Show More/ Show Less:
Extended Voice, Technology, and Presence

Kristina Warren
Liverpool, UK

B.A., Duke University, 2011

A Dissertation Presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Music

University of Virginia
April 2017

ABSTRACT

Extended voice is a vocal-electronic practice integrating compositional and performative, especially improvisational, methodologies. In extended voice, the same single practitioner does both vocal and technological work. This document is structured in three parts: (1) technical discussion of my hardware, software, and vocal work; (2) analysis of works by seven other extended voice practitioners (Antye Greie-Ripatti (AGF), Marie Guilleray, Stine Janvin Motland, Maja S. K. Ratkje, Andrea Pensado, Ami Yoshida, and Pamela Z); (3) and critical discussion of extended voice. Extended voice encompasses a wide variety of sounds, but crucial themes are the desire for complex, novel interaction between voice and electronics, and, toward this end, the dovetailing of 'recorded' and 'live' methodologies. Practitioners undertake *self-listening* to manage the vocal-electronic whole, which is a cyborg of sorts, and thus evince *presence*, or curation of one's own vocal-electronic sound. Research futures for extended voice center on its use of creative methodologies to subvert the notion of normative body, and its demonstrated ability to encourage more diverse participation than that which currently typifies electronic music as a whole.

keywords:

extended voice presence self-listening agency interaction cyborg technology digital MaxMSP granulation Arduino Bela Pd analog AGF Guilleray Motland Ratkje Pensado Yoshida Z recorded live disability normative inner voice body

CONTENTS

ACKNOWLEDGEMENTS

I am deeply grateful to my parents, who have supported my every academic inclination, and to Paul Turowski, whose support has kept me going. My heartfelt thanks go to my advisor, Ted Coffey, whose curiosity and humor have made me a better musician and thinker; to Judith Shatin for making my existence in this field possible; to Bonnie Gordon for her ongoing mentorship; and to Elizabeth Barnes for her invaluable perspective on my project. Thanks as well to Anthony Kelley, Kerry McCarthy, and Louise Meintjes, who knew how and when to push me further than I was ready for; to Rachel Devorah for her laughter, music, and perspective; and to Kevin Davis for his generosity, especially in improvisation. Thanks to Antye Greie-Ripatti, Marie Guilleray (who also graciously invited me to The Hague), Stine Janvin Motland, Andrea Pensado, Maja S. K. Ratkje, and Pamela Z for their thoughtful engagement with my questions; and to Tom Johnson, Christian Wolff, and Amnon Wolman for their text score permissions. Thanks to Emily Baker, Ged Barry, Michael Beiert, Amy Brandon, Peter Bussigel, Liz Dobson, Robin Fencott, Myra Melford, Ailís Ní Ríain, Kimberly Sutton, Travis Thatcher, and all of my friends and colleagues in Charlottesville. Finally, I am deeply grateful for support from the PEO Scholar Award, which has enabled my travel and research ventures this year.

For Grandma Dorie.

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

I walk onstage slowly, feeling their eyes on me. I breathe slowly, keep everything slow even as my mind and ears quicken with possibilities. I begin vocalizing meditatively. I think of Oliveros' Sonic Meditations, and of my diaphragm and intercostal muscles, as I listen to the room and try to sound with it. I bring up my hand and flip on granulation (all those lovely sidebands in just eight seconds), but almost instantly decide I need electronics following the voice instead. In a few minutes I find that spiky vocal shrieks are working, and I start to jitter the voice-following electronics for noise to balance my shrieks. This electronic base has a strong identity, which I push off of as I shriek more loudly and gradually transition to ragged breathing. On a dime I cut the electronics entirely. With my upper teeth on my lower lip, I inhale, yielding a high, very quiet, buzzing-squeaking sound with mid-length envelopes. I spend what feels like a long time varying the pressure of teeth on lip. Less pressure: only one or two voices. More pressure: polyphonic, even spectrally noisy. I tell myself to keep it slow, make them listen, find real depth in this little sound. I move closer to and further from the mic. I am obvious about including the room and the audience in the sound.

A. Extended voice

This dissertation addresses what I call 'extended voice', a term derived from existing research but by which I imply particular contemporary forms of sound and practice. An important historical influence on my category 'extended voice' is the modernist vocal music of the mid- to late-twentieth century, which highlighted extended vocal techniques. Composer-performers such as Cathy Berberian and Joan La Barbara used these non-normative acoustic

vocal sounds, sometimes with added electronics, in order to explore *what else* the voice can do. This stylistic approach went hand-in-hand with an attitude of objective, scientific exploration of voice. Some contemporary extended voice practitioners, as I call them, use a similar language of objectivity when describing their work.

In addition, recent research by contemporary composer-vocalists underpins my use of the term extended voice. Marie Guilleray (2012) writes of her musical goals in developing a practice of "electronically extended voice" using software and novel glove controller. Similarly, Donna Hewitt (2011) describes how technologies such as her novel instrument the eMic serve to "extend the voice." Such work inherits from (post)modernist vocal music an attitude of 'what else' exploration of voice, but significantly deepens the component of self-driven technological use and research.

I define *extended voice* as a vocal-electronic practice which synthesizes compositional and performative – especially improvisational – methodologies. To be an extended voice practitioner is to be a composer-vocalist-technologist, i.e., to be involved in all steps of actualizing a voice-electronics piece. Extended voice encompasses a wide variety of stylistic approaches, but all share an aesthetic goal (to achieve a rich, meaningful dialogue between voice and electronics) and a methodological approach (exploring novel juxtapositions of 'recorded' and 'live' sound techniques). In fact, the aesthetic and the methodological go hand in hand. Voice-electronics interaction is grounded in the trans-temporality of using live and recorded sound techniques together. Deep questions arise from this work: Are the electronics derived from a recording of voice? Do they shape the subsequent vocal or electronic improvisation? Has a certain sound been heard before? Is the sound being heard now vocal or electronic or both? If

late twentieth century vocal experimentation extended the timbre of voice, what I call 'extended voice' continues this thread by using electronics to innovate in the temporality of voice.

Among the many forms of musical activity that contribute to voice-electronics interaction and live-recorded overlap in extended voice, I find that *self-listening* is crucial. I define self-listening as listening to one's own vocal and electronic sounds during performance/recording, for the purpose of sculpting voice-electronics interaction. Self-listening is important for several reasons. Its frequent use in extended voice is a marked departure from the eschewal of self-listening typical of *bel canto* (and of those contemporary Western vocal practices which distill *bel canto* understandings of 'good' vocal sound and 'proper' performer engagement). Self-listening unites the soundworlds and musicking activities of performer and listeners, and features prominently in the creative process. Self-listening is a crucial component of my notion of *presence*, which I define as curation of one's own vocal-electronic sound.

The first paragraph of the Introduction uses my own practice as an example of some of the particular ways in which self-listening can unfold. Self-listening touches voice and electronics, along with space and audience, in an attempt to coax a complex, artful interaction among them. The notion of self-listening allows that everyone hears and listens differently (I hear myself one way, a listener hears me in a different way, another listener hears me in a different way still), and moreover treats personalized listening and resultant technology use as creative acts.

I expand on the concept of self-listening by drawing a connection between text scores and extended voice. Several practitioners identify their unique working vocabularies, which are at once documentation of their practice and ideas for compositional-performative action. This 'inner voice' or internal language use means that several key features of text scores – physical

organization, number, and listening – refract into the extended voice features of body, mixing, and agency or self-listening.

The final part of this dissertation considers ways in which disability studies can model thinking about the social work done by extended voice. Elizabeth Barnes' important text *The Minority Body* (2016) posits that defining a 'normative body' is a task fraught with individual exceptions and cultural variations. Likewise, by foregrounding self-listening – a physical-electronic and necessarily personalized activity – extended voice problematizes the notion of a normative body in sound. Because self-listening is a minority embodied practice undertaken for particular sonic reasons, and because the agency of extended voice manifests partly as a commentary on genre, extended voice is importantly a social category, like Barnes' view of disability. The conclusion to this essay ventures some possibilities as to the sonic and social futures of extended voice. I am optimistic that this practice will help normalize vocal experimentation and diversify electronic music as a field.

This dissertation is structured in three parts: description of my recent work (hardware, software, vocal), survey of several extended voice practitioners' work, and aesthetic and theoretical discussion of extended voice as a whole.

B. Defining the field

Seven extended voice practitioners whose work I value feature in this essay. Antye Greie-Ripatti (AGF) skillfully uses production techniques and lyrics to explore voice-machine identity. Marie Guilleray makes detailed lines and layers that demand repeated listening even as they flit from moment to moment. Stine Janvin Motland uses her extreme vocal stamina to blur the line between overpowering and absorbing her technology. Andrea Pensado's nasty noise converses

intelligently with playful patch control and keen attention to form. Maja S.K. Ratkje creates grand soundscapes whose precisely timed timbral synergy questions how voice can and should sound. Ami Yoshida's self-described 'howl voice' joins subtle but mind-altering loops to interrogate the sound and temporality of voice. Pamela Z fearlessly juxtaposes *bel canto*, spoken word, looping, and other techniques to yield a profound exploration of language and meaning.

Notably, all of these case studies are women. I did not initially intend to study only women, but elected to do so after noticing that most extended voice practitioners are women. In addition, I have frequently had the experience of being the only woman present at events in the larger field of electronic music (not necessarily voice-related), but I have never seen these billed as 'men's events'. By contrast, the women-led events I have participated in are viewed as non-normative and often billed as 'women's events'. Many women also feel intimidated by men's technical knowledge. It is not right that potentially half of the population feels directly or indirectly intimidated or excluded by the other half. In short, my discussion of women is an effort to counteract the perceived masculine monopoly on technical knowledge or access, which manifests first as physical normativity or majority presence in a space.

There are many musicians and sound artists making important voice-electronics work who I do not discuss here. Stylistic forebears such as Joan La Barbara, Laurie Anderson, and Diamanda Galás have already been extensively written about. Others complicate and enrich the category of extended voice in ways that are beyond the scope of this project. These musicians include Björk, Jaap Blonk, Paul Botelho, Nadah El Shazly, Imogen Heap, Holly Herndon, Donna Hewitt, Kyoka, Amy X Neuburg, Shara Nova, Nichola Scrutton, Amanda Stewart, Reggie Watts, Trevor Wishart, Andrea Young, and Jonathan Zorn. Continued research is needed to account for the diverse sonic output of these artists.

C. Notes

Audio/video documentation of relevant work can be found at:

http://kmwarren.org/dissmtl.html

Portions of this document have been published as “Notated Control as Composed Liveness in Works for Digitally Extended Voice” (2017). In attributing countries to extended voice practitioners, I list their country of residence and, if separate and listed in their professional bio, their country of birth. Text scores excerpts by living composers are published with their permission.

II. WORK BY THE AUTHOR

This section focuses primarily on my most recent hardware and software work: Bela running a Pure Data (Pd) patch (November-December 2016); Abacus versions 3 (August 2016) and 4 (February 2017); and Max patch Vox6 (March 2017). In actual chronology, Abacus 3 with Max patch Vox5 came first, followed by Bela-Pd, followed most recently by Abacus 4 with Max patch Vox6. But I have chosen thematic rather than chronological organization. I begin with Bela-Pd because I consider this laptop-free setup different in kind to the laptop-based Abacus-Max work. I continue by discussing the two most recent Abacus versions, which have driven my thinking about voice-electronics interactions in performance. I conclude with my current Max patch, Vox6. MaxMSP has continually been the at the heart of my voice-electronics sounds, and Vox6 includes my favorite elements of previous patches along with many new features. At times these sections refer to each other because of the shared ideas connecting them all.

Notes on terminology: ‘pre-sample’ indicates a pre-recorded audio buffer; ‘live sample’, audio recorded to a buffer during performance. I frequently treat live samples with ‘pulsar granulation’, a form of granular synthesis which streams an alternation of grain and short silence (Roads 2001).

A. Bela, Pd

In Fall 2016, I developed a streamlined road setup for solo and group improvisation using a Beagle Bone Black microcontroller with Bela audio shield, and breadboard-mounted physical control components. I premiered a preliminary version of this setup in an electronics-only set at Noisevember (London, November 2016) and used an expanded version with basic processing of

vocal input for Winter 2016-17 voice-electronics performances with the Merseyside Improvisers Orchestra and the AKA trio.

The Bela-Pd setup is an important lateral step from much of my practice because it does not use laptop. Using a small but multi-wired electronics setup felt very different (perhaps more hands-on or transparently tech-y) than performing with a laptop.

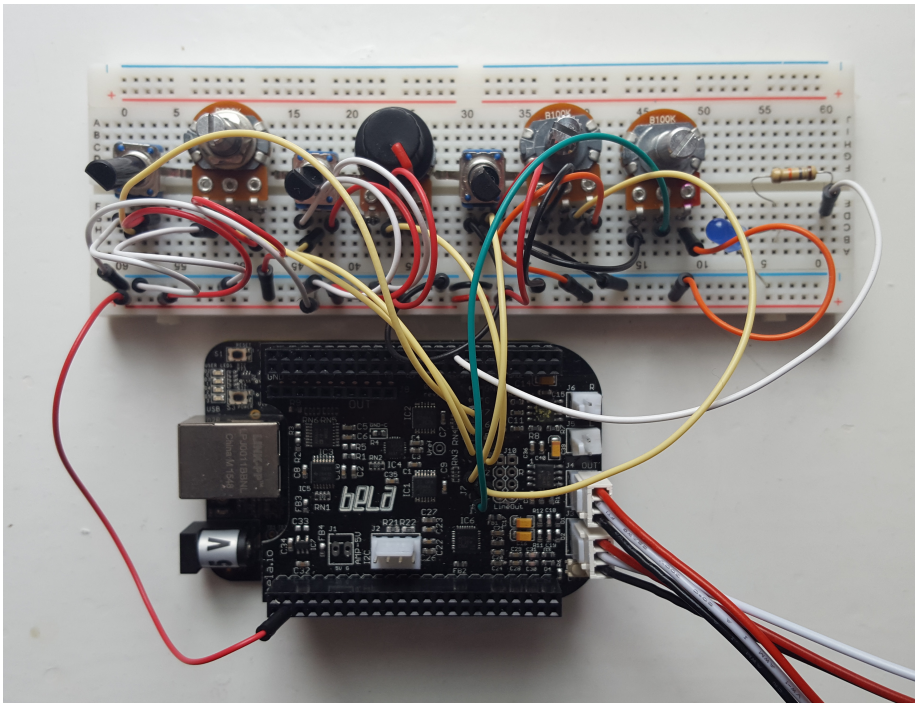


Figure 1. Laptop-free setup:
Bela running Pd, potentiometers on breadboard

My approach was to use pre-composition to simplify the granulation procedures which I had explored up until that point in Max. My previous Max patches required me to painstakingly craft parameters for rhythmic granulation and record the rhythmic granulation output to a buffer for pulsar granulation. By contrast, the Pd patch for Bela only does pulsar granulation. This patch pre-loads two sounds from my standard library: ‘inhale gliss’ (unprocessed) and ‘saliva’ (rhythmic granulated beforehand). In performance, I simply use breadboarded potentiometers to

control the three important pulsar granulation parameters Start Time, grain Length, and Wait time between grains. Thus, I pre-composed almost all of the process and skipped straight to pulsar granulation during performance.[1]

Pre-composing buffers for pulsar granulation made control much easier, but, while performing, the electronics initially felt less live to me. While I sometimes worry that the voice and electronics are too separate in my work, I felt that separation even more acutely during these performances. Nevertheless, these extreme limitations on control and liveness were creatively productive. I later concluded that the dialogue between voice and electronics was actually one of the most sophisticated I had achieved to date.

Thus, *presence* in extended voice, which I define below as curation of one's own vocal-electronic sound, can involve displacing or re-imagining so-called liveness. When using the Bela-Pd setup, I set aside my usual expectations that all synthesis procedures occur 'live' during performance, and this allowed me to achieve a more fluent dialogue between voice and electronics. This proved highly influential. In crafting my subsequent Max patch Vox6 and mapping control information from Abacus version 4, I made numerous compositional choices to prioritize speed and fluidity and thus to devote more time and attention during performance to voice-electronics interactions.

Finally, the streamlined Bela setup seems to suggest that solo and group performances require distinct setups. The current Max patch Vox6 is the fastest yet, but may still be too slow for group performance (testing required), whereas the Bela setup is less deep technically but more nimble, and therefore seems suitable for group contexts.

---

[1] In addition to the pulsar granulation unit, the Pd patch includes units for additive synthesis with ring modulation, and variable delay and bit-crushing on the live vocal signal. All three processing units receive control information from physical potentiometers.

B. Abacus

Building a novel performance interface to control vocal processing in Max had long been appealing, but this idea was sparked in Autumn 2015. In a private composition workshop at a November 2015 conference, I performed a short voice-electronics set (then using a Wacom tablet for control) for New York-based composer-vocalist Kate Soper. Soper provided the astute feedback that my voice-electronics performance seemed too small and intimate. Soper felt she should "give [me] privacy" in which to complete this personal vocal ritual of performing with electronics. This was apparently a way of saying that my vocal performance ought to include more big, assertive gestures and sounds.

Upon reflection, I decided there was indeed a mis-match between the bold performance implied by Western vocal tradition (particularly *bel canto*, in Soper's case) and the small, quiet affect of my performance. However, since I am especially interested in small, detailed vocal and electronic sounds, I elected to address this mis-match not by performing more dramatically, but by building a controller to emphasize these minuscule and mouth-centric sounds. The Abacus evolved through several drafts, with design advice from Peter Bussigel, a composer, media artist, and maker in residence at the University of Virginia. All versions have used an Arduino Teensy, which sends control data to MaxMSP via USB connection and Serial protocol.

I consider version 3 the first significant Abacus version. It consists of an Arduino Teensy, eight toggle switches, one potentiometer, and two LEDs. All components are affixed to the microphone clip using thermoplastic. I gave three public performances with this Abacus in September-October 2016: the 2016 conference of the Irish Sound Science and Technology Association (Derry, UK), Radiophrenia (Glasgow, UK), and Twisted Branch Tea Bazaar

(Charlottesville, Virginia). In addition, my album track “couldn’t” is centered on playing this version of the Abacus.

Figure 2. Abacus version 3

Achieving a manageable control mapping between Abacus version 3 and the Max patch Vox5 was a notable instance in which the physical object prompted evolution of the software. My early attempts at mapping failed to acknowledge what the patch did well or how it might effectively develop. For instance, I imagined one processing state in which Toggles 3-4 would control sounding diatonic interval, ranging from binary output value 0 = mostly intervals smaller than a minor third, to 3 = mostly intervals larger than a major sixth. The patch was never equipped to output particular intervals. When I attempted to introduce such precise pitch control to the patch, I felt I was going too far afield from the existing granulation and noise specialization.

Ultimately I settled on the following control mapping for Abacus version 3 and the Autumn 2016 version of the Max patch:

Toggle 1	**Toggle 2**	**Toggle 3**	**Toggle 4**
listen for undertone singing	solo Layer-2 voices cyclically	random vocal processing	input to Layer-1 = voice signal / Layer-2
Toggles 5-8 each address a different Layer-2 voice These toggles control one of two parameters, depending on Patch State: **State 1** rhythmic/ pulsar granulation **State 2** (always pulsar) listen to/ ignore larger envelope rhythm			
Toggle 5	**Toggle 6**	**Toggle 7**	**Toggle 8**
Layer-2 voice 1	Layer-2 voice 2	Layer-2 voice 3	Layer-2 voice 4

Table 1. Mapping: Abacus version 3, Autumn 2016 Max patch

My memory of this mapping became something of a mantra: "listen – solo – voxproc – route || 1-4."[2] The decision to map Toggles 5-8 to the four voices doing pulsar granulation proved especially important for both musicality and mnemonic purposes. That decision coordinated with the creation of two overarching patch states which differed in timbre, rhythm, and tempo.

The physical experience of interacting with the Abacus in several performances provided much useful information about the patch and the sound, and influenced both how I performed and how I later developed the patch. Most significantly, using the Abacus to alternate rhythmic/pulsar granulation brought to my attention that I often preferred the sounds of pulsar granulation over rhythmic. As a result, the current patch, Vox6, now achieves pulsar sounds quickly. Likewise, because Abacus 3 allowed rapid alternation between Patch States 1 (fast rhythmic envelopes) and 2 (slow rhythmic envelopes), I readily noticed that the fast rhythmic envelopes sounded awkward and nervous relative to the patient, contemplative sound of the slow

[2] See ensuing discussion about intersections between text scores and the internal language of extended voice practice.

envelopes. This prompted development of a probabilistic rhythm envelope system, nuanced by variations in tempo and envelope length, in Vox6.

Finally, it is noteworthy that in performances with Abacus version 3, I also engaged with the laptop. I often looked at the screen, which presented much useful information, including values of rhythmic granulation and filtering, metering volume level, and recording start/ end. I frequently hit laptop keys ‘r’ (record), ‘e’ (trigger parameter sets), and ‘p’ (generate pulsar sounds), and I dragged the trackpad to control gain. Although it was interesting to perform as a laptopist, I desired more visual interaction with the audience and more manual interaction with the Abacus.

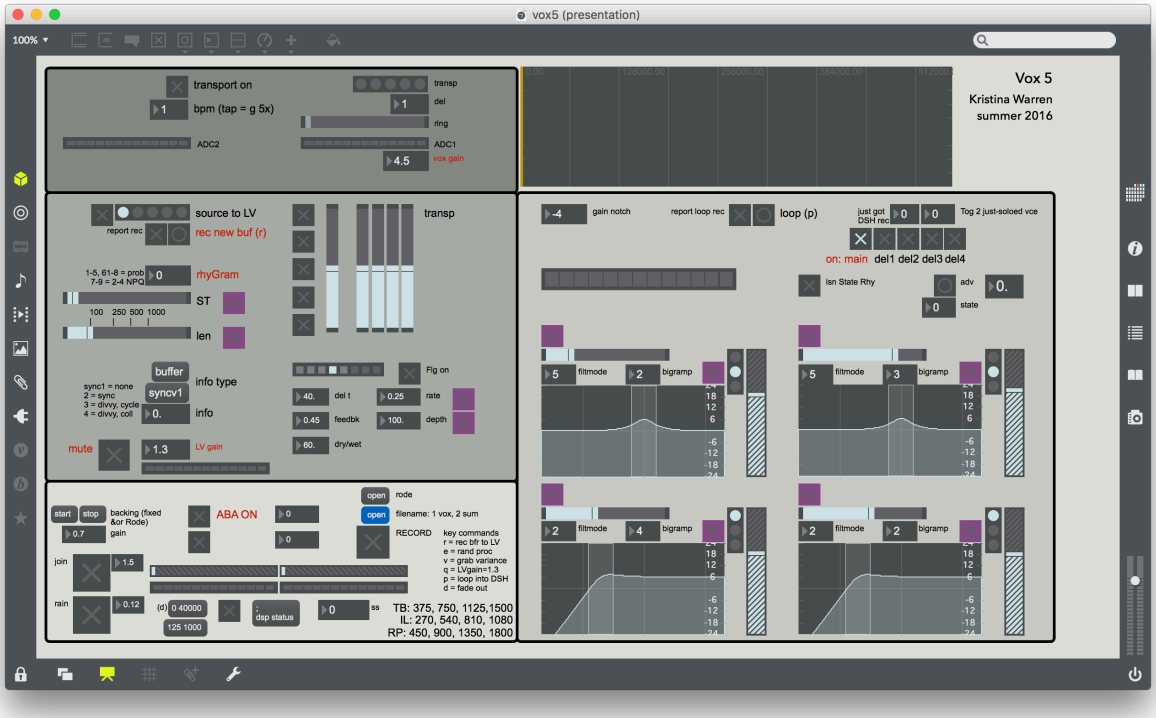


Figure 3. Vox5 Max patch (with Abacus 3): viewed during performance

Abacus version 4 still includes eight toggles and an Arduino Teensy. It adds a button and three potentiometers (now four potentiometers total), and removes an LED (now only one present). It also uses less thermoplastic coverage so that some of the wiring is visible from

below. I gave two public performances in March 2017 using this version: the International Women's Day event of the Yorkshire Sound Women Network (Huddersfield, UK), and the 2017 Guthman Musical Instrument Competition (Atlanta, Georgia).

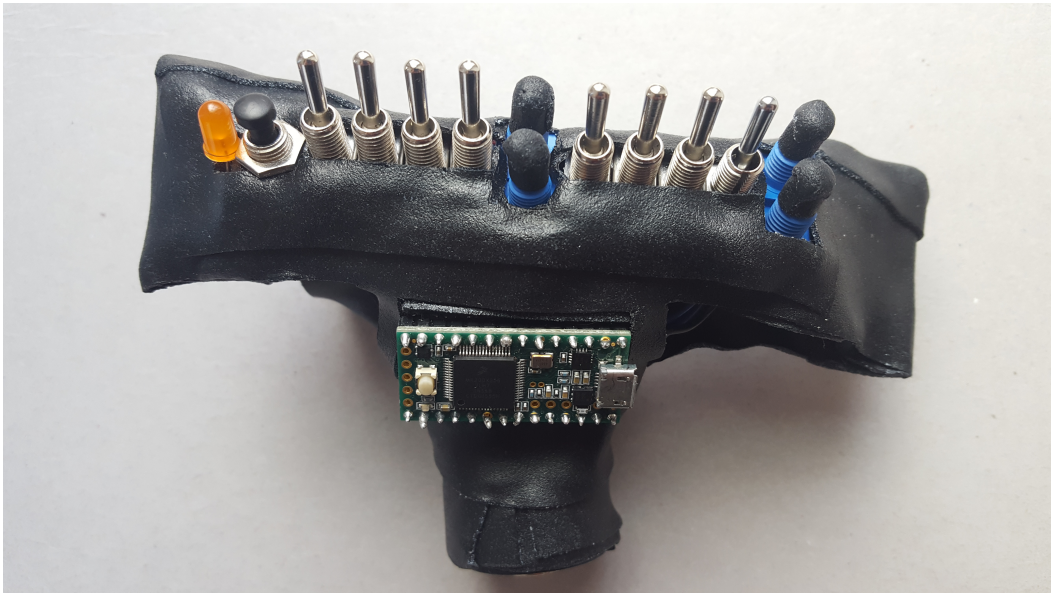


Figure 4. Abacus version 4

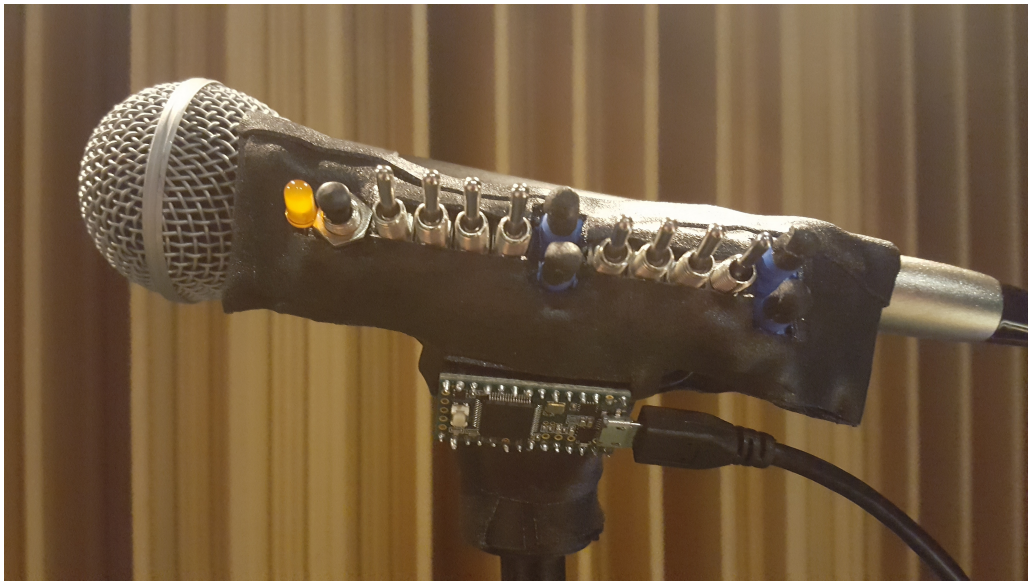


Figure 5. Abacus version 4, on mic stand

In the control mapping of Abacus 4, control information is variably routed to four distinct areas of the patch. I flip toggle 1 one, two, three, or four times to attend to the Voice signal, Rhythmic granulation, Pulsar granulation, or 2d.wave~ processing respectively. There are several key points about this mapping. First, redundancy aids memory. For instance, potentiometer 4 always controls gain, whether of voice, rhythmic/pulsar granulation, or 2d.

Toggle 1 routes control data to: Voice, Rhythmic gran., or Pulsar gran. Fast-flip Toggle 1 # times to specify routing: 1 flip = Voice 2/3 flips = Rhythmic/Pulsar gran. 4 flips = Wavetable						
	Btn.	**Toggle 2**	**Toggle 3**	**Toggle 4**	**Pot. 1**	**Pot. 2**
Vox	hit = rec. live sample hold = automate ctl.	process/ clean	rand. processing	RESET		
Rhy.		"	"	un/mute	ST	len
Puls.		"	"	"	"	"
2d	master fade in/out	"	"	"	freq.	
	Toggle 5	**Toggle 6**	**Toggle 7**	**Toggle 8**	**Pot. 3**	**Pot. 4**
Vox	transpose	pong~	delay			gain
Rhy.	solo	no env., i.e. pops	ctl. var.	rec. to buf. for puls.	tempo	"
Puls.	"	rhythm & density	"	change buffer	wait	"
2d	listen ctl. freq.		LED on/off	"		

Table 2. Mapping: Abacus version 4, Vox6 patch

In addition, I cluster multiple parameters onto a single component whenever possible. For instance, when attending to Pulsar granulation, Toggle 6 controls "rhythm and density," which in the patch is a constellation of random tempo, tutti versus soli rhythms, frequent or infrequent

notes likely, and random ratio of sound to rest. The decision to cluster parameters together is based on my frequent previous experiences of having made too many parameters controllable during performance, and thus having difficulty remembering all parameters while still attending to artistry. Clustering parameters and using randomization reduces what I have to remember in performance, and gives me something interesting and partially unpredictable to respond to. Finally, this mapping includes blank space, i.e., room to add more parameters.

The updates in Abacus version 4 have greatly influenced my performance practice. The simple addition of a button and three potentiometers greatly increased control ability. The accompanying Max patch, Vox6, shows little information in presentation mode, and the Abacus mapping is deep but memorably structured. Thus, I spend more performance time and energy engaging with voice, Abacus, and audience than with laptop.

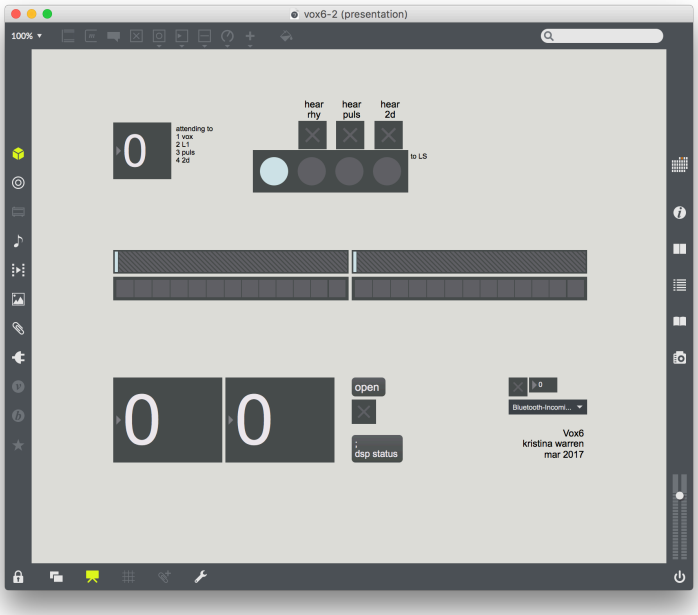


Figure 6. Vox6 Max patch (with Abacus 4): viewed during performance

When performing with Abacus version 4, I have the strong sense that I am playing the Abacus as instrument, and moreover improvising with the voice-electronics as cohesive whole. If

I have a particular sonic idea, I can usually execute this easily through only a few toggle flips. But equally, in the process of getting to this desired state, I may happen upon a more suitable or more interesting sound, and I may stop there.[3] The Abacus helps me discover unpredictable sonic possibilities, and demands that I continually work to coax a dialogue between voice and electronics.

In addition, I increasingly seek out unique choreography for using the Abacus. For instance, pressing and quickly releasing the button triggers recording and granulation. But pressing and holding causes the patch to "listen and emulate control variance." For instance, if I hold the button and continually vary the Start Time potentiometer, the patch will register this and execute similar Start Time variation (rate, depth) once I release the button. Notably, this demands particular choreography: I must bring my left hand up and over to hold the button while the right hand manages the control component. On an immediate level, this is a visually noticeable gesture which signals the onset of automation. In a larger sense, it is, reflexively, a configuration of body that highlights the dialogue between technology and body in my work.

The Abacus requires further development. My primary long-term goal is to shift the focus even more toward the mouth. This could involve replacing manual control with, for instance, a short-range infrared proximity sensor to read embouchure as control information.[4] And/or, I might remove the laptop from the setup, i.e., replace Arduino with Bela. This standalone Abacus would be more portable and would permit continued exploration of the controller as instrument in voice-electronics performances.

---

[3] For instance, when I change patch state, potentiometer information updates to the new state but toggle information does not. This is a source of interesting unpredictability in navigating the Abacus/Max setup.

[4] An important objection to the infrared sensor/embouchure approach is that it may be too ocularcentric. Scholars including Adriana Cavarero (2002) and Annette Schlichter (2011) have written of the pernicious ocularcentrism which has for centuries dogged Western culture and, in particular, perceptions of the vocalizing female body.

C. Vox6

Vox6, my current MaxMSP patch for voice-electronics performance, consolidates the musically successful parts of previous patches and expresses several of my musical goals.[5] First, I desire deep, recursive processing that retains something of the input timbre and temporality but also suggests new timbre and temporality. Second, I desire a rich sonic field, i.e., the ability to make the sound more or less dry, distant, incisive, etc. I achieve these two goals through a combination of live sample manipulations (granulation and two-dimensional wavetable) and processing of the un-recorded signal (distortion, transposition, delay).[6]

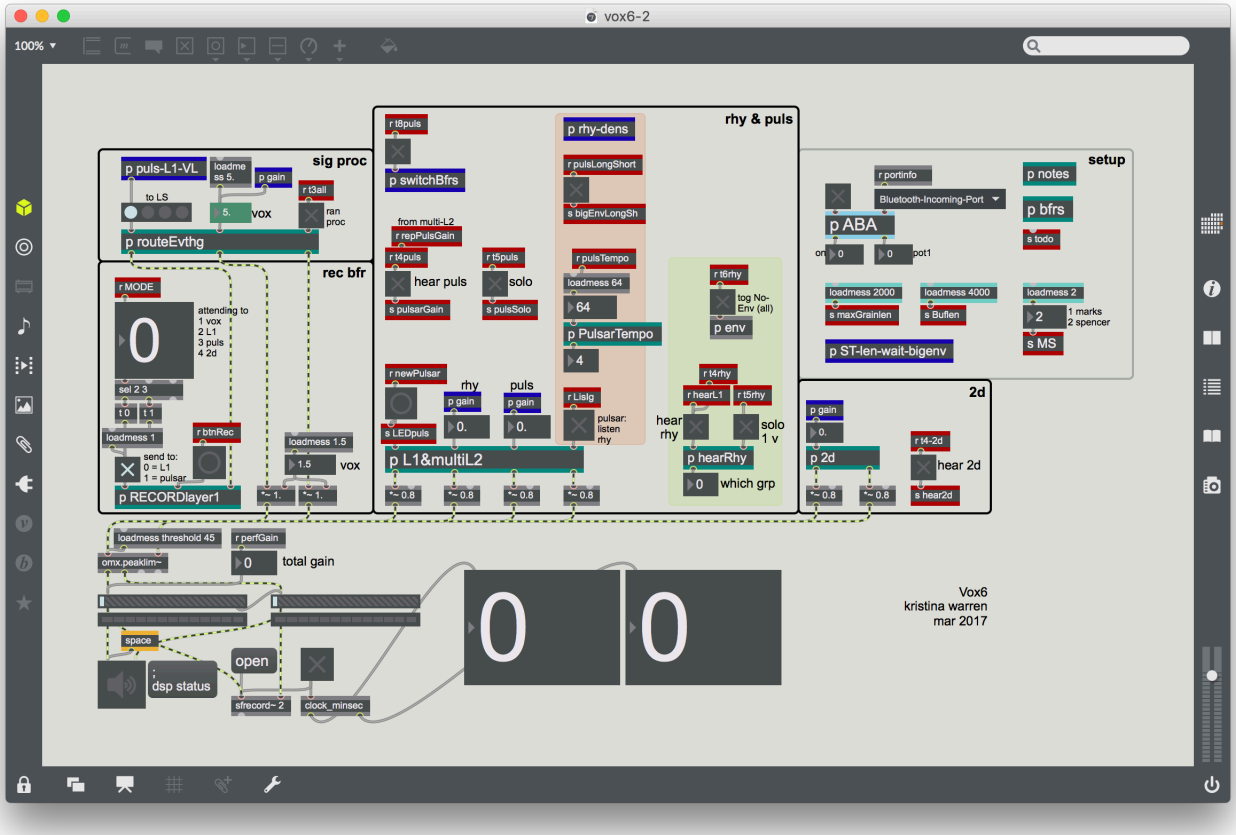


Figure 7. Vox6

[5] See Appendix E for screen shots of the full patch.

[6] From a digital perspective, the difference between 'live' and 'recorded' sound is trivial or nonexistent. But from an aesthetic perspective, I use live-signal treatment to give depth of field to complement granulation's forward, dry sonic output. This combination of live-signal and sample-based techniques contributes to my argument that 'live' and 'recorded' methodologies overlap in extended voice practice.

Inspired by Sam Pluta's (2012) discussion of modularity as a key component of his composition and improvisation practice in SuperCollider, I made Vox6 modular and recursive. Any of the four modes (voice, rhythmic granulation, pulsar granulation, wavetable) can be granulated, wavetabled, or live-processed. Since all electronic sounds are live sample-based, the digital output is fundamentally reliant on input from my vocal body. But I aim for complex dialogue between voice and electronics: the voice feeds the system but also does much more, and the electronics have a distinct sonic identity despite their derivation from voice. The electronics are a prosthesis to my vocal body because they expand and inform what I the composer-performer can physically (sonically) do, and the resultant vocal-electronic body is a new cyborg whole. But within this whole there are also differences between voice and electronics, particularly in what each does well, so there is complex dialogue between them.

The ensuing discussion attends to granulation (mainly pulsar, with brief discussion of rhythmic for context) and treatment of the live signal. I supplement these two primary functionalities with 2d.wave~, treating live samples. I use this wavetable object as a source of continuous sound to contrast the sound of granulation, which on the whole is stuttering and non-continuous. This object easily accepts scaled potentiometer input as a sort of frequency/timbre control. The wavetable module in Vox6 resembles my Bela-Pd work in that it is computationally lightweight and easy to perform with.

1. Pulsar granulation

Pulsar granulation sounds best to me when it is treating a buffer consisting of rhythmic-granulated voice, i.e., a buffer which already includes many sidebands and digital audio pops. Pulsar-granulating a straight voice sample, by contrast, sounds too plain. Thus, I always step

through rhythmic granulation on the way to pulsar. Because the probabilistic structure of my rhythmic granulation can yield occasional silences, I use four voices doing the same type of rhythmic granulation to sum to one buffer, whose consistent volume is effective for pulsar. In other words, in order to achieve four complex and sonically interesting pulsar-granulation voices in parallel, I use 16 rhythmic-granulation voices divided into four groups of four.

I begin each performance with vocal improvisation. After some amount of time, I trigger recording-and-pulsar (Abacus button). From trigger to pulsar takes eight seconds: four seconds to record a live vocal sample, and four more to do rhythmic granulation,[7] which, as it occurs, is recorded into a new live sample to be pulsar granulated. By default, rhythmic granulation does

t = 0 s	hit Abacus button to trigger record-and-pulsar start recording voice (buffer 1)
4 s	stop recording voice (buffer 1) send this buffer to rhy. gran. (muted) random parameters of rhy. gran. start recording rhy. gran. #1 (buffers 2a-d)
8 s	stop recording rhy. gran. #1 (buffers 2a-d) send this buffer to pulsar gran. hear pulsar gran. new random parameters of rhy. gran. start recording rhy gran. #2 (buffers 2e-h)
12 s	stop recording rhy. gran. #2 (buffers 2e-h)

Table 3. Events after triggering record-and-pulsar

not send to audio out. Thus, the first electronic sounds, at eight seconds after trigger, are pulsar granulation of rhythmically granulated voice. Also at eight seconds after trigger, a second round of rhythmic granulation occurs with different parameters (i.e., yielding different timbre), and this

[7] See Appendix C for description of rhythmic granulation parameters.

records into a new live sample for pulsar. Rhythmic granulation is still muted, so these new samples are a silent timbral alternative on hand for pulsar.

Once pulsar is running, I often enter a freely navigable phase of using the Abacus to sculpt the sound. To sculpt pulsar, I flip Abacus toggle 1 three times to enter pulsar mode, meaning the Abacus components communicate only to parameters of pulsar granulation. For instance, if I want a more sustained buzzing sound, I lower the value for grain Length (potentiometer 2). This becomes the center length around which each of the four Pulsar voices jitters (every 8-14 seconds). If I want a new rhythm/density profile, I flip toggle 6. If I dislike the current pulsar timbres, I can switch to the extra set of rhythmic-granulation live samples (toggle 8), which were silently recorded starting eight seconds after recording was triggered.

Meanwhile, I can easily switch to another mode, perhaps sculpting the vocal signal. I flip Abacus toggle 1 once to enter voice mode. If I want transposition-delay, I flip on toggle 5. This randomly selects one of two configurations: short bits of signal, short or long delay, large transposition (arpeggiated vocoder-like); or long portions of signal, short delay, small transposition (flange-like). Then I can quickly get back to sculpting pulsar by flipping toggle 1 three times to return to pulsar mode.

In addition to sculptability, the main advantages of the pulsar functionality are its speed and depth. In Vox6, it takes only eight seconds to record a source buffer, then do rhythmic granulation while recording this to a new buffer for pulsar. In previous patches, this process took several minutes. Initially, I had to vary all fourteen rhythmic granulation parameters by hand. I found this was far too much to manage while still vocalizing artfully, so I simply composed a control routine and stopped bothering with most live control.[8] In a later patch iteration, I introduced the ‘hit e’ functionality, i.e., a single laptop keystroke to trigger a random cluster of

[8] See Appendix B for sample control routine.

rhythmic granulation parameters. But I still had to manually approve each rhythmic granulation and record it to a new buffer for pulsar. In both of these patch versions – sculpting a parameter set by hand, or blasting random parameter sets until I found one I liked enough to record to pulsar – it took several minutes to build up a choir of four pulsar voices.

This slow speed in earlier patches caused an idiosyncratic performance structure. For instance, in Vox5 (Autumn 2016, corresponding with Abacus version 3), the first few minutes of each show were necessarily a preface section in which I established the pulsar voices. This felt so clunky that I began to experiment with creating the pulsar sounds immediately before a performance, or doing all the pre-work in my home studio and then performing with only the last part of the process, pulsar granulation (Bela, Pd). But along with the performative awkwardness, the slowness of getting to pulsar raised fundamental questions about liveness. Is it more 'live' to make the pulsar sounds at home when rehearsing, or right before the performance, or during the beginning of the performance? Do certain of these methods particularly enhance or hinder my ability to connect with this digital performance partner in real time?

Vox6 is purposefully fast, a fact which is directly linked to the depth of the granulation system. I use 16 rhythmic-granulation voices in parallel in order to achieve four pulsar-granulation voices in parallel. If I do not like a certain sound, my options are to mute it or move to the next sound. I no longer waste valuable mental energy[9] on sculpting the perfect intermediary sound. This speed-depth combination also informs the vocal improvisation I can do. When the patch was clunkier, I had to record a weirdly varied vocal improvisation – with the full pitch and timbral gamut represented – so I could be sure to wind up with a wide range of pulsar sounds. This was not always aesthetically pleasing. Now that pulsar is achieved relatively

[9] It is inherently political for a woman to discuss using her "valuable mental energy" for artful technological manipulation of her voice. This gets to the heart of my term *presence*, which I discuss in detail later.

quickly, I can vocalize in a way that more naturally fits the musical moment. If pulsar become stagnant, it only takes eight seconds to change this.

Eight seconds to a significant change is certainly fast enough for my solo performances. To my satisfaction, my solo sets to are increasingly 'stream of consciousness', moving smoothly from one idea to the next, in large part because of the speed and depth of pulsar. I am curious to try this patch in a group setting. My previous attempts at group improvisation (various groups, various earlier patch versions) were challenging. It was difficult to complement – much less match – other players' pitch and rhythmic material. But with Vox6 I can more quickly achieve variations in timbre, texture, and density. I can, for instance, vocally match a pitch from other players, and eight seconds later get this pitch out of pulsar in digitally refracted form (i.e., noisier and/or with added harmonics). Though Vox6 is mainly oriented toward solo performances, an important upcoming research step will be to test it in group contexts.

2. Live signal treatment

LiveSig, the part of the patch which processes live signals in Vox6, consists of four modules: distortion (pong~), transposition-delay (2 settings), reverb (dense delay), and sparse delay (1-5 taps). Any module can be used singly, or a pair can be used together: either pong~ or transposition-delay, plus one of the delay settings.

The transposition-delay module, which uses a poly~ object (48 voices), has two settings with distinct envelope length, delay amount, and transposition amount. The first setting, which vaguely resembles a vocoder with arpeggiation, treats short bits of signal (300 ms) with large transposition (several steps up or down) and short or long delay (50-100 ms, or 1000-1050 ms). The second setting, which resembles a flanger, treats long bits of signal (2200 ms) with small

transposition (20-50 cents up or down) and short delay (15-100 ms). In both settings, a single trigger to poly~ actually sends in triplicate, yielding a dense, clustered sound. This module achieves transposition and delay by sending a V-shaped ramp to tapout~.[10]

The transposition-delay module, in other words, achieves transposition via time stretching rather than by using FFT. Previous patch versions used FFT for transposition without time stretching, but I disliked the tinny timbres that resulted, and FFT-based work, especially when proliferated in numerous abstractions, was computationally expensive. I generally prefer the timbres of transposition-through-time-stretching, particularly when transposing by only a small amount. In the first setting of transposition-delay, which transposes by a large amount, I combat the timbre issue by using a short envelope length, so the focus goes more to clustering than to individual timbres.

I included the pong~ object for several reasons. First, I admire the vocal distortion sounds of, for instance, Maja Ratkje and Andrea Pensado, and pong~ outputs similar timbres. Second, I felt that the parallel processing in my previous Max patches – EQ, ring modulation, delay, FFT-based transposition – yielded sound output too similar to the input. By contrast, distortion yields what I consider more complex and interesting sounds, and thus helps me push the boundaries of what voice is and how it can sound. Finally, on a conceptual level, I am interested in the impact on listeners of female vocal timbre. In the 2016 US Presidential Election campaign season, Hillary Clinton was criticized for having a "shrill" voice (Khazan 2016). I disagree with this assessment; pong~ helps me demonstrate how I define 'shrill' and how this sound can be aesthetically valuable.

[10] Ramp values are calculated using empirically derived regression equations. See Appendix D.

D. Liveness, voice-electronics interaction

Liveness is at issue in my work because live samples and live signal are processed to create a digital improvisation partner of sorts. In my recent work I consider whether granulation, specifically the actions of recording and sculpting granulation parameters, does or does not contribute to liveness. If neither the initial vocal utterance nor the live processing of this live sample is included in the track, is it really 'live'? Or, does advance preparation of the buffers-for-pulsar, then sudden un-muting of these buffers during the piece, give a greater sense of spontaneity? In other words, there is a temporal remove or distorted liveness implicit in my Autumn 2016 Abacus3-Vox5 setup. I have responded to this in my live work with the Pd-Bela setup and the Abacus4-Vox6 setup, and in my album tracks with their varying balances of improvisation and composition.

My vocal sounds, ranging from singing to extended techniques, occupy a spectrum of normativity. In my early voice-electronics performances, pitched diatonic singing functioned as a way of contextualizing, or rendering acceptable, the non-normative vocal sounds. But further performance research (building on extant compositional scholarship, e.g. Wishart 1996) has revealed additional audio-rate oscillators in the vocal tract, such as air against the hard palate, or teeth on lips. In other words, the vocal folds are far from the only source of pitch in the vocal mechanism.

Nevertheless, normative pitched singing and speech do serve an important rhetorical function in extended voice, suggesting archetypes such as "opera singer" and "militant leader" (Ratkje, quoted in Elektronski 2017). Most extended voice practitioners, including myself and others whose vocal and electronic sounds are often noisy, utilize speech or singing at least occasionally. This use of a wide range of vocal sounds complements the electronic extension of

voice. Technology is an indispensible prosthesis to voice, and, by treating body as a kind of technology, variation in vocal techniques in fact highlights the deep, complex relationship between voice and electronics.

Pitched	
Harmonics (vowel)	slow alternations [u] (harmonics 1-2 above fundamental)
Undertone singing	false vocal folds (inverted harmonics 1-2 below fundamental)
Ululation	repeated chest-modal register alternation, avg. rate 4 Hz, avg. depth P4
Sing on inhale	pure or noisy (resembles bit-crushing with harmonics 1-2 above fund.)
Lip squeak	moist lips; upper teeth rest on lower lip; inhale; polyphonic
Pursed lips	pitched squeak, (sub)audio rate
Pressed exhale squeak	oscillation high in vocal tract/ post-nasal

Unpitched / noisy	
Duck call	exhale, air resonates in mid-back of hard palate near right molars
Glottal stop	beginning/end of note
Lip buzz	usually sub-audio rate, occasional audio rate (octave 3)
Pressed voice/fry	effective with discrete notes; can merge into noise oscillator
Epiglottal click	in- / egressive, single click or several in rapid succession

Table 4. Extended vocal techniques (selected), Warren

Several of my frequent thoughts while performing, which guide my realization of voice-electronics dialogue, are: “needs more/ different,” “enough,” “stop,” and “copy.”[11] It is difficult to pinpoint exactly how these frequent self-instructions translate to sound, but I will attempt to explain by providing several performance examples.

(1) I have been doing short lines (3-4 seconds) of egressive epiglottal clicks, separated by pauses for breath (5 seconds), for about thirty seconds. Slight timbral variety is needed, so I introduce ingressive epiglottal clicks which occasionally break to short pitched inhalations (half a second each).

(2) I am experimenting with delay and distortion on the voice (Bela, Pd). I notice the pulsar sounds have gotten a bit stagnant, so I quit vocalizing for the moment. I turn pulsar to treat only the inhaled sample, and I vary start time. The last 250 ms of the buffer, I quickly recall, were almost silent, so this yields a break in the sound output which I use for phrasing.

(3) I trigger record-and-pulsar, but when pulsar starts eight seconds later it is too busy relative to the voice. I flip rhythm-density (Abacus toggle 6) several times until finding a calmer rhythmic state, and lower grain length and wait time to give low-pitch buzzing sounds.

(4) Rhythmic granulation is at a high tempo, yielding a clicking, distorted sound, so I attempt a messily structured line of consonants and saliva to emulate the digital noise. This vocalization sounds a bit lackluster relative to the electronics, so I start doing loud popping “dah” sounds instead.

(5) I have granulated a low-range sung pitch, and the grain length is long, so the original vocal timbre is audible. I lower the granulation volume and loudly screech a sustained high pitch

[11] See subsequent discussion of inner voice and text scores.

(circa C6) for contrast. The second time I screech, I flip on delay. I stop screeching and instead turn on the 2d module for a similarly high frequency but more markedly digital timbre.

(6) I have a nice, diatonic (roughly), arpeggiated pad going by continuously singing one pitch and sending this to transposition-delay (first setting). I refine the [u] vowel I am singing, making this tone especially pure. I start sending rhythmic granulation to the same transposition unit. It sounds as though the digital sounds are clumsily trying to keep up with the vocal. I alternate flipping the send on and off a few times, every few seconds, to foreground the halting quality of this poly~ setting.

(7) The whole thing is too much and I want a clean break. I trigger Reset to go immediately to dry voice only, everything else muted.

After performances in which I use noisy vocal techniques, listeners frequently report feeling either mesmerized by the similarity between vocal and electronic sounds, and/or concerned that vocal injury was likely. These attitudes reflect a Western cultural paradox: non-linguistic (i.e., noisy or non-laryngeal) voice has an entrancing, almost meditative potential, and yet, almost instinctively, it reads as unhealthy or disordered. There is very little scholarship on voice-electronics intersections, and popular reviews of extended voice, particularly its noise potentials, heavily favor demeaning or de-humanizing descriptions of practitioners, such as aliens or intruders (Attn:Magazine 2014), primal or feral creatures (Moliné 2007), or menstruating women (Neset 2003). Much of the impetus behind my own extended voice work is to normalize a wider range of vocal sounds, and to invite inclusivity by suggesting that vocal

experimentation can be both healthy (Borch *et al* 2004) and reminiscent of our daily interactions with sound (Demers 2010).[12]

Our impoverished vocabulary for describing our experience of listening to extended voice runs parallel to frequent ocularcentrism in assessment of the vocalist. For instance, a listener (non-vocalist computer music composer) once commented that they initially felt concern that my use of extended techniques would be damaging, but on observing my shirt billowing in and out, they realized I was using good diaphragm-centric technique to effect bursts of sound, and injury was unlikely.[13]

Since my medium is sound, I prefer to be assessed aurally. I dislike musical assessments based on my appearance, which I believe reinforce erroneous notions of the normative female body and what can be done with it. One approach I have begun to attempt is didacticism: the sound itself somehow demonstrates to listeners that they should engage with my work aurally, not visually. But I continue to have performance experiences in which, despite my best efforts, much of the audience seems to care more about how I look than how I sound. Thus, I am wary of over-reliance on audience perspectives. Instead, I keep returning to the personalized, sonic goal of achieving deep, meaningful dialogue between voice and electronics.

E. *filament*

My album *filament* began as a compilation. I imagined that each track would feature distinct vocal techniques, technical methods, or compositional approaches. But gradually my

---

[12] Demers (2010) argues that electronic music is especially conducive to *aesthetic listening*, i.e., intermittent listening in which attention may freely turn to extra-musical sound. In fact, electronic music encompasses seemingly external sounds by acknowledging their aesthetic character. Because we use our voices in everyday contexts, extended voice has the unique potential to connect to daily sound experience and participation.

[13] Laura Mulvey's (2009, orig. 1975) now well-known theorization of the male gaze is germane here.

thinking about the album evolved. A significant inspiration was Ami Yoshida's album *Tiger Thrush*, whose 99 "howling voice" miniatures contain and articulate pauses which question not only the sound but also the temporality of musicking. I concluded that artful variation of sounding and silent durations would help frame the wide stylistic variety I desired for *filament*.

Another key inspiration was Ratkje's album *Voice* (2001), which she co-produced with experimental duo and frequent collaborators Jazzkammer [NO]. Tracks 1, 5, 7, and 11 (of 11) each use one or both of two similar materials, which are distant and ethereal, with a looping organ-like melody line. This multiple recurrence of identical or similar sounds acts as a linking device from beginning to end of the album. Similarly, I worked to develop sonic continuity across my album, though not in quite as literal a fashion as in Ratkje's album. For instance, my tracks "ul" and "For one" both use dense layering of voice, without additional electronics. Other tracks represent distinct versions of the Max-Abacus setup, or various uses of the Bela-Pd setup.

An important consideration in my album is the idea of voice and electronics as source material for each other at the DSP level. For the track "quantum," I did a noise improvisation and a separate vocal improvisation using the Bela-Pd setup, then made several subsequent iterations using each as the modulator signal for the other, with various smoothing algorithms. For "eager to die," my Bela-Pd solo performance (with AKA trio, December 2016, Manchester, UK) became the input to a 2d.wave~ object. I dynamically varied input phase values to the wavetable in order to control timbre. In sum, composition, improvisation, 'live', and 'recorded' are multiply combined and re-combined in this album.

III. SURVEY OF EXTENDED VOICE

A. Antye Greie-Ripatti (AGF)

Antye Greie-Ripatti (AGF) [FI, DE] creates work engaging with, in her words, "voice, noise, off-beats and radical sounds and this weird listening aspect involved with club aesthetics" (Taylor 2016). She is prolific, collaborates often, and works in a variety of media including albums, installations, and field recording trips.[14] This discussion contrasts tracks from two of her albums, and provides brief context about her recent practice as a whole. AGF's live setup generally includes laptop, mixer, handheld microphone, and one or more control surfaces. She has two primary modes of performing, which I term "sound poet" and "DJ," that involve different proportions of vocalizing and controlling. In sound poet mode, AGF mostly vocalizes (sing-song speech) but frequently taps a touch pad to control delay and reverb on the voice, for instance. In DJ mode, AGF vocalizes little and undertakes complex choreography to achieve live mixing, rhythmic triggering, etc. Both modes can occur in the span of one performance.

The third track on AGF's 2003 album *Westernization Completed*, "PRIVATEbirds," is a recorded manifestation of what I term her "sound poet" mode of performing live. AGF's acoustic voice is prominent throughout the track as she sings-speaks the lyrics. The pointillistic soundscape moves quickly but mostly monophonically from one melodic sound to another. As the piece progresses, gradually more and longer pauses intervene.

Verses and choruses bear slightly different harmonic material. Verses emphasize electronically synthesized pitches C and F in a seeming dominant-tonic relationship. There is a registral division, such that most Cs occur in a middle or treble register, while Fs occur in a bass

[14] For instance, AGF writes: "I facilitated the Sonic Wilderness camp in Hailuoto [Finland] 2016" (Greie-Ripatti 2016).

register. Choruses, on the other hand, emphasize diminished fifths. For instance, a progression of minor thirds after each repeated half of the chorus – E4-G4, G4-B♭4, A3-C4 – creates a sense of harmonic instability which complements the contrasting desires for isolation or companionship expressed in the lyrics.

Textural and temporal choices give "PRIVATEbirds" a feeling of suspended animation. This track has few sounds at once, and the tonic pedal is elaborated chordally rather than timbrally. Electronic manipulations of voice, such as granulation and distortion, are infrequent and soloistic. This too aligns with AGF's "sound poet" mode, which puts voice in greater relief from electronics than does "DJ" mode, which values the holistic soundscape.

Track 9 of AGF's 2016 album *Kon:3p>UTION to: e[VOL]ution*, "Dis Dance," resembles "PRIVATEbirds" but partly bridges the gap from "sound poet" to "DJ" performance mode. This track contains a low tonic pedal of about 169 Hz (E quarter-sharp 3). Occasionally, background percussive sounds and syllables of AGF's vocal performance are tuned to the mediant at 214 Hz (G 3/4-sharp 3), articulating a major key area. The first minute includes the tonic pedal and various background sounds, primarily clicks and scrapes, often with 2-3 seconds of audible reverb. Enveloped bits of voice singing [ü] occur from 0:07-0:12 at the high subdominant, about 452 Hz. Various scraping ostinati, all at differing speeds, occurs starting at 0:25, growing out from similar rhythmic action in the pedal tone. AGF's vocal performance begins at 0:52 and, though spoken like much of her vocal work in previous albums, contains an entirely new optimism. Perhaps this is because she tunes her voice, both technologically and acoustically, to the prevailing major key. I can imagine her (or my?) face upturned to the sun while vocalizing.[15] Rumblings of the lower tonic, two octaves down, enter at about 3:00.

[15] Ocularcentrism is, unfortunately, a deeply entrenched response to voice. For instance, Connor (2000) suggests that listeners customarily imagine the vocalist's body when listening to recorded voice. I am

AGF's recent work seems to demonstrate particular investment in recorded media and discussion-based live formats, rather than 'live performance' per se. For instance, the performances page of her poemproducer website notes, "I try to move less. fly less. love more. but i still sing for you. …" (Greie-Ripatti 2017c) Likewise, another of her sites, antyegreie.com, categorizes her work using many types of projects seemingly distinct from strict performance, including Interpretation, Radio Commission, Performance Score, Curation, Single, and Voice (Greie-Ripatti 2017b). Though the album format has consistently been crucial to AGF's output, it is important to acknowledge that her work also innovates in myriad intersections of composition and liveness.

Relative to extended voice as a whole, AGF's work contains much more obvious political and social valence. In the past this has taken a more aesthetic path. For instance, the 2003 album *Westernization Completed* poetically considers intersections of language, identity, and technology. But as AGF's social work has become more incisive, for instance through her female:pressure blog featuring women using music technology, so too have her albums become more politically direct.

B. Marie Guilleray

Marie Guilleray [NL, FR] is an expert in creating space for subtle vocal sounds and using electronic extensions to imbue each vocal sound with a unique temporal profile. Her minimal setup – laptop and microphone, and sometimes her novel glove controller – achieves an impressively wide range of sounds. Guilleray stands relatively still, looking slightly upward while singing or gazing intently at the laptop when changing presets (laptop interaction occurs at

---

interested in how extended voice re-values listening instead of seeing. But body is still germane; even I can't avoid thinking about it sometimes while making or listening to extended voice.

least once per minute). Her streamlined approach to the electronic prosthesis draws attention to the temporal interactions between acts of vocalizing and acts of controlling, and to the timbral similarity of their resultant sounds.

Guilleray's solo composition-performance *Entre chien et loup* (2012; voice, tape, live electronics) is a deconstruction of harmonic ideas, vocal timbres, and the act of communication. The piece is characterized by distinct formal sections, many of which are introduced by a struck metal sound and/or a recording of crickets. Each section emphasizes either harmonic or percussive content: possibly a formal-level creative interpretation of vowels versus consonants in voice. Most of the piece foregrounds the indistinguishability of live and pre-composed sound. The application of memory technologies to the live vocal signal (for instance, delaying enveloped bits of sound of varying lengths, thereby creating a morphing collage) brings voice into the digital realm, emphasizing the similarity and cooperation between vocalizing body and electronic prosthesis.

Early sections are prominently intervallic. For instance, a plucked B4 begins the piece and quickly decays to a gong-like dyad, during which the voice sings [a] and [u] sounds on F#4, G4, and G#4. At 0:40, a struck metal followed by a recording of crickets, apparently tuned to F# and G# in octave 5, marks a new gesture. The next gestural section, beginning at 1:30, lasts about as long as the first two combined, and uses a sharply prepared by a struck metal both to prepare and elaborate this new material. Gentle vocal hisses and exhalations, multiply layered, enter after the first metal strike. Panning plosive [k] sounds prepare another struck metal sound, this one thinner in timbre (i.e., more spectral energy concentrated in higher frequencies) than were the previous metals. Text seems present but is unintelligible, and each consonant or

whispered vowel is given unique delay-based treatment. Some delay clouds are brief, while others adopt a quasi-pedal function after settling into a quiet feedback loop.

Middle sections become more spectral in that they are less confined to the chromatic scale. The quieter dynamic level around 3:00 is a subtle marker of a new section. Vocal and piano material take up new pitches and slightly faster rhythms in similar registers as before. At 3:47, a B4 pluck and crickets initialize a gesture similar to the first gesture of the piece. The voice soon returns, now (4:50) singing [a] on B4 and [u] on A#4, a tonic-leading tone oscillation which suggests some gravity but, in its spectral rather than functional harmonic context, is more unsettling than cadential. A high sine tone with tremolo, C quarter-sharp 6, underscores the microtonal character of this section.

The conclusion of the piece is unique. Ingressive and egressive breathing through cavernous embouchure, paralleled by continuous salivated lines, occurs during the last minute of the piece. The cavernous mouth is evocative but mysterious, suggesting something evil, or perhaps a state of injury or illness. Meanwhile, the saliva sounds are less emotive: primarily a signal of the diverse and unusual sounds the body can make. This embodied juxtaposition of emotion and quasi-scientific demonstration is powerful, and possible only through the electronic prosthesis which expands the temporal and timbral capabilities of the vocalizing body.

Guilleray frequently improvises with other electronic musicians. Her performance “Après la mer” (2014) with frequent collaborator Johan van Kreij is a prime example. In the beginning of the set, Guilleray’s contribution consists of sung downward glissandi beginning around E4. Variable delay and transposition create a gentle microtonal cloud of digitally processed voice around her live performance. Around 1:15, the voice becomes more of a cloud, consisting of distant delays around pitches B3-C4. From the audio recording alone, it is unclear whether this

material derives from live vocalization or pre-samples. Guilleray's novel glove controller, designed and built in collaboration with van Kreij, potentially enables either live or sampled work. (Guilleray's Masters thesis (2012) discusses both granulation techniques and live signal treatment.) This sustained texture, which has discernibly vocal timbre and [u] vowel content but digital-sounding continuity, breaks into enveloped bits of sound at the end of the line.

Next, Guilleray whispers quiet but forceful consonants (e.g., "sah!"), which are delayed; she then moves to humming B3, then singing pitches around B5 in a thin tone. Chorusing of these high pitches begins, and this material again dissolves in short bits of sound. Guilleray returns to [s] sounds, which are active across the stereo field, for a short time. Pressed-voice quasi-speech, almost glossolalia, begins about 4:45. Consonants, slightly more salivated than before and now punctuated by occasional screeches, complement van Kreij's choppy, quiet, noise-like material. After this, there are about two minutes of material which is not discernibly vocal; from the audio recording alone, it is difficult to know if this is a solo by van Kreij, or if Guilleray contributes some sounds.

At 8:00, Guilleray takes up sung pitches again, beginning with A4-G#4 lines, both live and delayed, using vowels [i] and [a]. The slightly whiney timbre of this singing alludes obliquely to the characteristic sound of much vocoder work, a subtle but sophisticated gesture which instantly puts the voice and the digital into close dialogue. This material transforms into a vocoded line, which has a slightly sad or introspective character. This line jumps up an octave and is soon replaced by lip trills and duck call-like hisses produced by breath and saliva. The conclusion is frenetic, with screechy interruptions to whooshy saliva sounds complicated temporally by variable delay.

C. Stine Janvin Motland

Stine Janvin Motland [DE, NO] is outstanding in her use of extended vocal techniques; her stamina and range are exceptional. She deploys electronics with a light touch in order to foreground the voice itself. Repetition, either acoustic or looping, is significant. In addition, Motland often uses field recordings, which for her act as a point of engagement with voice as solo versus ensemble instrument. Her live electronic setup includes a mixer and one or more analog delay pedals; wires proliferate, suggesting that she has pre-composed processing states and possibly feedback loops. Motland stands facing the audience, singing into the microphone and moving rhythmically with the music. At important formal moments (at most once per minute), Motland turns to the table of gear at her side and hits a button or twists a knob to change delay values. That Motland keeps this techy-looking arrangement at lateral arm's length is suggestive of her desire to treat voice as central.

In her solo voice-electronics "pop experiment" project Stine II, Motland creates loop-and-variation textures using vocal samples, synthesized loops, and field recordings. Though Stine II is a solo project, field recordings contribute to a collage of which the voice is only one part. For instance, in the track "Born in An Early Age," the looped text "I was born in a very early age" gradually draws attention to the melodic content of speech. New vocal loops enter, containing similar but nonequivalent text, and gradually the voice becomes a rhythmic-melodic texture unto itself. Though Stine II alludes to work by (broadly speaking) 'pop' producers who use pre-existing vocal samples, Motland's work has a very different valence because the sampled voice is her own. In other words, Motland's work as Stine II alludes to the working methodology of the independent producer, but has more presence (as I will later define it) because, in sampling

her own voice, there is no temporal or personal distance to be inferred between producer and vocalist.[16]

> Motland's piece *Fake Synthetic Music* (2016; voice, delay pedals) is, in her words:
>
> a driving imitation of melodic synthetic sequences and an outspoken tribute to past and present pioneers of electronic music and rave deconstruction. A game of true or false, psychoacoustic and high-frequency exhaustion for voice, echo, and spatial distribution. (Motland 2016a)

Figure 7 gives a partial transcription of the NMASS 2016 version (approx. 12') but Motland's website also includes a shorter version, Cynteart 2016, with similar pitch and rhythmic material. In my transcription, regular noteheads represent acoustically sung pitches; x-noteheads, delayed taps.

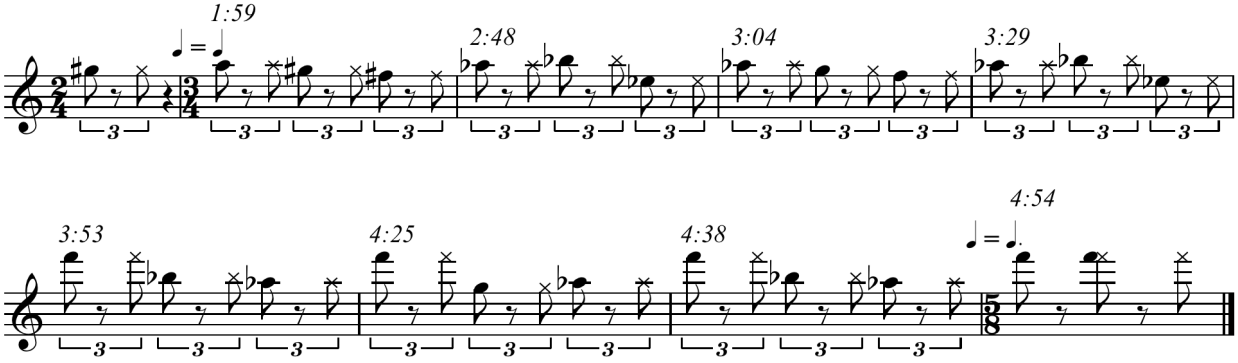


Figure 8. *Fake Synthetic Music* (Motland 2016), excerpt

Timbre and repetition are of the utmost importance in *Fake Synthetic Music*. The very notion of singing breaks down in such a high register and for such a prolonged duration. The variable delay acts as an aid to listeners' imagination. Although the audience sees Motland vocalizing and observes from her stage position that she considers her vocalization central, aural acclimation over time leads to confusion about whether that sound is really voice, and whether it

[16] Motland's album *In Labour* (2014) takes a related but distinct approach to voice use and authorship. This album was produced by Lasse Marhaug [NO], frequent collaborator of Maja Ratkje and co-producer of Ratkje's pivotal album *Voice*. *In Labour* unites field recordings and extended vocal techniques (sometimes looped). Here, though field recordings are used in collaboration with a producer, they take a background position while voice, the soloist, engages with the various indoor/outdoor recording spaces.

is emerging from her mouth in real time. Especially in octave six, the timbre of Motland's controlled, skillful shrieks is almost indistinguishable from synthesized sound. Though Motland writes of exhaustion, her stamina is the more salient feature; her treatment of voice more closely resembles a synthesizer with slight human 'imperfections' (microtones, glissandi, etc.) than, for instance, a vocoded sound which is still centrally vocal in origin. Motland's nuanced dovetailing of electronic and extreme vocal timbres is a unique instantiation of electronic prosthesis to the vocalizing body.

D. Andrea Pensado

Andrea Pensado [US, AR] often performs using – along with laptop, mixer, and one or more hand/foot controllers – a headworn microphone, which few other extended voice practitioners use. Her vocal performance is highly expressive, ranging from whispering to screamed outbursts, so the fixed distance between microphone and mouth has practical utility in following her motion; it also blurs the difference between voice, electronics, gesture, and control even more than in other extended voice performances.

In her 2015 album *Without Knowing Why* and her recent live performances, Pensado explores a variety of voice-noise interactions. The titular character in "Rondo con Andreita" is likely the doll with which Pensado sometimes performs live (e.g., Back Alley Theater performance, Washington DC, Sep 2014) and which is mentioned in the album liner notes. In fact, "Rondo" and the Back Alley performance seem separate versions of the same material, where the A section of the rondo form consists of whispering or quiet speech, little processed, while the alternating contrasting sections are much noisier. These noise sections are characterized by rapid timbral changes (often 2-4 times per second). The noise initially follows

the melodic and/or rhythmic contour of Pensado's voice but in later non-A sections grows more independent and increasingly masks the voice. Around 4:30, Pensado foregrounds the relationship between left and right channels, which in one moment are quite similar (perhaps filtered, delayed versions of the same material) and in the next moment totally distinct.

Every rhythmic line seems hand-touched, as though Pensado is manually controlling synthesis or deep processing of samples (e.g., granulation). Nevertheless, harmonic structuring and rapid register changes within the noise material (6:30) imply that pitch- or spectrum-following of the live vocal input may partially drive the electronic sounds.

Near the end of "Rondo" and Back Alley, there is a lengthy sample of Argentine tango music with orchestral and sung parts (perhaps from the 1950s, judging from recording quality). Pensado sings along, and her singing and the recording are variably noise-ified. This recalls Hollywood tropes of garbled communication perhaps through weak radio signal or transmission from outer space.

The humor and weirdness of Andrea Pensado's use of a ventriloquist's dummy apparently named "little Andrea" augment the questions of control raised in both studio and live versions of the piece. The voice can be used for both gesture and control, and these functions sometimes become indistinguishable in Pensado's sound. The apparent spectral following which drives synthesis is evidence of hands-free vocal-gesture work. Yet Pensado also uses her hands to animate the doll as though it is speaking or intently listening (visible in the Back Alley video), and these motions sometimes seem to trigger sonic changes. The doll may contribute to the sense of hand-sculpted rhythmic content, or this audio-visual link may be completely imagined. The rapidity of timbral changes further obscures control source.

Notably, although the prevalence of and frequent changes in harsh noise timbres suggest a kind of anarchy, compositionally-trained phrasing lends a held or structured quality. Gestures are not random, but rather occur within some structure. For instance, the garbled, chorused speaking voice beginning at 3:07 in the left channel uses quantized, almost vocoded, frequencies within the range of about a minor sixth. Noisy sections do not begin abruptly, but are instead prepared by brief, growing interruptions often panned centrally.

Another track from *Without Knowing Why*, "¿gnunülulúk!," exhibits freer formal structure than the straightforward alternating sections of "Rondo," but shares a similar noise aesthetic. The low click pulse beginning at 0:50, panned center, grows organically into a soloistic line, then returns to a pulse, and alternates in this fashion several times. Though much of this track is effectively a voice-electronics duet, the electronics in particular are inflected with great detail so they variably draw more or less attention away from voice. Initially at 2:10, and more assertively at 2:20, low synthesized lines are introduced with varying harmonics emphasized. At first these sounds seem to provide stasis against the actively pitched, texted but unintelligible voice. But within seconds this line begins to be interrupted, for a few seconds at a time, by bursts of distinct noise including "eh! eh!" upward glissandi, which are seemingly bitcrushed vocal samples. In other words, the electronics line seems deeply conflicted about whether to remain static or to burst forth into a more dynamic profile; meanwhile, Pensado's live vocal performance is arguably upstaged even as her sampled voice is incorporated into the electronics.

Pensado's frequent changes in electronic timbre are unique in the landscape of extended voice. They may indicate that she is less beholden to her material than are other extended voice practitioners, who tend to allow each sound to last longer, during which time the audience arguably becomes more deeply immersed. In personal correspondence with the author, Pensado

states, “I try to avoid mannerisms that are associated with the ‘noise aesthetics’, but of course, sometimes I do them, because I can’t help but being influenced by the material I’m working with.” From a cybernetics perspective, Pensado’s work may be considered more noisy than Ratkje’s (who also self-identifies as working with noise) because it contains more unique information in less time.

This protean quality may arise in part from deep familiarity with the capabilities of her technology. Though Pensado identifies as an improviser who takes a “highly intuitive” approach to "using Max as her main programming tool” (Pensado 2017), there is nonetheless the sense that her utterances are tailored to what the patch does well, for instance speech with larger than normal frequency range to provide interesting fodder for noise synthesis. Of course, this is not to assert a value difference between different rates of timbral change or different software uses. But for Pensado and other extended voice practitioners, the relationship with one’s technology influences the formal and vocal gestures that are accessible and prevalent.

E. Maja S. K. Ratkje

Maja Solveig Kjelstrup Ratkje [NO] is prolific in solo and group contexts. Her work involves slow sculpture of form through gradual sonic evolution of the vocal-electronic body. Ratkje uses a relatively large setup, often including laptop and microphones (stand-mounted and in-mouth), analog synthesizers and midi controllers, and a mixer. Her performances display impressive choreography to control of this wide range of equipment and sound. Ratkje engages most frequently with the midi controller and the in-mouth mic, which yield either immediate or delayed changes in sound. Ratkje’s rich, playful approach to the electronic prosthesis complicates the temporality of vocalization and relates gestural and control actions.

The percentage of performance time spent vocalizing (broadly speaking, that is, including inter-phrase pauses) varies markedly among Ratkje's solo voice-electronics sets. For instance, in her Kontraste (2012) and Punkt (2013) sets, Ratkje vocalizes for about 75% of the time, while in her Kongsberg (2015) set, only about 45%. This parallels variation in gear: Ratkje took more analog equipment to the Kongsberg performance and was excited about incorporating these instruments into the performance, whereas she took less gear to Kontraste and Punkt and thus utilized relatively more vocal sounds (Ratkje 2016).

Ratkje states that each voice-electronics performance is improvised (Kjus and Dansielson 2016). However, some broad rules demarcate the sonic and technical basis of Ratkje's improvisation in solo versus group contexts. For instance, the chirping sample which loops for the first ten (of forty) minutes of the Punkt set also loops for several minutes midway through the Kongsberg set two years later. To my knowledge, Ratkje has not employed this sample in group improvisations; it may not be flexible enough to suit an ensemble's potentially rapid changes in musical direction. In other words, improvisation is always crucial to Ratkje's voice-electronics work, but the character of this improvisation varies by context.

Ratkje's solo set at the Punkt festival exemplifies her typical strategy of building up a detailed loop (often from live vocalization), allowing this to loop for several minutes, and meanwhile adding more loops or a vocal melody over top. Throughout, Ratkje gives compositional attention to form: for instance, a crucial break in the action occurs at 32:00, about three-quarters through the set. Though a section change is prefaced for several minutes with the removal of the bass and the addition of some scratches and high, LFO-gliss analog sounds, the departure of the A♮ [u] at 32:16 is the real indicator of a new section. Over two minutes, Ratkje

builds up a new looped background using a pair of finger cymbals and a high-pitch analog oscillator. Then she introduces a new mid-range (octave 4) sung vocal melody with unclear text. A pulsing synthesized line follows her vocal melody a perfect fifth below, suggesting that vocal pitch following drives this synth. Another synth with grinding, saw-like timbre remains at ominously low volume but bursts upward in volume and pitch content at 37:15 and 39:05.

The vocal melody reaches its local maximum volume and pitch, E5, at 39:40 while the pulsing synth – which has combined timbral and rhythmic profiles of the earlier voice-following and contrapuntal synths – vies for attention. This pulse quickly becomes an almost audio-rate buzz at a low, ominous place in the mix, a transition covered by delayed saliva sounds. In the last two minutes, there is an ambiguous resolution as the music moves from the previous minor to a related major key; finally, the electronics drop out and Ratkje ends the set by singing [sa] on the upper tonic, B5. These last ten minutes, which recall the ethereal, longer-lasting loops of the first three-quarters but form a significant departure through more sensitive voice-electronics interactions, illustrate something of the formal and communicative sophistication which Ratkje repeatedly mentions as important to her music.

Ratkje is also prolific in group contexts. The album *Rasaka* (2016), a collaboration between Ratkje and the jazz trio Säkä (saxophone, drums, bass), consists of two twenty-minute tracks. Each track is apparently a one-shot recording trimmed out of a longer studio performance. Ratkje employs two primary techniques: non-exact loops (mostly sampling unpitched noisy sound, for instance saliva, consonants, or hisses) and processed live vocal signal (either speech/yelling or non-laryngeal pitched material such as inhaled squeaks). Ratkje uses a controller mapped to a Csound patch (Kjus and Danielsen 2016) which likely effects the

detailed, non-exact loops, while bussing to distortion pedals permits real-time mixing of processing on the live voice.

Improvisation with non-signifying, often rapid clusters of phonemes, such as "gada-gada-ga," is an important technique in Ratkje's repertoire. Ratkje undercuts the traditional associations between speech, intimacy, and the present, however, through looping, such that speech sometimes overlaps with noisy hisses. Within the recorded medium it is difficult to tell which vocal sounds are live and which are recently sampled. This speech and looping work is especially effective in *Rasaka* since the trio's often noisy sounds further obscure live versus sampled voice source (e.g., 6:15 in Track 2).

Some similar details apply to Ratkje's varied loops and processed live signal, including spectral variety and distortion, LFO gain envelopes or ring modulation, density variation, and control-rate hard pan alternation (for instance, rate 4 Hz at about 2:00). Arguably the most skillful part of Ratkje's work on this album is the cohesion between her polyphonic and monophonic sounds, which helps her blend effectively with the trio. In solo and group sets, Ratkje uses both dense looping and more soloistic processing of live voice, but continuous looping predominates in solo work, while mono- or homophonic real-time processing is effective in group contexts because the voice and its parallel processing can respond rapidly to the other players.

F. Ami Yoshida

Ami Yoshida [JP] uses a highly streamlined setup in live performance: usually only a microphone, and occasionally one additional piece of gear such as a looper or effects pedal. Her

detailed use of amplification reveals her expertise, particularly since her quietest sounds occur through closed or only slightly open mouth.

Arguably Yoshida's best-known output is *Tiger Thrush* (2003), an album of 99 untitled miniatures. Yoshida's self-defined style of "howling voice," or quiet squeaking and screeching sounds often produced through inhalation or high-pressure exhalation, joins found objects and environmental background sounds in this work. The album lasts 1:07:51, giving an average track duration of 41 seconds. Timbre and time, particularly with respect to listeners' musical memory, are highly important in this album.

Tiger Thrush manifests the timbral alchemy common, in one form or another, in extended voice. Though Yoshida's "howl voice" employs non-normative modes or sites of vocal production, these sounds still register as vocal. Yet several tracks on this album use miscellaneous non-voice objects as sound sources. For instance, Tracks #24-27 use (perhaps) a shaken container of mints, delayed clicks as a pencil strikes a table, etc. Furthermore, many tracks are purposefully 'poorly' recorded, so that background sounds including lighting and air systems are audible. All of this begins to blur the traditional boundaries between voice and non-voice, leaving some ambiguity as to whether voice is instrumentalized or non-voice sounds come under the umbrella of 'extended voice'.

The vocal tracks of *Tiger Thrush* explore memorable or monolithic utterance. Some tracks, such as #8, are single vocal utterances of only a few seconds' duration. Though such tracks initially seem to function as non-structural palate cleansers, their presence throughout the album suggests that something more is at play. By contrast, other tracks, such as #5, loop a single vocal phrase, either mono- or polyphonically, for several minutes. At first blush, these longer tracks seem important formal anchors within the album. Nonetheless, the stillness and

persistence of *Tiger Thrush* gradually suggest an inversion: micro-tracks come to read as intense sound bytes which eschew embellishment, while longer looped tracks provide a sense of familiarity and pad the more concise vocal statements.

Yoshida's exploration of voice-electronics relationships is also apparent in the non-vocal tracks, of which #45 is a prime example. This track, consisting of looped, high-frequency background noise, follows shortly on the heels of the previous track and lasts almost five minutes. This should be interpreted as a statement about album construction and musicality. Though long pauses (5-8 seconds) typically separate consecutive tracks, some tracks incorporate pauses of nearly this length, perhaps 3-5 seconds. Track #45 draws from the same background noise present in Track #44, which contains several substantial pauses, but articulates a new musical statement because of its length and subtle variation of loops. Through slight manual adjustments to the time bounds of the loop, the listener begins to question whether the sound is changing or if this is in fact a perceptual mirage. Yoshida's intermingling of silence and continuity serves to make timbre strange and to unsettle short-term sonic memory.

Delicate reverb and some reversed sounds subtly remind listeners of the presence of technology. Though many extended voice practitioners are keen to exhibit control over their technology, Yoshida is minimal in her demonstration of technological control. The subtle compositional decision of whether to conclude a track partway through a loop or at the end of a cycle is one point of engagement with the technology. Other tracks underscore the sonic role of the technology itself. For instance, in Track #41 Yoshida's singing provides a gentle, lower-register counterpoint to a whistling synth texture; and in Track #49 Yoshida performs a melody consisting of distorted microphone pops.

G. Pamela Z

Pamela Z [US] is a renowned composer and performer of voice-electronics works, often employing digital looping software and novel digital controllers. Her stage work is inflected by both performance art and *bel canto*-style engagement with the audience. The strangeness of technology is apparent in her sound and body language, for instance when gestures of typing on a typewriter enter into dialogue with sound, or when looping takes on new resonance because of her hands' varying proximity to a novel controller. Z makes clear that she considers performance crucial to her practice (Z 2017).

Analysis of her composition "Badagada" in its album version versus live performance is a useful entry point to her creative work.[17] This piece occurs on Z's album *A Delay is Better* (2004). On the album, it begins with a mid-range, spoken/ lightly belted "bada-gada-gada-gada-ga" loop including rest, soon joined by a lower pitch, murmured loop of similar text without rest, along with a loop of rhythmic breathing in eighth-notes. Distant wailing glissandi in a higher register follow, preparing the next line of *bel canto* singing in a mid-high range, text "I know you're not in there." The piece continues in this fashion until 1:26, when the melodic singing begins a new "bada-gada" line about an octave higher than the yelled speech which began the piece. At 2:27, low, grunt-like, multiply delayed "huh, huh" sounds enter, only to stop a few seconds later. The piece ends at 3:41 after having built up numerous mid- to high-range melodic and ambient layers, several of which are phonemic (i.e., apparently texted but not clearly signifying) and seem to pull away from the rhythmic loop structure.

[17] Z (2017) states that she considers performance and fixed media separate streams of practice. Though "Badagada" exists on an album and is meaningfully different from the version I heard live, the fact that Z conceives of her fixed media work in a still separate category suggests that the album is closely related to live performance. In short, 'live' and 'recorded' methodologies overlap.

In a February 2016 performance I attended, "Badagada" lasted over nine minutes. This near-tripling of album duration occurred in part because Z used no pre-samples, but rather built up all loops real-time. She also took more time with improvisational sections, such as the last third of the piece, which lasts only from 2:30-3:40 in the album version. Some loops were less precise rhythmically in live performance than in the album version, which lent a vibrant buzz to the performance as rhythms phased in and out of synchronization.

Much electronic music scholarship treats looping as a simplistic or unskilled technique; Baars (2015), for instance, writes that Z makes "straightforward use of technology." One key exception to this trend of marginalizing looping is George Lewis' (2007) argument that Z's work constitutes a meaningful intersection of musical, technical, and identity practices, including the Black aesthetic technique of the cut, the Japanese notion of *gaijin* or foreigner, and novel linguistic-sonic experimentation. In its album and live versions, "Badagada" impresses as highly skilled and not at all straightforward. Z is one of few extended voice practitioners to purposefully vary the timbre of her normative acoustic voice,[18] and her utilization of looping technology underscores this unique practice. Hughes (2015) also argues for the skilled quality of looping in some popular music, an argument which could be extended to Z's work, particularly in its blurring of popular and 'high art' idioms.

Z's large-scale multimedia work *Voci* (2003) further illustrates her nuanced and skillful use of voice and technology. Vocal clicks of various sorts are significant in several movements of this work. The first movement uses an ostinato of a whispered "ah" sound with an initial glottal click. Though the human ear tends to focus on vowels as important conveyors of

[18] Some pop singers use a different vocal timbre in each song for expressive reasons; see Lady Gaga, *Joanne* (2016). Other extended voice practitioners achieve variety in acoustic vocal timbre primarily through extended techniques, but their normative singing/speaking usually adheres to a narrow range of sound comprising their "natural voice" (Looser 2005). This raises complex questions about authenticity and second-order social matters such as racial coding of 'appropriate' vocal timbre (Eidsheim 2015).

linguistic meaning, repetition gradually renders the non-signifying clicks even more obvious than the whispered vowel content. This whisper-click line coincides with sampled narration (a scientific text explaining the larynx) and live singing.

In the next movement, “Qwerty Voice,” Z types on a typewriter while theatrically stumbling through a text containing only consonants. Here, the typewriter’s clicks become almost an acoustic vestige to her speech, whose consonants become vowel-like through purposeful addition of sing-song pitch. Another movement, “Voice Studies,” combines sampled tongue clicks and live narration of a text of acoustic-sociological voice research.

Z’s intentionally sparse aesthetic is noticeable throughout – few sounds exist besides voice, whether pre-sampled or live-performed. But, far from evincing a lack of technological skill, this sparseness underscores the dual communicative-sonic character of voice. As Lewis (2007) notes, Z uses repetition technology to interrogate the musical traits of language, which are normally ignored, and in so doing directs listeners’ attention toward her playful re-examination of contemporary life.

IV. LISTENING, AGENCY, AND PRESENCE

The voice has historically been situated at cross-purposes with performer self-listening. *Bel canto* pedagogy, for instance, discourages vocalist self-listening (Wakefield 2003) and instead imagines performer listening as embodied feeling (Madaule 2001). Embodied feeling also underpins vocal idioms and practices seemingly remote from *bel canto*.[19] This trans-idiomatic insistence that vocal music cultivate a *felt* body is part and parcel of the historical and problematic association between voice, nature, Other (often woman), and inarticulacy, which treats voice as non-agential instrument (Bosma 2013, Weber-Lucks 2003, Weidman 2014).

Practitioners of *extended voice*, by contrast, utilize compositional and improvisational methodologies to create electroacoustic works which entail vocalizing, recording and processing voice, and – crucially – listening, all in real time. Extended voice practitioners use analog and digital technologies as a prosthesis to the acoustic vocal body. This prosthesis demands agential self-listening and response. Extended voice works manifest ever expanding body-technology configurations and vocal-electronic timbres. The use of technology as prosthesis fundamentally changes the character of voice, demanding listening where otherwise only feeling is permitted, and encouraging vocal timbral exploration in a dialogue with the electronic prosthesis. Analog and digital technologies articulate a new cyborg body[20] marked by novel acoustic and technological configurations.

---

[19] For instance, Girilal Baars, who performs using materials such as folk singing techniques and live processing in MaxMSP, writes: "It is important for me to retain the physicality of the voice, *the body itself*, and I have come to minimize my use of audio convolution and granulation and their characteristic shimmering and fluttering sounds, which I feel often detract from *the bodily experience* of processed voice." (2015, 52) [emphasis added]

[20] I am most interested in understandings of 'cyborg' that speak to cultural-technological modifications of body, which I consider crucial to extended voice (cf. Gordon 2011, Gourlay 2012).

Close examination reveals the contingent, gendered, and power-based character of the disembodiment presumed to occur when recording voice in particular. Conscious manipulation of technology through self-listening, however, questions this disembodiment and cultivates agency in voice, which ordinarily occupies a sonic-cultural matrix that privileges feeling the body and minimizes performer self-listening and agency.

A. Gender

In the Western vocal-electronic composition tradition, a great many works are composed by men for women performers, where much of the vocal content is non-verbal singing. Extended voice stands apart from this tradition by unifying composer and performer roles and by exploring fundamental questions of what voice can be and how it can sound, beyond mere inarticulate singing.

Hannah Bosma (2013) surveys a body of voice-electronics compositions from the late twentieth century. All works had a division of labor between composer and vocalist, where all composers were men, and vocalists were men or women. The compositions for live vocalist and electronics "show a strong pattern… there are no compositions for *singing* male vocalist and electronics, while female vocal parts have a substantial amount of non-verbal singing" (60, emphasis original). Similarly, in the compositions that use vocal pre-samples instead of live vocalist, "four male voices utter non-verbal sounds while not singing, while none of the female voices are doing this without singing too" (62). In short, in the surveyed academic electroacoustic compositions, *the female vocalist, both live and pre-recorded, most often sings non-verbally*. Bosma argues that this correlates with traditional associations of the feminine with

emotion, body, and powerlessness, whereas the masculine is linked to reason, language, and power.

Extended voice practitioners, many of whom are women, diverge from academic compositional approaches to voice. Extended voice practitioners define language more broadly than do most academic composers (all sung material has some linguistic valence, since embouchure maps so closely to vowel content) and therefore use non-normative vocal and electronic sounds in order to subvert language. Moreover, extended voice practitioners do their own compositional and technological work. Bosma, by contrast, analyzes compositions in which composer-technologist and performer were distinct roles, and the composer-technologists were all men. This division of labor corroborates my experience of academic and non-academic electronic music. By comparison, extended voice practitioners' unification of composition, technology, and performance is indeed methodologically novel and agential.

B. Recorded voice and disembodiment

Recording the voice is often thought to result in disembodiment: the acousmatic voice.[21] "Technology is capable of separating voice from speaker, conversation from community" (Davidson 1997, 103) is a common attitude. Likewise, discussions of tape-recorded voice interpret the distinct visual appearance of body and technology as a sign of the gulf between the two. Voice is uniquely subject to expectations of embodiment and thus susceptible to accusations of disembodiment when technologies are used for vocal processing. Embodiment is far less at stake when non-voice instruments are technologically processed; technology complicates the

[21] Cf. Auner (2003), Kane (2014), Zorn (2012). A complementary perspective (e.g., Young 2015) suggests that recording the voice resituates embodiment in a different but still material medium. Both are problematic because they treat body primarily as something received, rather than as a malleable technology curated by the practitioner.

sound output but does not fundamentally re-define the instrument or its relationship with the performer. And while some vocalist-technologists still choose to experience the voice primarily through embodied feeling, extended voice practitioners' use of technology demands self-listening in order to manage the sonic whole. This agential expansion of the vocal body interacts with and subverts the notion of disembodiment.

Vocal disembodiment carries creative and subversive potential in spite or perhaps because of its relationship to larger questions of power. Weheliye (2002) writes that "the recorded voice in contemporary mainstream R&B… reconstruct[s] the black voice in relation to information technologies" (30) where Black people were previously excluded from vocal-technological agency. Arguing that "[p]erceived vocal gender is the primary generator of meaning in listening to voices" (134), Vágnerová (2016) suggests that Wendy Carlos' work with vocoders is especially subversive because it disrupts gender. Anxiety around disembodiment, therefore, is at heart a patriarchal anxiety that the vocalist is feeling, and expressing, a different body than their own. If technology use obfuscates the vocalist's race and/or gender, the vocalist cannot be (seen to be) engaging in proper embodied feeling of their own sound – thus, the negative valence of disembodiment is in fact a patriarchal listening bias. Extended voice, then, mobilizes posthuman ideology in a practice-oriented counter to disembodiment fears. Personal identity and aesthetic motivations, such as timbral-temporal innovations and voice-electronics dialogue, articulate agency (Bell 2016). The composer-performer mind-body undertakes real-time listening to the cyborg body with the intent of achieving of a rich, meaningful dialogue between voice and electronics. In sum, extended voice does not altogether dispose of the notion of body, but fundamentally re-frames body as something to be technologically curated through agential self-listening.

Historical studies of recorded voice provide some precedent to the timbral exploration, in dialogue with electronics, which is central to extended voice. Jacob Smith (2008) contends that, since the inception of the medium, recorded voice has helped push the boundaries of acceptability, with narrative content used to justify non-normative vocal sounds, including grunts, screams, and screeches. The 1905 phonograph mini-drama *Dr. Jekyll and Mr. Hyde* is one example. Most scholars, including Smith, assume that recorded voice is necessarily separate from the audience in time and space. Yet extended voice practitioners often record and process the voice during live performance, directly in front of the audience. Extended voice practitioners heighten the real-time character of recording, and question its disembodying effect, by exploring the spectrum between voice and electronics. For instance, a noisy sound which initially seems very remote from voice may gradually be revealed as vocal in origin, and meanwhile the acoustic voice may take up extended techniques which are timbrally similar to the electronics. This fundamentally alters the vocalizing body, and demands compositional/ improvisational self-listening in order to control the vocal-electronic whole.

C. Prosthesis

Extended voice practitioners subvert the traditional and problematic association of voice, woman, body, and inarticulacy by positing that the vocal body (any gender) can be artfully extended using an electronic prosthesis. Timbral and temporal exploration, and in particular dialogue between voice and electronics, are considered worthy goals in themselves. The prosthesis and the vocal body together comprise a new vocal-electronic body. Extant scholarship makes some provision for this agential adoption of a prosthesis to self-articulate a new vocal-electronic body. For instance, Hughes (2015) coins the term “autonomized vocals” to indicate

vocalists' agential use of technology, principally loop pedals, for creative purposes. Extended voice practitioners express agency by juxtaposing voice and technology to create complex sound works in which boundaries between vocal and electronic timbres are broken down and genre implications questioned. In this agential use of technology for vocal processing and genre exploration, there is an underlying assumption that the voice itself is somehow insufficient or unsatisfying; technology fills this void.

Despite the "autonomy" increasingly granted to voice in some scholarly treatments, many studies of recorded voice and body equate 'author' and 'composer', and largely overlook the unique category of 'composer-vocalist' and its agential exploration of body. Other scholars address voice from a primarily linguistic perspective, highlighting text-sound composition and related styles which feature recordings of speech. On the other hand, extended voice entails a feedback loop between body and electronic prosthesis. Electronic extensions demand acoustic extensions; extended vocal techniques often supplement or replace normative speech and singing. Live recordings, delay, granulation, and other memory technologies take on a unique valence, indicating and questioning the sound of the body more than if they were simply applied to normative speech or singing without extended techniques. In so doing, these technologies destabilize the ocularcentrism of many vocal theories by underscoring audition, not vision, as the sense which weighs the efficacy of the electronic prosthesis and its relationship with the vocal body.

Electronic prosthesis in extended voice bears some resemblance to literary understandings of prosthesis. Narrative treatments of the non-normative body, argue Mitchell and Snyder (2000), are important precisely because they unite literal and abstract "registers of meaning-making" (62). Similarly, the text can indicate itself as prosthesis, yearning for the

"personating subjectivity" (802) of the photographic medium or the writerly body (Wallace 2008). This subjectivity – an ability to participate in, comment on, and direct the action as it happens – occurs in extended voice in particular because customary embodied feeling is replaced by performer self-listening to the extended vocal-electronic body. There comes to be a symbiosis between the prosthesis and the prosthetized; neither could exist without the other. Lakoff (2015) understands prosthesis as corporeal but also prominently temporal. "The prosthesis is familiar, but also different: it demarcates the limits to the organic body, but also challenges the solidity of those limits" (3), and this demarcation and challenge occurs as a temporal oscillation. In extended voice practice, the electronic prosthesis questions the bounds of body, doing so in markedly temporal fashion through timbral and genre exploration unfolding over time.

Although extended voice subverts disembodiment by agentially articulating a new cyborg body, the underlying dissatisfaction with the timbres of the body, which motivates use of the electronic prosthesis in the first place, is not necessarily resolved by the prosthesis, nor is the prosthesis even the central issue. As a technique that demands performer self-listening, using electronics as prosthesis to voice serves mainly to increase performer agency. Gordon (2011) makes a similar argument when she re-figures the seeming disempowerment experienced by early-modern era castrati as a contemporary "phallocentric" "squeamishness" (111) about the culturally modified body. Voice is not customarily an agential instrument, but extended voice practitioners' use of technology as prosthesis re-imbues agency into voice. Listening becomes the embodied method by which the composer-performer manages voice-electronics dialogue and thereby posits their own creative agency.

D. Presence

I define *presence* as, simply, curation of one's own vocal-electronic sound. What is present is a posthuman mindbody (Herndon 2010) which undertakes agential self-listening in order to sculpt the sound. Importantly, presence allows spatio-temporal coincidence of the performer and listeners (extended voice practitioners often perform live) but does not require it (extended voice practitioners also create much recorded sound work, including albums). Presence, by my definition, derives solely from the performer's compositional-improvisational curation of their own vocal-electronic sound, and this curation occurs independent of audience location. The notion of presence has two main theoretical advantages: (1) it sidesteps the patriarchal listening information implicit in (dis)embodiment, and (2) it accommodates the increasing overlap between recorded and real-time practices in both extended voice in particular and electronic music as a whole.

In live performance, presence has a certain range of temporal manifestations. Extended voice practitioners take various approaches to musical form, but all are inflected by the distinct performance activities of vocalizing and controlling. Impett (2010) writes, "If the past (memory) and the future (invention and projection) are acts of imagination, then perhaps motion – physical rhythm on some level – has a vital role in the mediating and structuring of both" (87). The body is deeply involved in presence, not in the traditional way of embodied feeling of voice, but rather undertaking motions to control the sound, and this bodily engagement manifests temporally. Instead of vocalizing continuously throughout a performance, extended voice performers intersperse moments, or sometimes long sections, devoted solely to listening and controlling the vocal-electronic body. This listening is crucial to management of the electronic prosthesis.

Common parlance would seem to correlate vocal 'presence' with vocal sound, so it is perhaps ironic that my definition of presence includes vocal silence when self-listening occurs (of course, curational listening and vocalizing can coincide temporally, but often they do not). But I am concerned instead with the presence of an agential mind cooperating with an expanded body in order to explore timbre and temporality in service of rich voice-electronics interactions. Or in other words, because of the extremely close dialogue in extended voice work between vocal and electronic sounds, presence must be defined in terms of performer agency, not simply and solely correlated with the moments when sound emerges from the performer's mouth.

Contrary, again, to colloquial understandings of vocal 'presence', presence as I define it necessarily involves re-evaluation of liveness. The agential self-sonic curation of extended voice is by nature trans-temporal. Technology articulates a cyborg body whose utterances include but are not limited to those sounds emerging from the mouth now. Vocal sound is traditionally considered 'live' only in the moment when it emerges from the mouth, but, by my definition, a looped vocal sample is still 'live' in that it emerges from the self-curated cyborg body. In other words, liveness is associated with diverse embodied actions, including acoustic vocalizing as well as self-listening and controlling. To put it yet more simply, I argue that 'live voicing' could include, for instance, turning a knob to control processing of one's own vocal-electronic sound.

Importantly, this version of liveness accommodates the many creative variations within extended voice practice, including use of extended vocal techniques, balance of voice and electronics, self-identified position along the composition-improvisation spectrum, and primary output medium (I am thinking especially of albums versus live performance).[22]

[22] Extended voice practitioners are equally fluent in recorded media (produced full-length albums, self-posted short tracks on Bandcamp and Soundcloud, etc.) and live performance. As a result, I have treated recorded media and live performance as roughly equivalent. Nonetheless, there are important differences between live and recorded sound work. To elucidate the spectrum incorporating liveness and recording, a

I do not expect that simply informing skeptical listeners about power imbalances inherent to vocal "disembodiment" would instantly convert them to lovers of extended voice. Though I hope presence-as-agential-curation prompts some people to reconsider the social stakes of embodiment, I admit that taste is unlikely to change overnight. Nevertheless, presence is valuable as a descriptive project: it accounts for current trends in extended voice which are taking root despite that many listeners find these sounds strange or distasteful.[23]

E. Inner voice and text scores

Presence, or curation of one's own vocal-electronic sound, also manifests in the 'inner voice' use of internal language by some extended voice practitioners. I draw on Brandon LaBelle's (2014) conception of inner voice as a personating rehearsal with alter-ego: "Self-talk is precisely the making of conversation, as if *in preparation*" (94, emphasis original). This alter-ego is akin to the technologized performance partner. To practice extended voice is to engage with this cyborg alter-ego and, moreover, to do personating work, because the electronic prosthesis is a cultural-technological modification of body.

Text scores are characterized by contemplative interpretation of instructions provided by a composer, who composes ranges of sonic possibilities rather than particular sounds. This approach is mirrored in the agential dialogue within extended voice practitioners in their distinct but related capacities as composers and performers. In other words, linguistic material in text

---

fuller taxonomy of extended voice practitioners' memory technology work is needed. On-stage recordivity, for instance, provides valuable insight into the varying manifestations of liveness in sound art (Knowles and Hewitt 2012).

[23] Further research is needed into two disparate tendencies in extended voice: it is highly personalized, yet has unique communicative potential, i.e., is listener-driven. Every practitioner I asked said they did not consider 'beauty', 'ugliness', or 'noise' – admittedly subjective terms – when making work. Yet in my experience, beauty (aural, visual, positive, and negative) is an almost inescapable metric in many listeners' relationship to my work. Guilleray's (2012) discussion of vocal emotion as a deeper layer of meaning, for instance, may help craft a new metric of quality, other than beauty, for extended voice.

scores and extended voice alike aids sonic realization of compositional concept while also underscoring the sonic uniqueness of the performer.

Extended voice encompasses diverse approaches to inner voice and text. Guilleray, Motland, and I use particular verbal material as a sort of documentation of creative practice. AGF considers lyrics themselves a sort of working method, while Z values lyrics but does not treat them as documentation. Ratkje and Pensado disavow thought altogether.

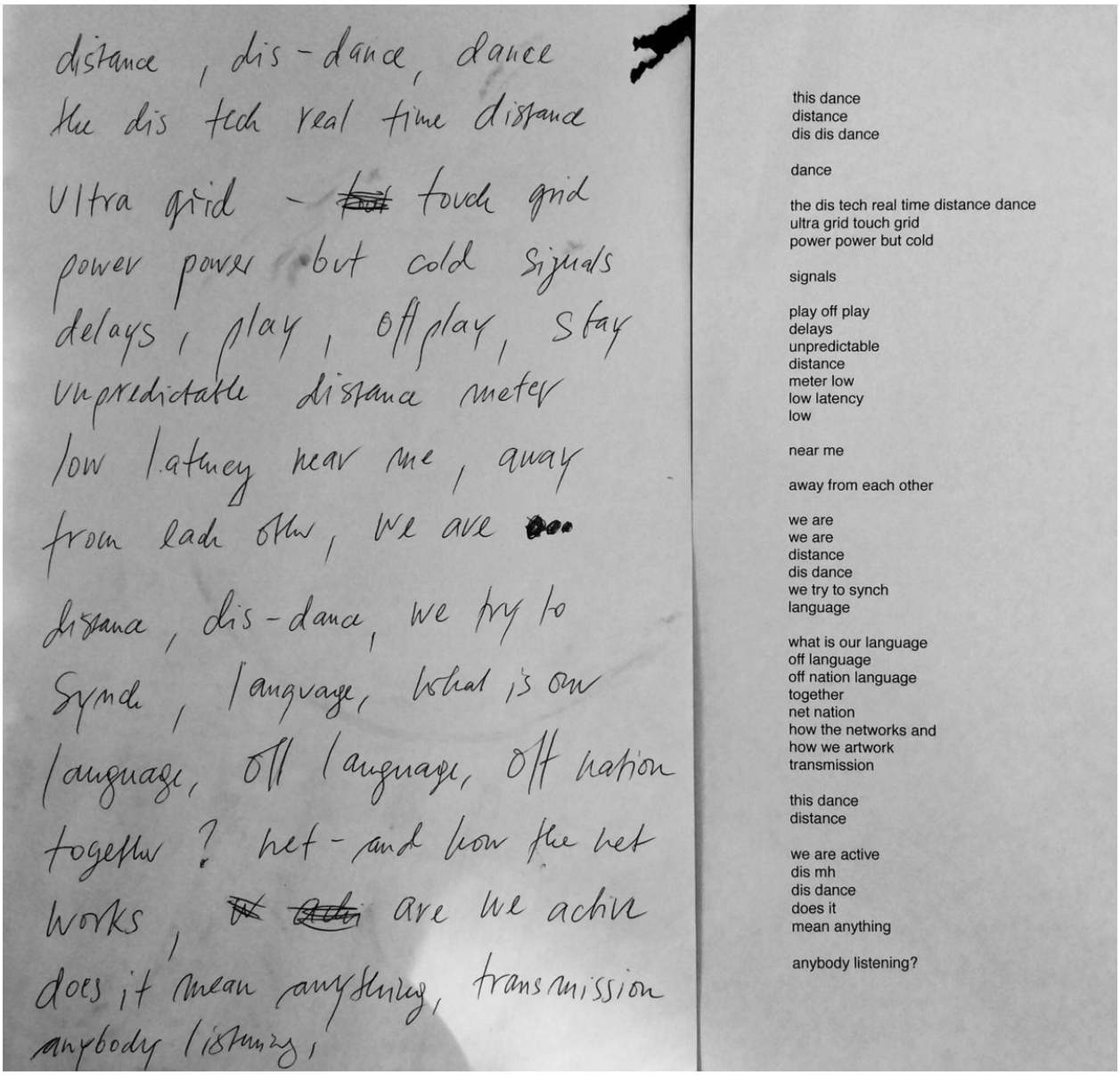


Figure 9. "dis dance" Bandcamp track image (AGF 2016)

Guilleray	texture metallique, piquante, veloutée, rude, douce,sombre, granulée, rocailleuse registres soufflé, air hybrid, unité, dualité, coexistence densité dynamics (volumes) intensité perception, immersion structure, blocks, sections, superimpositions	
Motland	"internal monologue while performing" *Fake Synthetic Music*: inhale, support, close A.T tempo, don't drop stay on pitch embody forward energy B.T embody check eq pitch feedback down tempo down wait silence A.T go! tempo, don't drop pure tones backwards stop	
Warren	diaphragm, intercostals, listen [hit r → hit e → LV → hit p] x 4 listen – solo – voxproc – route || 1-4 unh ST, len, wait needs more, needs different, enough, stop, copy	(vocalizing) (Max patch) (Abacus toggles) (undertone singing) (pulsar granulation params.) (improvising, sound mix)

Table 5. Extended voice practitioners' personal vocabularies (selected)

Like text scores, extended voice involves a correlation between certain configurations of natural language and particular ranges of sonic results. In both text scores and extended voice, there is overlap between descriptive and prescriptive language, and between instructional and

abstract language. One important difference is that text scores are usually printed and viewed by many, while the language of extended voice is often unprinted and known only to the practitioner. It might be hidden as a comment in a software patch, or might only be thought and never voiced aloud. (AGF's published poetry is a notable exception.)

This 'inner voice' status of extended voice language accords with the non-division of labor in this practice. Composers are vocalists are technologists, so scores need not be printed for realization by other performers. Indeed, this 'inner voice' manifestation of language pushes text score goals a step further. Text scores question genre, challenge the status of the so-called 'concert piece', and shift some traditionally compositional work (shaping form, making parts, balancing sound and silence) to the performer.

Text scores and extended voice share an emphasis on material exploration of sound. Below, I discuss three important features of text scores (physical organization, number, and listening) and how they refract into features of extended voice language (body, mixing, and agency).

Notably, internal language in extended voice is in service of sound. Text scores capitalize on the distinction between natural language and music-making acts, and this distinction is similarly important in extended voice. In the moment of sound-making – whether in front of an audience or not – sound is the foremost consideration. Yet language is undeniably one of the things voice does well, so internal language is an indispensible part of extended voice. Furthermore, because inner voice is a personating rehearsal that spans studio, in-home, and performative musical moments (to name a few), it contributes to the argument that 'recorded' and 'live' methodologies dovetail in extended voice.

1. Physical organization – body

I argue that text scores' physical organization parameter refracts into the parameter of body in extended voice. This line of reasoning first requires a re-examination of how best to approach text scores. These scores' physical organization on the page is, I contend, a more useful analytical approach than the customary method of categorizing verbal content as "instructional" or "allusive" (Anderson 2014) because these categories in fact overlap greatly. For instance, the imperative grammatical mood in text scores, which at first seems instructional, can house quite abstract content:

> Play or sing extremely short events
> until
> each one
> seems like an eternity
> Karlheinz Stockhausen, *Elongation* (1976)

Similarly, Daniel Goode's *I* ⟷ *Ou* (1980) begins with instructions for phonemic vocal chanting, but quickly turns to matters of performer attention. Goode suggests that performers might consider the universe or God, which I take to suggest contemplation of the act and the meaning of sounding. This somehow renders sound both more and less immediate. In short, instructional language in text scores can be quite abstract, philosophical, and/or personalized. Inversely, seemingly abstract language can, depending on performer decisions, yield consistent sonic results with each performance and thus take on an instructional quality.

Instead of instructional versus abstract, a more direct method of categorization deals simply with the physical organization of the score on the page: Is text all in one paragraph? in multiple paragraphs? in several single lines? divided by category (instructions, materials, etc)?

Single-paragraph text scores include many of the scores in Christian Wolff's *Prose Collection*, such as *Pit Music*, *X for Peace Marches*, and *Stones*. Wolff values the economy of

this approach: “a short paragraph of prose to bring about a large variety of performances” (1998, 494). Multi-paragraph text scores include Alvin Lucier’s *Vespers* (1968) and Pauline Oliveros’ *The Grand Buddha Marching Band* (1981). Each of these examples gives a progression of information (respectively: number of players, suggested performance space, materials, task, process, way of ending, inspirations; and formal shape, who may play, goal, wardrobe, allowable instruments, player roles) but do not label information by category. Sometimes a new piece of information is marked by a paragraph break; sometimes multiple pieces of information occur in a single paragraph. Text score composers who use this style may consider it a casual, direct way to connect with players.

Text scores comprised of distinct lines often use seemingly abstract language to foreground performers’ role in the compositional process. Single-line works, such as those in Brecht’s *Water Yam*, force performers to get to the crux of how to realize the score. Brecht’s *Word Event* has yielded a wide variety of performances, but, scholars note, many of these share a distinctive performance profile consisting of a singular, decisive gesture (Kotz 2001, Pisaro 2009, Robinson 2002).

- EXIT
  George Brecht, *Word Event* (1961)

In multi-line text scores, stanza breaks aid interpretation:

> When you are thinking only of your own playing,
> play as quietly, gently and as long as possible.
>
> When you sense that another player is thinking of you,
> play moderately loud, rather agitatedly and moderately long. …
>
> Karlheinz Stockhausen, *Communication* (1968)

Stockhausen's stanzas suggest a formal structure (players first think of their own playing, then try to detect whether another is thinking of them, then perhaps attempt both at once, etc.). In other cases, stanzas are paragraph-like in function. For instance, each stanza in Michael Pisaro's *fragile being, hopeful becoming* (2013) carries different information (timbre/ dynamics/ duration; formal structure).

In sum, the boundary between abstract and instructional language in text scores is highly permeable. The criterion of text's organization on the page acknowledges the interactions between the abstract and the instructional, yet also provides valuable clues as to how composers present and structure information within text scores.

Text scores' redistribution of compositional labor is a crucial part of the analogy to the internal language of extended voice. That is, the page is to text scores as the body is to the internal language of extended voice. The page and the body are the physical vehicle and the testing ground for compositional ideas, as well as the interface between compositional and performative labor. Different arrangements of text on the page, like varying deployments of body in response to inner voice, suggest different sonic value systems and desired sonic outcomes.[24] To realize a text score, or to respond to inner language in extended voice, is crucially to engage with relationships between compositional ideas and their physical medium.

One strike against this analogy is its transgression of boundaries between materials. The body in extended voice may be the medium for compositional ideas, but it is also (a large part of) the sounding instrument, whereas realizing text scores usually requires separate instruments, i.e., the page is arguably not equivalent to the sounding instrument. Nonetheless, the urgent need to

[24] In non-text scores that inhabit the page, page organization has varying importance. For instance, in graphic scores, arrangement of material on the page can be extremely important. On the other hand, in scores using Western staff notation, arrangement of staves on the page generally matters little, apart from practical performance considerations such as ease of page turns.

re-examine what the body is and does, along with the deep creative potential of likening text score page and inner voice body, are ample reason to continue with this analogy. Furthermore, in extended voice internal language, as in text scores, there is great overlap between the seemingly abstract and literal. For instance, the imperative mood is not directly correlated to action. My "wait" is not a command to wait, but instead refers to a pulsar granulation parameter, and thus functions more as an invitation to interact with a physical potentiometer to control the wait value. Motland's "go!," "don't drop," and "pure tones" importantly permit artful micro-variations in tempo and pitch, rather than prescribing exactly the same outcome each time.

Where the instructional-versus-abstract binary breaks down, we can make sense of complex verbal information by considering how text scores inhabit the page, or by considering how extended voice inner language inhabits the body. One useful notion which is gaining traction in several scholarly subdisciplines is that spoken language is a set of material configurations of the body (Faudree 2012, Painter 2008, Shankar and Cavanaugh 2012). I contend further that the descriptive and prescriptive internal language of extended voice influences and is influenced by bodily configurations.

Of course, many performing musicians, whether acoustic or electronic, develop an idiosyncratic internal vocabulary for their own working methods. But the stakes are different in extended voice, in just the same way that the notion of embodiment is uniquely problematic for voice. Though bodies figure into all instrumental/electronic playing, the burden of 'embodiment' in its current usage is that vocalists must feel rather than listen to their own sound and must make their bodies normatively legible. Likewise, probably most or all performers use some sort of internal language for learning and mnemonic purposes. But internal language is uniquely

important in extended voice because it describes and prescribes curation of one's own of vocal-electronic body.

2. Number – mixing

Of crucial importance to text scores is John Cage's correlation of rhythm, time, and percussion:

> The composer (organizer of sound) will be faced not only with the entire field of sound but also with the entire field of time. The 'frame' or fraction of a second, following established film technique, will probably be the basic unit in the measurement of time. No rhythm will be beyond the composer's reach. …Percussion music is a contemporary transition from keyboard-influenced music to the all-sound music of the future. Any sound is acceptable to the composer of percussion music; he [sic] explores the academically forbidden 'non-musical' field of sound insofar as is manually possible. Methods of writing percussion music have as their goal the rhythmic structure of a composition. (1969, 5)

In text scores, these considerations often manifest as number and pulse. Michael Pisaro writes that the notion of number is important because it illustrates "how indeterminacy and exactness in the writing of music can co-exist" (2009, 36). Some text scores use number to unite parameters of formal shaping and sonic articulation. Pisaro often uses numbers to represent section duration, material choice, or melodic motion. It is noteworthy that number in text scores is a convenient method of grouping multiple musical parameters together while still allowing for a wide range of sonic outcomes. For instance, temporal density has implications for dynamic level and articulation. Playing six breath-length tones in thirty minutes will probably lead to quiet dynamic and smooth articulation, whereas playing twelve short tones in thirteen seconds is more conducive to loud volume and abrupt or staccato articulation.

Others use number as an instantiation of rhythmic pulse and prescribed performer choice. Pauline Oliveros' *Six for New Time* (1999) instructs guitarists to "Start a pulse with another player or players using a pattern on the common tone of G of 3, 5, 7, or 11 in a common tempo." Christian Wolff's *Fits and Starts* (1970) asks each player to keep a unique pulse in mind, then to enact any of the available "sequences," for instance: "One sound or articulation of a sound underway every twenty-one beats, omitted every sixth time the twenty-first beat comes round" (1998, 474).

Text scores employ a Cagean understanding of rhythm as events in time. By using numbers, a single text score might permit anything from a very dense realization to a very sparse one. This approach encourages silence to be treated as a musical material, and allows complex rhythms to emerge that might be very difficult to notate or even conceive. By emphasizing numbers of sounds, composers of text scores implicitly ask performers to think deeply about each individual sound. This mindset is also closely connected to experimental percussion music of the twentieth century, which, unlike the heavily pitch-centric music of the past, places great emphasis on timbre and time. Indeed, extremely few text scores give any specific information about pitch; far more require performers to consider the number of sounds played and/or these sounds' arrangement in time.[25]

References to number in extended voice vocabularies address various sound design activities related to time and texture. My "x 4" and "1-4" refer to Layer-2 buffers in the patch and Abacus control thereof. Guilleray's "registres," "coexistence," and "superimpositions" comprise a slightly more abstract use of number to address arrangement of layers and

[25] I can think of only one text score with specific information on pitch and voicing: Gavin Bryars' *The Heat of the Beat* (1972) (never performed) requires a D♭ major chord with particular voicing.

management of harmonic and spectral content. Motland alludes to number in her variation of feedback values in order to articulate form.

Whereas number in text scores generally gives information on when to play a single improvised or pre-planned sound, number in extended voice refracts into texture and mixing. I sculpt and orchestrate clusters of four sounds at a time, and Motland's "feedback down" yields few discernible delay taps and, thus, formal definition. Mixing in extended voice – like its text score equivalent, number – permits co-existence of indeterminacy and exactness. While the technical is substantially composed, live mixing is improvisational.

Improvisation is a key consideration. I argue that improvisation occurs in different moments in text scores and extended voice.[26] Performers of text scores generally make number-related choices beforehand, then their realization simply consists of executing these choices. Numbers give text scores the veneer of process piece. Everything is decided beforehand, and the piece consists solely of spinning out the process. By contrast, number in extended voice permits improvisational mixing during performance. Although pre-performance choices do occur in software/hardware design, equally or perhaps more important are in-performance/ in-studio choices about how to use the numbers implicit to the setup to vary texture and timbre.

Another perspective on number in text scores versus extended voice is that a single extended voice practitioner manages a complex soundscape, which is analogous to the total sound of an ensemble performing a text score. Number in text scores typically treats each player monophonically, while number in extended voice is fundamentally about density management

[26] I wish to bear in mind George Lewis' (1996) now famous argument that in late-twentieth century music, what has been termed 'improvisation' is Afrological, while what has been termed 'indeterminacy' is Eurological. Furthermore, my own ability to identify as an improviser is evidence of various academic communities' growing, if not uniform, incorporation of improvisation. In addition, non- or partial text scores (e.g., graphic scores, combined graphic-text scores, etc.) which engage with improvisation and/or indeterminacy complicate and challenge my argument but are beyond the scope of this study.

work by a single performer. To put it very simply, text scores assume a direct correlation between ensemble size and sonic density. Number, which prescribes action, is more or less a scalar multiplier to this correlation.[27] By contrast, extended voice values texture variety, and number is a means for complex real-time management and documentation of this variety. Number is both descriptive and prescriptive because it relates to both performer action and the design affordances of the electronic setup.

3. Listening – agency

Many text scores emphasize listening, imagining, and remembering as much as playing:

> Keep the next sound you hear
> in mind
> for at least the next half hour.
>
> Pauline Oliveros, *For Annea Lockwood and Alison Knowles* (1975)

> The last vibrations of a friend inhaling and exhaling his sighs. It's a resounding sound comparable to a wheeze but emanating from his mouth. It is climactic and dramatic, full of exertion yet effortless. Time dissolves. …
>
> Amnon Wolman, *January 4, 2001*

> Play a soft low note and listen to it. …
> Imagine that you are playing a soft low note, and listen to it in your imagination. …
>
> Tom Johnson, *Imagining* (1967)

> **Field Trips Thru Found Sound Environments**
> An audience expecting a conventional concert or lecture is put on a bus, their palms are stamped with the word **listen,** and they are taken to and thru an existing sound environment.
>
> Max Neuhaus, *LISTEN* (1966)

[27] Though this claim about number in text scores is admittedly broad, it is supported by many text scores' lack of specification of number of players or even distinction between mono- and polyphonic instruments. Some text scores specify particular instrumental forces, but these are the exception rather than the rule, and even in these cases number still functions as a compositional method of scaling textural density up or down.

Oliveros' and Wolman's text scores often feature listening and the memory of listening; Johnson's, playing and the memory of playing. Neuhaus' text scores similarly address listening and more explicitly involve the audience as participants in sound.

In many text scores, listening and imagination come to replace the act of physical sounding. Of course, this is partly related to the understanding, inherited from Cage, that true silence never exists. Yet text scores also display a unique emphasis on the conceptual: often, the underlying concept or experience of a piece is more important than the resulting sound. Text scores might well be rehearsed several times and perfected, but subsequently do not need to be performed and can instead be thought or experienced through their conceptual content. Many text score composers explore and develop similar concepts across numerous independent works, interrogating the very notion of organizing principles in music. James Saunders (2011) writes:

> …serial work presents multiple articulations of a central formative principle or group of principles. …Serial approaches challenge the need to address why only one prioritized result of the creative process is required… (498)

Though text scores place a heavy burden of interpretation onto the performer, this interpretive work aims to get at the conceptual heart of text scores, and frequently involves practices of listening and imagining.

If the distance between compositional and performative work is shortened in text scores, it is all but erased in extended voice. Text scores' use of natural language promotes listening, prioritizes interaction in and with the moment, and increases the authorial status of the performers. These functions are refracted into extended voice practitioners' agential, listening-based work in performance. Inner vocabularies of extended voice mention listening either explicitly, as in my own vocabulary, or implicitly, as in Guilleray's "perception" and

"immersion." Even Pensado, who disavows conscious thought about or during her practice, evinces listening-via-intuition through statements such as "My way of using the voice is very intuitive. I just let it go" (Pensado 2017).

In short, listening is a tool for performer agency. Listening increases performer agency to some extent in text scores, and to an even greater extent in extended voice, aided by the non-division of labor between composer and performer. Even for extended voice practitioners who do not rely on inner voice or internal language, listening is an important thematic link to the music-making world of text scores.

4. *For soloist…* text score, "For one" album track

The score of my piece *For soloist and any number of musicians* (2015) is reproduced below in its entirety. The composition process was valuable in understanding some of the motivations behind text scores, how these scores can drive specific ranges of performer action, and how this can in turn relate to extended voice practice. This piece was premiered in April 2015 by the University of Virginia New Music Ensemble (flute, oboe, clarinet, piano, guitar, violin, cello, pipa, voice), with guest performer Paul Botelho as vocal soloist. The piece was performed again several months later, with Botelho again performing the soloist role, and a smaller subset of the UVA New Music Ensemble performing as "Others."

This piece evolved over several months through rehearsals with the ensemble. Early drafts used "If-then" verbal structures (e.g., "If the river runs over the moor, babble") and improvisation." Middle drafts retained the "If-then" structure and experimented with various means of graphically representing text. In its final version, the piece uses columns instead of "If-

For soloist and any number of musicians

Soloist: Begin. Vocally improvise on the italicized text below.

Others: Listen. If you notice that any condition is true, you may choose to enact (in any way) a condition from the other column.

Duration: c. 8’

river running over the moor	*babbling*
ridge exceeding the scar	*getting twangier*
arm being underneath the quiet	*bellowing (really bellowing)*
one person being diligent	*another being vigilant*
listener looking bored	*buffering*
the time being long	*going easy*

then” structure, and replaces the “group improvisation” statement with the more direct instruction to listen for and enact the various conditions.

Creating and refining this piece clearly illustrated the conceptual nature of text scores, and that the concept can be a uniting force despite a wide range of performer interpretations. For instance, each of the New Music Ensemble performers stated that they considered a few conditions ‘dead zones’: some performers loved “arm being underneath the quiet,” while others hated it and chose never to hear or play it. But everyone understood the general sort of reactivity that lay at the heart of the score, and in order to achieve this central concept, each player developed a unique bank of sonic materials to facilitate enactment of the verbal conditions.

Performing this piece also exemplified the audience confusion that often occurs in text score performances. The premiere performance ended compellingly but abruptly, after which there was a long silence (maybe 15 seconds) before one of the performers began the applause. In

re-figuring genre and the very status of the ‘concert piece’, text scores often lack or render unnecessary the formal and behavioral elements audiences have come to expect (e.g., recapitulation of early material in the final section of the piece, or an audible exhalation at the end). As composers such as Neuhaus, Oliveros, and Wolman demonstrate, many text scores lie at the boundary of the ‘concert work’ category, and their Cagean incorporation of ambient sounds makes it difficult to distinguish a performance from its surroundings.

To imbue a creative perspective into considerations of how text scores refract into extended voice, I created an album track, “For one,” using this score. First I recorded 16 separate vocal improvisations on the italicized text of the score. I did not listen to previous recordings or look at their waveforms while improvising (as I did during “ul”), so these 16 tracks are mostly disparate. Any moments of cohesion derive solely from my memory and intuited timing from previous improvisations.

Next, I layered and mixed all these recordings. I purposefully retained most of the material, adding detailed gain automation for balance. I removed only a few highly aberrant moments from particular layers. This summed sound represents the Soloist, who is instructed only to improvise on the text, not necessarily to listen to anyone else. Subsequently, I recorded three more vocal improvisations, now playing the role of the Others, i.e, ensemble members. In the first of these three Other improvisations, I listened only to the Soloist part (no Others existed yet); in the second and third Others, I listened to the Soloist and to the previous Other(s). This corresponds with the original score’s emphasis on the listening done by the Others.

Although “Soloist” seems to imply louder dynamic level and more forward mix position than that of the Others, “For one” inverts this relationship. In treating myself as a polyphonic soloist, I magnified the almost complete freedom of this improvisation, which is limited only

slightly by the text. This also necessitated a quieter mean volume level so that the Soloist could be heard texturally. By contrast, the Others come forward in the mix and seem almost clone-like in their similar mic placement and EQ treatment.

Alternatively, I might have used a more intuitive balance: one Soloist foregrounded in the mix, and many more Others in the background. But this, I believe, would have contradicted the central concept, and concept is crucial in text scores. The Soloist needs to have a very obviously unique role – improvising and not necessarily listening – and I achieved this by multiplying the soloist many times over and sending it back in the mix. The Others, by contrast, are listening and responsively sounding according to rules, which is somewhat different from improvising. The Others' awkward, non-improvsational timing and timbral logics are obvious because they are few in number and forward in mix position. If I had mixed a lone Soloist forward and several Others backward, the difference between improvising and following rules would have been less clear.

"For one" engages with electronics primarily through digital studio methodologies. Though I could have performed the Other role using electronics, I felt more fluidly able to sound out the italicized conditions when using only voice. I made two central observations in realizing this album track: (1) A 'small chamber' realization of the score (polyphonic Soloist plus three Others) feels more nimble than the previous large chamber realizations, but shares a similar feel of sonic conversation. (2) The number two, which is implicitly important in the text score (Soloist versus Others, left column versus right), came to bear heavily on performance and mixing. For instance, the mix uses left and right channels to point up and play with the conversations among various voices. In several Soloist run-throughs I found myself articulating two materials and exploring similarities and distinctions between them. Although more research

is needed to clarify how number in text scores refracts into mixing in extended voice, it is noteworthy that this refraction occurred across media and through a seeming re-imagining of the forces of the *For soloist…* score.

F. Disability

Thus far, the idea of a normative body has been implicit in my discussion of (dis)embodiment, self-listening, and presence. I will unpack this idea using insights from disability studies, bearing in mind both the potential similarities and the crucial differences between disability and extended voice. Elizabeth Barnes (2016) argues that being disabled is "a way of being a minority with respect to one's body" (6). This is, I take it, a way of questioning body normativity while re-valuing social aspects of non-normativity. Self-listening and agency are two primary features of extended voice which problematize the notion of a normative body and figure extended voice as a social category, similar to Barnes' view of disability.

Many Western vocal traditions, including *bel canto* and its stylistic descendants, discourage performer self-listening. Self-listening is an embodied practice but, importantly, a minority embodied practice. Furthermore, I have argued that self-listening renders voice agential, but have avoided defining agency. Myriad scholarly precedents exist, but few if any account for the ways that vocal embodiment has historically diminished the agency of certain voices. Thus, it is useful to ask: What exactly is vocal agency?[28]

[28] Of course, there is some debate about the value of agency. John Robert Ferguson (2013) writes, "A relevant question might therefore be: to what extent does a musician perform the technology or does the technology perform the musician?" He goes on to quotes Hayles (1999): "Conscious agency has never been 'in control'. In fact the very illusion of control bespeaks a fundamental ignorance about the nature of the emergent processes…" (141) I am reminded of Hannah Bosma's (2013) incisive question: "isn't it ironic to be declaring the composer 'dead' just at a point in history that more female composers are coming up?" (43). Just as the 'death of the author' excluded women authors at a convenient moment,

Voice studies has begun to define vocal agency as thoughtful, personalized play. Eidsheim (2012), who is interested in questioning ostensible givens about race and vocal timbre, writes, "a person's voice is an accumulation of experiences which allows us to find and articulate individual agency within a structure that itself consists of many nuances" (23). Computer music composers, by contrast, tend to think of agency in terms of sonically parametric decision-making. Simon Emmerson (2007) writes of live compositional work by DJs: "The participants might then judge that there are indeed intentional decisions being taken (on the basis of *choices*) and ascribe this to an agent. The studio has thus moved out on stage" (26; emphasis original).

I aim to synthesize vocal-theoretical and compositional definitions of agency. I define *vocal agency* as both articulation of sonic self and commentary on genre: knowledgeable articulation of 'me' relative to 'us'. I draw from Joanna Demers' (2010) definition of genre as "a way of cultivating an us-and-them attitude that divides the public into those who know something about a given music and those who do not" (136).[29] This genre-bending aspect of agency has, for instance, driven a change in the dialogue around Cathy Berberian (1925-1983). Berberian has for decades been famous as an avant-garde vocalist, but she is increasingly understood as a composer because of her various genre-questioning projects including the piece *Stripsody*, the album *Beatles Arias*, and the patently campy recital series "À la recherche de la musique perdue."[30] I do not intend to imply that 'agents' and 'composers' are equivalent. Nevertheless, it is largely on the merit of its genre-bending character that Cathy Berberian's work is being re-assessed. Likewise, extended voice practitioners evince vocal agency because

---

agency's teeth are being drawn just as vocalists are starting to gain real agency. Thus, I am cautious about caution about agency.

[29] I argue that simply doing the genre well is not equivalent to commenting on it. For instance, a knowledgeable listener may have special appreciation for an opera singer's skillfully executed *messa di voce*, but the singer has not commented on the operatic genre (if there is such a thing) simply by executing this technique well.

[30] See Karantonis *et al.* (2014)

their quest for voice-electronics dialogue gives rise to novel soundworlds and performer-audience relationships that cross, question, and re-draw genre boundaries.[31]

AGF unites myriad vocal and electronic sounds under a banner of fractured club aesthetics in order to engage with political themes, such as the complexities of global capitalism. Guilleray's detailed vocal and electronic gestures seemingly vie to out-subtle each other, troubling perceptions of the boundary between live and pre-composed. Motland deploys electronics with a light touch in order to highlight the quasi-synthesized timbre and duration of her extended vocal techniques. Pensado lets noise overwhelm or mask the voice at times, and instead relies on the rhetorical strength of whispering (Li 2011), found recordings, and doll-based performance as a counterpoint to noise. Ratkje engages in complex vocal-control choreography to deploy two techniques to articulate form: loops (sometimes deeply transformed) and direct following of voice. Yoshida's miniatures – some containing loops (with or without micro-variations) and others not – combine howling voice and silence in a way that confuses listeners as to when exactly she is musicking. Z juxtaposes *bel canto* and spoken word delivery with composed loop textures using variably clear arrangement of phonemic content, exploring the typically unheard musicality of language and voice. My own work explores various types of continuity, noise character, textural simplicity, and forwardness in sonic field in an effort to make body at once strange and somehow similar to the machine. These distinct creative approaches are all vocally agential: they articulate a distinct vocal sound while navigating among and commenting on various genre spaces.

[31] The experimental idiom and particular uses of technology in my case studies of extended voice are not relevant (or, unfortunately, accessible) to everyone. But I think my definition of vocal agency has broader relevance. It could potentially include, for instance, vocoderist-multinstrumentalist Casey Benjamin and cellist-vocalist-electronics artist Audrey Chen. In short, I think it's the frame-alluding latter part of my definition – "commentary on genre" – which in future research has the greatest potential for inclusivity.

Like Barnes' argument for the social character of disability, the social (self-listening, genre-bending) view of extended voice withstands scrutiny where customary body-based notions do not. She writes, "What is interesting, worthwhile, useful, etc. about disability as a category is, I contend, that it's a social category that people have found useful when organizing themselves in a civil rights struggle" (41). Both disability and extended voice upend the notion of the normative body. Extended voice, which is highly personalized, works directly on the creative and technical possibility space opened up by self-listening. No longer a one-size-fits-all rule circumscribing sound and action, the body now becomes one of several technologies curated by the practitioner. The voice-electronics dialogue resulting from this curation is importantly social in that it crosses and re-imagines genre space.

Of course, it is important not to overstate the similarities between disability and extended voice. Broad-scale ableism and lack of physical accessibility profoundly impact the lives of many disabled people; extended voice practitioners cannot be said to face comparable prejudice. It is crucial to engage practices of empathetic listening and dialogue to begin to understand complex issues of intersectionality, for instance the distinct class profiles of disability and extended voice. Ongoing study will further illuminate both the similarities and the important differences between disability and extended voice, insofar as both are social categories which take important steps toward destabilizing the notion of the normative body.

V. CONCLUSION

Through technical description of my work, listening-based analysis of other extended voice practitioners' work, and critical discussion of embodiment and agency, I have argued that presence – crucially enabled by self-listening – accounts for the sonic and social advances of extended voice. I hope that, in time, vocal-electronic experimentation will be popularly normalized. Though 'extended voice' per se is mainly a creative phenomenon, technology increasingly inflects our functional communicative acts. Thus, it is not too difficult to imagine that vocal-electronic self-curation, i.e., presence, will diffuse into popular activity.

My upcoming voice-electronics research will incorporate analog electronics. I have begun to investigate sending digital voice-electronics output as a control voltage input to analog synthesizers, e.g., Buchla and Serge. I am interested in how the physical understanding of voice as voltage can advance the twin goals of enriching expressivity and empowering voices of the otherwise silenced. Furthermore, I am curious to explore both sparser soundscapes, and control that is simpler yet more precise. In various scenes and subcultures, including electronic music, instrument building, and extended voice, I perceive what I call a 'pressure for polyphony'. Yet I increasingly understand my sound as effervescently monophonic. I am curious to explore this potential. Similarly, I will continue to refine the Abacus, with the aim of developing a precise yet fluid body language for control. Proximity sensors to read embouchure, locking-position (e.g., springless) joysticks, and accelerometer and/or spectral feature data are several potential methods for control precision-yet-ease, which might in turn contribute to detailed monophony.

I have argued that extended voice, like disability, is importantly a social category which subverts the notion of the normative body. The next critical and creative step will be to delve

deeper into the notion of normative body, especially its dual sonic-social character and how it contributes to and learns from theories of intersectional identity.

Thinking adjacently about normative technological extensions of voice will help clarify what 'normative body' means in sonic contexts. To simplify and summarize, we can assume that exceeding or questioning the temporal limitations of breath and vocal stamina is an important goal in electronically extending the voice. But if we follow a disability-conscious perspective, we must acknowledge the vast range of so-called 'disabled' and 'non-disabled' bodies in the world, and thus accept the fallibility of blanket assessments of breath and vocal stamina. Possible creative and critical interventions include: (1) treating electronic sounds' potentially infinite duration as a drawback, which can be overcome through addition of human-like phrasing and pauses; (2) giving greater attention to humans' varying breath lengths, to better understand similarities and differences among bodies; and (3) re-valuing the sonic changes attendant to the inevitable depletion of stamina in long-term vocalization.

Through self-listening, which is necessarily highly personalized, and resultant presence, extended voice undermines normative notions of body and foregrounds posthuman mindbody instead. Motland's use of whistle tones and my use of noisy harmonics, for instance, point instantly and immediately to questions of bodily weirdness and aesthetic motivation. Nuanced and challenged by electronics, these sounds further gesture to second-order matters of genre-crossing. Extended voice takes crucial steps toward creatively manifesting Cavarero's (2002) argument in favor of vocal individuality. This is itself a posthuman endeavor, underscoring that the mindbody is at work making choices to electronically extend the voice.

Extended voice is a valuable sonic practice because of its unabashedly personalized experimentation with 'live' and 'recorded' methodologies, its innovation in timbral and temporal

interactions between vocal and electronic sounds, and, therefore, its demonstrated ability to welcome and cultivate more diverse participation than that which currently typifies electronic music overall. Continued research and creative work will help elucidate the ways in which extended voice fundamentally re-calibrates the relationships between listening, sounding, and valuing.

BIBLIOGRAPHY

Auner, Joseph. "'Sing It For Me': Posthuman Ventriloquism in Recent Popular Music." *Journal of the Royal Music Association* 128(1) (2003): 98-122.

Baars, Girilal. "Ballads and Ohms: Vocal traditions, electronics and compositional strategies." PhD dissertation. University of Huddersfield, 2015.

Barnes, Elizabeth. *The Minority Body: A Theory of Disability.* Oxford: Oxford University Press, 2016.

Bell, Sarah. "The dB in the .dB: Vocaloid Software as Posthuman Instrument." *Popular Music and Society* 39(2) (2016): 222-240.

Blonk, Jaap, with Tomomi Adachi, Owen F. Smith, Duane Ingalls. *4 Contemporary Sound Poets*. Pogus Productions 21083-2, CD, 2016. Liner notes: http://www.pogus.com/21083.html

Borch, D. Zangger, J. Sundberg, P.-Å. Lindestad and M. Thalén, "Vocal fold vibration and voice source aperiodicity in 'dist' tones: a study of a timbral ornament in rock singing." *Logopedics Phoniatrics Vocology* 29(4) (2004): 147-153.

Bosma, Hannah. "Playing Loudspeakers, Unsettling Concerts: Gender and Performance in Interdisciplinary Electroacoustic Music." Proceedings of the Electroacoustic Music Studies Network Conference, Electroacoustic Music Beyond Performance, Berlin, June 2014.

----. "The electronic cry: Voice and gender in electroacoustic music." PhD dissertation. University of Amsterdam, 2013.

Brecht, George. *Water Yam*. New York: Fluxus Editions, 1963.

Brown, Linda. "The Beautiful in Strangeness: The Extended Vocal Techniques of Joan La Barbara." PhD dissertation. University of Florida, 2002.

Cage, John. "The Future of Music: Credo." *Silence*. Middletown, Connecticut: Wesleyan University Press, 1969.

Cavarero, Adriana. *For More Than One Voice: Toward a Philosophy of Vocal Expression.* Stanford, California: Stanford University Press, 2005.

Connor, Steven M. *Dumbstruck: A Cultural History of Ventriloquism.* New York: Oxford University Press, 2000.

Cormier, Ken. "Writing the Tape-Recorded Life." *A/b: Auto/Biography Studies* 27, no. 2 (2012): 402-26.

Danson-Faraday, Jacob. "The Composition and Performance of *Hyste*: A Work for Voice and Electronics." Masters thesis. University of Edinburgh, 2010.

Davidson, Michael. "Technologies of Presence: Orality and the Tapevoice of Contemporary Poetics." In Morris, Adalaide (ed.), *Sound States*. Chapel Hill and London: University of North Carolina Press, 1997. 97-125.

Demers, Joanna. *Listening Through the Noise: The Aesthetics of Experimental Electronic Music.* Oxford and New York: Oxford University Press, 2010.

Dolar, Mladen. *A Voice and Nothing More.* Cambridge, MA: MIT Press, 2006.

Eidsheim, Nina. "Race and the aesthetics of vocal timbre." In *Rethinking Difference in Music Scholarship*, edited by Olivia Bloechl, Melanie Lowe, and Jeffrey Kallberg, 338-365. Cambridge: Cambridge University Press, 2015.

----. "Voice as Action: Toward a Model for Analyzing the Dynamic Construction of the Racialized Voice." *Current Musicology* 93 (Spring 2012): 9-33.

Eileraas, Karina. "Witches, bitches & fluids: Girl bands performing ugliness as resistance." *TDR* 41(3) (1997): 122-139.

Elektronski Studio III Program Radio Beograda. Facebook. 11 Jan 2017. Accessed 20 Jan 2017. https://www.facebook.com/elektronskistudio3program/?hc_ref=PAGES_TIMELINE&fref=nf

Emmerson, Simon. *Living Electronic Music.* Aldershot, England and Burlington, Vermont: Ashgate, 2007.

Faudree, Paja. "Music, Language, and Texts: Sound and Semiotic Ethnography." *Annual Review of Anthropology* 41 (2012): 519-36.

Ferguson, John Robert. "Imagined Agency: Technology, Unpredictability, and Ambiguity." *Contemporary Music Review* 32(2-3) (2013): 135-149.

Gaburo, Kenneth (ed). *Allos*: 'Other' Language, 41 Writers of 41 'Writings.' La Jolla, CA: Lingua Press, 1980.

Gordon, Bonnie. "The Castrato Meets the Cyborg." *Opera Quarterly* 27(1) (Winter 2011): 94-122.

Gourlay, Lesley. "Cyborg ontologies and the lecturer's voice: a posthuman reading of the 'face-to-face.'" *Learning, media and technology* 37(2) (2012): 198-211.

Greie-Ripatti, Antye, email message to author, February 14, 2017.

----. "AGF aka poemproducer" [homepage]. http://antyegreie.com/ Accessed 16 March 2017.

----. “Performances/ 05 live and appearances.” Poemproducer. http://poemproducer.com/dates.php Accessed 16 March 2017.

----. Tumblr. 2 September 2016. http://poemproducer.tumblr.com/post/149826940489/i-facilitated-the-sonic-wilderness-camp-in?is_related_post=1. Accessed 4 January 2017.

Griffin, Farah Jasmine. “When Malindy Sings: A Meditation on Black Women’s Vocality.” In *Uptown Conversation*, edited by Robert G. O’Meally, Brent Hayes Edwards, Farah Jasmine Griffin, 102-125. New York: Columbia University Press, 2004.

Guilleray, Marie, email message to author, January 24, 2017.

----, email message to author, September 2, 2016.

----. “Towards a fluent electronic counterpart of the voice.” Masters Thesis. Institute of Sonology, Royal Conservatory of The Hague, 2012.

Halfyard, Janet. “Text and authority: Issues of gender and ownership in the work of Cathy Berberian.” Paper presentation, Cathy Berberian conference, Univ of Amsterdam, 2006.

Haraway, Donna. “A Cyborg Manifesto.” In *Simians, Cyborgs, and Women: The Reinvention of Nature*, 149-181. New York: Routledge, 1991.

Hayles, N. Katherine. “Voices out of Bodies, Bodies out of Voices: Audiotape and the Production of Subjectivity.” In *Sound States*, edited by Adalaide Morris, 74-96. Chapel Hill and London: University of North Carolina Press, 1997.

Herndon, Holly. “Embodiment in Electronic Music Performance.” Masters thesis. Mills College, 2010.

Hewitt, Donna. “Choreographic approaches to music composition for a new musical interface: the eMic.” *Proceedings of the International Symposium on Performance Science*, Toronoto (2011). 169-174.

----. “Compositions for Voice and Technology.” PhD dissertation. University of Western Sydney, 2006.

Hughes, Diane. “Technologized and autonomized vocals in contemporary popular musics.” *Journal of Music, Technology and Education* 8(2) (2015): 163-182.

Impett, Jonathan. “Let’s Dance Architecture: Improvisation, Technology and Form.” *Leonardo Music Journal* 20 (2010): 85-88.

Jarman-Ivens, Freya. *Queer Voices: Technologies, Vocalities, and the Musical Flaw*. New York: Palgrave MacMillan, 2011.

Johnson, Tom. “Imagining” in *Private Pieces*. New York: Two-Eighteen Press, 1967.

Kane, Brian. *Sound Unseen : Acousmatic Sound in Theory and Practice*. New York: Oxford University Press, 2014.

Karantonis, Pamela, Francesca Placanica, Anne Sivuoja-Kauppala, Pieter Verstraete P (eds). *Cathy Berberian: Pioneer of Contemporary Vocality*. Farnham, UK: Ashgate Publishing, 2014.

Khazan, Olga. “Would You Really Like Hillary More If She Sounded Different?” The Atlantic [online], 1 August 2016. https://www.theatlantic.com/science/archive/2016/08/hillarys-voice/493565/. Accessed 3 March 2017.

Kjus, Yngvar and Anne Danielsen. “Live mediation: performing concerts using studio technology.” *Popular Music* 35(3) (2016): 320-337.

Klein, Julieanne. “Voice and live electronics: An historical perspective.” *eContact!: Online Journal for Electroacoustic Practices* 10(4) (October 2008). http://cec.sonus.ca/econtact/10_4/klein_livevoice.html#recent. Accessed 6 January 2017.

Knouf, Nicholas Adrian. “Noisy Fields: Interference and Equivocality in the Sonic Legacies of Information Theory.” PhD dissertation. Cornell University, 2013.

Knowles, Julian and Donna Hewitt. “Performance Recordivity.” *Journal on the Art of Record Production* 6 (June 2012) [online]. http://arpjournal.com/performance-recordivity-studio-music-in-a-live-context/. Accessed 6 January 2017.

La Barbara, Joan. “Voice is the original instrument.” *Contemporary Music Review* 21(1) (2002): 35-48.

LaBelle, Brandon. *Lexicon of the Mouth*. New York and London: Bloomsbury Academic, 2014.

Lakoff, Jeremy. “‘A Variable and yet Perfectly Precise Mechanism’: Temporal Prosthesis in *The Secret Agent*.” *The Conradian* 40(1) (Spring 2015): 1-14.

van Leeuwen, Theo. *Speech, Music, Sound*. London: MacMillan Press, 1999.

Lewis, George. “The Virtual Discourses of Pamela Z.” *Journal of the Society for American Music* 1(1) (February 2007): 57-77.

Li, Xinghua. “Whispering: The Murmur of Power in a Lo-Fi World.” *Media, Culture & Society* 33, no. 1 (January 1, 2011): 19–34.

Looser, “Freeing the Natural Voice?” Masters thesis. University of Canterbury, 2005.

Madaule, Paul. “Listening and Singing.” *Journal of Singing of the National Association of Teachers of Singing* 57(5) (May 2001): 15-20.

Mitchell, David T. and Sharon L. Snyder, "Narrative Prosthesis and the Materiality of Metaphor: Literare and the Undisciplined Body of Disability." In *Narrative Prosthesis: Disability and the Dependencies of Discourse*. Ann Arbor: University of Michigan Press, 2000. 47-64.

Moliné, Keith. "Hear me roar." *Wire* 278 (April 2007): 24-29.

Morrison, Toni. "Toni Morrison on the 'Spoken Library'" (excerpt of speech at NCTE National Convention). *English Journal* 67(2) (February 1978): 29.

Motland, Stine Janvin, email message to author, January 24, 2017.

----, email message to author, September 12, 2016.

----. "Fake Synthetic Music." Stinesthetics.com (professional website of S. J. Motland). 2016. https://stinesthetics.com/2016/09/12/fake-synthetic-music/. Accessed 6 January 2017.

Mulvey, Laura. "Visual Pleasure and Narrative Cinema." In *Visual and Other Pleasures*, Second edition. Houndsmills, Basingstoke, Hampshire: Palgrave Macmillan, 2009. 14-30.

Neset, Anne Hilde. "Maja Ratkje: Woman behaving madly." *WIRE* 229 (March 2003). 12.

Neuhaus, Max. *LISTEN*. In Kahn D, Austin L (eds). *Source: Music of the Avant-Garde, 1966-1973*. Berkeley: University of California Press, 2011.

Neumark N, Gibson R, van Leeuwen T (eds). *VOICE: Vocal Aesthetics in Digital Arts and Media*. Cambridge, Massachusetts: MIT Press, 2010.

Oliveros, Pauline. "Six for New Time." In Golter, Samuel and Lawton Hall (eds.), *Anthology of Text Scores*. Kingston, NY: Deep Listening Publications, 2013.

Painter, Desmond. "The Voice Devoid Of Any Accent: Language, Subjectivity, And Social Psychology." *Subjectivity* 23(1) (2008): 174–87.

Pensado, Andrea. "Bio." http://andreapensado.com/Andrea_Pensado/Bio.html. Accessed 16 March 2017.

----, email message to author, January 25-26, 2017.

----, email message to author, September 2, 2016.

Pisaro, Michael. *fragile being, hopeful becoming*. Mettmann, Germany: Edition Wandelweiser. 2013.

----. "Writing, Music." In Saunders, James (ed.), *The Ashgate Research Companion to Experimental Music.* London: Ashgate, 2009.

Pisaro, Michael. *Ascending Series (4)*. Mettmann, Germany: Edition Wandelweiser. 2008.

Pluta, Sam. "Laptop Improvisation in a Multi-Dimensional Space." PhD diss., Columbia University, 2012.

Ratkje, Maja S. K., email message to author, February 13, 2017.

----, email message to author, August 17, 2016.

"Review: Stine Janvin Motland – In Labour." Attn:Magazine [online], 4 May 2014. http://www.attnmagazine.co.uk/music/8120. Accessed 6 December 2016.

Roads, Curtis. *Microsound*. Cambridge, MA: MIT Press, 2001.

Rodgers, Tara. *Pink Noises: Women on Electronic Music and Sound*. Durham, NC: Duke University Press, 2010.

Saunders, James. "Testing the Consequences: Multipart Series in the Work of Wandelweiser Composers." *Contemporary Music Review* 30(6) (2011): 497-524.

Schlichter, Annette. "Do Voices Matter? Vocality, Materiality, Gender Performativity." *Body and Society* 17(1) (March 2011): 31-52.

Scrutton, Nichola Jane. "Hearing voices: electroacoustic composition portfolio and commentary." PhD dissertation. University of Glasgow, 2009.

Shank, Barry. *The Political Force of Musical Beauty*. Durham, NC: Duke University Press, 2014.

Shankar, Shalini, and Jillian R. Cavanaugh. "Language and Materiality in Global Capitalism." *Annual Review of Anthropology* 41(1) (2012): 355–69.

Smith, Jacob. *Vocal Tracks: Performance and Sound Media*. Berkeley: University of California Press, 2008.

Stamper, Chloe. "Our bodies, ourselves, our sound producing circuits: feminist musicology, access, and electronic instrument design practices." Masters Thesis. Mills College, 2015.

Stockhausen, Karlheinz. *Für Kommende Zeiten: 17 Texte für Intuitive Musik*. Kürten, Germany: Stockhausen-Verlag, 1976.

Sunabacka, Karen. "She is the voice, she is the sound: Women's voices in electroacoustic music, and, 'And then there was great calm.'" PhD Dissertation. UC Davis, 2008.

Taylor, Gregory. “An Interview with Antye Greie-Ripatti (AGF).” Cycling ’74. https://cycling74.com/2016/02/23/an-interview-with-antye-greie-ripatti-agf/. Accessed 3 Dec 2016.

Vágnerova, Lucie. “Sirens/Cyborgs: Sound Technologies and the Musical Body.” PhD dissertation. Columbia University, 2016.

Wakefield, Gregory. “Vocal Pedagogy and Pedagogical Voices.” *Proceedings of the International Conference on Auditory Display*, 2003, Boston. 2-11.

Wallace, Maurice. “Print, Prosthesis, (Im)Personation: Morrison’s Jazz and the Limits of Literary History.” *American Literary History* 20, no. 4 (December 21, 2008): 794–806.

Warren, Kristina. “Notated Control as Composed Liveness in Works for Digitally Extended Voice.” *Proceedings of the Third International Conference on Technologies for Music Notation and Representation*, A Coruña (forthcoming 2017).

----. “Composing and Performing Digital Voice Using Microphone-Centric Gesture and Control Data.” *Proceedings of the International Computer Music Conference*, Utrecht (2016). 439-442.

Weber-Lucks, Theda. “Electroacoustic voices in performance art – a gender issue?” *Organised Sound* 8(1) (2003): 61-69.

Weheliye, Alexander. “‘Feenin’: Posthuman Voices in Contemporary Black Popular Music.” *Social Text* 20(2) (2002): 21-47.

Weidman, Amanda. “Anthropology and Voice.” *Annual Review of Anthropology* 43 (2014): 37-51.

Wishart, Trevor. *On Sonic Art*. Amsterdam: Harwood Academic Publishers, 1996.

Wolff, Christian. “Prose Scores.” *Cues: Writings & Conversations.* Köln: MusikTexte, 1998.

Wolman, Amnon. *January 4, 2001*. In Möller T, Shim K, Stäbler G (eds). *SoundVisions*. Saarbrücken, Germany: PFAU-Verlag. 2005.

Woloshyn, Alexa Lauren. “The recorded voice and the mediated body in contemporary Canadian electroacoustic music.” PhD dissertation. University of Toronto, 2012.

Young, Miriama. *Singing the Body Electric: The Human Voice and Sound Technology.* London and New York: Routledge, 2015.

----. “Latent body – plastic, malleable, inscribed: The human voice, the body and the sound of its transformation through technology.” *Contemporary Music Review* 25(1-2) (2006): 81-92.

Z, Pamela, email message to author, January 26, 2017.

----. “A Tool is a Tool.” In Malloy, Judy (ed.), *Women, Art, and Technology*. Cambridge, MA: MIT Press, 2003. 348-361.

Zorn, Jonathan. “Voices in the Electronic Furnace: Towards an Analysis of Electroacoustic Vocal Music.” PhD dissertation. University of Virginia, 2012.

DISCOGRAPHY

AGF, *Kon:3p>UTION to: e[VOL]ution*. AGF Producktion, CD, 2016.

----. *Westernization Completed*. AGF Production, CD, 2003.

Guilleray, Marie. *Entre chien et loup*. SoundCloud. November 2012. https://soundcloud.com/marie-guilleray/entre-chien-et-loup. Accessed 17 December 2016.

Guilleray, Marie and Johan van Kreij. SoundCloud. *Après la mer*. October 2014. https://soundcloud.com/marie-guilleray/apres-la-mer-stereo-mix. Accessed 17 December 2016.

Motland, Stine Janvin (as Stine II). *Born In An Early Age*. SoundCloud. Februrary 2016. https://soundcloud.com/stinejm/stine_ii_born-in-an-early-age. Accessed 4 January 2017.

----. *Fake Synthetic Music.* 2016. https://stinesthetics.com/2016/09/12/fake-synthetic-music/. Accessed 4 January 2017.

----. *In Labour*. Pica Disk PICA033, LP, 2014.

Ratkje, Maja. “Maja S. K. Ratkje live at the Punkt festival 2013.” YouTube. 17 October 2013. https://www.youtube.com/watch?v=1VDc0ac6PLg. Accessed 17 December 2016.

----. *Voice.* Rune Grammofon RCD2028, CD, 2002.

Ratkje, Maja and Säkä. *Rasaka*. KBA002, CD, 2016.

Yoshida, Ami. *Tiger Thrush*. Improvised Music from Japan IMJ-504, CD, 2003.

Z, Pamela. *A Delay is Better.* Starkland ST-213, CD, 2004.

----. *Voci* (excerpts) [video]. PamelaZ.com. http://www.pamelaz.com/VociMov.html. ODC Theater, San Francisco (2003). Accessed 18 March 2017.

Appendix A Voice-electronics performances

2017

May 25 A Coruña [ES]
For ami

May 4 Charlottesville, VA
Solo improvised voice-electronics set

March 8 Atlanta, GA
Solo improvised voice-electronics set
https://soundcloud.com/kristinawarrencomposer/intlwomensday2017

March 4 Huddersfield [UK]
Solo improvised voice-electronics set
https://soundcloud.com/kristinawarrencomposer/yswn_intlwomensday

January 28 Liverpool [UK]
Merseyside Improvisers Orchestra: conduction, free improvisation, open works

2016

December 16 Manchester [UK]
Improvised set, AKA trio: Ailís Ní Ríain (prepared piano), Amy Brandon (guitar, electronics)
https://www.youtube.com/watch?v=z6LzcSySEXc&feature=youtu.be

December 15 Liverpool [UK]
Improvised set, AKA trio

November 30 London [UK]
Improvised voice-electronics sets
https://soundcloud.com/kristinawarrencomposer/30nov-perf

November 26 Liverpool [UK]
Merseyside Improvisers Orchestra: conduction, free improvisation, open works

October 1 Charlottesville, VA
Solo improvised voice-electronics set
http://medie.kmwarren.org/warren_wcw.wav

September 16 The Hague [NL]
Is There Time
http://medib.kmwarren.org/IsThereTime_loos_mix.wav

Appendix A (cont.) Voice-electronics performances

2016 (concl.)

September 10 Glasgow [UK]
DNR

September 8 Derry [UK]
Is There Time

May 15 - June 4 New Smyrna Beach, FL
Daily group improvisations (Atlantic Center for the Arts)

April 23 Charlottesville, VA
Improvised set, Aorist: Rachel Devorah (horn, electronics), Kimberly Sutton (cello, electronics)

April 22 Charlottesville, VA
Who Freed
https://www.youtube.com/watch?v=DwoAf2hlMNQ

March 23 Charlottesville, VA
Improvised set with Kevin Davis (cello, effects pedals), Rachel Devorah (horn), Marc Mazique (drums)

March 18 Charlottesville, VA
Solo improvised voice-electronics set
https://www.youtube.com/watch?v=b8796oDeluU

February 13 Statesboro, GA
Baltimore Sits
http://medie.kmwarren.org/BaltimoreSits.wav

2015

October 29 Charlottesville, VA
Solo improvised voice-electronics set
https://www.youtube.com/watch?v=FRm6ODYls4k

October 17 Charlottesville, VA
Schemer
https://www.youtube.com/watch?v=IbkX8SBE7l8

Appendix A (concl.) Voice-electronics performances

2015 (concl.)

September 30 Denton, TX
Look the Other Way
https://www.youtube.com/watch?v=LnVMHuVnm_E

July 3 Charlottesville, VA
Improvised set with Kevin Davis (cello, effects pedals), Rachel Devorah (horn)

June 24 New York, NY
Look the Other Way
https://vimeo.com/133399480 (excerpt)

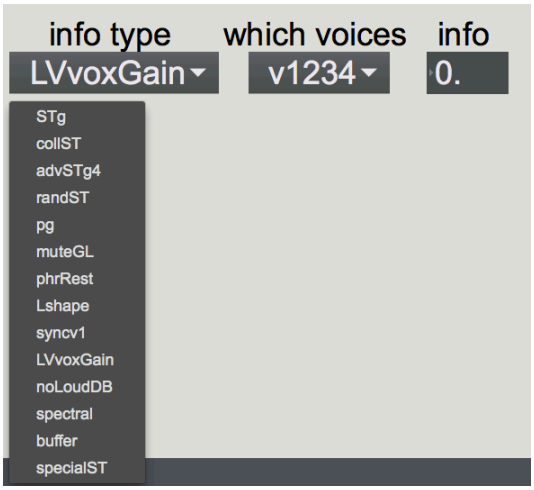
April 18 Charlottesville, VA
Look the Other Way

2014

November 14 Fredericksburg, VA
Look the Other Way

November 6 Charlottesville, VA
Look the Other Way
https://www.youtube.com/watch?v=xTORGvxyh-8

Appendix B Spring 2016 patch: Composed control routine



My Spring 2016 Max patch had 14 control parameters – ranging from spectral treatment, to grammar for rhythmic playing of grains – which could be sent to any combination of the (4) rhythmic granulation voices. It was too difficult to manage this complex system while performing, so instead I composed routines for automated control. The table below contains the routine used for my March 2016 performance at the Twisted Branch Tea Bazaar (Charlottesville, Virginia).

Time	Event
load (0:00)	voices 1234: STg = 4, collST = 5, no loud downbeats = 1, adv thru ST coll 11x
0:32	voice 2: gain = 2
1:11	voices 34: playgrammar = 6
1:31	voices 134: gain = 1.5
2:04	voice 1: playgrammar = 2, (2:07) 5, (2:20) 2, (2:26) 5
3:00	voices 14: playgrammar = 2, (3:10) 5
3:19	voices 13: playgrammar = 2
3:37	voices 1234: gain = 2
4:05	voice 3: playgrammar = 5
4:30	voice 4: playgrammar = 2
4:47	voices 1234: playgrammar = 2
5:19	voice 4: lineshape = 14, (5:25) 4, (5:30) 3, (5:33) 4, (5;46) 3
6:03	voice 2: lineshape = 12, (6:18) 3, (6:24) 12, (6:30) 13, (6:32) 12, (6:35) 13, (6:37) 12, (6:41) 3
6:52	voices 13: gain = 2.5
7:25	voice 2: playgrammar = 3
7:37	voice 2: gain = 3
7:43	voice 2: adv thru ST coll 1x, (7:46) 1x, (7:48) 1x, (7:54) 1x, (8:35) 1x, (8:37) 1x
8:43	voices 1234: adv thru ST coll 1x
8:59	voices 134: gain = 2.5, (9:01) 3
9:07	voices 1234: adv thru ST coll 1x, (9:09) 1x, (9:11) 1x
9:21	voices 1234: gain = 0
9:45	voices 1234: playgrammar = 4, spectral = 3
11:01	voices 14: gain = 1
11:08	voices 1234: phraseRest = 1
11:49	voices 23: gain = 1.5, (12:20) 2

Time	Event
12:24	voices 14: gain = 2
12:45	voice 3: playgrammar = 2
12:58	voices 1234: adv thru ST coll 1x
13:17	voice 3: spectral = 1
13:31	voice 3: playgrammar = 2, (13:46) 3
13:55	voice 3: muteGrainLen = 1, (14:26) 0, (14:57) 1, (15:28) 0, (15:59) 1
14:19	voice 3: adv thru ST coll 1x, (14:22) 1x, (14:24) 1x, (14:26) 1x, (14:29) 1x
14:41	voice 3: lineshape = 2
14:56	voice 4: playgrammar = 3
15:01	voice 4: spectral = 1
15:19	voices 12: gain = 3
15:53	voice 1: playgrammar = 3
16:11	voice 2: spectral = 1
16:26	voices 34: gain = 3
16:45	voice 12: gain = 4
17:03	voices 34: playgrammar = 1, (17:07) 3
17:17	voices 1234: trigger random ST 1x
17:28	voices 1234: playgrammar = 3
17:52	voices 14: lineshape = 5
18:09	voice 1: gain = 3.5, (18:10) 4, (18:12) 4.5
18:22	voices 34: gain = 3, (18:24) 2.5, (18:32) 2, (18:35) 1.5
18:47	voices 124: spectral = 3
19:00	voices 124: gain = 4
19:15	voices 34: gain = 3.5, (19:18) 3
19:24	voices 12: gain = 5, (19:27) 5.5, (19:45) 6
20:10	voice 3: spectral = 2
20:24	voice 3: gain = 3, (20:26) 4
21:03	voice 4: spectral = 1; (21:06) voice 2, spectral = 1; (21:10) voice 1, spectral = 1; (21:17) voice 3, spectral = 1

Appendix C Vox6 patch: Rhythmic granulation parameters

Vox6 has sixteen voices doing rhythmic granulation, divided into (4) groups of (4) voices each. In Groups 1-2 there is a mono flanger after each granulation voice, while in Groups 3-4 there is no flanger. Upon trigger (Abacus button), 4 distinct random parameter clusters are generated, i.e., one parameter cluster per group. There are four parameters which can be affected:

Start Time in source buffer
Length
Tempo
Flange preset (Groups 1-2 only)

(Vox5, an earlier patch, used these parameters along with two more: Rhythmic grammar and Transposition (FFT-based). However, Vox6 eliminates rhythmic grammar because variations were not clearly audible, and eliminates transposition because it was too computationally expensive.)

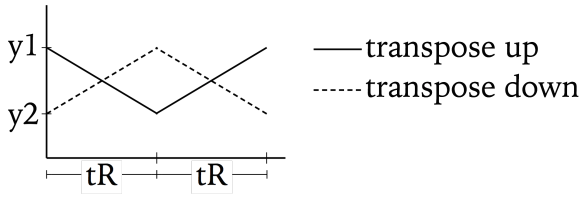
There is a 25% chance that one parameter will be affected, 50% chance two, and 25% chance three parameters will be affected. Start Time, Length, and Tempo can all receive either a single random value or one of a few pre-composed Low Frequency Oscillators (LFOs), while flange receives a random 1 of 5 pre-composed presets.

For instance, a single trigger might output the following four parameter clusters:

Group 1	Tempo = 1060 bpm Flange = preset 1
Group 2	Tempo = 675 bpm
Group 3	Start Time = 1300 ms
Group 4	Tempo = LFO #2 Start Time = LFO #3 Length = LFO #1

Appendix D Vox6 patch: Transposition-delay unit

When an input signal is sent to tapin~ and a V-shaped ramp is sent to tapout~, the input is transposed and delayed. I repeatedly triggered a bit of sound (sine wave) to tapin~ with different ramp values to tapout~, and measured tap delay time and fundamental frequency (using Tristan Jehan's pitch~ object). I observed the following for ramp format [y1, y2 tR y1 tR] :



Particular ramp slopes give particular transposition intervals:

ratio = tR/ | y1 – y2 | by transposition direction and interval								
transpose down					**transpose up**			
m2	20	**P5**	3		**m2**	20	**P5**	2
M2	8.75	**m6**	2.72		**M2**	8	**m6**	1.7
m3	6.3	**P6**	2.5		**m3**	5	**P6**	1.5
M3	4.55	**m7**	2.3		**M3**	4	**m7**	1.3
P4	4	**P7**	2.15		**P4**	3	**P7**	1.15
+4	3.5	**P8**	2		**+4**	2.3	**P8**	1

Initial ramp value and tap delay time are linearly correlated:

tD = a*y1 + b tD = delay time y1 = initial ramp value						
transpose down				**transpose up**		
	m	**b**			**m**	**b**
m2	1.052	14.425		**m2**	0.9535	-3.1243
M2	1.311	13.148		**M2**	0.8892	-1.2294
m3	1.1915	10.451		**m3**	0.8356	-3.2577
M3	1.2795	15.17		**M3**	0.7983	1.619
P4	1.3339	12.006		**P4**	0.7505	-0.1507
+4	1.3985	16.106		**+4**	0.6966	4.3436
P5	1.498	14.614		**P5**	0.6713	7.0204

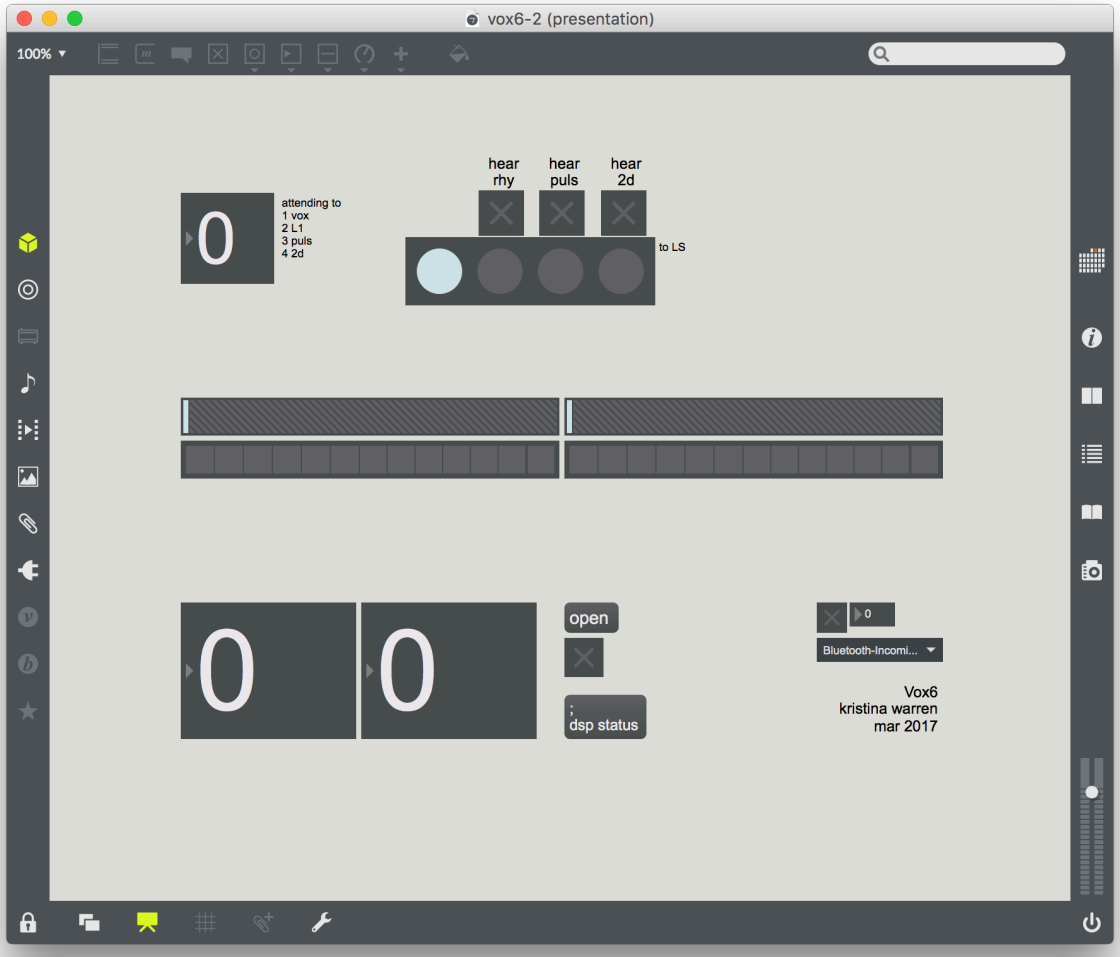
Appendix D (concl.) Vox6 patch: Transposition-delay unit

Finally, I summarized this information to allow calculation of ramp values using desired transposition (cents) and delay time (ms). Notably, since empirical measurements drove calculation of regression equations, ramp values yield approximate, not precise, transposition and delay values. I enjoy this slightly detuned and a-rhythmic sound, so I have maintained the imprecision.

Given c = transposition (cents) and tD = delay time (ms), calculate ratio = tR/ | y1 – y2 | and y1 = (tD – b) / a (let | y1 – y2 | = 200)				
Transpose down			**Transpose up**	
Ratio	**a, b for y1**		**Ratio**	**a, b for y1**
1505.2c ^ -0.956	a = 1.0037e^0.0006c b_avg = 13.703		4132.3c ^ -1.168	a = 1.0045e^-0.0006c b = 0.016c – 4.6711

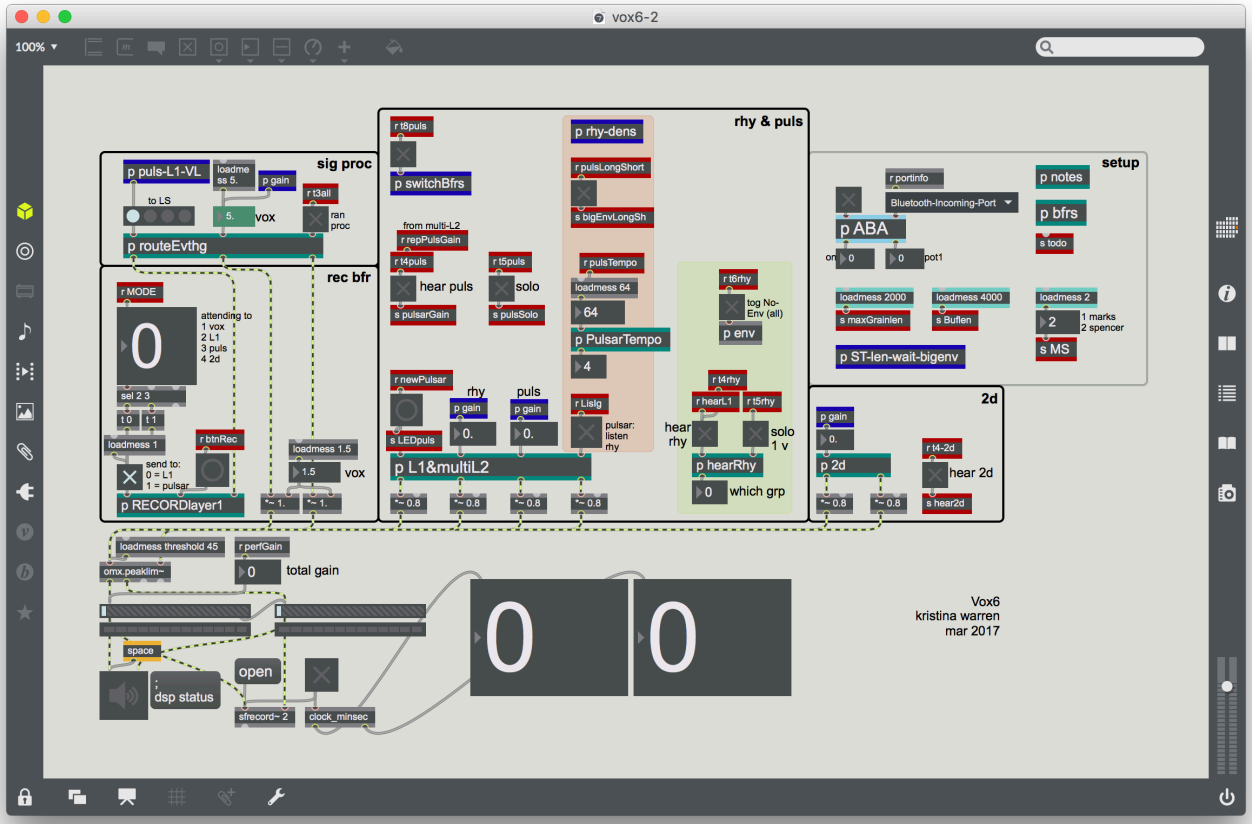
Appendix E Vox6 patch

Vox6 presentation mode (viewed during performance)



Appendix E (cont.) Vox6 patch

Vox6 patching mode (for editing)



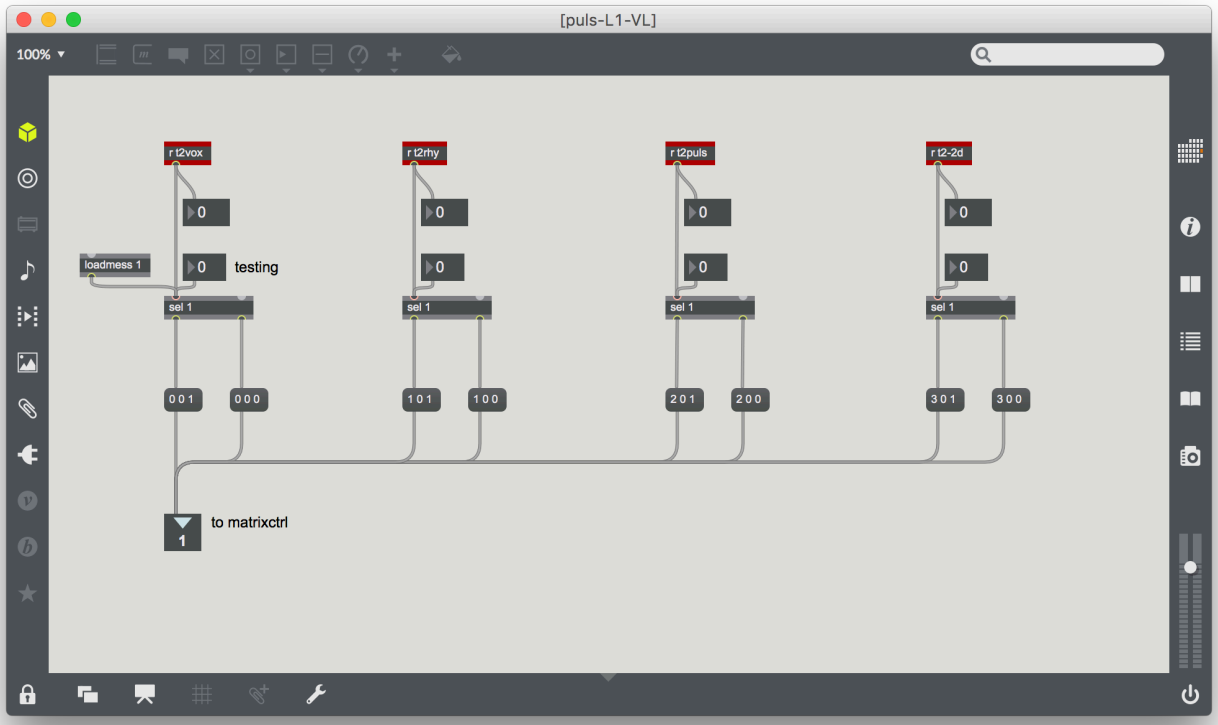
Appendix E (cont.) Vox6 patch

This chart shows nested subpatches within Vox6 in order to explain the ensuing screenshots. Redundant or self-explanatory subpatches (e.g., 8.3 grps2-4) are not pictured.

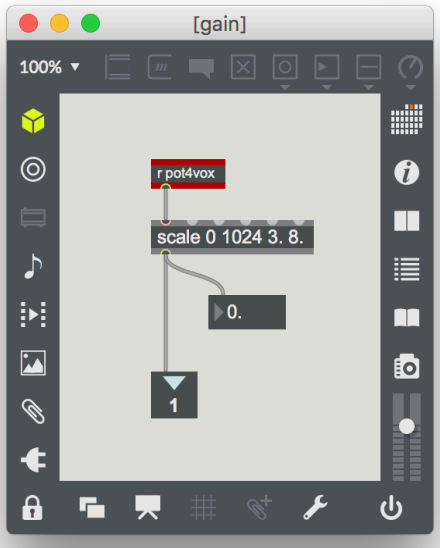
1	puls-L1-VL					
2	gain (vox)					
3	routeEvthg					
	1	voxCtl				
	2	t3-to-randVox				
	3	LiveSig				
		1	pong			
		2	trDel-Short-Long			
			1	trDelAll (poly~)		
				1	decideDel-sendDelTransp	
				2	ramp	
					1	transpDown
		3	verb			
		4	del			
	4	pan				
4	RECORDlayer1					
5	switchBfrs					
6	rhy-dens					
7	PulsarTempo					
	1	SoundingQtrs				
	2	RhyPhrases				
		1	rhyPhrase			
8	L1&multiL2					
	1	RecL1then4xProcessing				
		1	g1ranpro-flange			
			1	LFO-ST		
	2	grp1tempo				
	3	grp1				
		1	meteredPlay-v2-grp1			
			1	presetsQtrUnits		
				1	odds-gate	
				2	detgridlen-decideActualen	
					1	detgridlen
					2	decidelen
		2	kw-pulsarvoice3			
			1	jitterST		
				1	jitST	
			2	play		
				1	jitterLen	
		3	flanger_mono			
		4	route			
			1	grp1solo		
			2	toLS		
	4	vary-pgST				
	5	lsn-ignRhy				
		1	bigEnv1			
			1	maxgainTime		
	6	ROUTE				
		1	pulsSolo			
9	hearRhy					
	1	g1ranpro-flange-L1only				
10	ABA					
	1	serial				
	2	led				
	3	slowFast				
	4	RESET				
	5	fastFlip				
	6	t2route				
11	ST-len-wait-bigenv					
	1	ctlvar				
		1	ctlvar_store			
			1	mean-stdev		
		2	ctlvar_do			
	2	send-ST				
12	2d					
	1	2dvox				
	2	randBfr				

Appendix E (cont.) Vox6 patch

1. puls-L1-VL

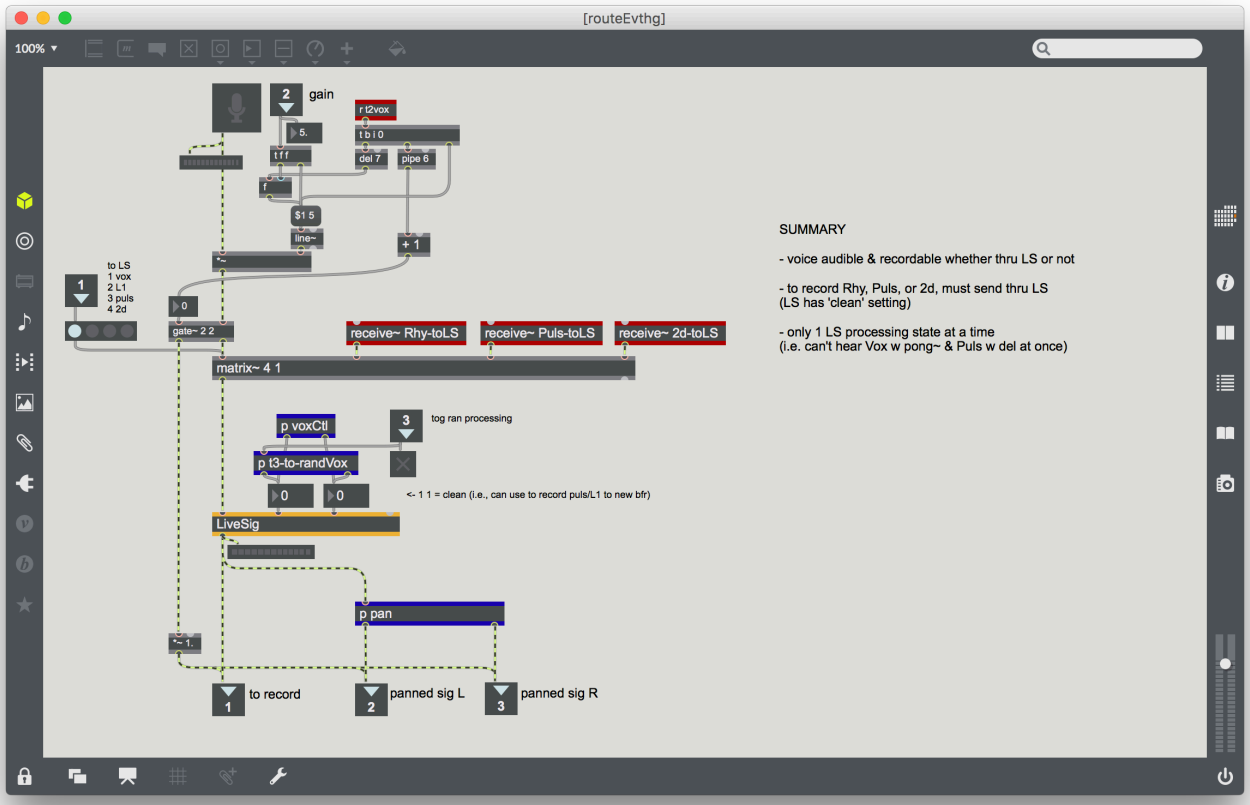


2. gain (vox)

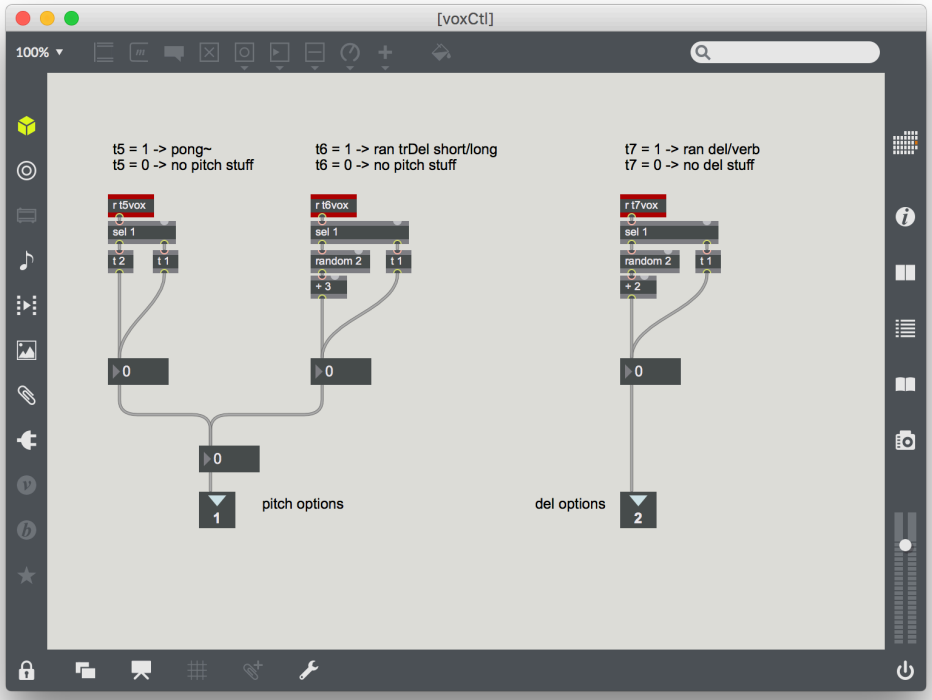


Appendix E (cont.) Vox6 patch

3. routeEvthg

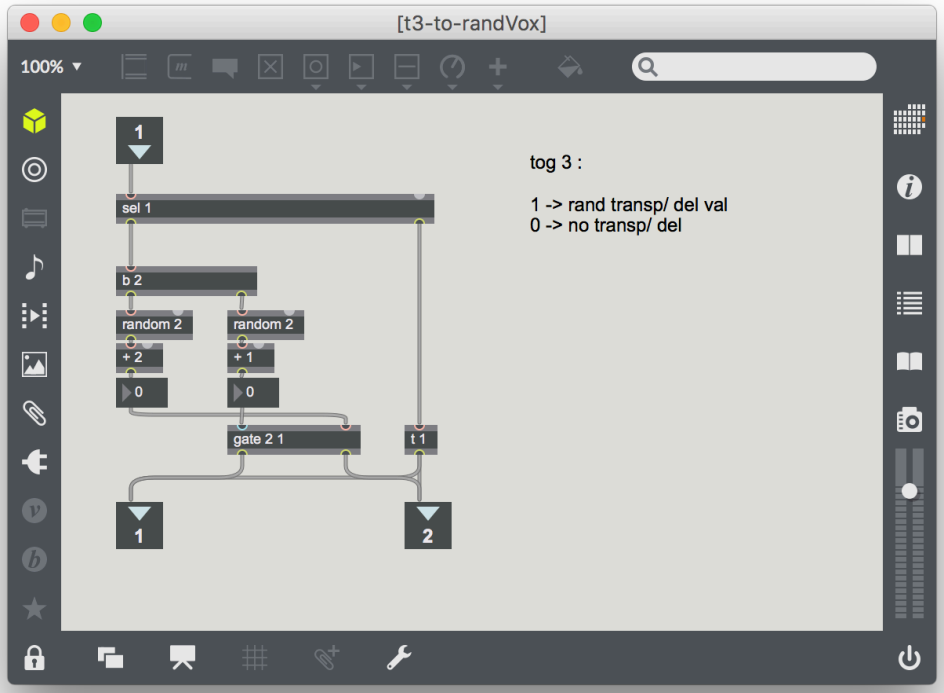


3.1 voxCtl



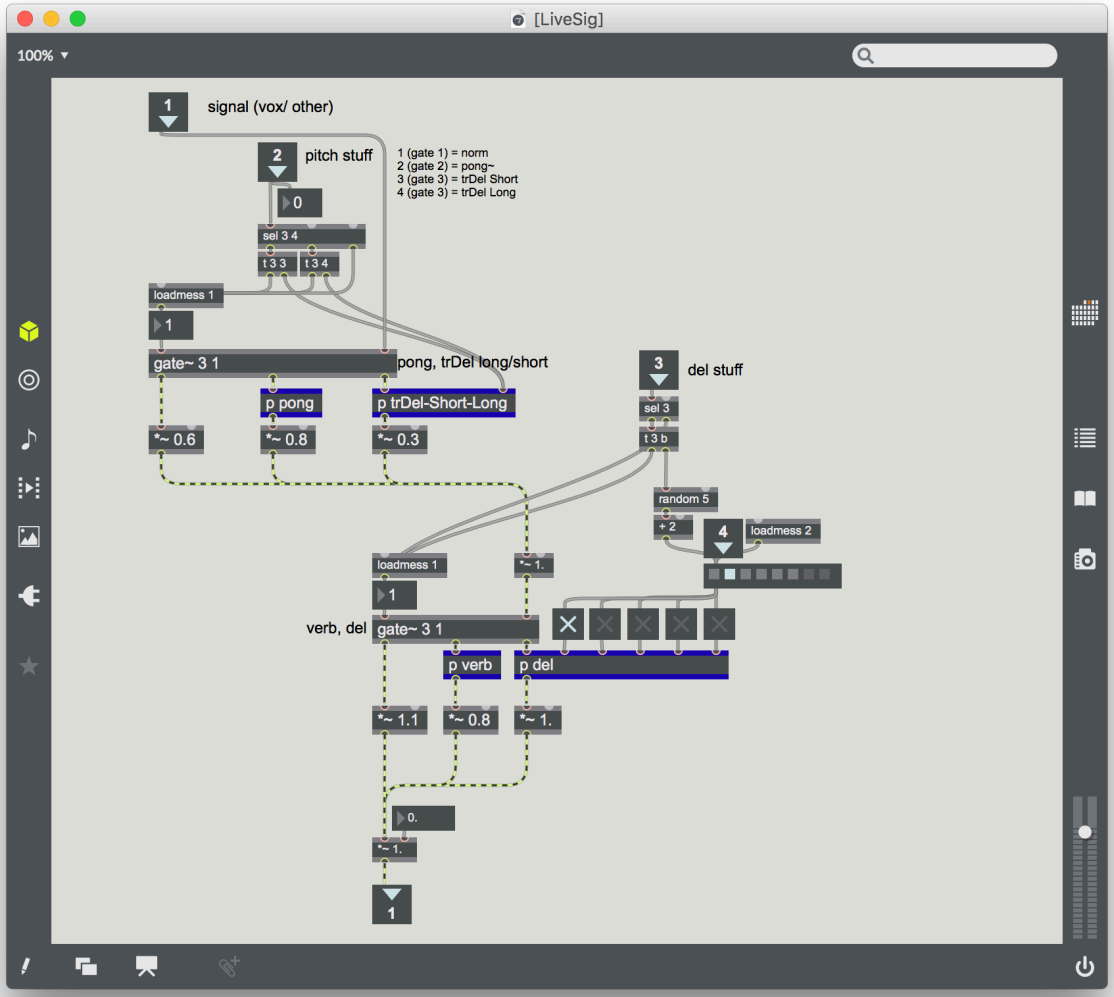
Appendix E (cont.) Vox6 patch

3.2 t3-to-randVox



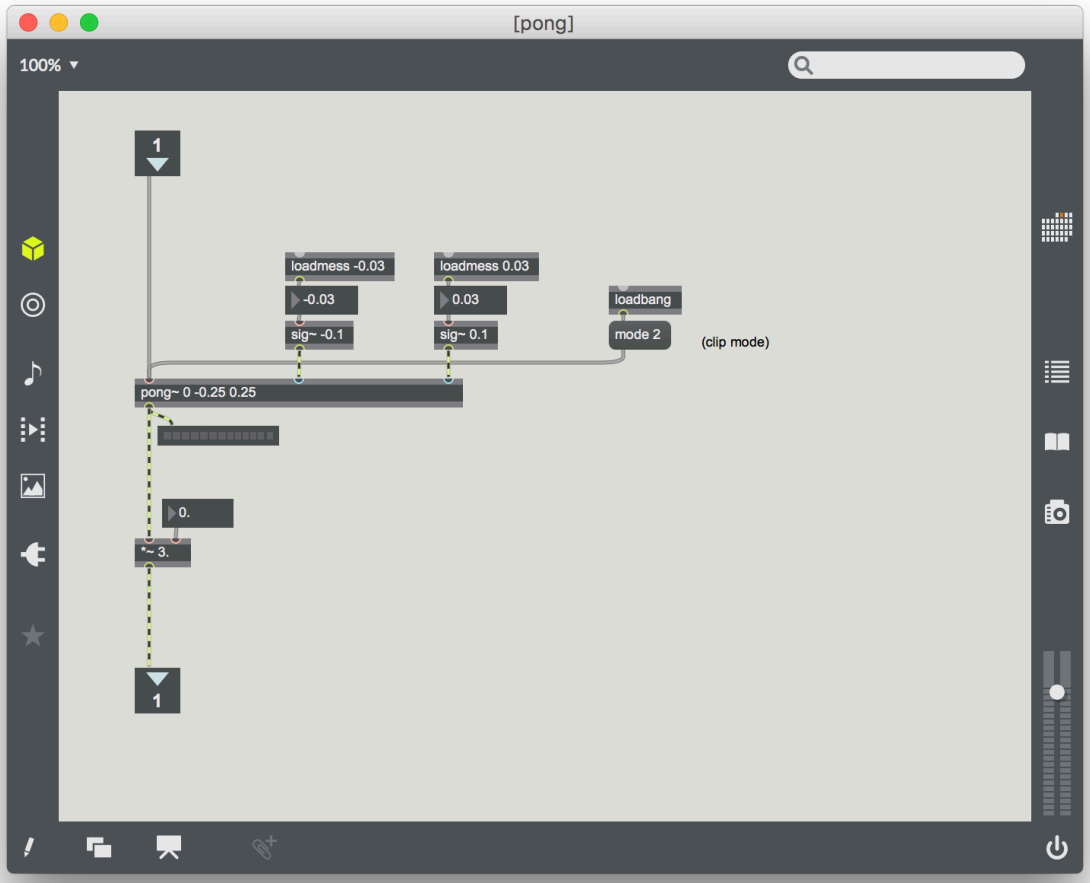
Appendix E (cont.) Vox6 patch

3.3 LiveSig



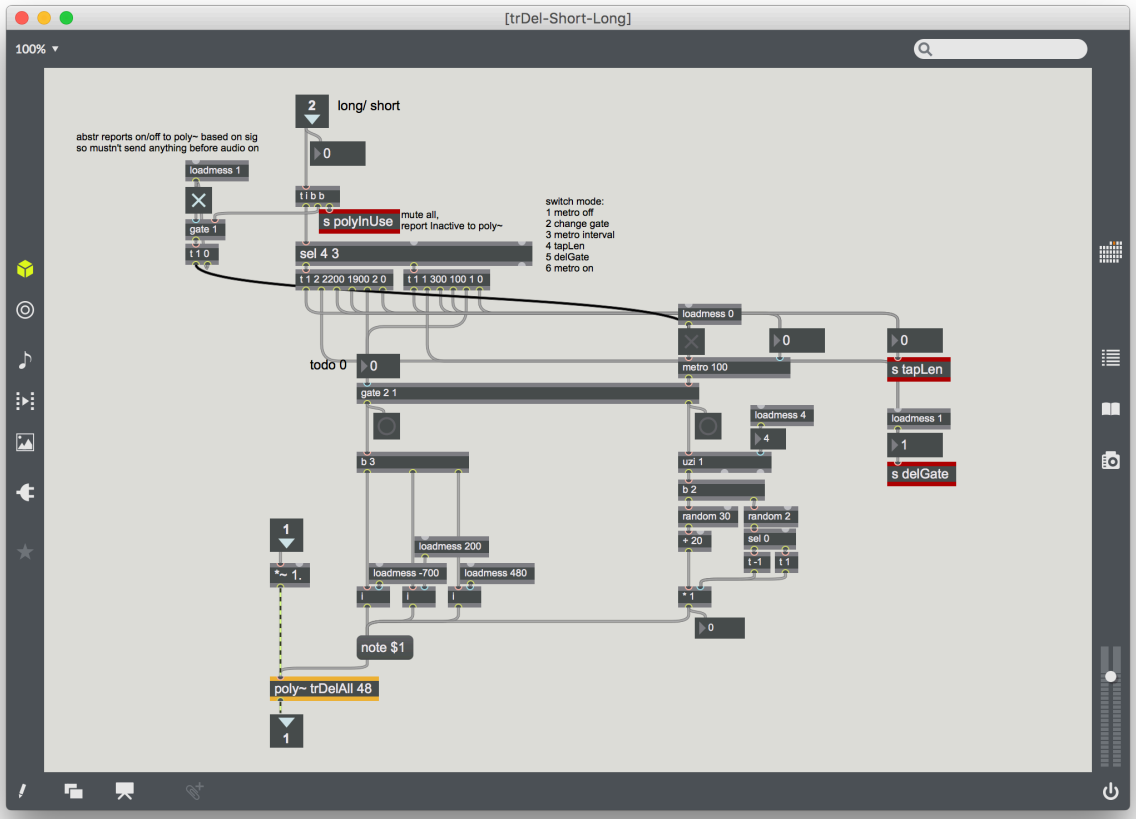
Appendix E (cont.) Vox6 patch

3.3.1 pong



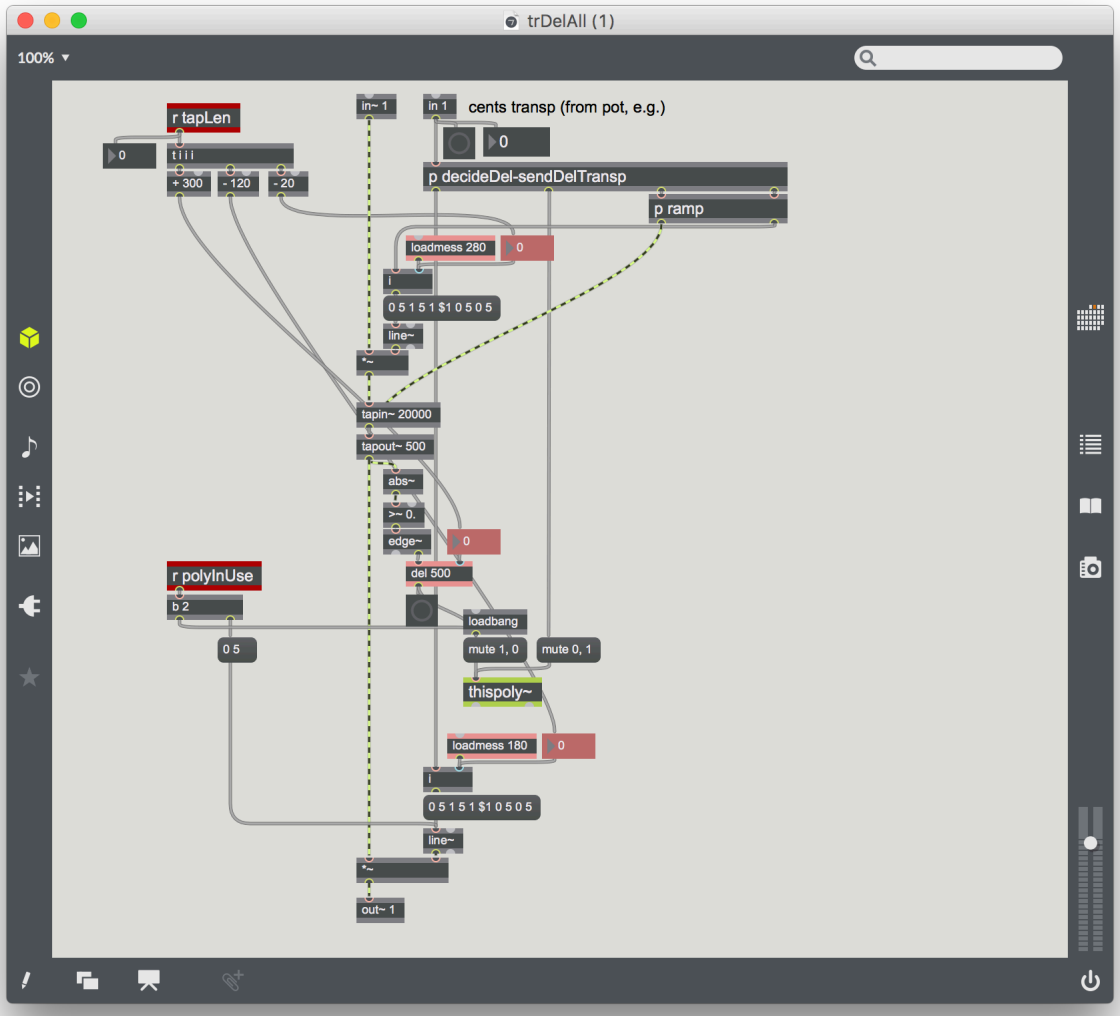
Appendix E (cont.) Vox6 patch

3.3.2 trDel-Short-Long



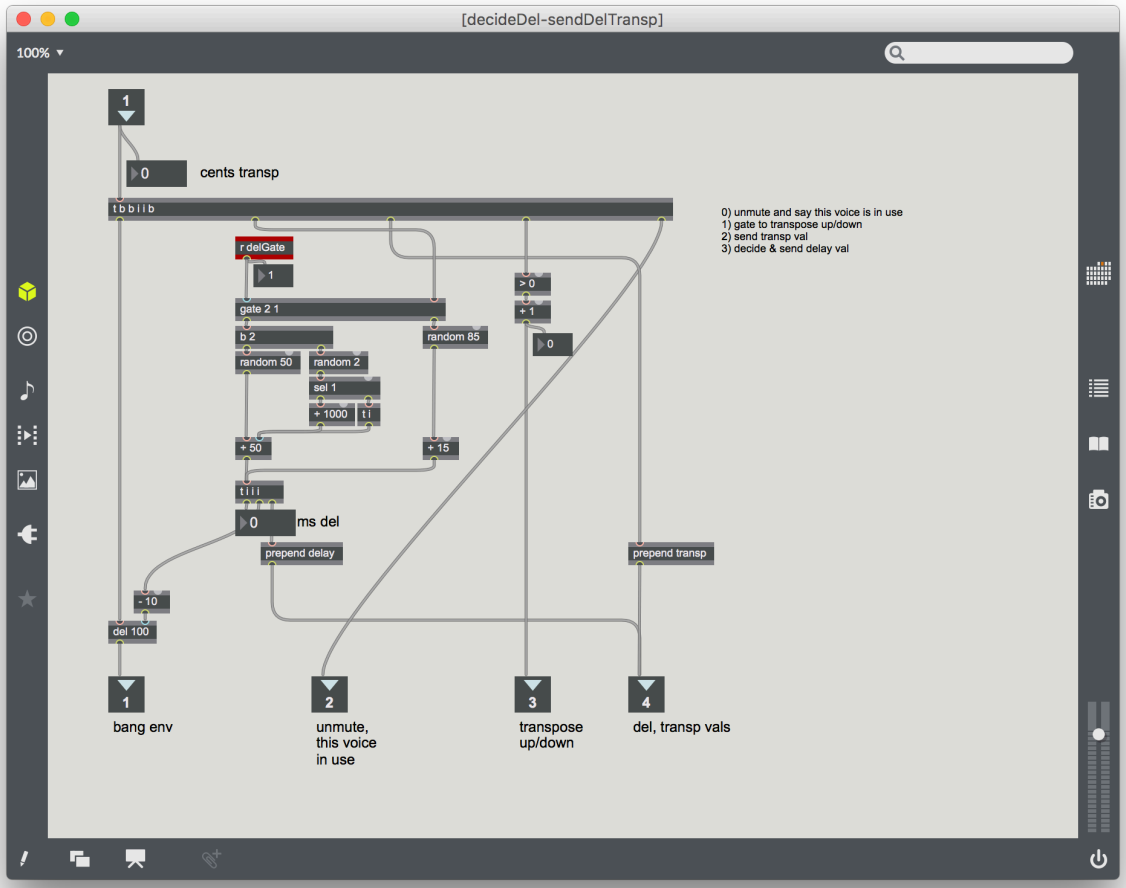
Appendix E (cont.) Vox6 patch

3.3.2.1 trDelAll (poly~)



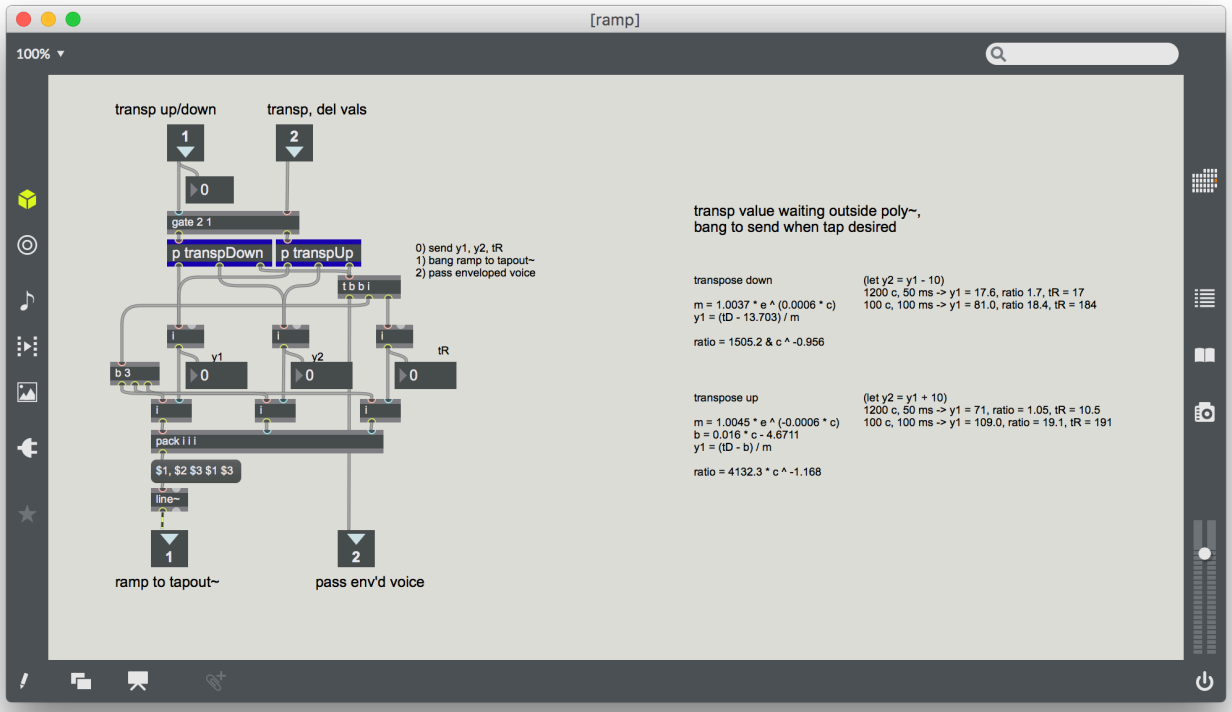
Appendix E (cont.) Vox6 patch

3.3.2.1.1 decideDel-sendDelTransp



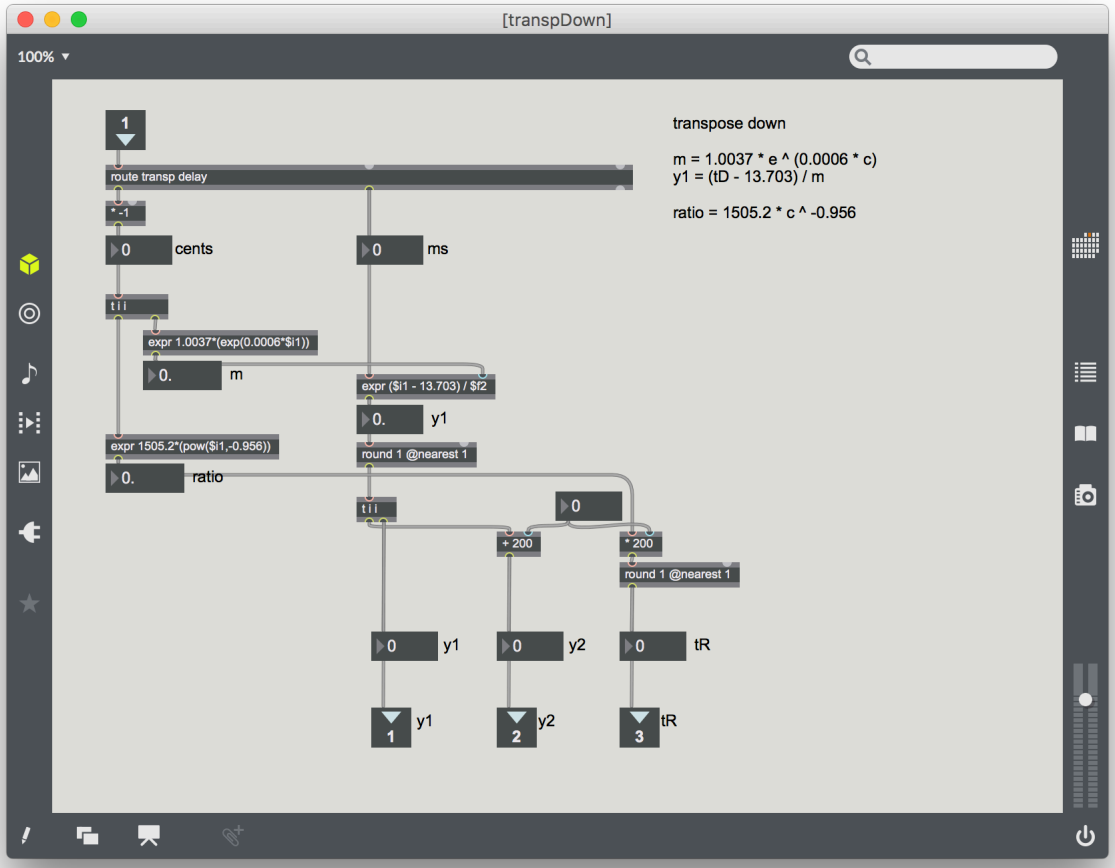
Appendix E (cont.) Vox6 patch

3.3.2.1.2 ramp



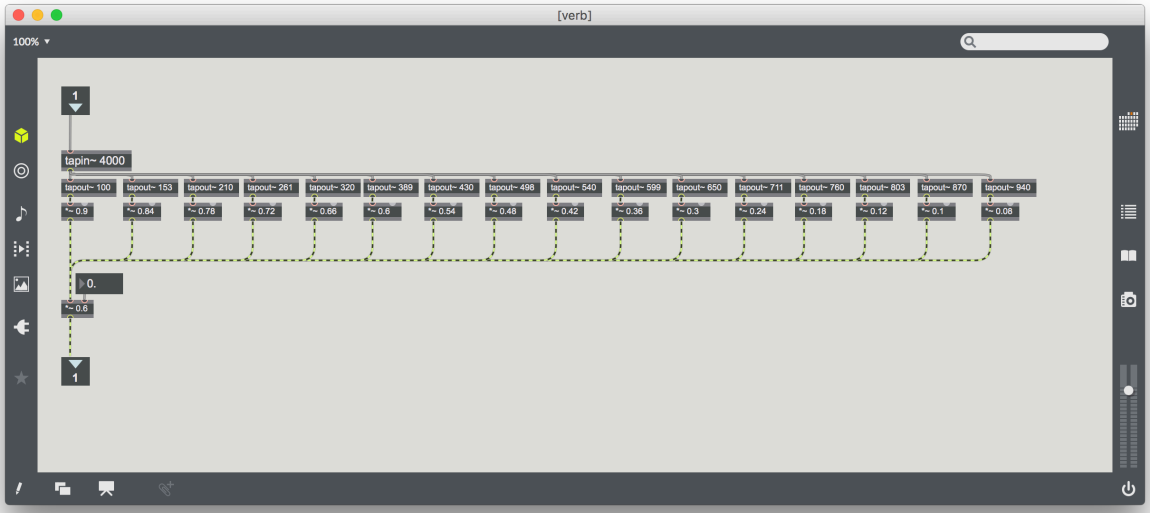
Appendix E (cont.) Vox6 patch

3.3.2.1.2.1 transpDown

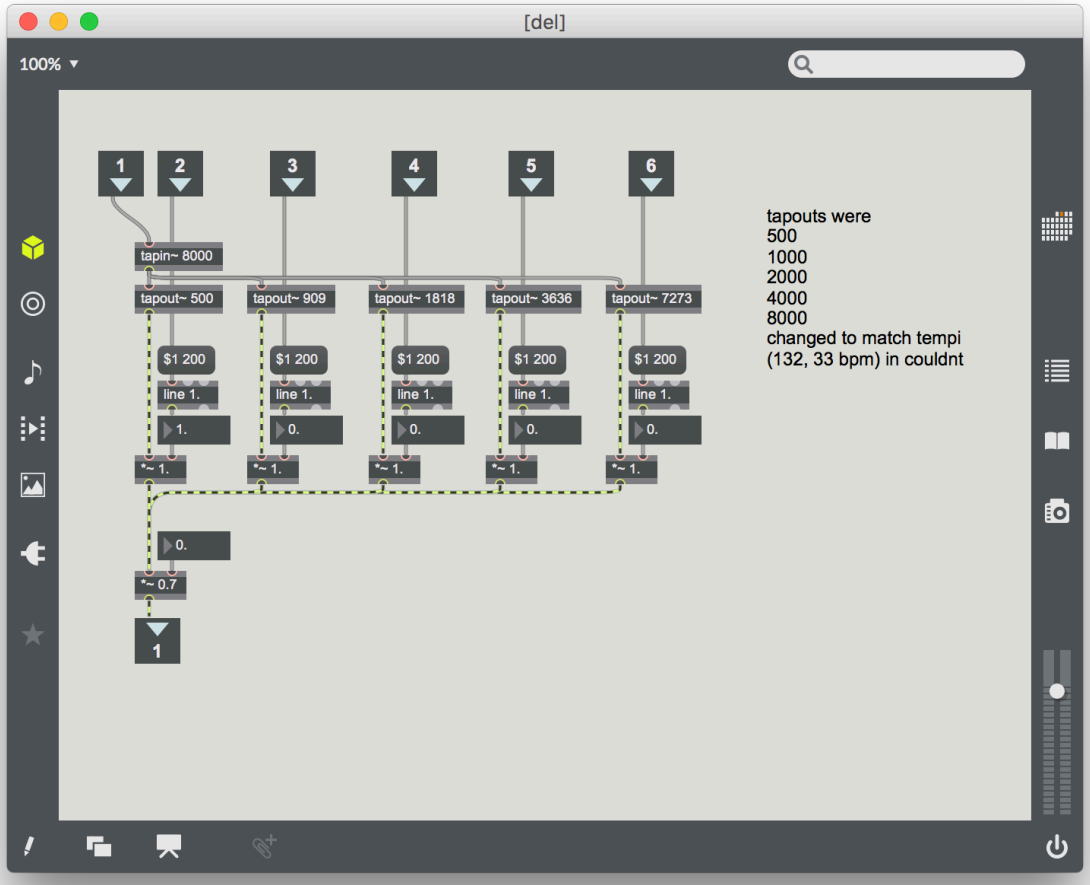


Appendix E (cont.) Vox6 patch

3.3.3 verb

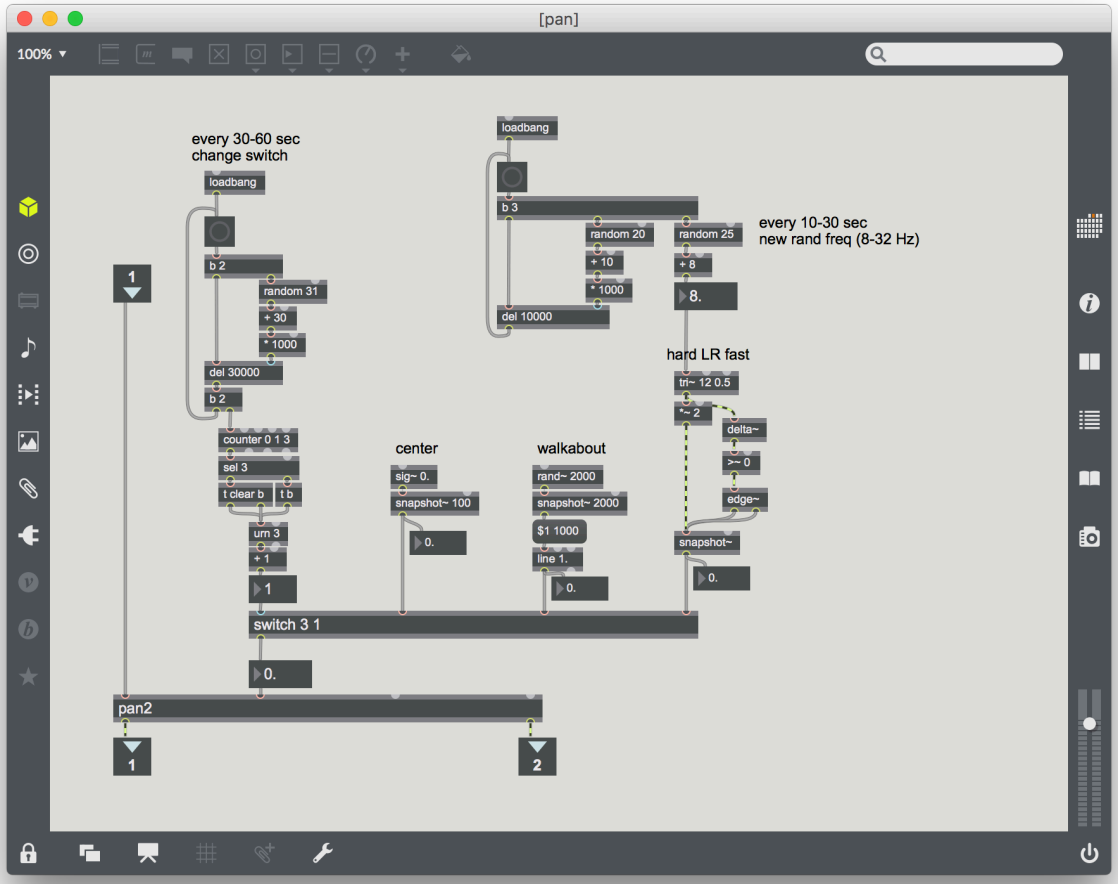


3.3.4 del



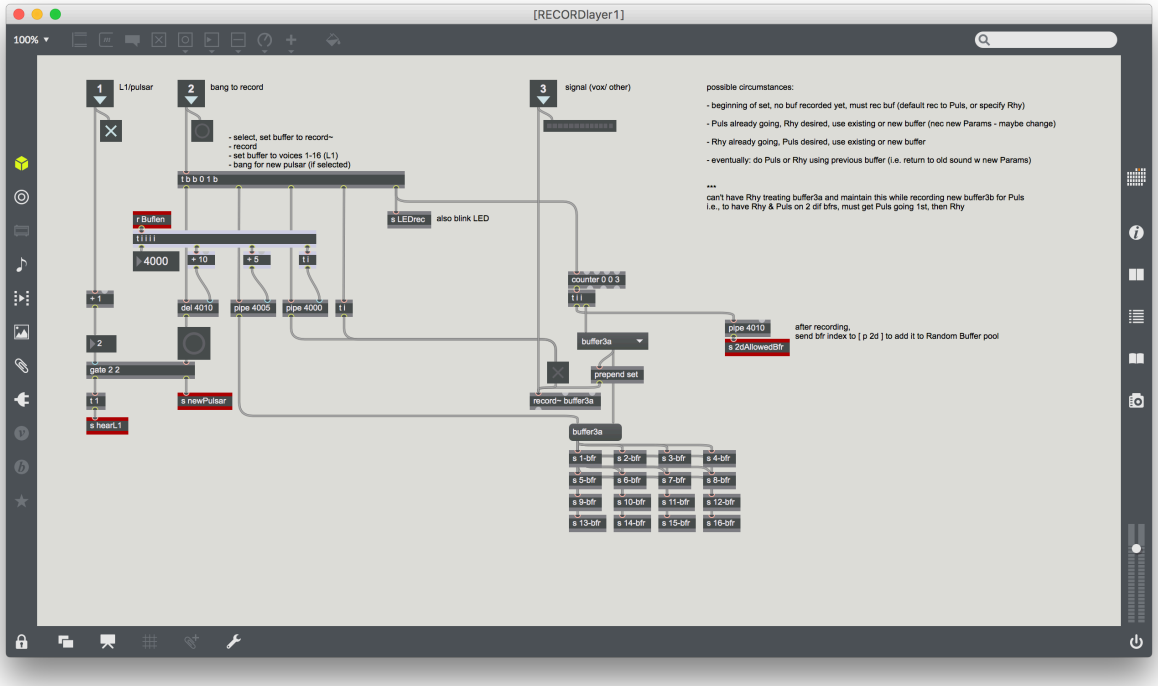
Appendix E (cont.) Vox6 patch

3.4 pan

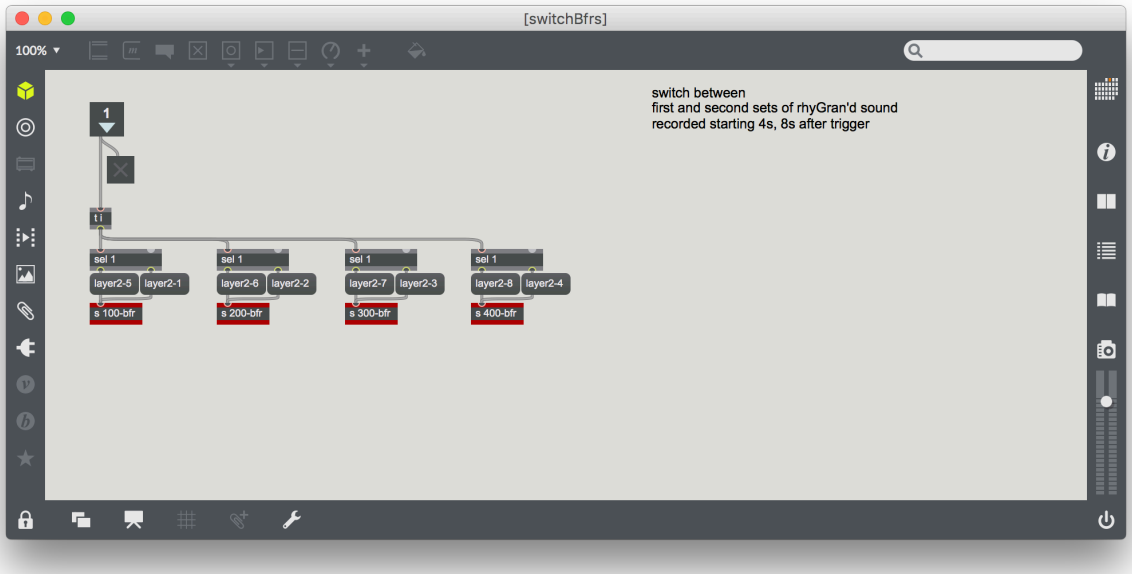


Appendix E (cont.) Vox6 patch

4. RECORDlayer1

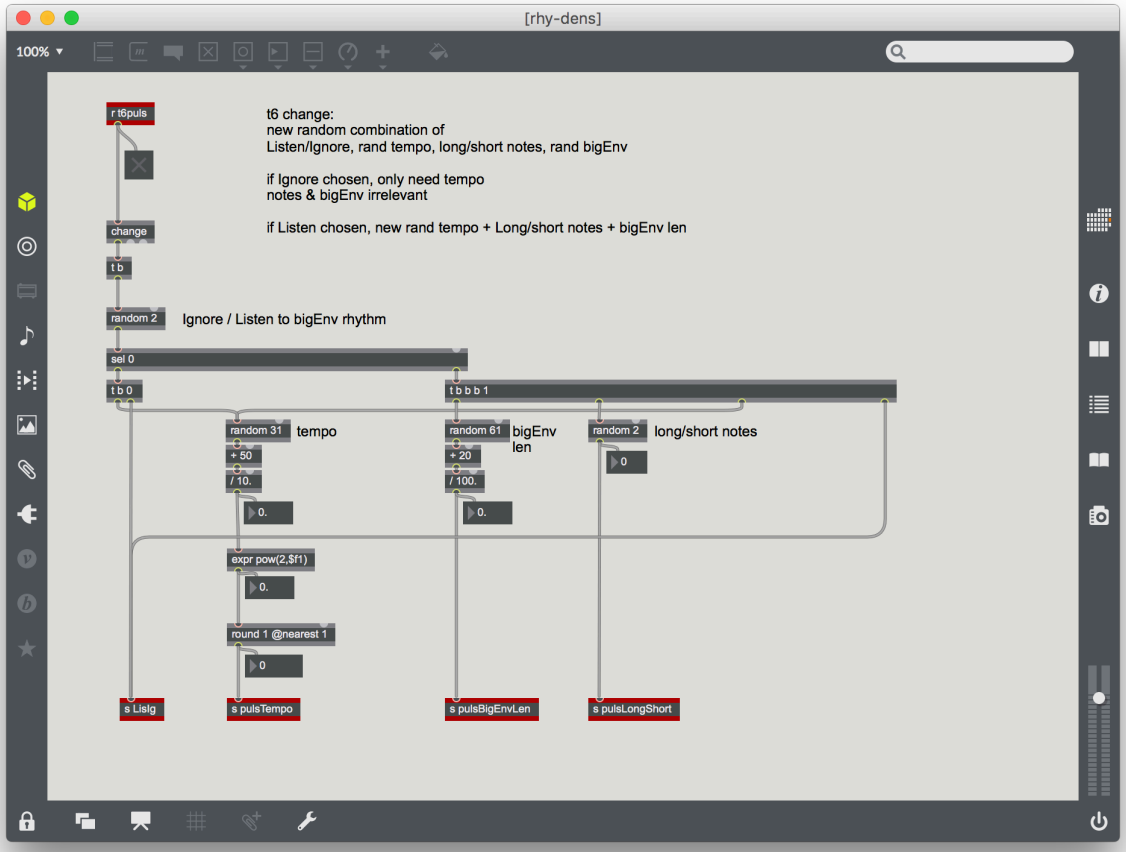


5. switchBfrs



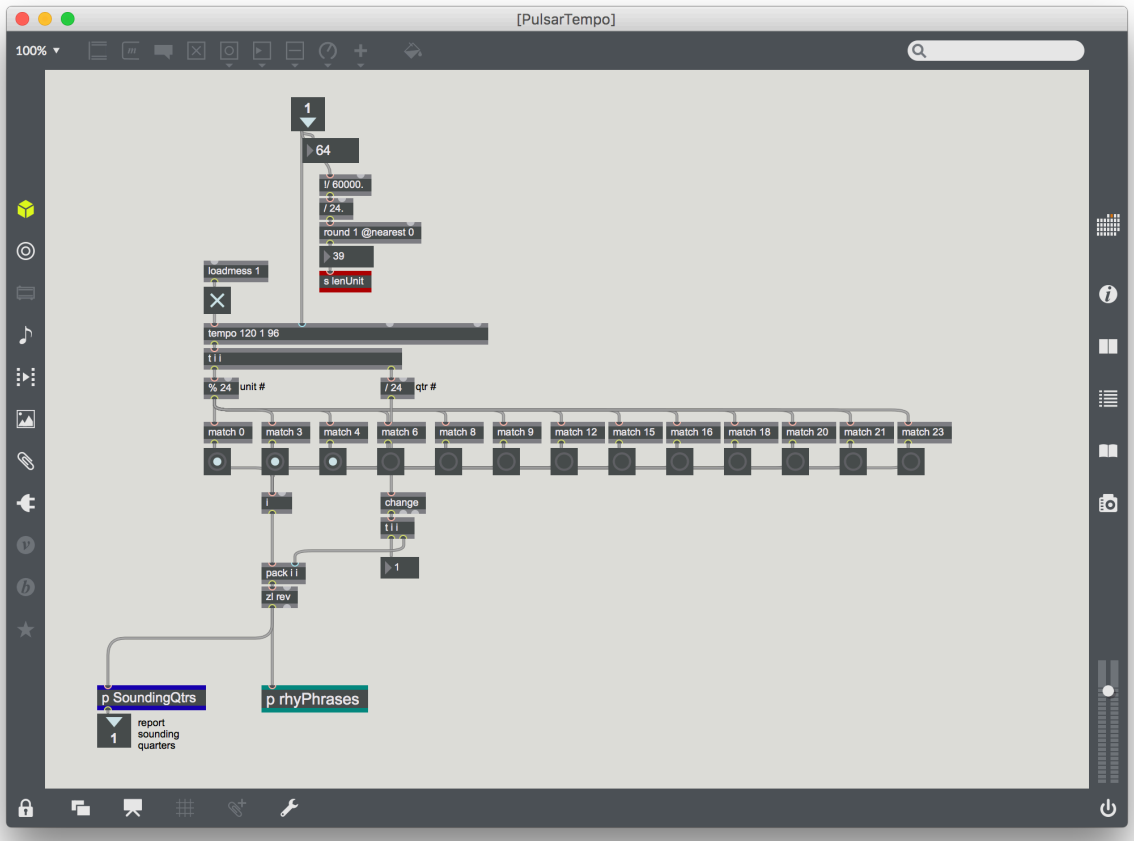
Appendix E (cont.) Vox6 patch

6. rhy-dens



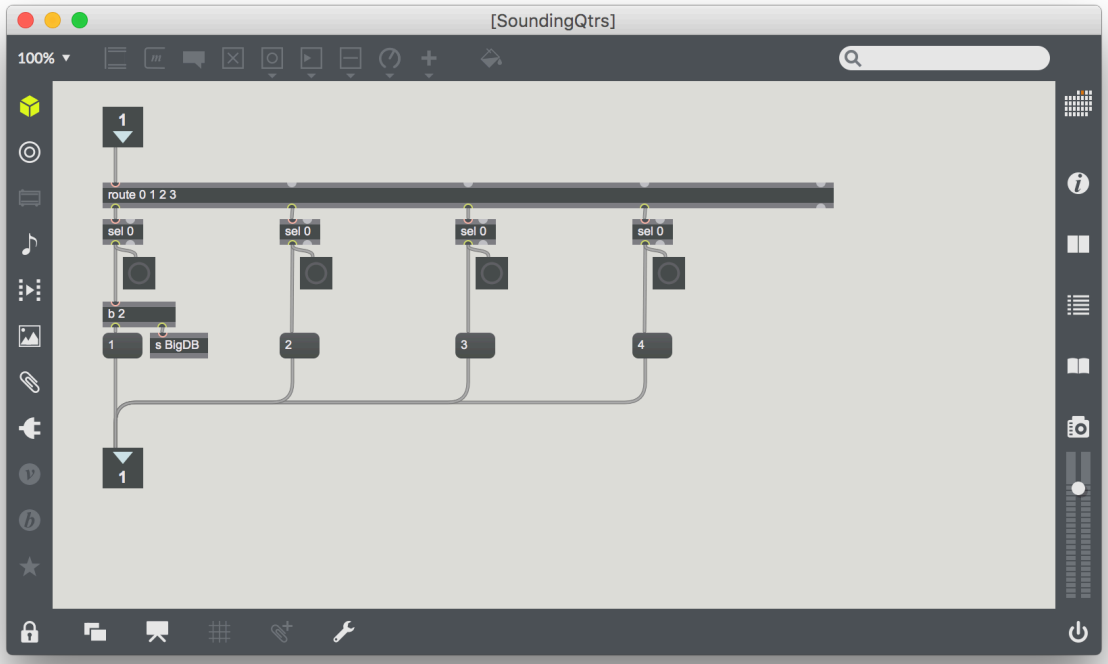
Appendix E (cont.) Vox6 patch

7. PulsarTempo

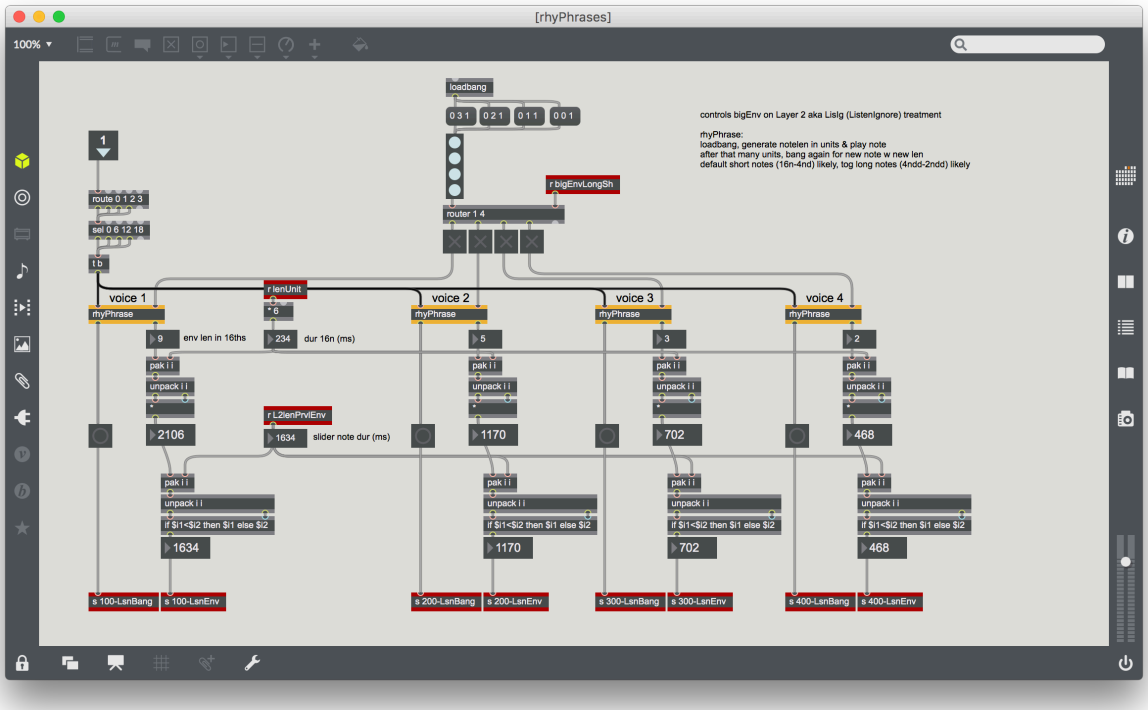


Appendix E (cont.) Vox6 patch

7.1 SoundingQtrs

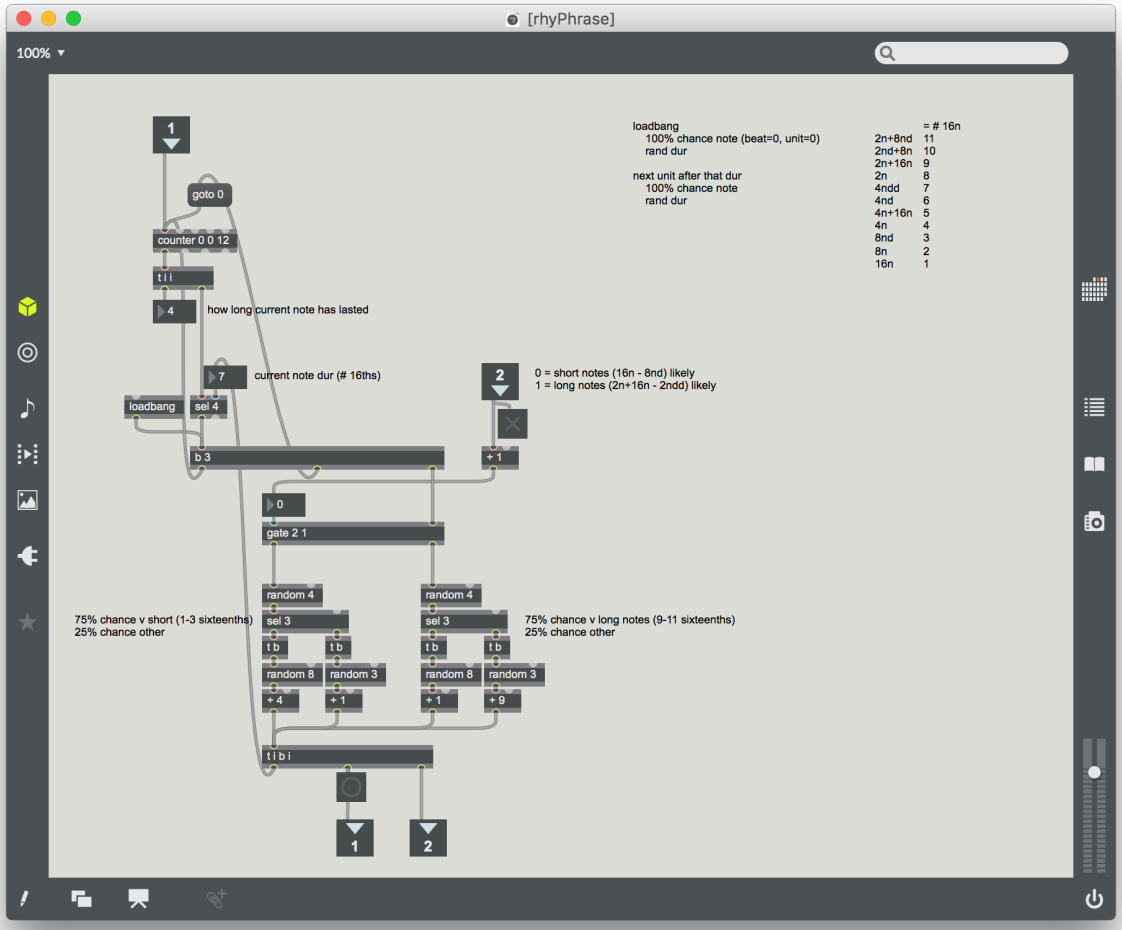


7.2 rhyPhrases



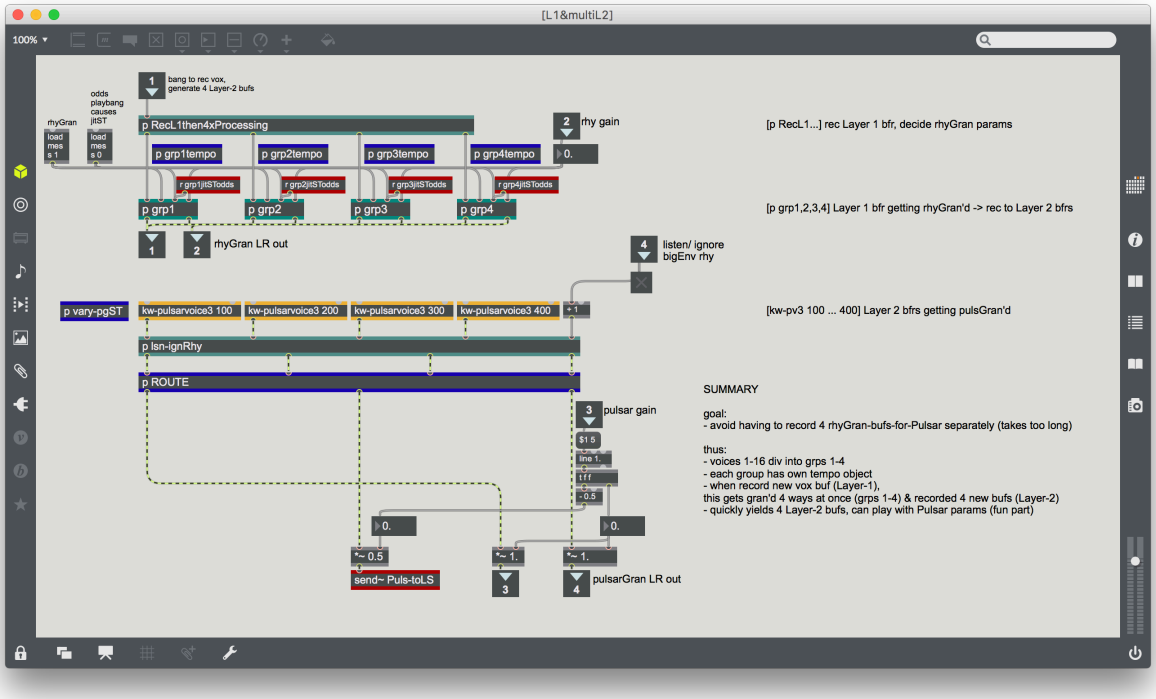
Appendix E (cont.) Vox6 patch

7.2.1 rhyPhrase

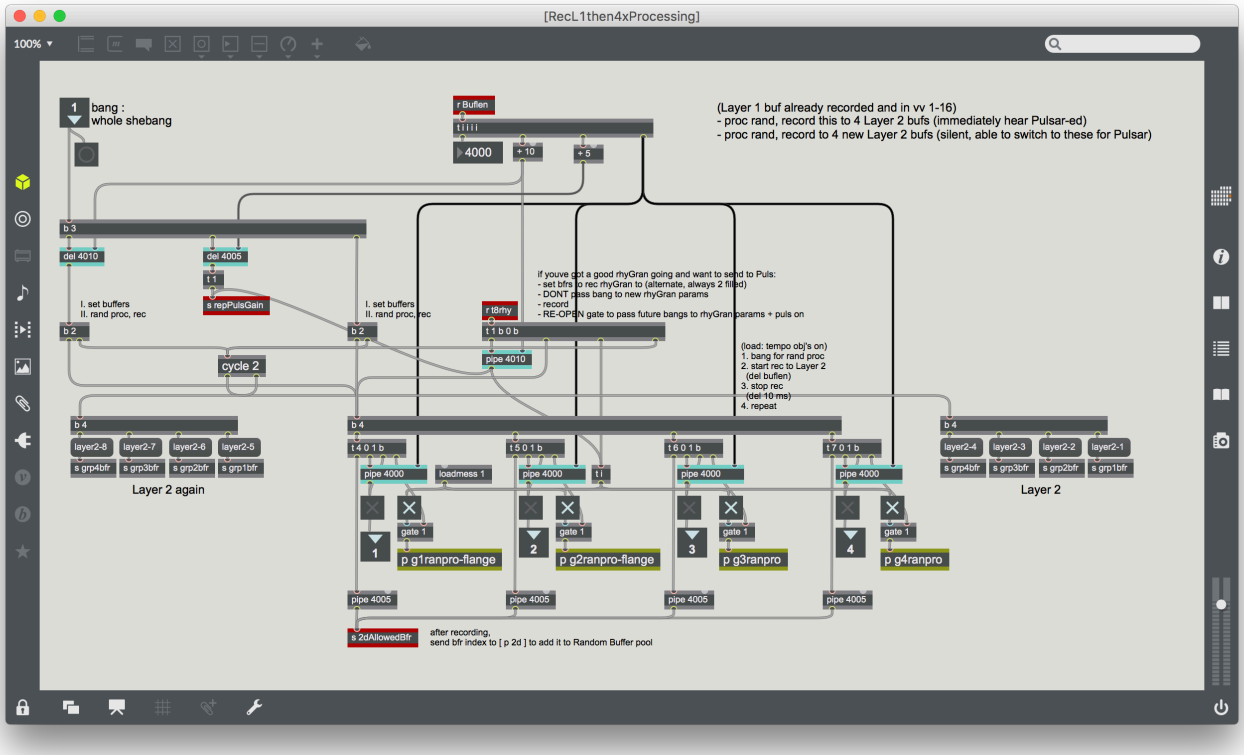


Appendix E (cont.) Vox6 patch

8. L1&multiL2

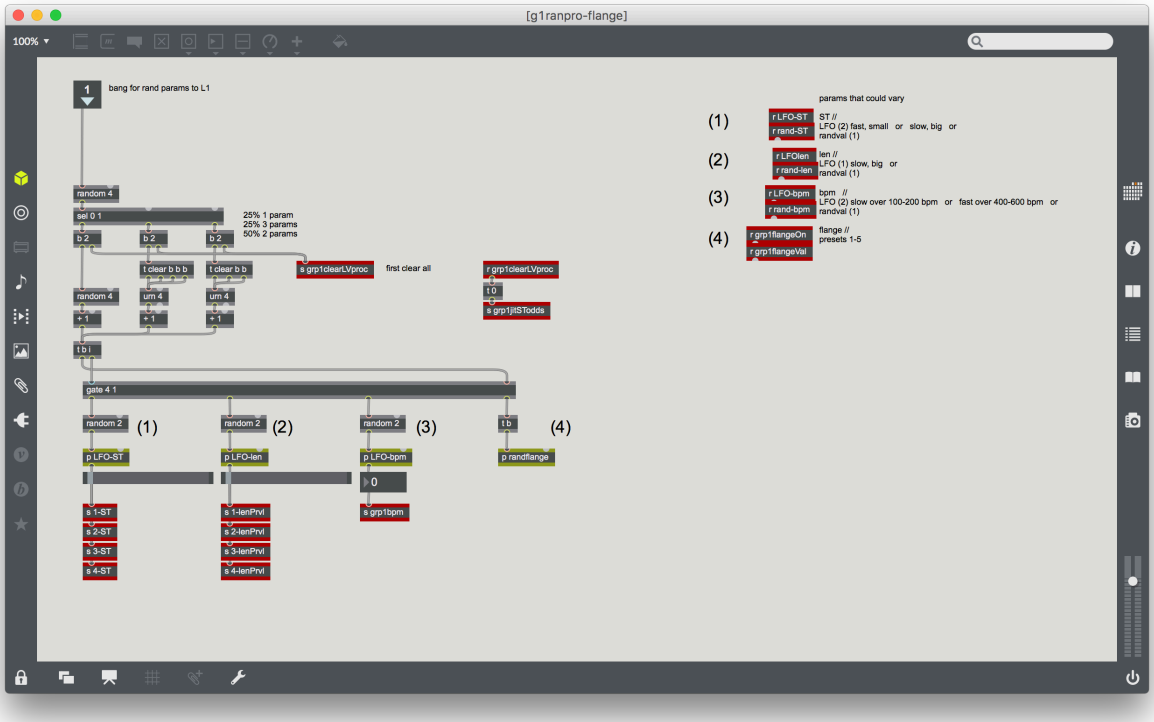


8.1 RecL1then4xProcessing



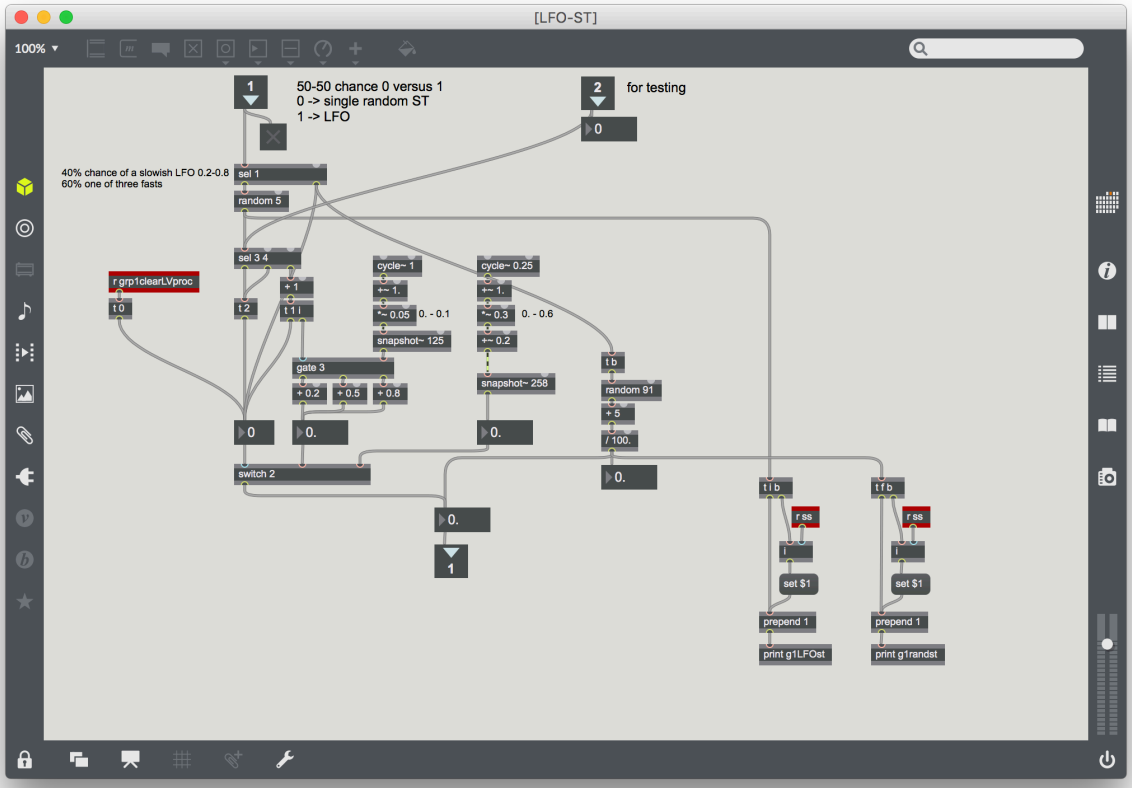
Appendix E (cont.) Vox6 patch

8.1.1 g1ranpro-flange



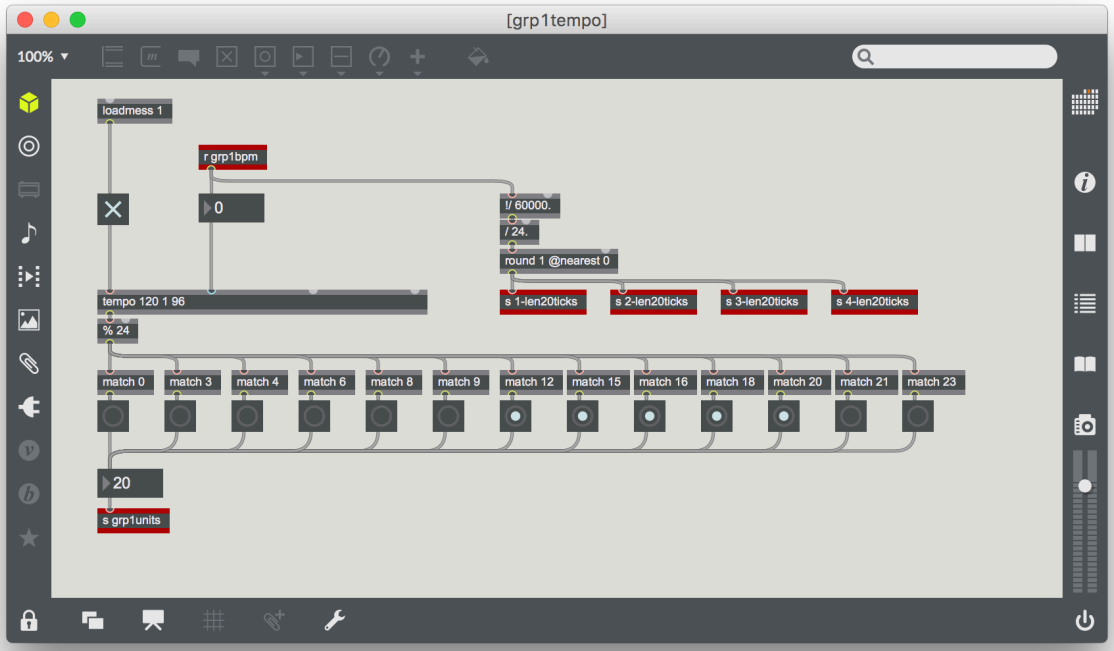
Appendix E (cont.) Vox6 patch

8.1.1.1 LFO-ST



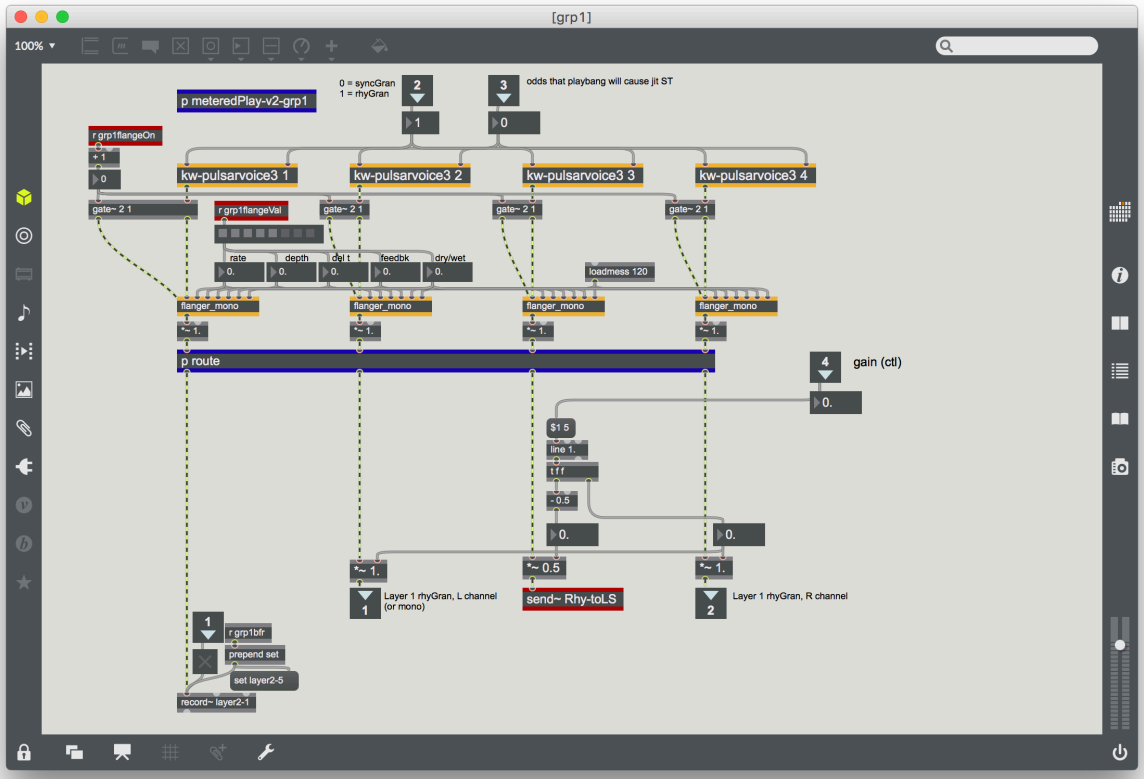
Appendix E (cont.) Vox6 patch

8.2 grp1tempo

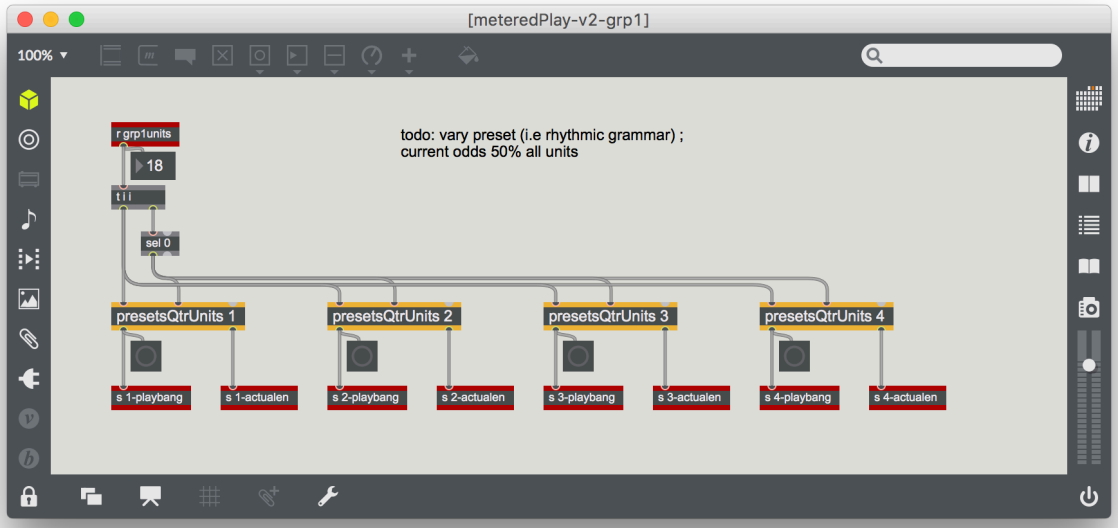


Appendix E (cont.) Vox6 patch

8.3 grp1

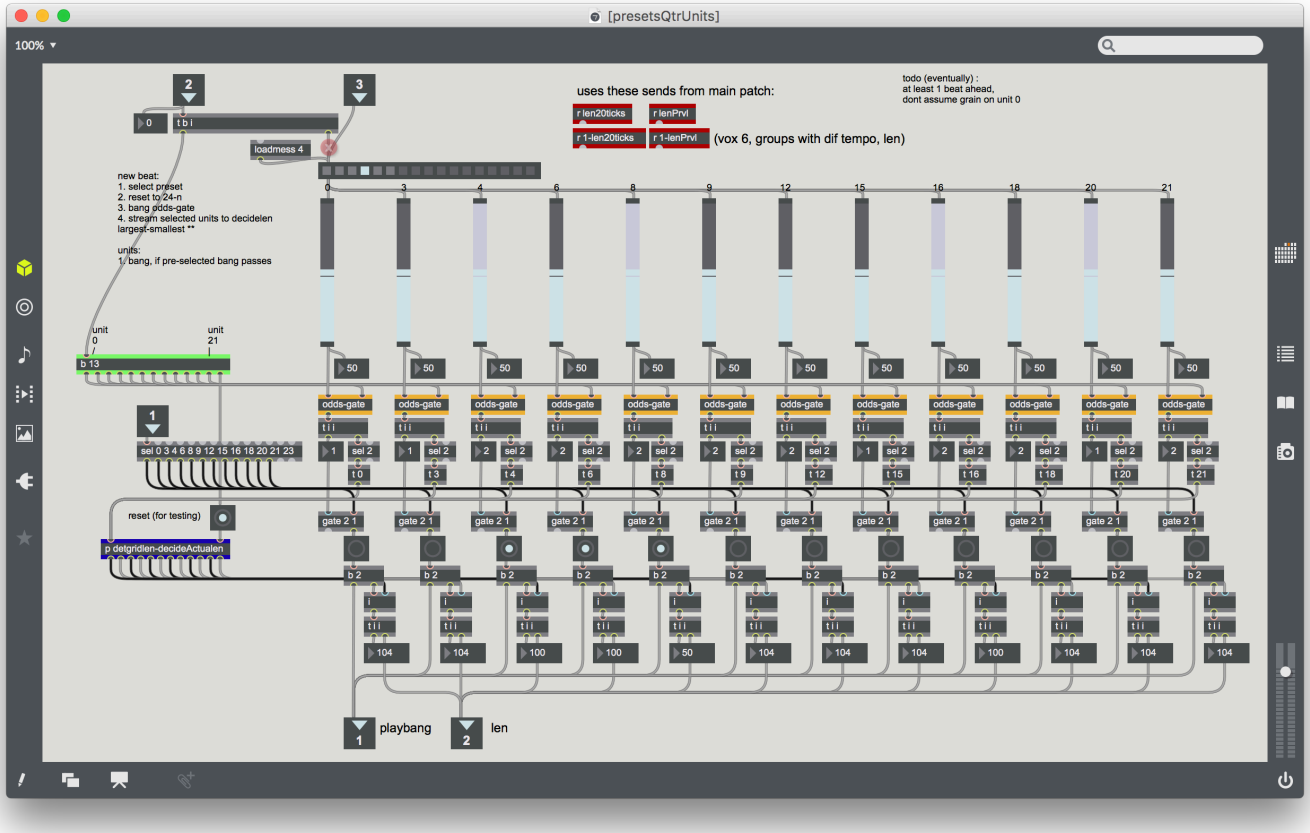


8.3.1 meteredPlay-v2-g1

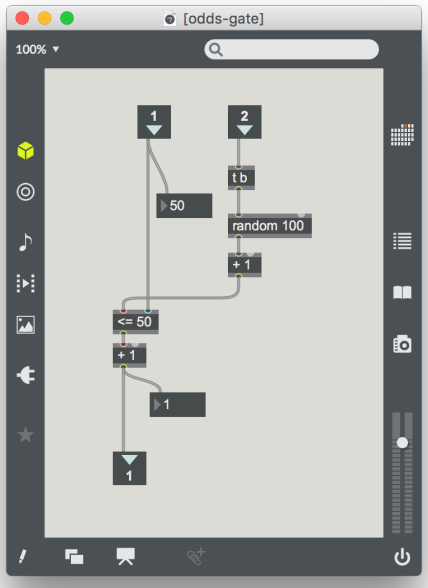


Appendix E (cont.) Vox6 patch

8.3.1.1 presetsQtrUnits

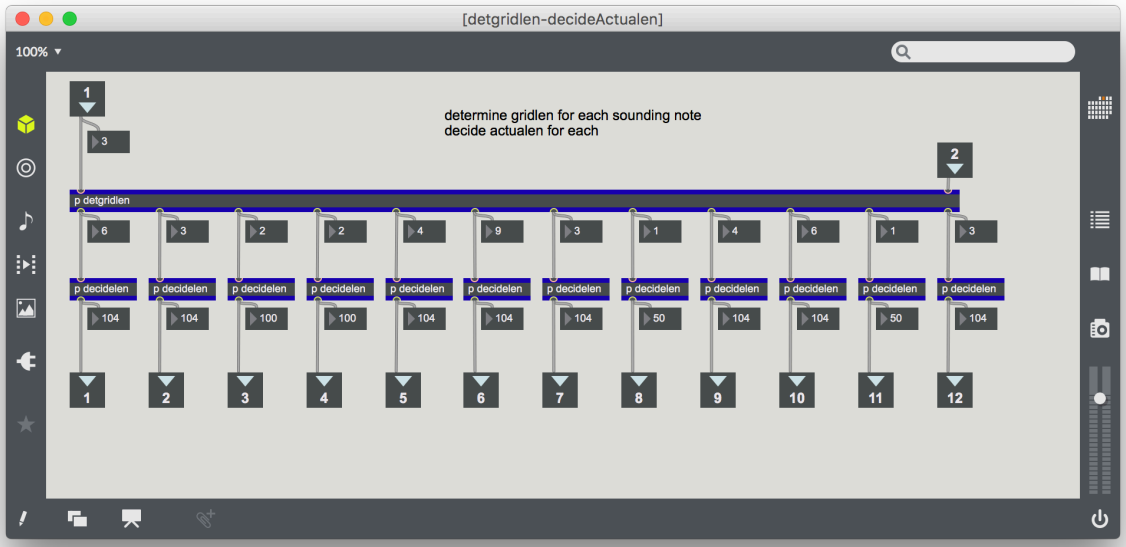


8.3.1.1.1 odds-gate

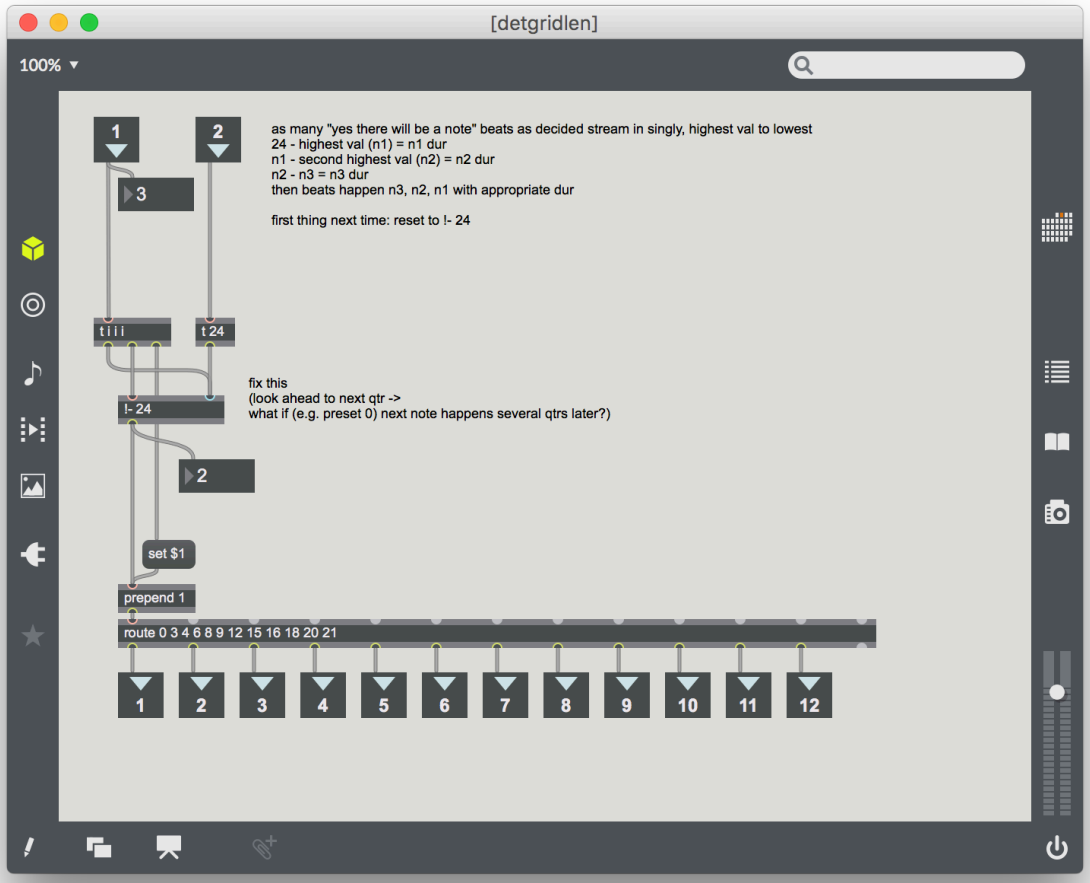


Appendix E (cont.) Vox6 patch

8.3.1.1.2 detgridlen-decideActualen

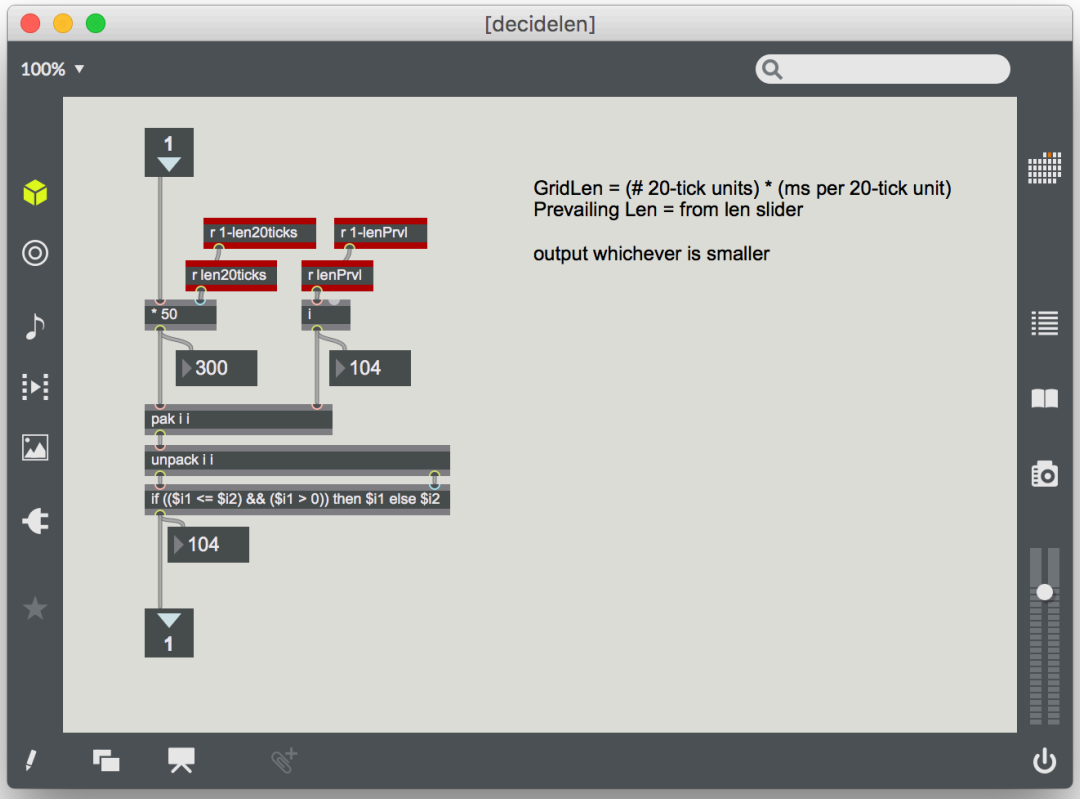


8.3.1.1.2.1 detgridlen



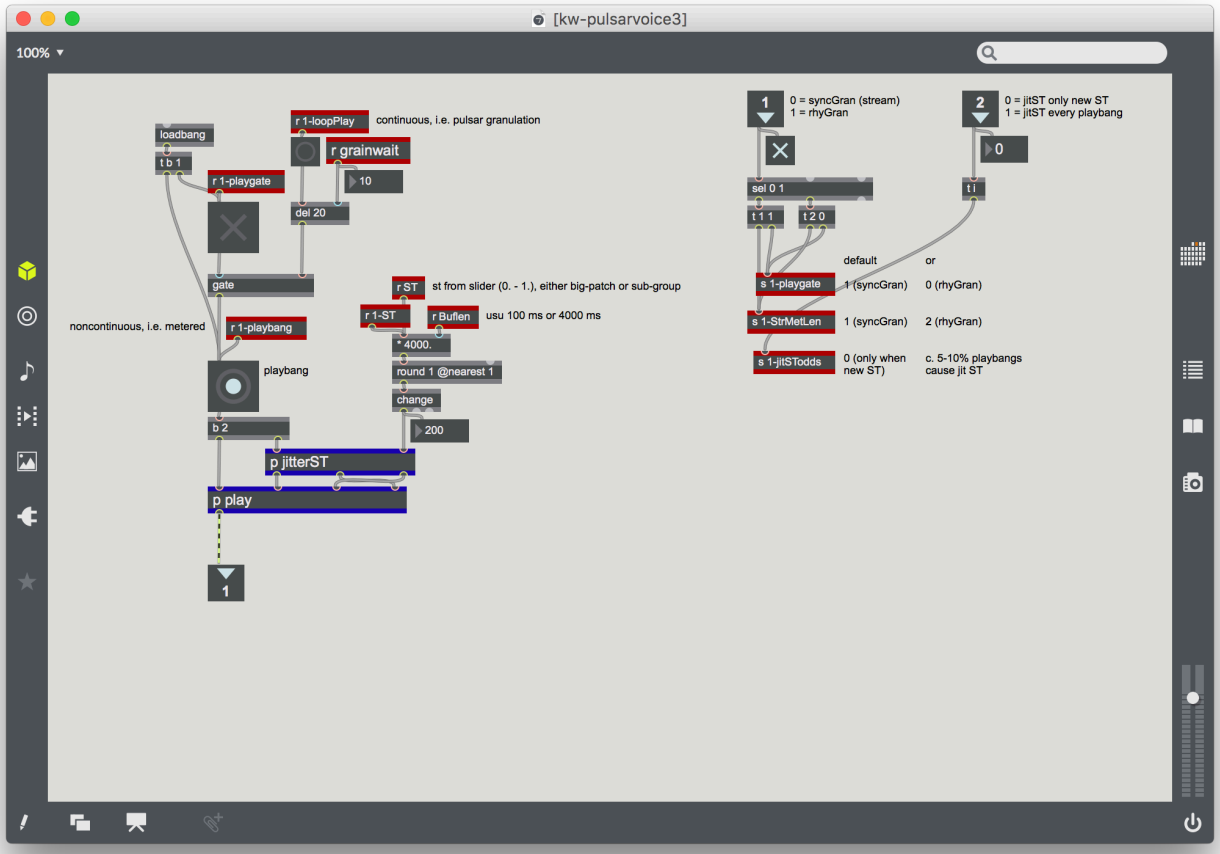
Appendix E (cont.) Vox6 patch

8.3.1.1.2.2 decidelen



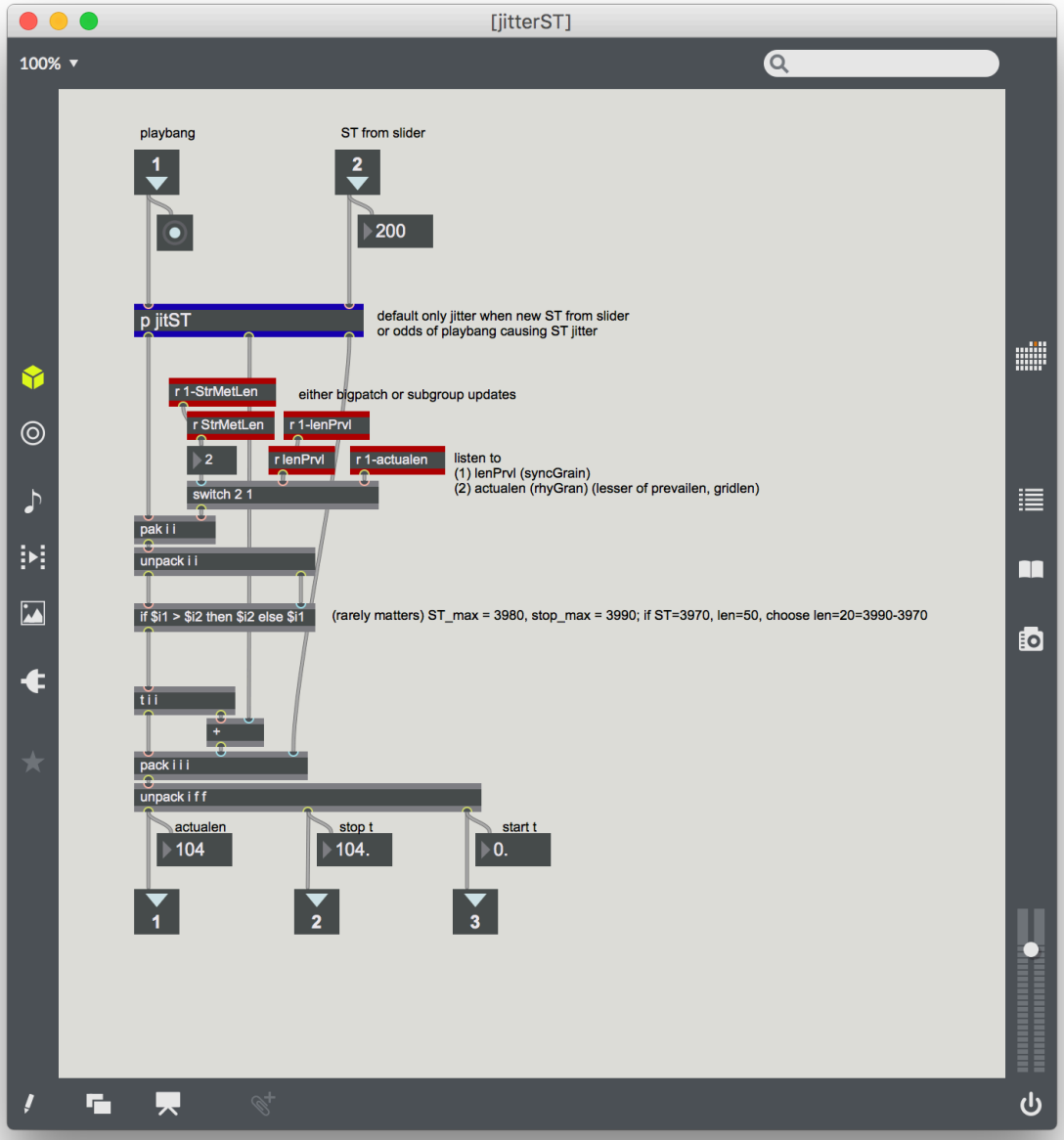
Appendix E (cont.) Vox6 patch

8.3.2 kw-pulsarvoice3



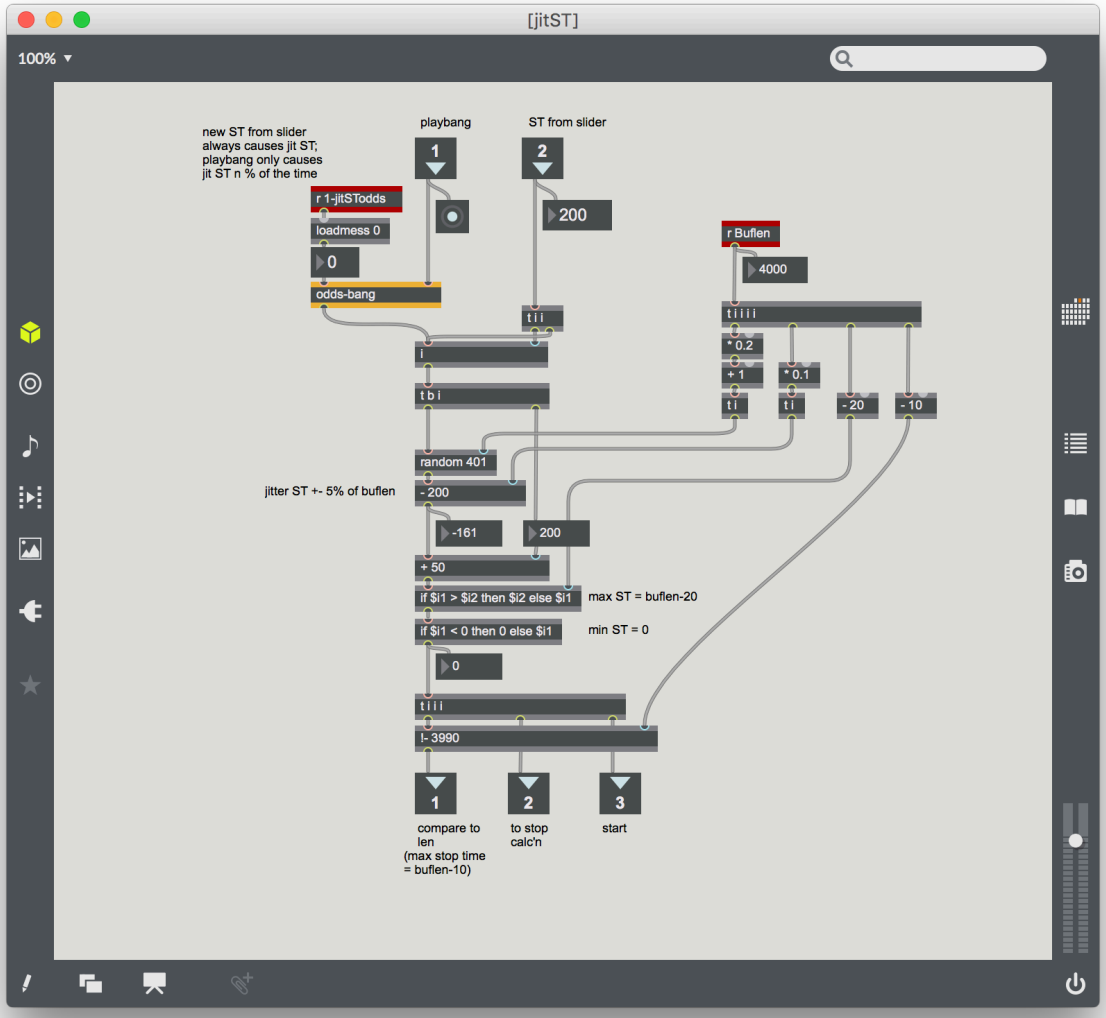
Appendix E (cont.) Vox6 patch

8.3.2.1 jitterST



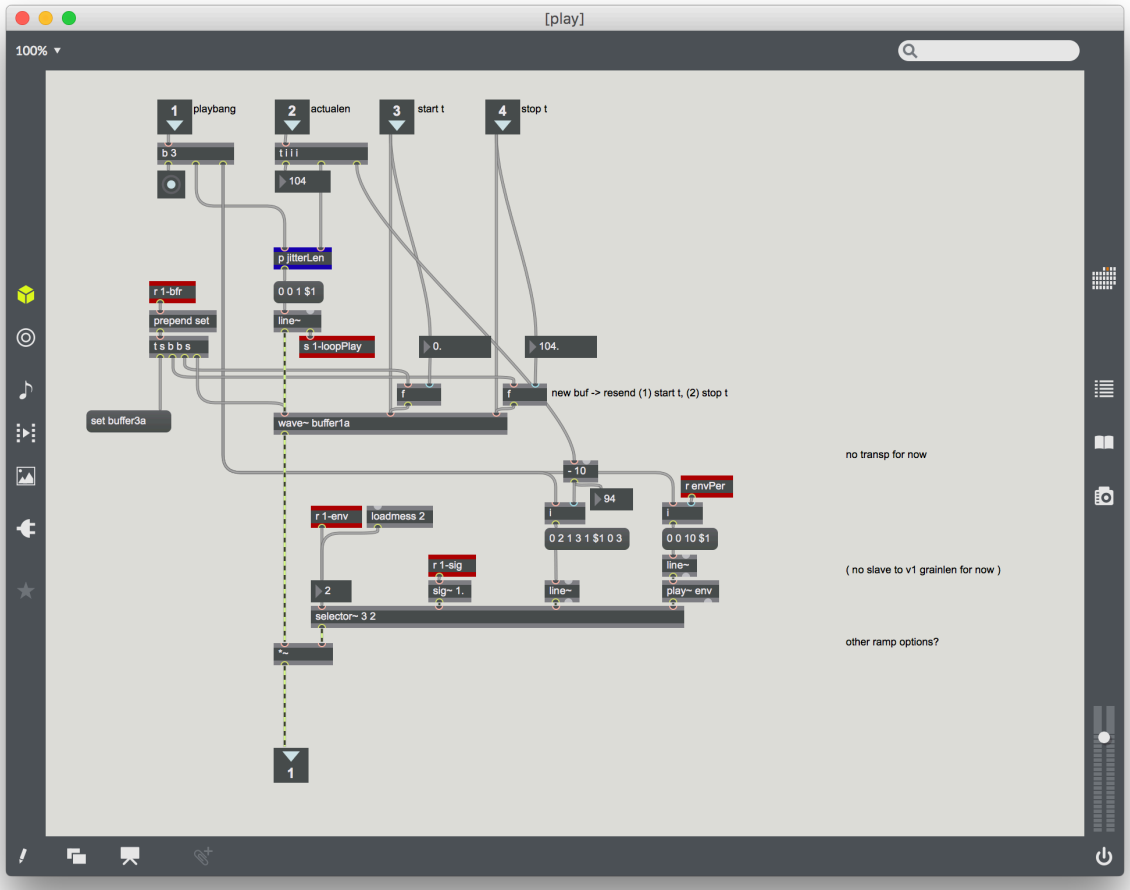
Appendix E (cont.) Vox6 patch

8.3.2.1.1 jitST



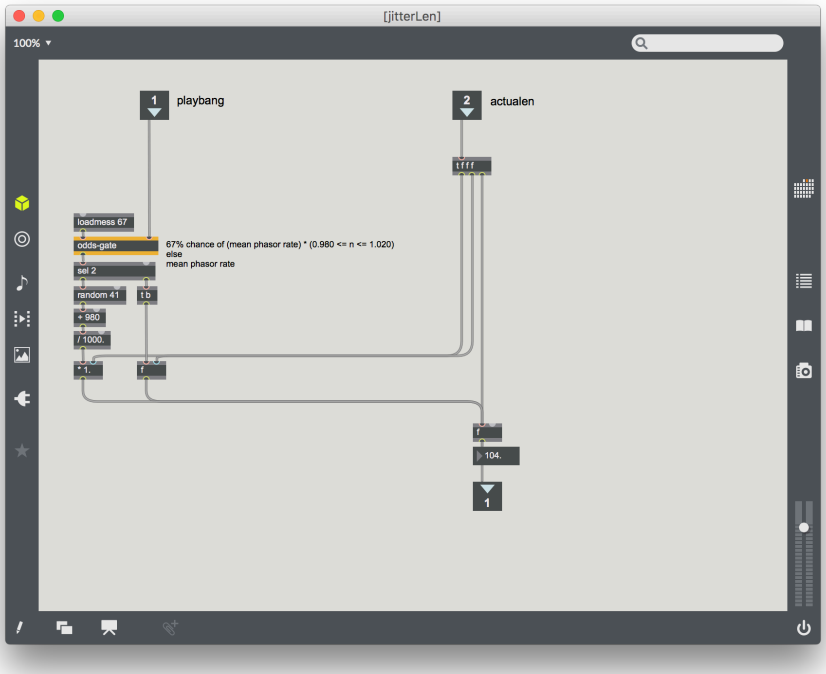
Appendix E (cont.) Vox6 patch

8.3.2.2 play

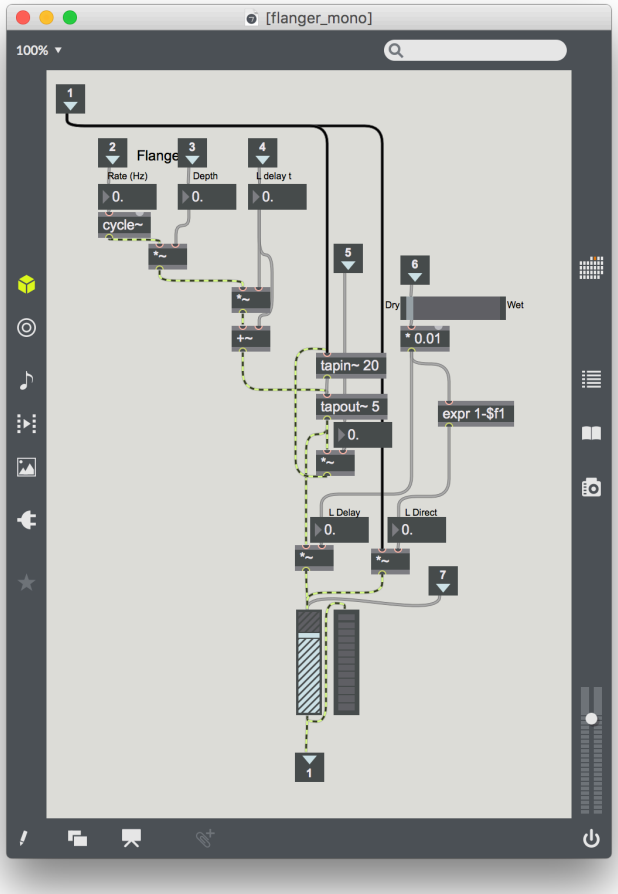


Appendix E (cont.) Vox6 patch

8.3.2.2.1 jitterLen

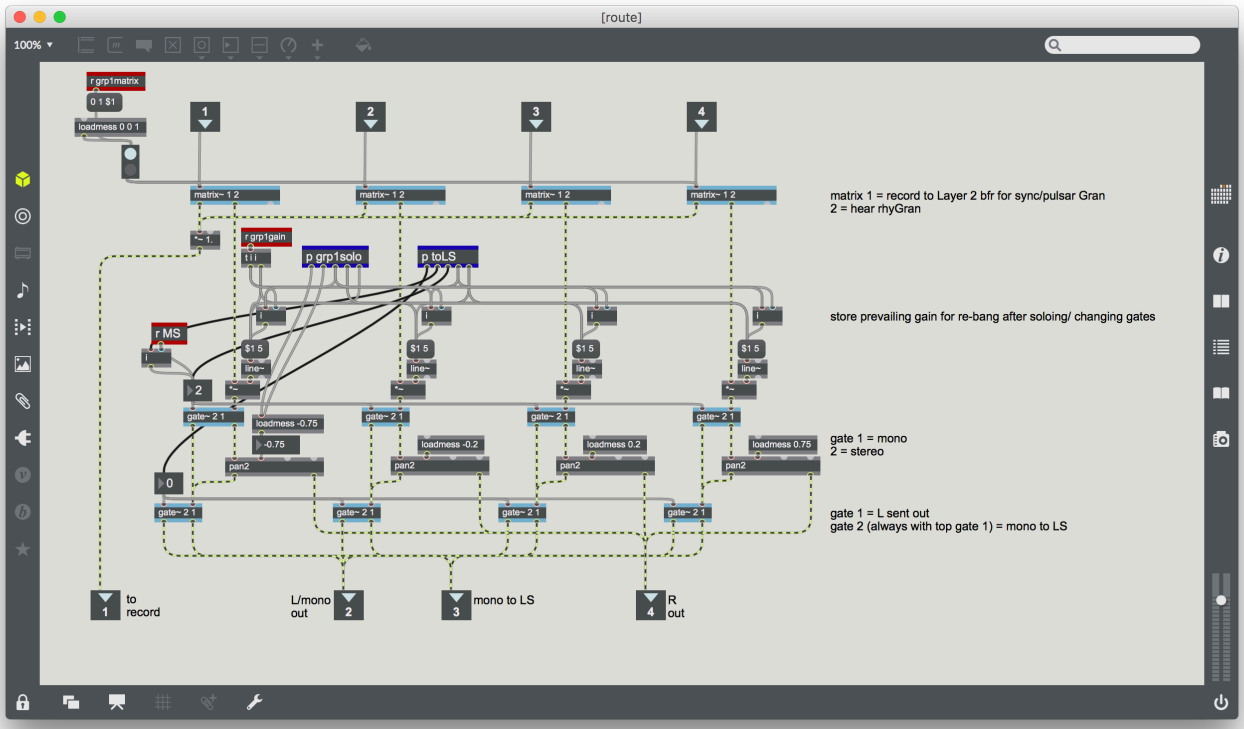


8.3.3 flanger_mono



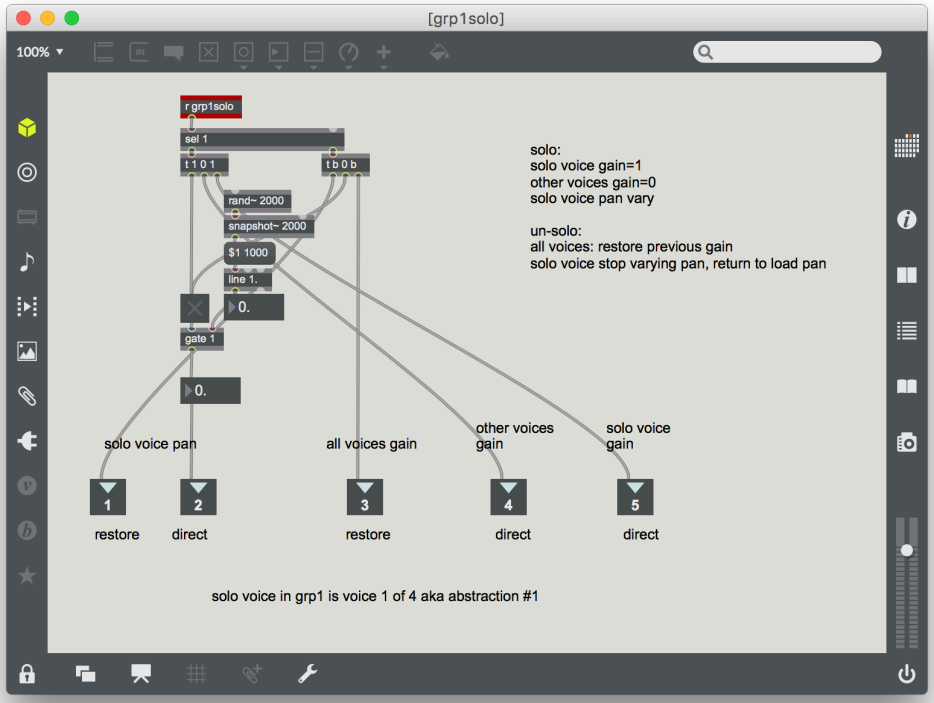
Appendix E (cont.) Vox6 patch

8.3.4 route

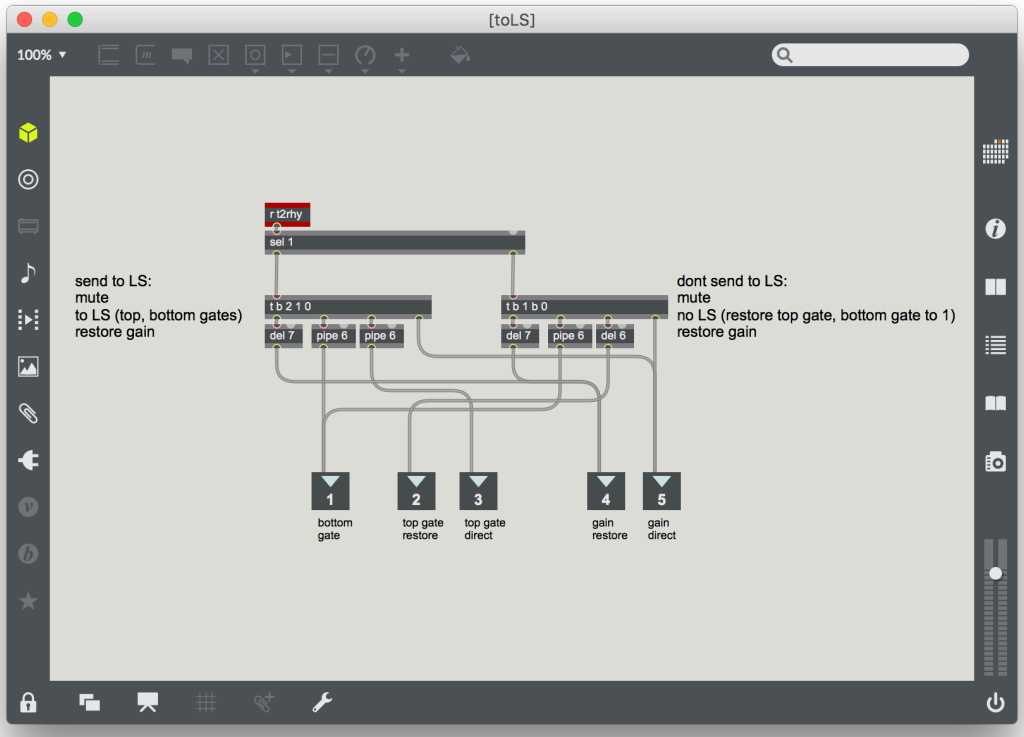


Appendix E (cont.) Vox6 patch

8.3.4.1 grp1solo

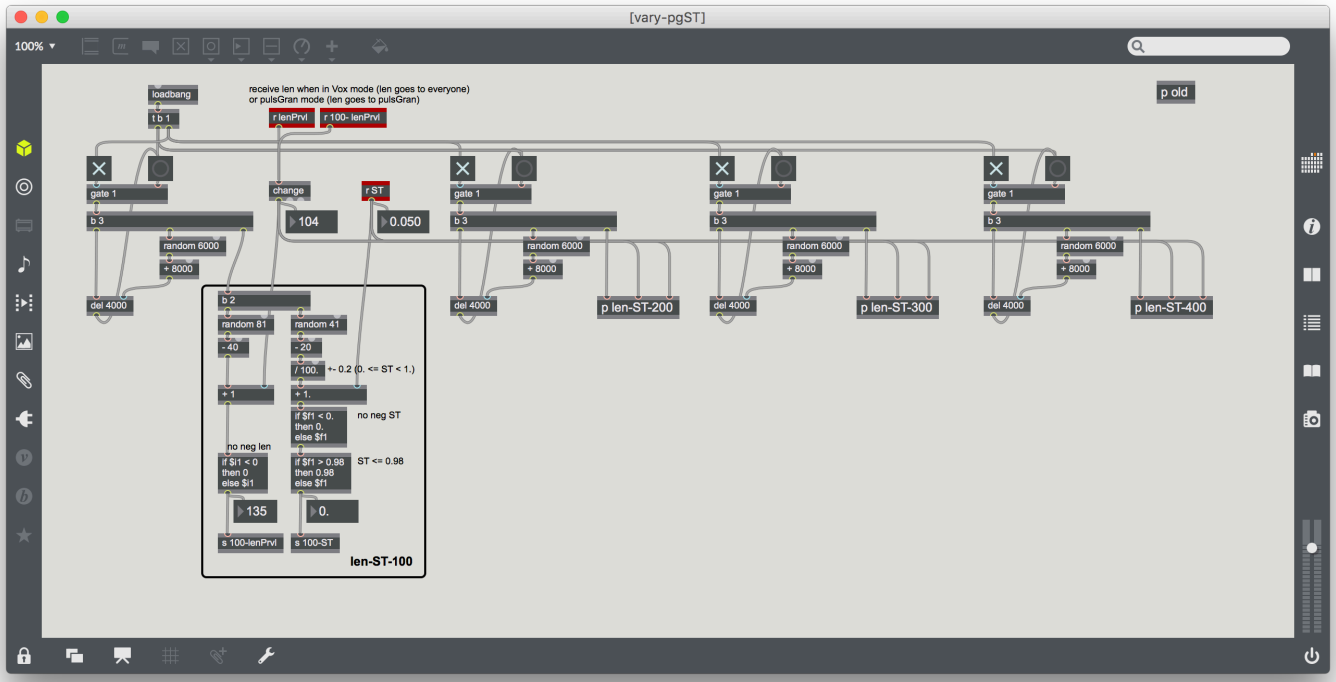


8.3.4.2 toLS



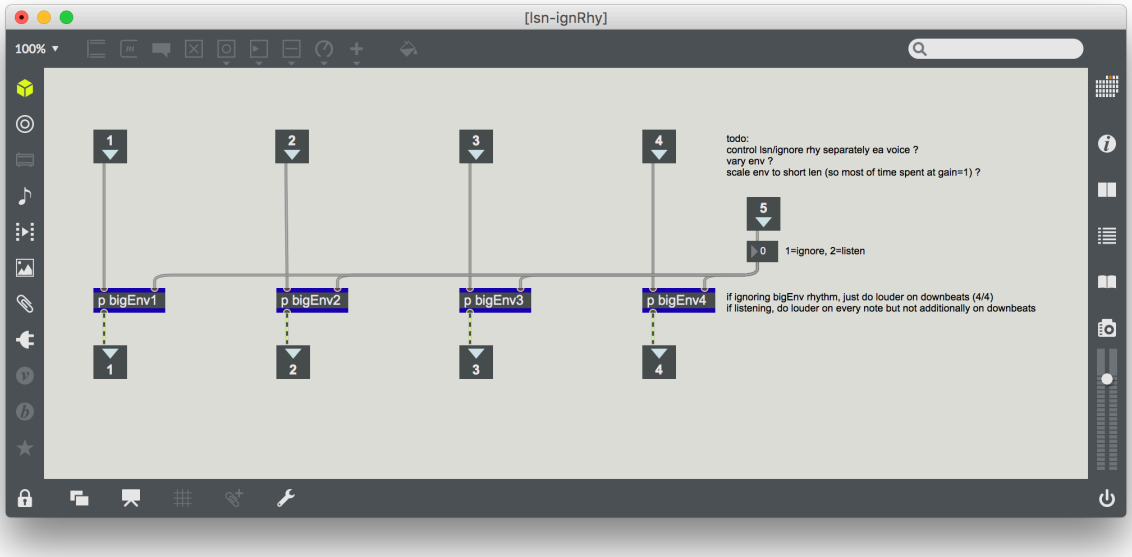
Appendix E (cont.) Vox6 patch

8.4 vary-pgST

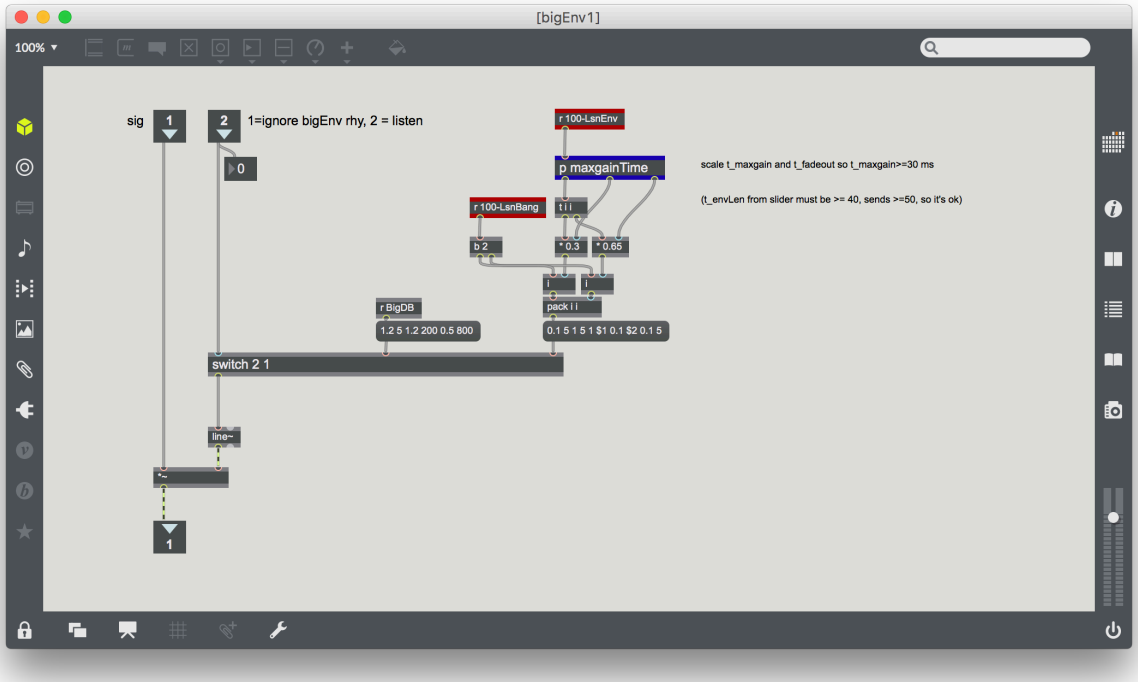


Appendix E (cont.) Vox6 patch

8.5 lsn-ignRhy

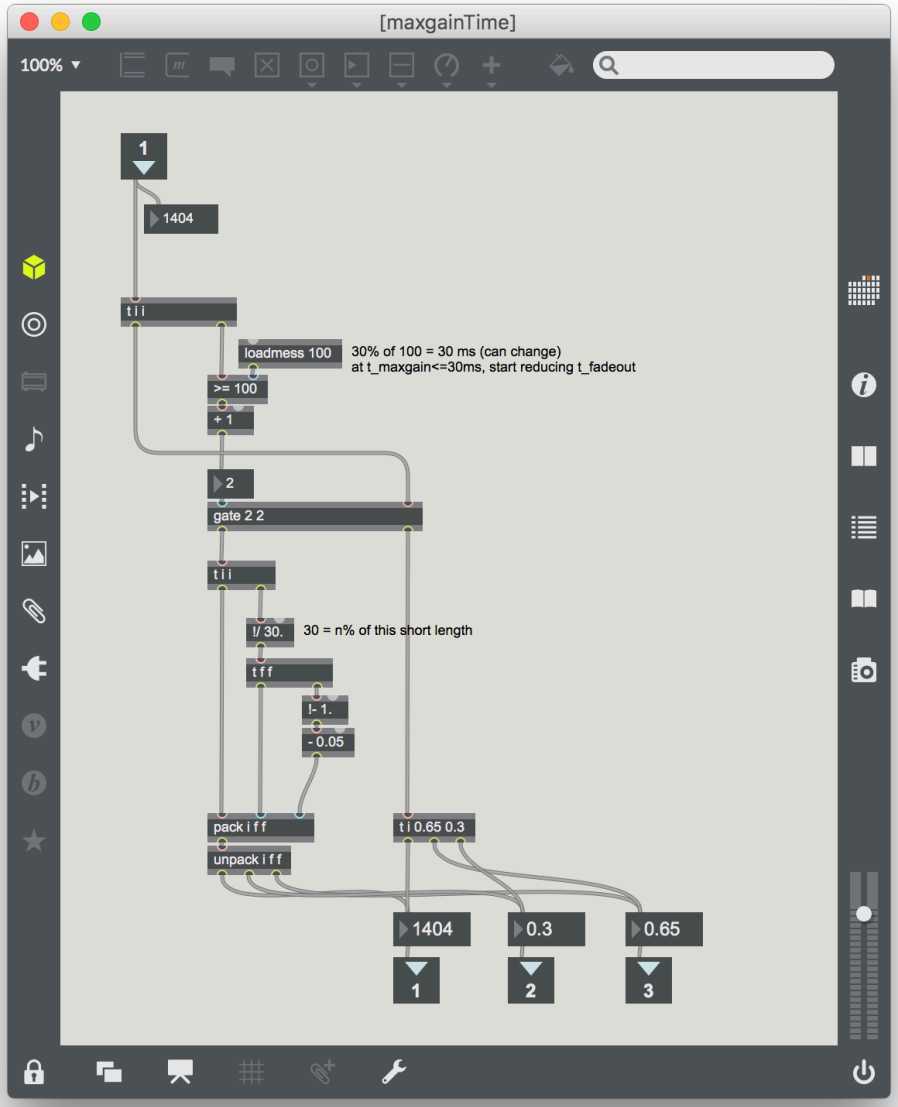


8.5.1 bigEnv1



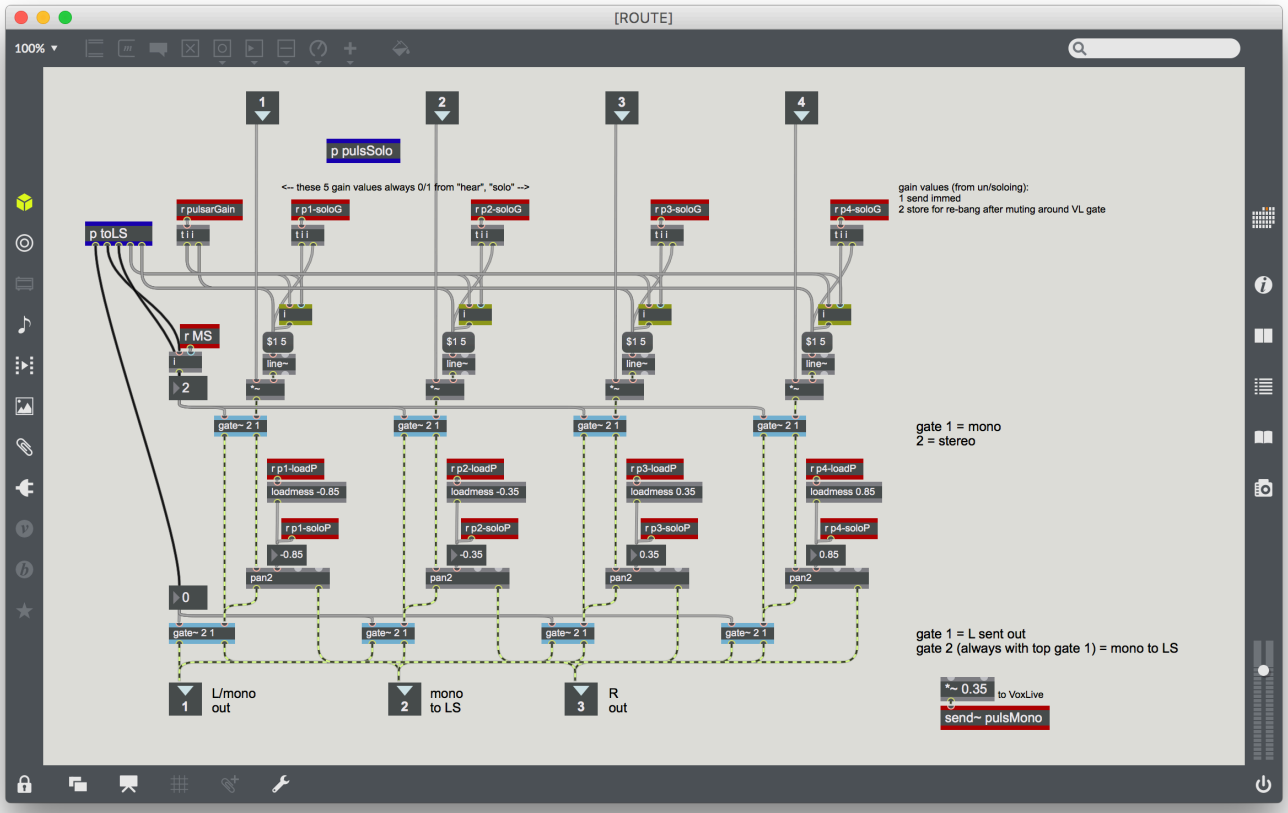
Appendix E (cont.) Vox6 patch

8.5.1.1 maxgainTime

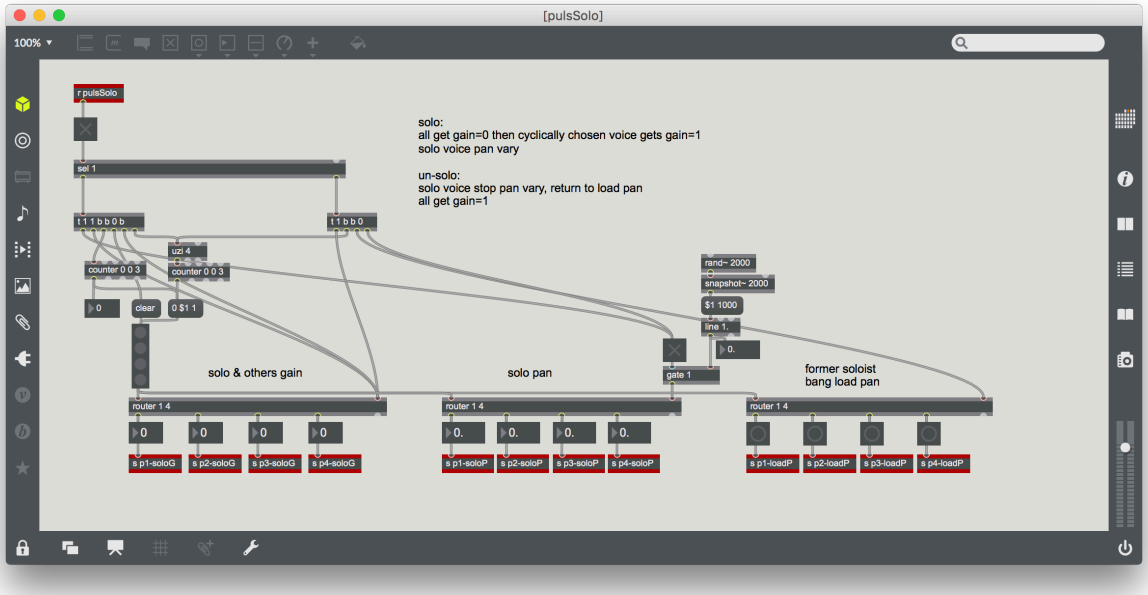


Appendix E (cont.) Vox6 patch

8.6 ROUTE

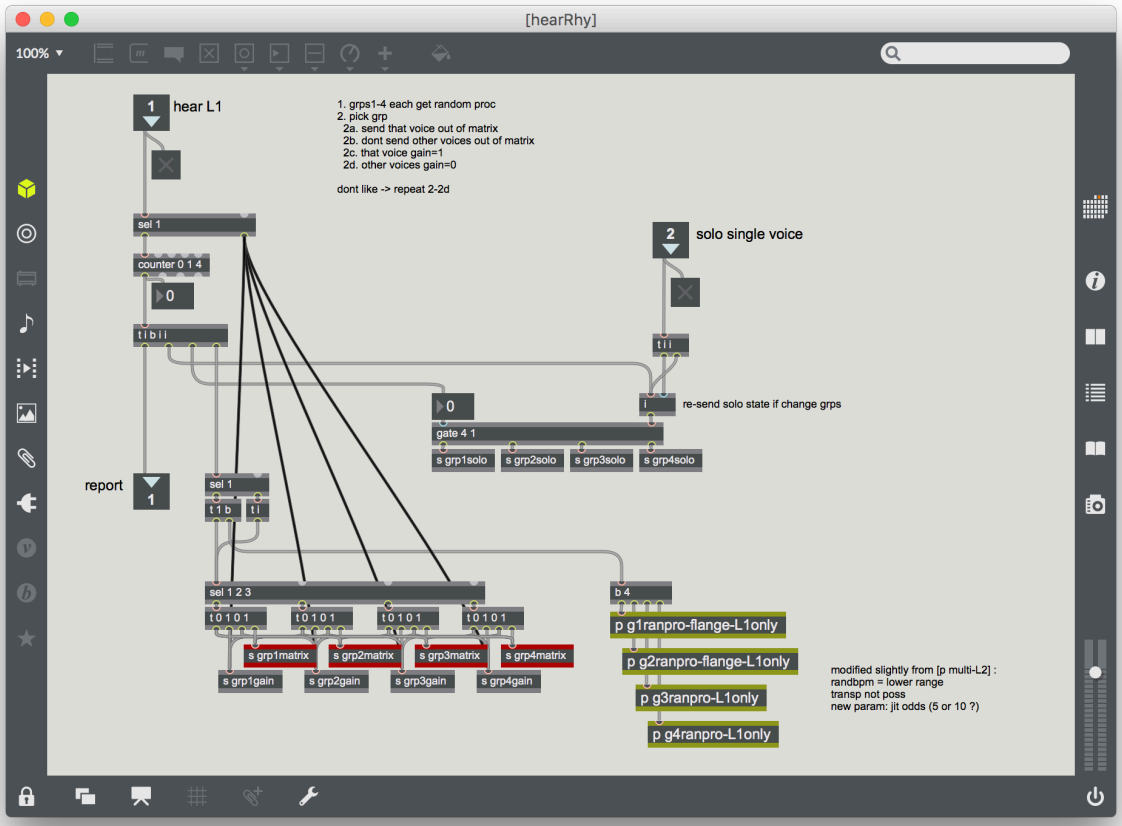


8.6.1 pulsSolo



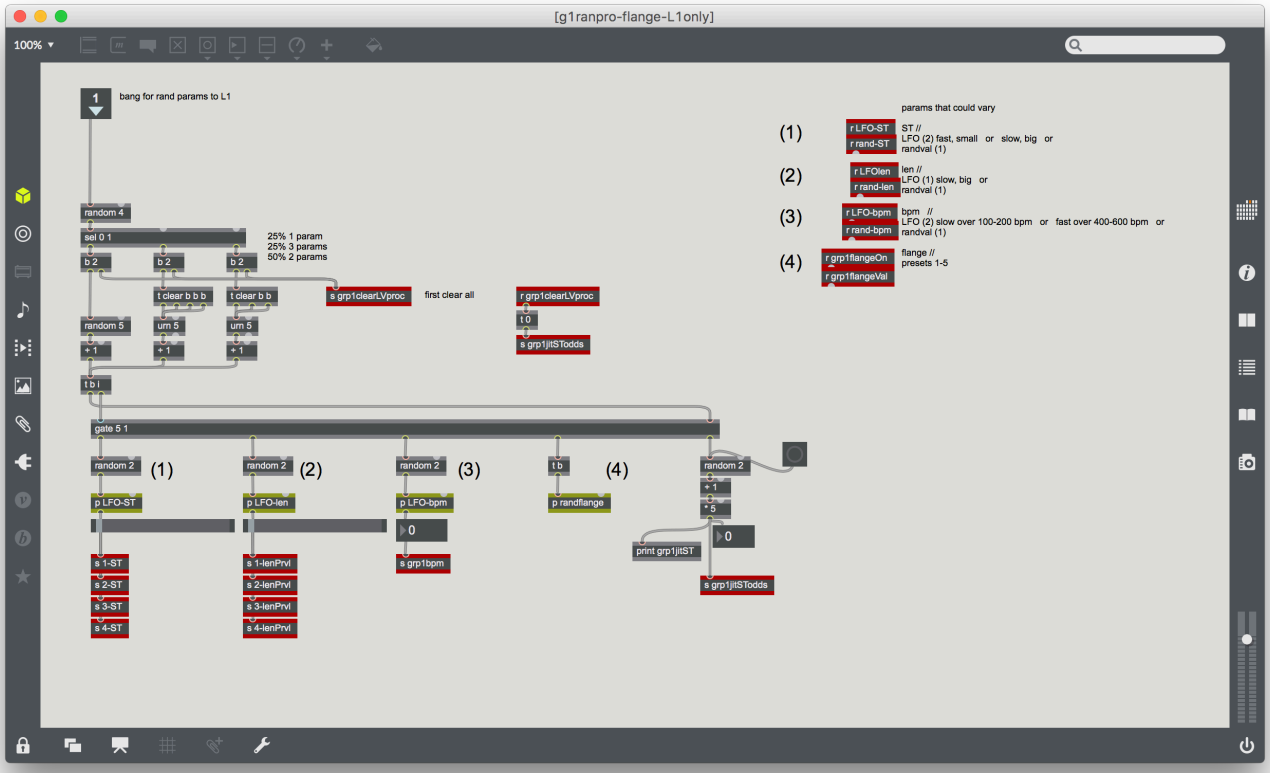
Appendix E (cont.) Vox6 patch

9. hearRhy



Appendix E (cont.) Vox6 patch

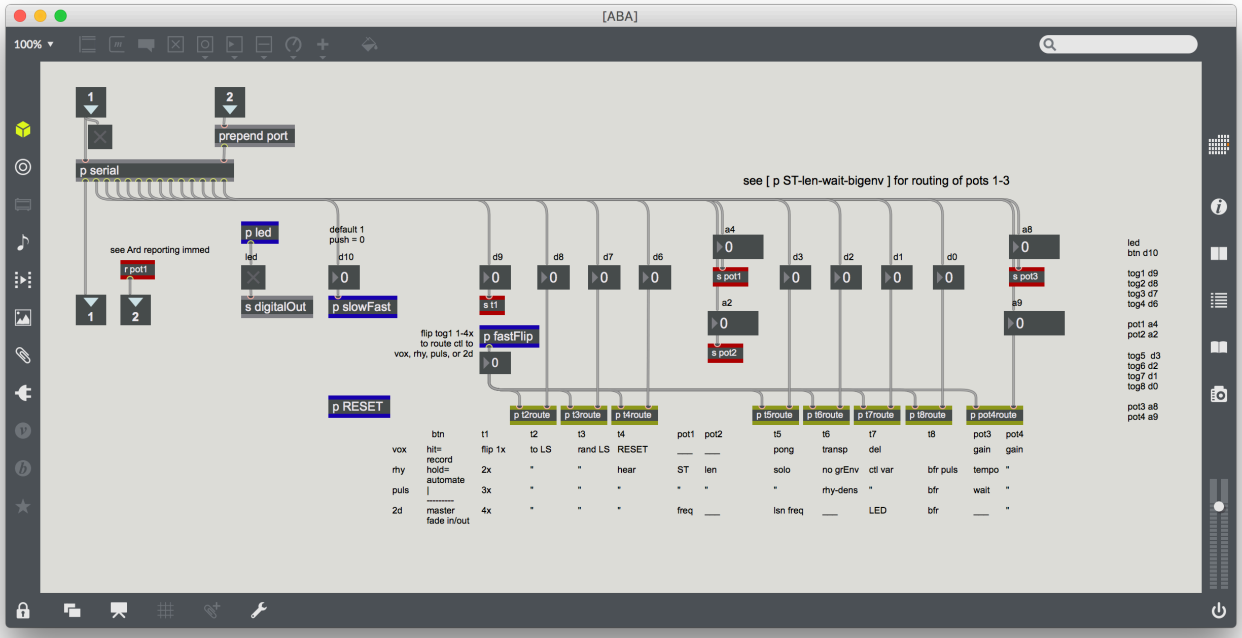
9.1 g1ranpro-flange-L1only



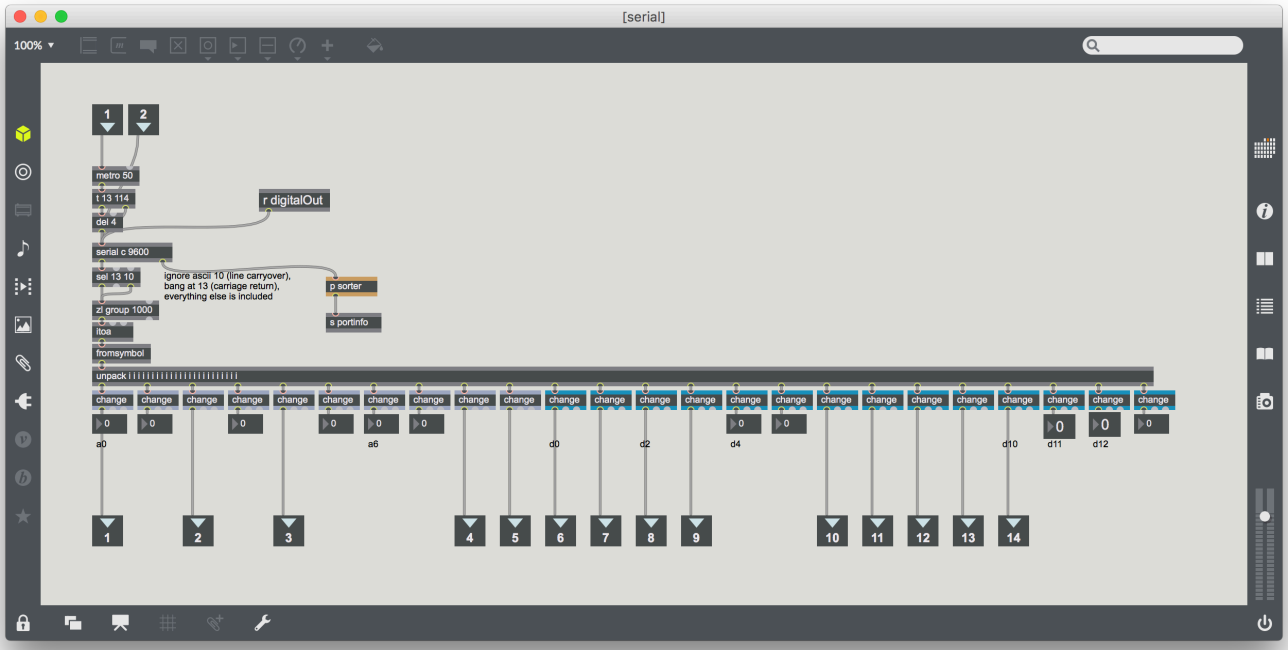
Unlike 8.1.1 g1ranpro-flange, this patch adds the parameter jitSTodds, i.e., the likelihood that a playbang will cause jitter around the Start Time value from the slider (values = 5 or 10%). This parameter is not available when creating buffers for pulsar granulation because if Start Time jitters frequently, audio sidebands are disrupted, preventing desirable low frequencies from being created by pulsar.

Appendix E (cont.) Vox6 patch

10. ABA

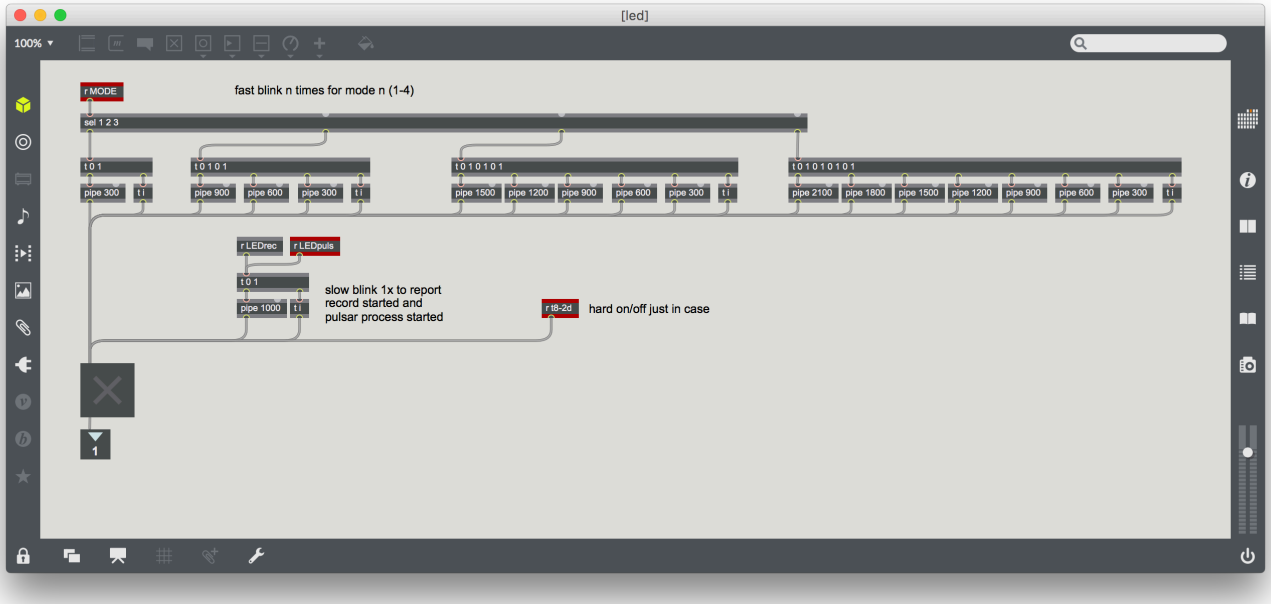


10.1 serial

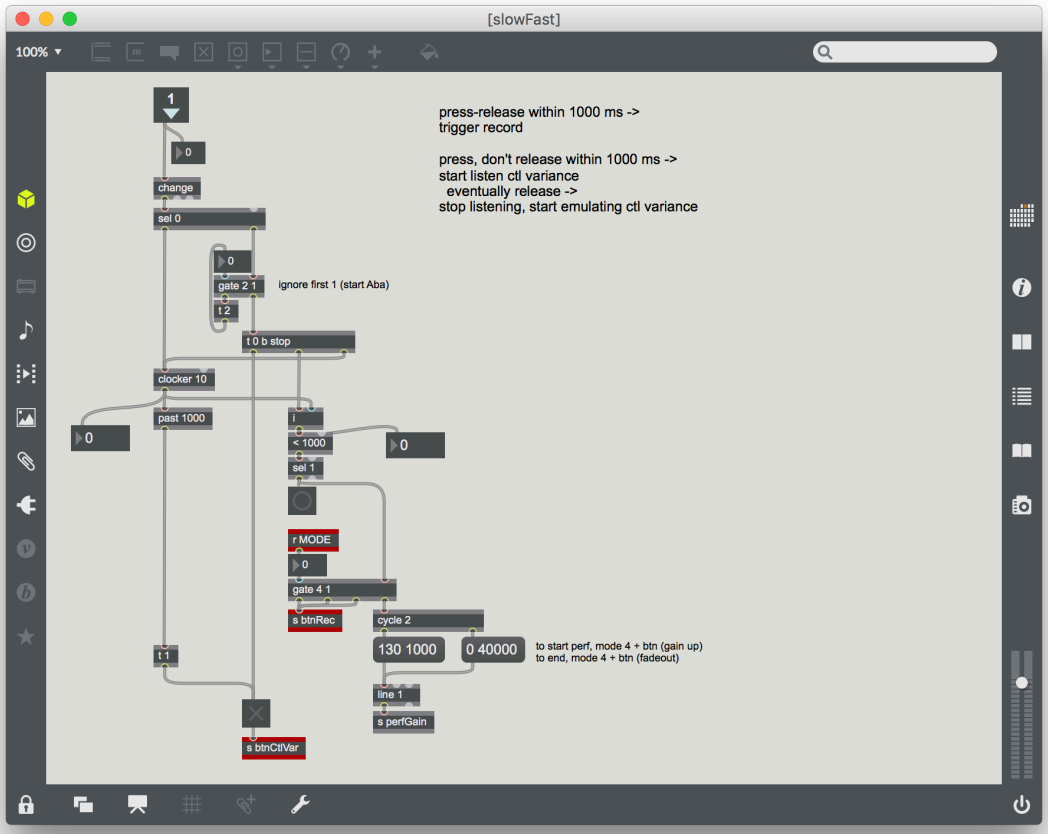


Appendix E (cont.) Vox6 patch

10.2 led

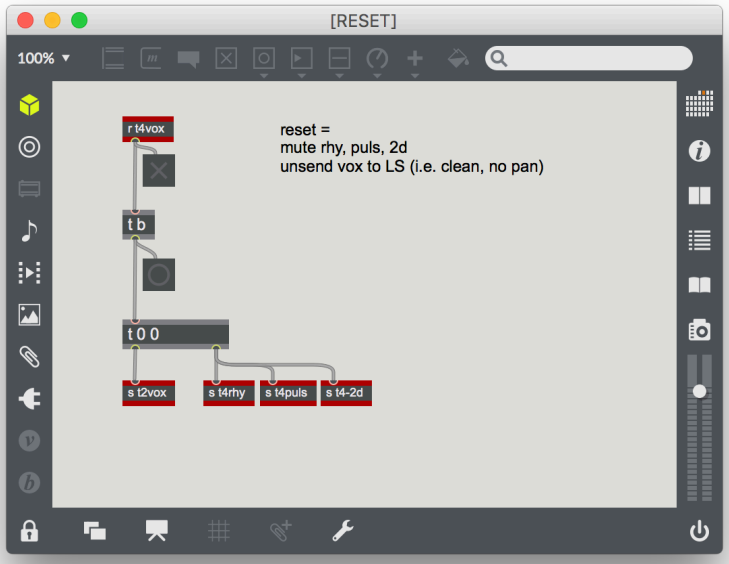


10.3 slowFast

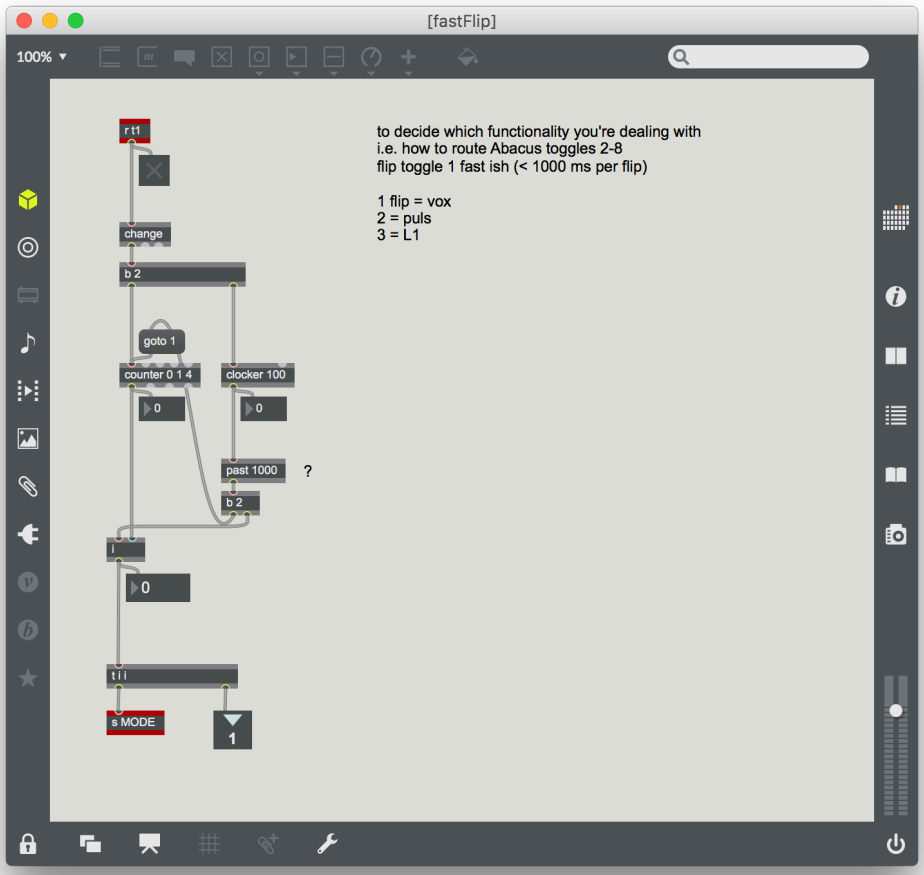


Appendix E (cont.) Vox6 patch

10.4 RESET

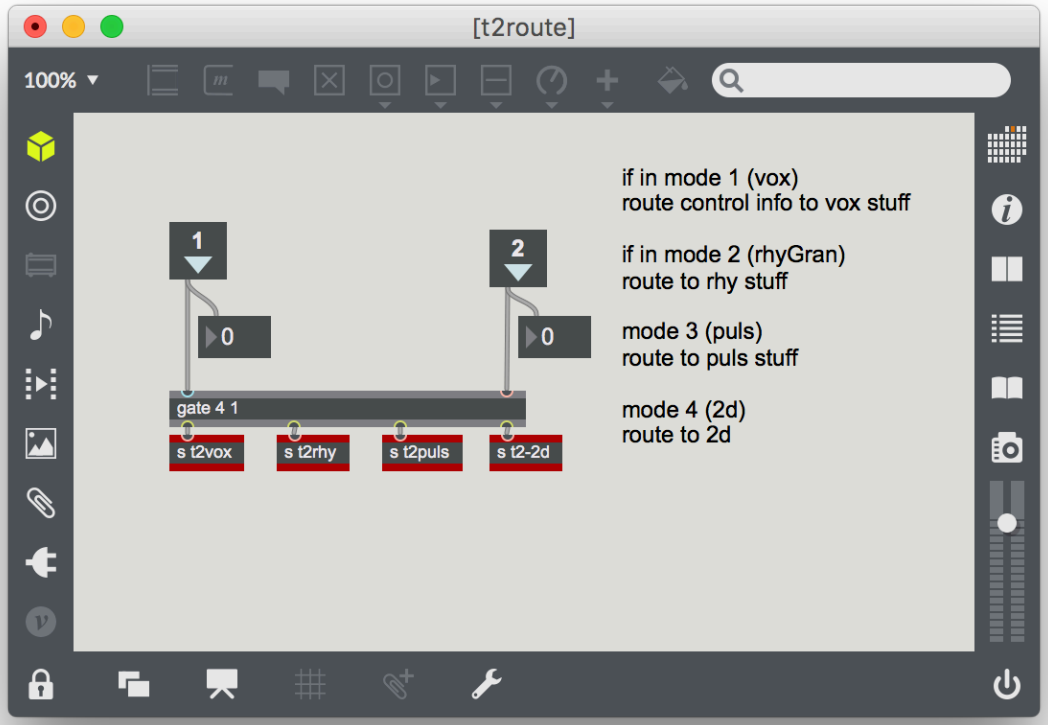


10.5 fastFlip



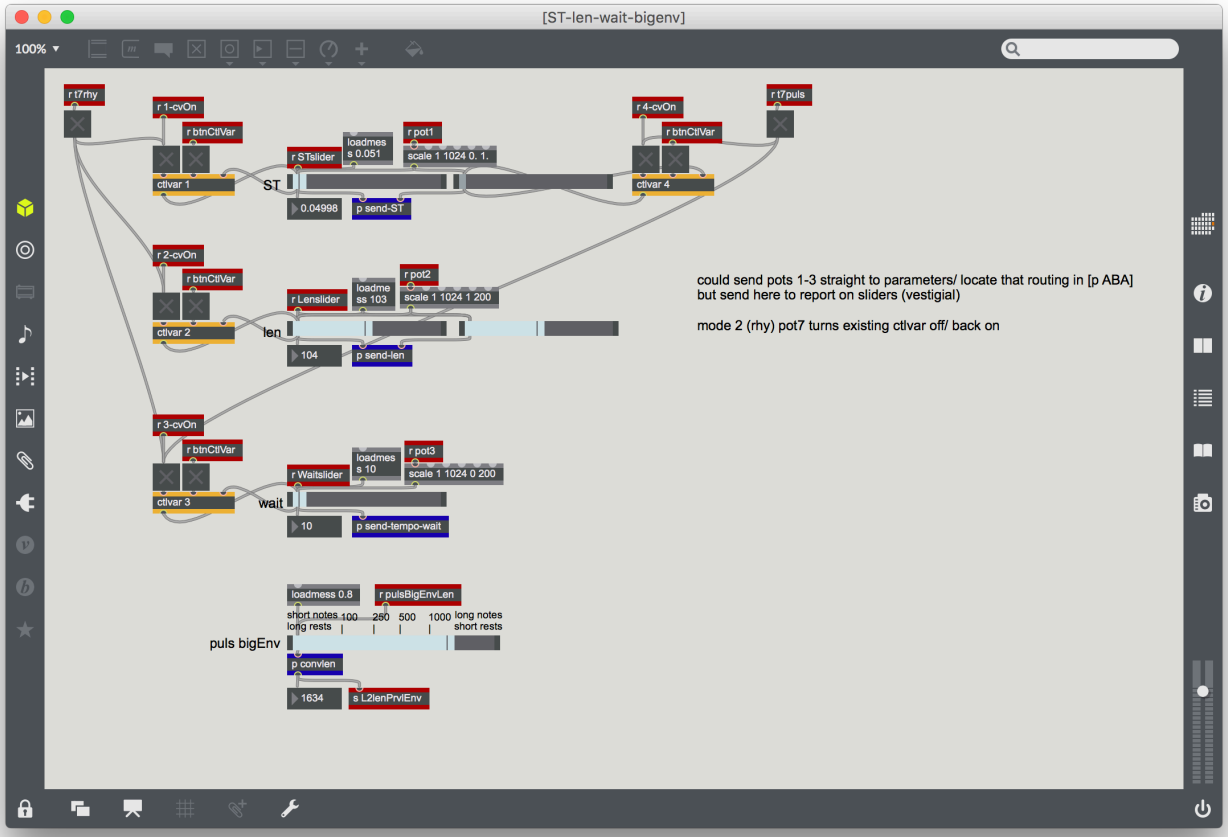
Appendix E (cont.) Vox6 patch

10.6 t2route

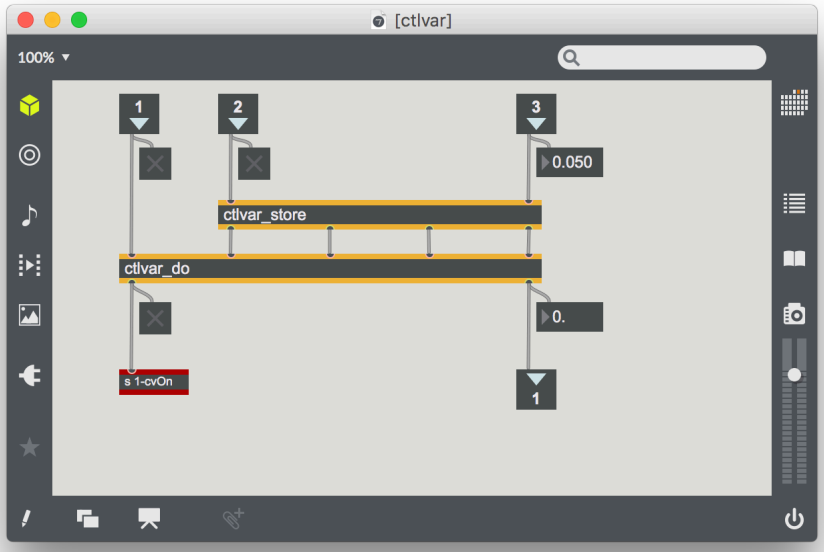


Appendix E (cont.) Vox6 patch

11. ST-len-wait-bigenv

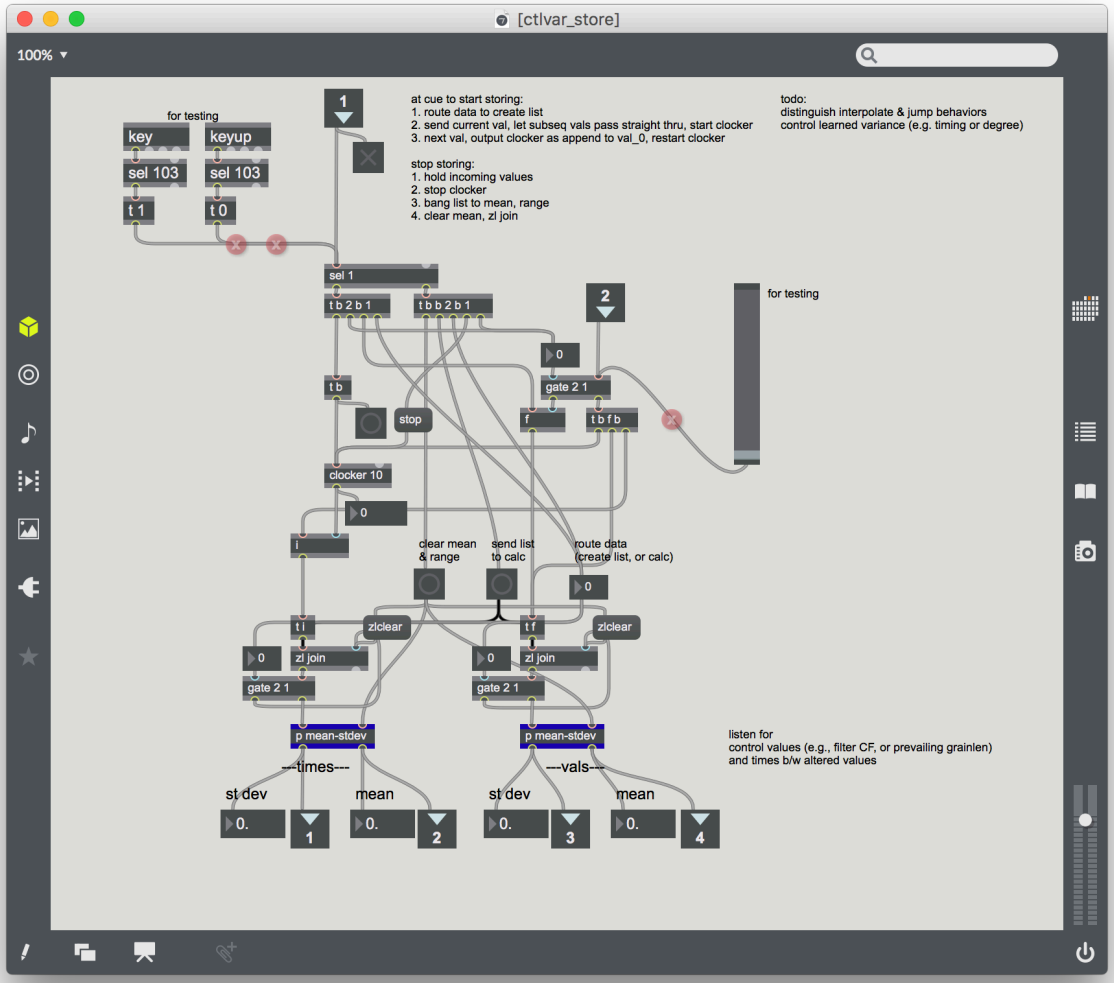


11.1 ctlvar



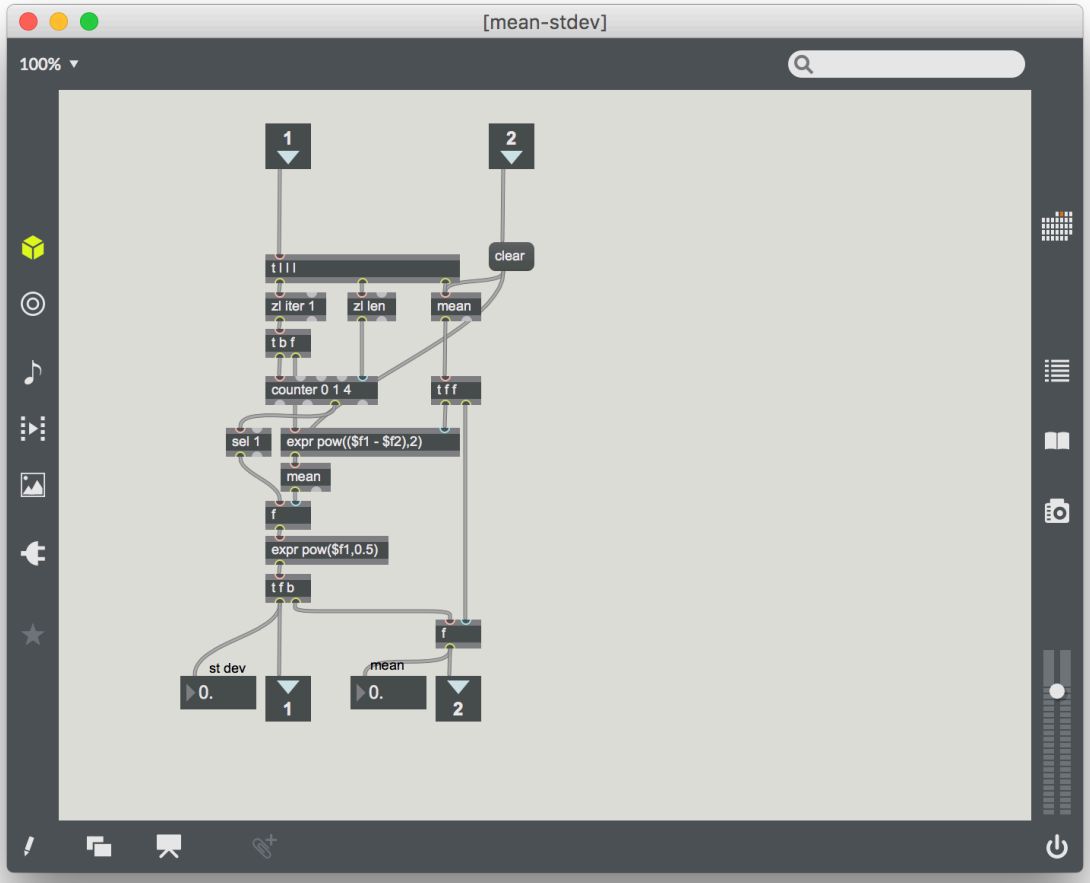
Appendix E (cont.) Vox6 patch

11.1.1 ctlvar_store



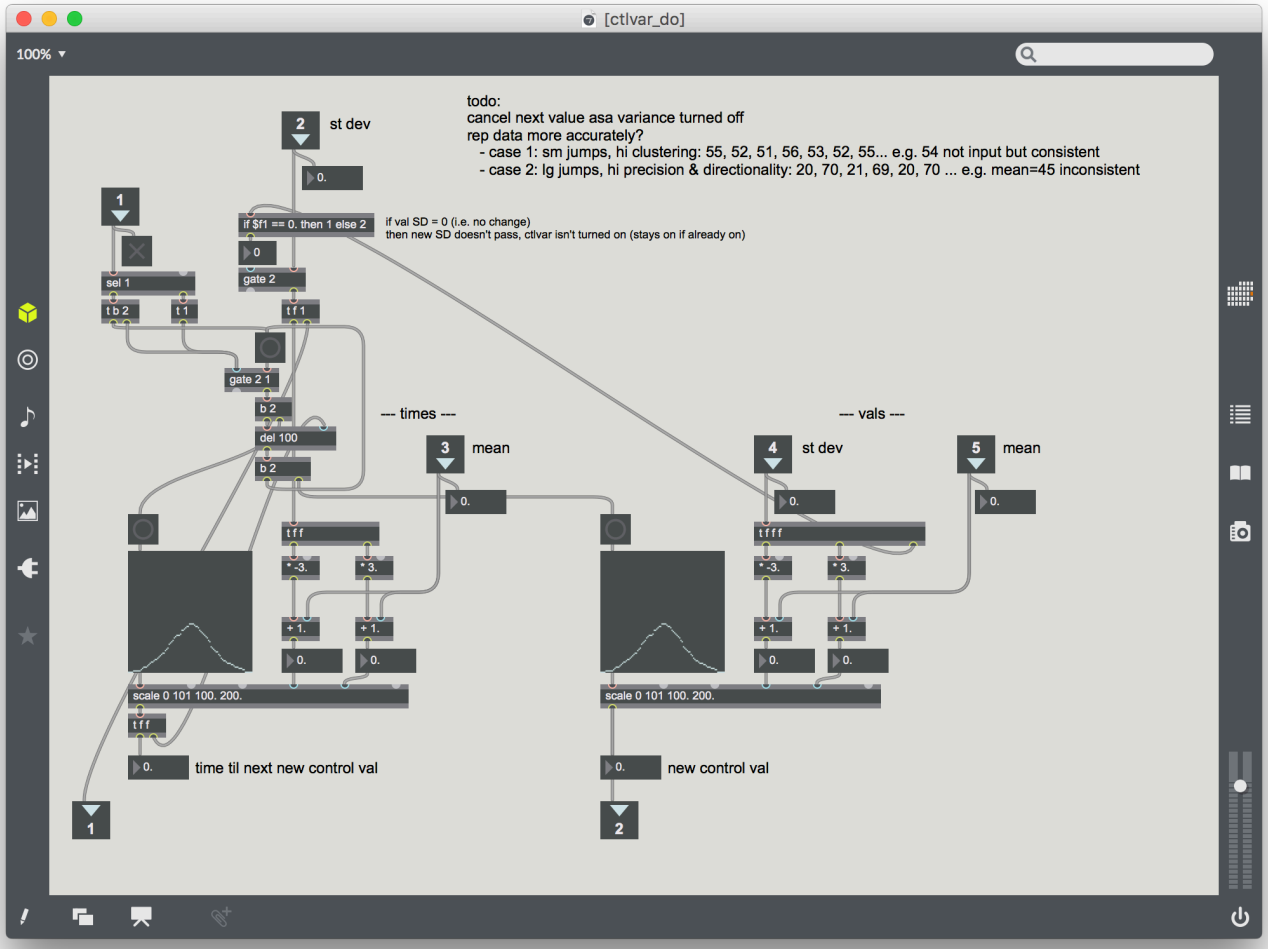
Appendix E (cont.) Vox6 patch

11.1.1.1 mean-stdev



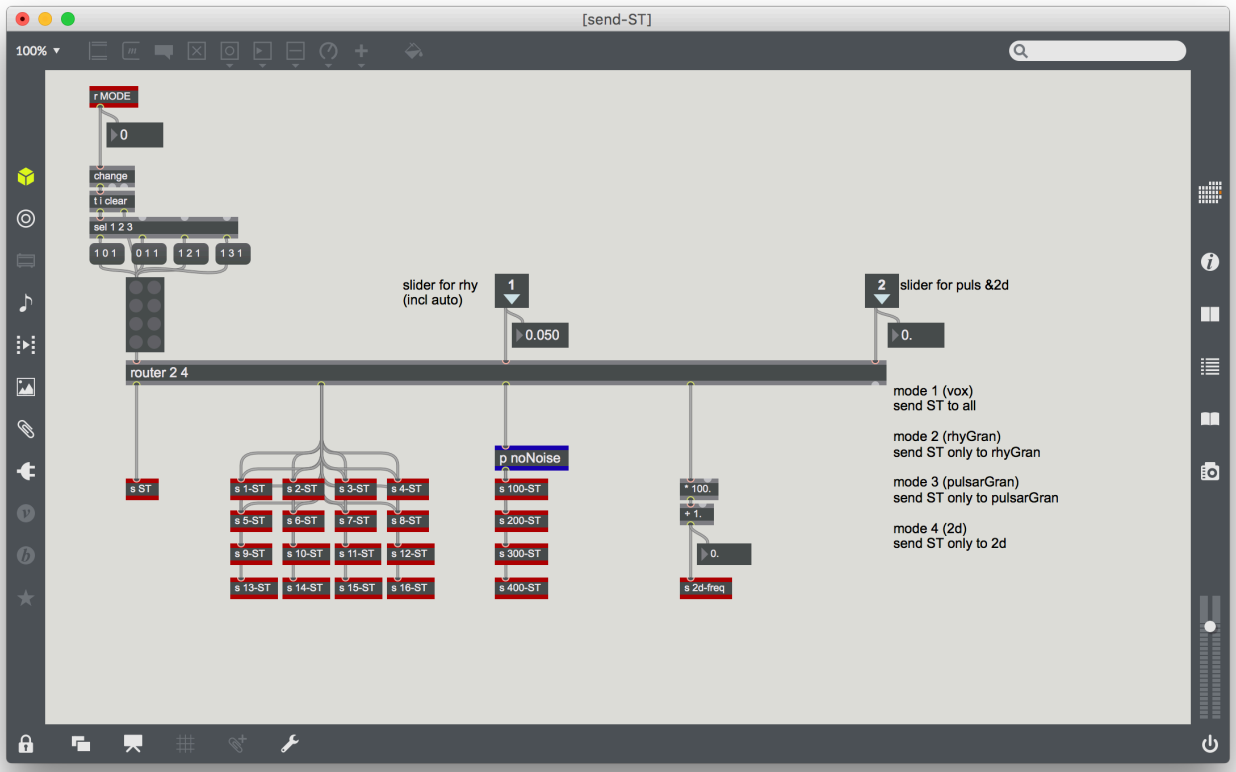
Appendix E (cont.) Vox6 patch

11.1.2 ctlvar_do



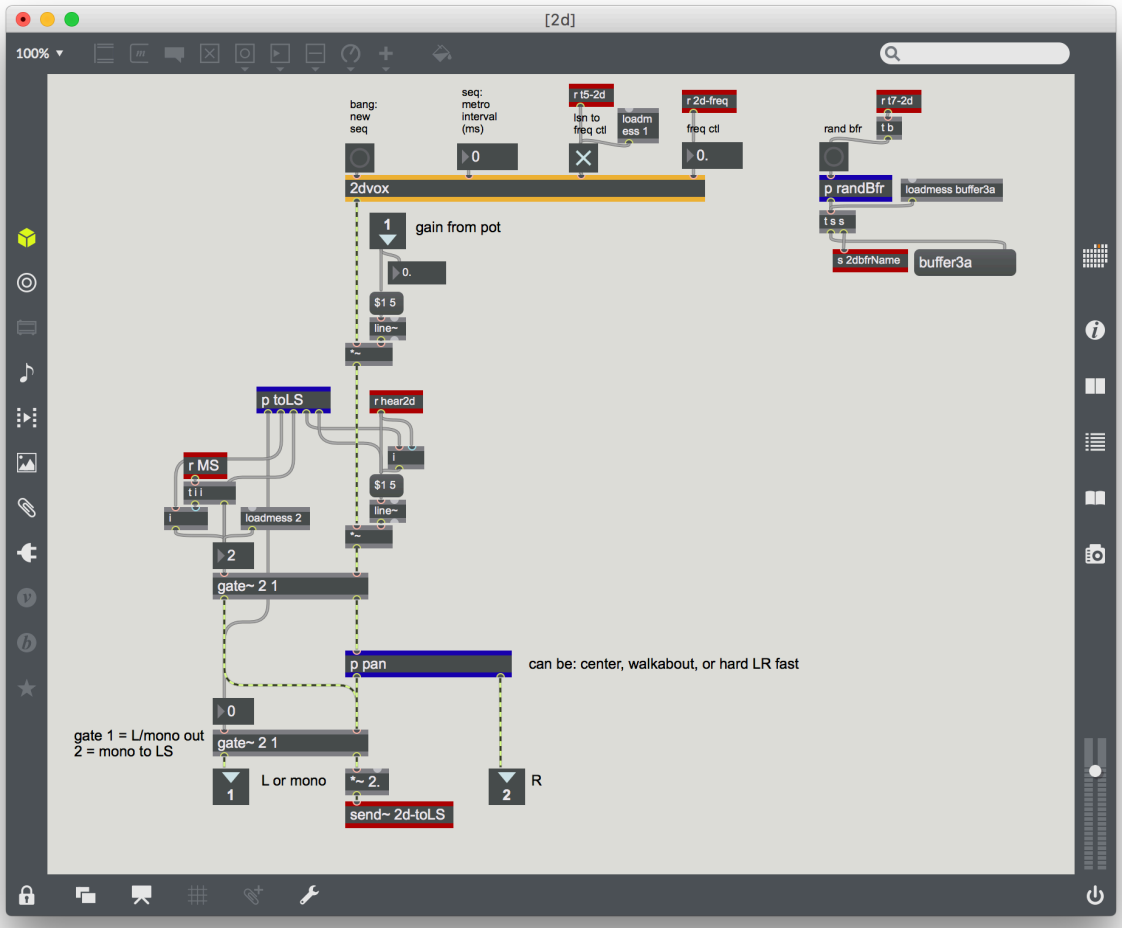
Appendix E (cont.) Vox6 patch

11.2 send-ST



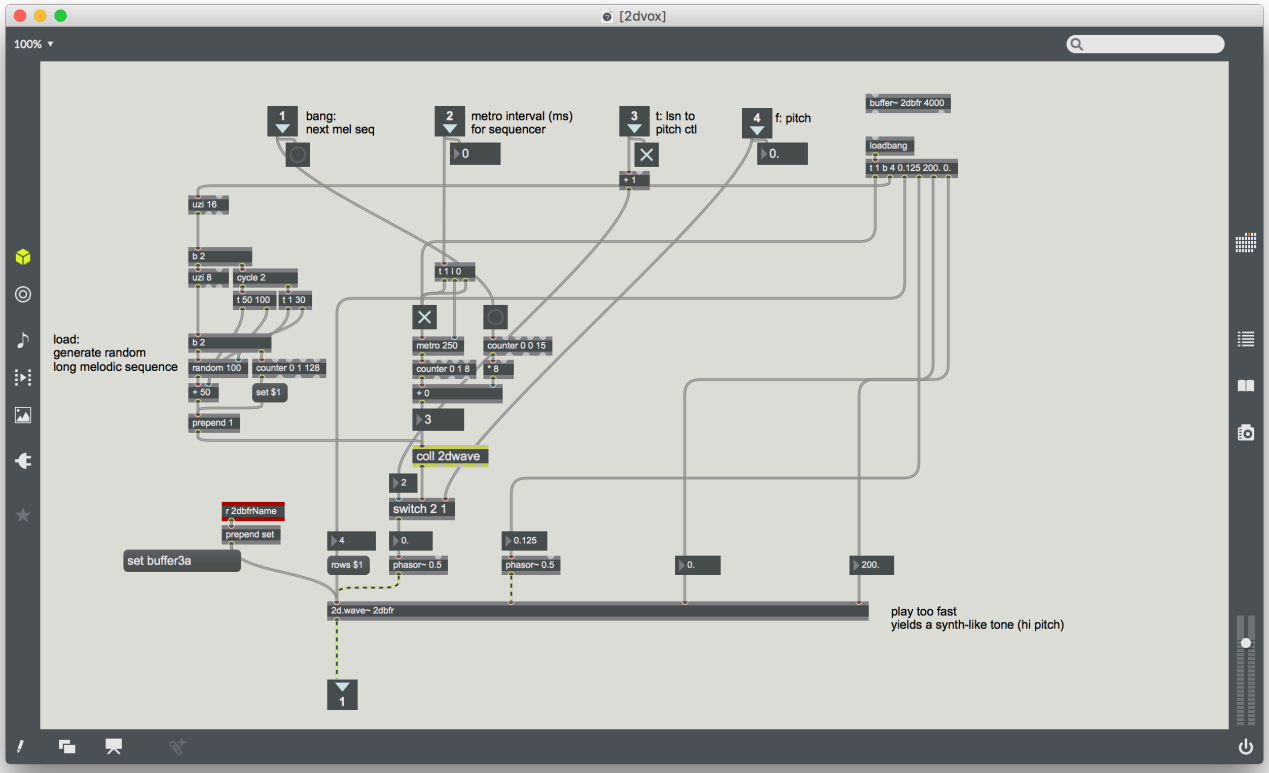
Appendix E (cont.) Vox6 patch

12. 2d



Appendix E (cont.) Vox6 patch

12.1 2dvox



Appendix E (concl.) Vox6 patch

12.2 randBfr