam technologies in 2008 support the broad spiritually-colored imagination of the composer.

Prof. dr. Bruno Bossis is Professor in musicology, analysis and computer music, and permanent researcher at the University of Rennes 2, in the music laboratory (Team: Arts : Pratiques et poétiques, EA 3208). He is director of the master Arts (including masters in Musicology, Movie studies, Theater study, Arts history and Visual arts), co-director of the master Arts and Digital Technologies. Author of the book *La voix et la machine, la vocalité artificielle dans la musique contemporaine* and of numerous papers on electroacoustic music, editor of several books on contemporary music, he is currently writing a book on Jonathan Harvey.

Hans Fidom

Vox Humana

It is a common misunderstanding that organ stops, as they have been developed ever since the 16th century, were meant to really imitate the instruments their names refer to. No ensemble provided such richness in flute-types, for example, as organs could. But even if they were meant to imitate, that would have been bound to remain an only theoretically interesting perspective: in practice, an organist would never ever be able to play his stops separately the way instruments are played in ensembles. It is, instead, the combination of stops that determines the (changes of the) organ's sound. Which invites us to develop ideas regarding synthesis when thinking about how organs sound in another direction than this colloquium is focusing on: organists construct overtone structures by combining stops. To make things even more complicated, one should not forget that each organ stop consists of many separate pipes, which allows, to mention only one aspect, for colour changes throughout a stop. This uncovers organs as synthesizers avant-la-lettre, which the huge advantage that they don't depend on just a few electronic sound generators (loudspeaker systems) but work with thousands of generators (i.e. pipes) instead.

Yet, although it is quite tricky to base hypotheses about the way former generations sang on the sound of the organ stop 'Vox humana', listening to organs seems to be connected to the way people make their languages sound: to put it mildly, it seems to make quite a bit of sense hat the classical French organ has a pretty nasal character, as it reminds of the way French people speak French. By presenting a few examples of Vox humana's throughout the world, I would like to demonstrate that, just like there is no such thing as THE organ, there is no such thing as THE Vox humana either.

Prof. dr. Hans Fidom is professor of organ musicology at VU University Amsterdam and leader of the Orgelpark Research Program. His main fields of research are the ontology of music as a sound(ing) phenomenon, history and development of new technology in organ building, and the way the organ repositions – if not invents – itself in 21st century musical life. He is member of the core-group that monitors the building of the New Baroque Organ at the Orgelpark, which will be a so-called 'hyper organ', as it will be uncompromisingly historical and include digital technology at the same time, thus inviting musicians to work with the organ concept developed in Johann Sebastian Bach's time and region in historically inspired as well as in radically new ways.

Jaap Blonk

Research by Practical Mechanics of the Mouth

Around 1990 I started to use the International Phonetic Alphabet (IPA) for notation of my sound poems and vocal parts in compositions. The classification and instructions in the IPA are mainly given in terms of the 'mechanics' of the mouth: position and movement of the lips and tongue, and air pressure.

This gave me a new and useful mode of research: instead of hunting for new sounds directly by ear, I took to experimenting primarily with these mechanics, with the resulting sounds as a secondary phenomenon. As a result of this, sounds found were sometimes quite surprising. I will give a few examples of this.

A natural next step was using the hands to manipulate the face and throat, yielding a category of sounds not represented at all in the IPA, for obvious reasons.

Lastly, I will give a few examples of my 'cheek synthesizer' technique: stereo mouth sound produced in various ways, inspired by mechanics as well as my experience with electronic sound.

Jaap Blonk (born 1953 in Woerden, Netherlands) is a self-taught composer, performer and poet.

His unfinished studies in mathematics and musicology mainly created a penchant for activities in a Dada vein, as did several unsuccessful jobs in offices and other well-organized systems.

In the 1970s he took up saxophone and started to compose music. Soon after, he discovered his potential as a vocal performer, at first in reciting sound poetry and later on in improvisations and his own compositions. For his vocal parts in the latter, he has developed an idiosyncratic notation system, based on the International Phonetic Alphabet, but with many added symbols.

As a vocalist, Jaap Blonk is unique for his powerful stage presence and almost childlike freedom in improvisation, combined with a keen grasp of structure. He has performed, lectured and taught workshops around the world, in widely different contexts: contemporary composed music, improvised music, literature and visual art.

With the use of electronics and visuals the scope and range of his work has acquired a considerable extension.

From his sound poetry scores he developed an independent body of visual work, which has been published and exhibited.

Martin Riches

Martin Riches was born on the Isle of Wight in 1942. He is best known as a sound artist.

Four speaking machine projects:

Talking Machine (1989-1991): A computer-controlled wind-blown organ with 32 phoneme pipes inspired by Chiba and Kajiyama resonators (1942) and X-ray data by Gunnar Fant (1960). Current vocabulary: about 350 words.

MotorMouth (1994-1999): A mechanised vocal tract with a blower, 8 computer-controlled stepping motors driving the lips, teeth, tongue, tip-of-the-tongue, air-supply, glottis and pitch. It can count to 10 in English and German and speak a few simple sentences: How are you? I love you? etc.

Singing Machine (2010-2013): A tubular vowels-only vocal tract with moving tongue and lips and a range of one octave. Among other works: the poem hitonkiesari (in the synthetic singing language, Gyakku) by Sadakazu Fujii, set to music by Masahiro Miwa and accompanied by 9 musicians of the Kammerensemble Neue Musik Berlin.

Four Voices (2015 - work-in-progress): Soprano, alto, tenor, bass, similar in concept to the baritone *Singing Machine* (above).

For these and a variety of other projects see: http://martinriches.de

For *MotorMouth* we thank the Berlinische Galerie - museum for modern art, photography and architecture.

BERLINISCHE GALERIE

Ute Wassermann

Organ Pipe Chitter (2016), for voice, small organ pipes and bird whistles

In *Organ Pipe Chitter* I de-familiarize the sound of my voice using palate whistles, bird whistles and small organ pipes. I mask my voice and create a hybrid vocal instrument. In the multifaceted area between voice and whistle-instrument a sound world takes form within chirps, trills, wave flows and fluctuating particles.

Pneuma (2016)**,** for voice and *Talking Machine*

In *Pneuma* (breath, or soul) I, the singer, explore the relationship between a human voice and that of the *Talking Machine* constructed by Martin Riches.

With his help I have explored the machine and discovered that it is capable of human sounds quite like breathing, stuttering, sighing and yodelling as well as its own purely mechanical utterances and noises. It can also generate rapid glissandi across its rows of pipes which, to a willing listener, might be interpreted as cryptic words or phrases. The singer, for her part, finds herself in a mechanical world that generates layers of multi-phonic sound intermingled with human sounds and words.

I have used its capabilities as a context within which a human voice may imitate, exaggerate, translate and re-interpret the phrases of a machine. In this, I have been inspired by Laurie Anderson's *O Superman* and also by Salomo Friedländer's *Goethe speaks in the phonograph - Goethe spricht in den Phonographen* (1916) in which a miraculous artificial larynx recreates the original vibrations of Goethe's voice.

Expanding the voice

Since many years I have worked on de-familiarizing the sound of my voice exploring many-voiced and extended vocal techniques. I shift, distort and reinvent the connection between the body parts used in vocal production. Also I use different resonators, speaker objects, palate whistles and bird whistles in order to create hybrid vocal instruments. My performances redefine in different ways the threshold at which the voice makes contact with the outside world. The boundary between self and environment is differently redrawn each time.

This lecture / performance will include a live demonstration of a catalogue of extended vocal techniques organized by timbre, register and articulation. The audience will be invited to participate.

Ute Wassermann is known as a vocal soloist and composer/performer for her extraordinary, many-voiced and extreme vocal sound-language, which she has brought into contemporary music in diverse ways. She has developed vocal techniques in which the voice goes beyond itself into a much les identifiable vocal space. She de-familiarizes the sound of her voice using birdcall-whistles, palate whistles, speaker objects or field recordings. Ute Wassermann performs her own works on international festivals and collaborates with internationally known improvising musicians, composers, dancers, visual artists. She is an integral part of the experimental music scenes and performs regularly with many musicians in different groups like "Speak Easy" (with Phil Minton, Thomas Lehn and Martin Blume), with Aleks Kolkowski (Stroh instruments and musical saw), Birgit Ulher (trumpet), as well as in larger formations such as fORCH or ECIO and is a member of the Berlin-based composer/performer ensemble "Les Femmes Savantes".

She gives lectures and workshops at various universities. She has given premieres of numerous works & music theatre pieces composed specially for her voice, for example by Richard Barrett, Chaya Czernowin, Henning Christiansen, Hans-Joachim Hespos, Matthias Kaul, Michael Maierhof, Michael Finnissy, Ana Maria Rodriguez, Gerhard Stäbler, Pia Palme with the ASKO Ensemble, Elision, Munich Chamber Orchestra, KNM Berlin, L´ART POUR L´ART.

http://femmes-savantes.net/en/lesfemmessavantes/ute-wassermann/
https://vimeo.com/user20410741
https://soundcloud.com/utewassermann
http://www.auditionrecords.com/ar063.php

Michael Edgerton

Controlled use of nonlinear phenomena in contemporary vocal music

Complex and multi-phonic voice signals of vocal performers will be presented within the framework of nonlinear dynamics. Evidence will be given that nonlinear phenomena are extensively used by performers associated with contemporary music. Narrow-band spectrograms of complex vocalizations used to visualize the appearance of nonlinear phenomena (spectral bifurcation diagrams), will be paired with audio and video documentation. Possible production mechanisms are discussed in connection with previous research, personal performance and pedagogical experience.

Examples for period doubling, biphonation and irregular aperiodic phonation in vocal sonorities of contemporary vocal improvisors will be given, and recent research focused on the glottal whistle (M4) production encompassed with biphonation and triphonation will be shown. Furthermore, coincidences of harmonics-formant matching associated with abrupt transitions to subharmonics and biphonation in the vocal output are provided. This also shows the recurrent use of nonlinear phenomena by performers. It is argued that mechanisms such as source-tract coupling or vocal fold de-synchronization due to asymmetry are used in a reproducible way for musical tasks.

Dr. Michael Edgerton is currently Associate Professor at the Guangxi Arts University. He is a composer whose work elides the boundaries of complexity with practical applications of physical and perceptual models. Since the mid-90's he has been pioneering work with Multidimensionality and Nonlinear Phenomenon applied to sound production and composition.

His work has been recorded on Alles – Auβer – gewöhnlich (Auditivvokal Dresden, 2012), ELECTRONIC VOICES (2001), and more; and in Lena Giovanazzi' film *Stimmig – 10 Vokalexkursionen* (Mainz, 2012).

Michael is engaged with research into voice, acoustics and perception. This work has been supported by numerous grants, including a Postdoctoral Fellowship with the National Center for Voice and Speech. The results have been published in diverse scholarly journals, as well as book chapters in *Perspectives on Teaching Singing: a celebration of Vocal Pedagogy in the 21 Century* (ed. Scott Harrison and Jessica O'Bryan; Springer Publishing), and *The Oxford Handbook of Singing* (ed. John Nix; Oxford University Press). Edgerton has two books available: *The 21st Century Voice* (New Instrumentation Series, Rowman & Littlefield, 2nd ed. 2015), *Music within the Continuum* (Lambert Academic Publishing, 2015).

Nicolas d'Alessandro

Get objects to talk and sing: Creating and controlling artificial voice

Speech and singing are gestures, physiologically, cognitively, perceptually and socially. When voice is used in performing arts, it brings humans in very emotional and intense experiences, through principles that we still not fully understand nowadays. In this talk, we will look back on ten years of scientific, technological and artistic research on voice, largely outside the mainstream academic preoccupations, and closer to the work of an instrument maker, who empirically tries to understand and extend the expressivity of the sonic body. Ten years making digitally-enhanced objects that simulate speech and singing, but offer their guts to real hands. The performance will demonstrate this thanks to ten years of practicing the HandSketch, a singing tablet.

There are many different ways to produce an artificial voice on the computer, from knitting every single sinusoidal sound of it, to transform pre-recorded speech and singing samples, and now even going to the cloud and tap into online speech synthesis web services. In the workshop, we will approach the creation of synthetic voices in two ways. First we will put our hands into actually making vocal sounds in Max and/or Pure Data, browsing the 4-5 most common synthesis techniques. Then we will discuss and demonstrate how to control vocal sounds in performance. Voice is not an easy medium to "make sound good", so we will explore how the choice of the controllers and the gestures can affect our ability to get where we expect with artificial voice synthesis.

Dr. Nicolas d'Alessandro obtained his PhD in Applied Sciences from the University of Mons in 2009. From a lifelong interest in musical instruments and his acquired taste in speech and singing processing, he will incrementally shape a research topic that aims at using gestural control of sound in order to gain insights in speech and singing production. Late 2009, he moves to Canada, to take a postdoc position with Prof. Sidney Fels at the MAGIC Lab, University of British Columbia, Vancouver, where he will work on the DiVA project. Since early 2015, he has left academia and founded Hovertone, a startup standing at the crossing between arts and technology.

Jannie Pranger

Objectification of sound and sound source, modelling of emotions: singing with care

By singing a fragment of Luciano Berio's *Sequenza III*, I want to demonstrate how affective forces are intimately entwined in processes of creation. This may seem obvious in an art environment, however, I hope by introducing the concept 'care,' to contribute to a wider discussion that seeks to intervene in creative practices that tend to devaluate materiality and embodiment.

Jannie Pranger is Affiliated Researcher, member of the Research Group Musicology at the Utrecht University. Currently she is preparing her PhD. Based of the writings of Carolyn Abbate and Karen Barad, she investigates the onto-epistemology of musical performance. Her promotores are prof. dr. Karl Kügle and dr. Iris van der Tuin.

Pranger is an accomplished singer in the field of New Music and collaborated with outstanding composers and performers. She teaches at the Utrecht University of the Arts at the faculty of Music and Technology.

Peter Pabon

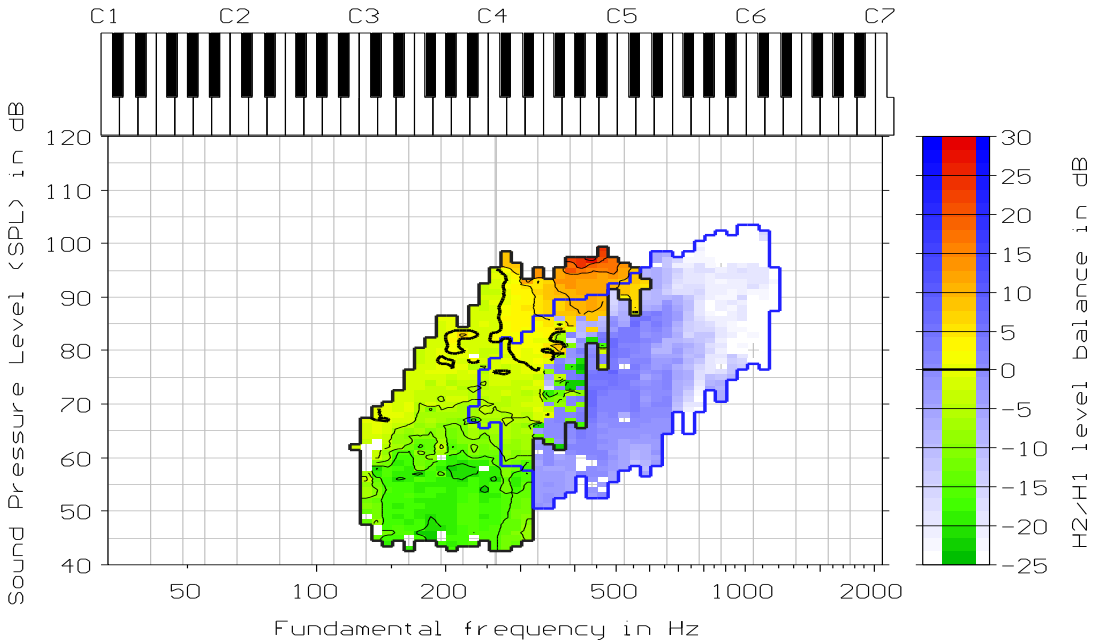
Catching the "grain" of the voice: the search for acoustic textures and patterns that make a voice unique

Roland Barthes introduced the concept: “the grain of the voice” and characterizes it as “the body in the voice as it sings, the hand as it writes, the limb as it performs”. It is the embodiment of a certain physicality of the singer or singing as it comes to live in the head of the listener. To Barthes the grain is also the thrilling aspect of a voice; a concept that cannot be fully caught in words and that is bound to lose its essence the moment it is objectified. Following his thinking, we will never catch the grain, as every artificial singing voice is inevitably an objectification, the mock-up of the real thing. But what if we love the thrill or the grain of the mock-up? What if, by just being playful and inventive, we do catch the grain of the real voice, or an element of it? I love that thought. This philosophical self-destruct clause, and the experience how difficult it is to catch the grain of the voice only make the exploration more challenging.

Science understands the voice via models, where the source-filter model still has the leading role; it is aging but not yet at the end of its singing career. It is a versatile and effective model that has a straightforward processing scheme. With only a limited set of control parameters an acceptable singing voice quality is readily synthesized, (which I will certainly demonstrate). The source-filter model is very general and knows no natural constraints, which is its strength and its weakness. The model is typically good in reproducing the acoustical ingredients that we know from an earlier analysis. It can convey the grain of the voice without ever grasping where it came from. Although its parameters have a fixed physiologic and physical interpretation, the connection to the perceived physicality of the singer or the singing remains intricate and variable. So, in what acoustical structures is the grain hiding?

The voice range profile (VRP) – a mapping of the intensity range of the voice as a function of the fundamental frequency (the pitch) – can give some interesting answers. Modern interactive VRP recording systems allow a very detailed and reproducible imaging of the full progression of dynamic quality variation that is characteristic to a person’s voice (see Figure). It is possible to drive an abstract voice synthesis model directly from the analysed parameters, all without an immediate source or filter interpretation, which makes this method much more open-minded to the real acoustic variations that differentiate voices. A large study of the spectral variations over the VRP resulted in a “wish list” of characteristic dynamic quality contrasts that any singing voice model should exhibit, if it actually wants to be as peculiar as the real thing. The list scores distinct spectrum (co)variations and state dependent phenomena that the real

voice is very well capable of, but that existing models fail to show or highlight. It suggests an alternative view on the acoustical features that hold the grain of the voice and challenges some ideas on what controls our perception. During the presentation, a few selected phenomena from this list will be explored, each time with dedicated synthesis examples to elucidate the contrasting aspects.



Peter Pabon (*1956) studied biochemistry, signal processing and sonology at Utrecht University. His professional career started in 1983 as a part-time researcher on a project called Objective Recording of Voice Quality with Professor Plomp at VU University in Amsterdam, and he worked at Utrecht University as a teacher/researcher on (singing) voice analysis and speech and music acoustics from 1983 until 2011. He currently teaches at the Institute of Sonology at the Royal Conservatoire in The Hague.

He initiated a project for singing voice synthesis and analysis at the Royal Conservatoire that later resulted in a cooperative project with the singing department to monitor voice change as an effect of voice training. In 2002, he founded Voice Quality Systems, a company in which he develops the voice quality recording system Voice Profiler, which is nowadays still in use at many clinical centres, conservatories and schools for professional voice training. Peter Pabon is currently writing a PhD thesis at KTH Stockholm, which has generated several papers and presentations on Voice Range Profile (VRP) recording methodology and the effects of voice training. A final paper will deal with a description of spectrum change over the VRP.

Arthur Dirksen

Concatenative Speech Synthesis: Playing the Imitation Game

The ultimate test for a speech synthesizer is very simple: to what extent does it sound like a human being reading a text or speaking a sentence? To meet this test, the speech generated by a synthesizer needs to sound completely natural, perhaps to the extent that it includes breathing, creaky voice, and other artifacts of natural speech.

While early synthesizers attempted to create speech from scratch, it soon became clear that the rather robotic sounds produced by these synthesizers would not be able to play the Imitation Game successfully. This has resulted in a move away from actually modeling the speech production process to a more data-oriented approach.

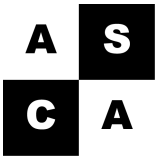
Second generation synthesizers made use of diphones, cut from recorded speech, which were manipulated in terms of pitch and duration to obtain the desired intonation and rhythm. But, although much easier to listen to, these synthesizers still leave something to be desired with respect to naturalness of the voice.

Almost all commercial synthesizers nowadays are based on unit selection, which, in its purest form, does not use signal processing at all. Instead, small (or larger) fragments from a large database of linguistically annotated speech are strung together to form new utterances. These newer synthesizers often sound surprisingly natural.

In his talk, Arthur Dirksen will discuss these three generations of speech synthesizers in some more detail, illustrated with examples from his own work.

Dr. Arthur Dirksen studied Dutch Language and Literature at Utrecht University. After his PhD he became involved in speech synthesis research. In 1996 he founded Fluency, a small company, now based in Amsterdam, that develops text-to-speech software for Dutch and Frisian.

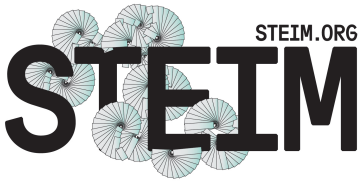

UNIVERSITEIT VAN AMSTERDAM


A S
C A


KONINKLIJKE NEDERLANDSE
AKADEMIE VAN WETENSCHAPPEN


ORGELPARK


VU
VRIJE
UNIVERSITEIT
AMSTERDAM


STEIM.ORG
STEIM