Journal SEAMUS

Spring/Fall 2020

The Society for Electro-Acoustic Music in the United States

Volume 31, Number 1–2

ISSN 0897-6473

Contact Information
Drake Andersen
Department of Music
Vassar College
124 Raymond Ave.
Poughkeepsie, NY 12604

journal@seamusonline.org

Journal SEAMUS
Volume 31, Number 1–2
Spring/Fall 2020

ISSN 0897-6473

Please visit us at www.seamusonline.org to find out more about SEAMUS

Information for Contributors

Interested contributors should submit manuscripts electronically. Microsoft Word is the preferred format. If another word processor is used, files should be saved in rich text format (RTF) with an accompanying PDF version. Main articles are generally 2,000 to 6,000 words.

Editing Guidelines Please use Times New Roman fonts with font size 12. Manuscripts should be formatted and prepared using *The Chicago Manual of Style*, 17th edition (2017) as a guide. References should follow Author-Date format. Specific citations should be provided in text in parentheses. Footnotes should be used sparingly and reserved for explanation beyond the text of the article. All references should be listed after the text of the article in a section labeled "References." Any computer code should be placed in fixed-width format to facilitate readability. Images, figures, musical examples, and other graphics should be sent as separate attachments for ease of layout. The approximate location of each graphic should be indicated in the text by a (sequentially numbered) label and a brief caption.

Graphics Any artwork, graphics, photos, and flowcharts should be sent as separate individual files. We recommend uncompressed graphic files such as TIFF at 300 dpi.

Submission All submissions, including articles, reviews, review proposals, and items for *Tips and Tricks* should be emailed to the Editor-in-Chief, Drake Andersen: journal@seamusonline.org.

About SEAMUS

Founded in 1984, The Society for Electro-Acoustic Music in the United States (SEAMUS) is a non-profit national organization of composers, performers, and teachers of electro-acoustic music representing every part of the country and virtually every musical style. Electro-Acoustic music is a term used to describe those musics, which are dependent on electronic technology for their creation and/or performance. Many members of SEAMUS, like Jon Appleton, the guiding light in the conception of the Synclavier, are recognized world leaders in their fields. All are dedicated to the use of the most advanced technology as the tools of their trade.

SEAMUS seeks to provide a broad forum for those involved or interested in electronic music. Through its journal, newsletter, national meetings, and its national archive at the University of Texas, SEAMUS seeks to increase communication among the diverse constituency of the relatively new music medium.

The Society's objectives include:

To encourage the composition and performance of electro-acoustic music
To develop a network for technical information and support
To promote concerts and radio broadcasts of electro-acoustic music both in the US and abroad
To create an exchange of information through newsletters and other means of communication
To establish and maintain a national archive and information center for electro-acoustic music
To attract a wide diversity of members and supporters
To advocate licensing and copyright concerns

SEAMUS strives to address not only relevant technology but also the non-technical issues pertinent to the electro-acoustic music community. In a field usually dominated by technical concerns, it is refreshing to hear paper sessions devoted to aesthetics, collaboration, education, and of the ethical and social issues facing electro-acoustic musicians. The provocative sessions provide fuel for lively discussions during the national meetings.

SEAMUS Board of Directors

President	Elizabeth Hoffman
Vice President of Programs	Leah Reid
Vice President for Membership	Abby Aresty
Member at Large	Eli Fieldsteel
Member at Large, Outreach	Ralph Lewis
Treasurer	Heather Mease
Secretary	John Gibson
Editor, SEAMUS Newsletter	Jiayue Cecilia Wu
Editor, SEAMUS Journal	Drake Andersen
Diversity Officer	Elizabeth Hinkle-Turner
Director, Recordings	Scott Miller
Director, Tech. Development	Becky Brown
Director, Communications	Mark Vaughn
Immediate Past President	Ted Coffey

Contents

Journal SEAMUS

Volume 31, Number 1–2, 2020

From the Editor

I am pleased to announce the publication of Journal SEAMUS Volume 31. This issue showcases a wide range of work and scholarship, from Elizabeth Hoffman and Elizabeth Frickey's exploration of themes of "radical play" in the music of Stuart Dempster to William Thompson's machine learning-based compositional system to an extended interview with the composer Yi-De Chen. In addition, we continue to pursue new initiatives, including community-sourced cover art. For Volume 31, the Journal team has selected a graphic score by the composer and visual artist Julie Zhu.

Another new initiative is *Where We Create*, an ongoing community-based project in which SEAMUS members are invited to share photographs of their workspaces. After the initial call in December 2023, we received dozens of wonderful photographs, both current and historical. The first batch is presented in this issue, featuring Buchlas, trumpets, keyboards, patch cables, historically-accurate hairstyles and, of course, cats. We hope that you will enjoy perusing them as much as we have! And by the way, if you have not yet contributed a photograph, please consider doing so by sending us an email (journal@seamusonline.org).

I am also pleased to introduce a new member of the Journal SEAMUS team, Assistant Editor Kramer Elwell. Kramer has jumped in enthusiastically and has already played a crucial role in realizing this latest issue. Speaking of Journal matters, I'll add that we're always interested in hearing your ideas for new articles, projects, initatives, and the like. We're especially interested in articles that take advantage of our new digital format. For more information—and the latest additions to our growing digital archive of back issues—check out this page: https://seamusonline.org/journal-seamus/.

As always, I would like to close by expressing my appreciation to the Journal SEAMUS staff, SEAMUS leadership, SEAMUS members, and all of the members who continue to support the Journal's mission.

Drake Andersen, Editor-in-Chief

Articles

Synthesizing Music and Speech as MIDI Files Through Machine Learning

William Thompson

The University of Southern Mississippi
w.a.thompson@usm.edu

Abstract

This research addresses the development of machine learning techniques used to create musical scores and performances that are inspired by the intersection of speech and music. Machine learning models are created from MIDI files that are transcribed from datasets of musical audio recordings and human speech audio recordings. Through the creation of succinct models, model based cross synthesis is possible. Models trained on musical MIDI data are asked to replicate MIDI data that approximate human speech. Alternatively, models that have been trained on MIDI data that approximate the spectral content of speech are asked to replicate musical MIDI data. The product of these developed techniques is a collection of piano music, *Seven Piano Etudes Speaks the Moody Machine*. These etudes are intended to be performed on one Yamaha Disklavier piano with two performers, one human pianist and one machine player piano. For context, an accompanying video example has been submitted that includes the musical score and the Disklavier piano performance.

Introduction

Human speech is an intrinsically musical phenomenon. The musicality of speech has aroused the interest of such scholars as Plato (Levman 2000), Charles Darwin (Farrar 1865), and modern researchers like Aniruddh Patel (Patel 2003) (Patel 2006). As perceived by the human ear, spoken-word contains extended meaning and emotional expression. The expressive qualities of music and speech overlap. Rhythmic, melodic, intervallic, and spectral qualities of speech contain a power to convey human expression strikingly similar to musical expression.

Music and speech are tools used for communication. Music is a means of communicating expression that is subjective in interpretation. Language exists as a tool for conveying information. When language becomes spoken word there is meaning beyond the literal interpretation of words. This non-literal aural expression exists in music and speech.

Musical qualities in speech and speech-like qualities found in music are valuable sources of inspiration and creation. As a result, techniques to create musical scores and performances that are inspired by the intersection of speech and music are fruitful means of composition. In the history of electroacoustic music speech has played an important role in composition. These roots date back to such works as Luciano Berio's *Thema (Omaggio a Joyce)* and Karlheinz Stockhausen's *Gesang der Jünglinge*.

In recent years the advancement of machine learning as a tool for creation has provided a powerful means for exploration of this unique relationship. Machine learning models are trained using recurrent neural networks to generate musical material inspired by speech. The corpora of these models are composed of symbolic music data in the form of MIDI files. These models fall into two overarching categories: those trained on corpora containing traditional MIDI files used to represent musical information and models trained on novel MIDI files that have been created to imitate the rhythmic, melodic, and spectral characteristics of speech. A wave to MIDI file conversion tool creates MIDI encodings from audio files of speech. Several moods to include amusement, anger, and neutral speech are trained into corresponding models, each with unique speech characteristics. The musical models are based on curated corpora to include the works of various modern composers and improvisers as well as the author's own improvisational style.

Through the creation of models based on speech and those based on music, cross synthesis is possible. These trained models can then be assigned MIDI primers which inspire models to generate MIDI material that resembles the primer file as seen through the lens of each model's training examples. Speech primers fed into models trained on music and music primers fed into models trained on speech, allow for the examination of sonic possibilities that exist in the middle ground. Additionally, interesting MIDI files can be generated by interpolating between these various models within a single piece of music. These techniques, as represented in figure 1, allow for the generation of sonic landscapes that hybridize music and speech in the form of musical scores for compelling human and machine performances alike. With the exploration of the aforementioned methods, a body of curated scores are created and presented as a portfolio of piano works entitled, *Seven Piano Etudes Speaks the Moody Machine*. These works are intended to be performed by a unique piano duo that features one human pianist and one self playing Yamaha Disklavier Piano performer.

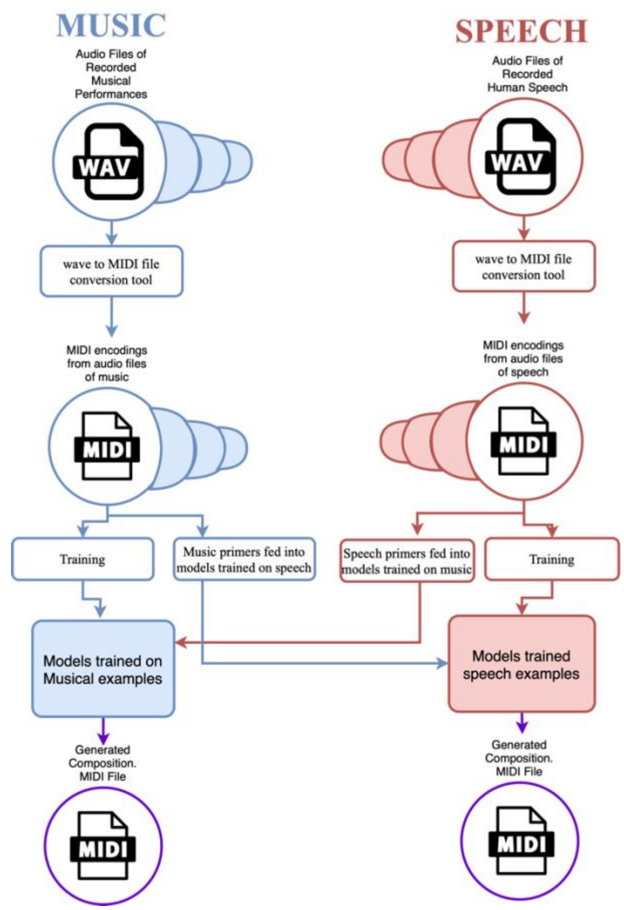


Figure 1. MIDI models trained on music and speech.

Methods

Magenta

Magenta is an open-source Python library designed to investigate the possibilities of machine learning in artistic creation. It was designed by the Google Brain team and uses TensorFlow, another google library, to ease the process of building datasets, training models, and creating predictions (Dubreuil 2020).

Arguments and Parameters

The generation algorithm works by iteratively predicting each generation step in a sequence based on what the model has learned during training. Because we are interested in generating symbolic musical information this means that each note or rest is predicted sequentially until an entire score of a desired length is created. The scalable features that determine how the model will generate a sequence are: beam size, branch factor, and the number of steps per interaction. The following is a list of flags that can be used to set the parameters of any of Magenta's RNN architectures:

- –num outputs. Specifies how many generations to be created per execution. These generations are MIDI files.
- –num steps. Specifies the length of the generated MIDI file in steps. Each step equals one 16th note. Because of this a measure of 4/4 time will have 16 steps. If a primer file is used these steps are in addition to the number of steps contained in the primer file.
- –qpm. Specifies the number of quarter notes per minute (QPM). In Magenta QPM is a way of controlling tempo. If a

primer file is used the QPM of the primer file will be used and this flag will be ignored.

- –primer midi. Specifies the path to a primer MIDI file. The number of steps of the generation must be longer than the number of steps in the primer file.
- –temperature. Specifies the randomness of a generation. A temperature of 0.1 uses softmax possibilities, a number higher than 0.1 makes the generation progressively more random and a number lower than 0.1 makes the generation less random.
- –beam size. Specifies the beam size to use for mean search.
- –branch factor. Specifies the branch factor to use for bean search.
- –steps per iteration. Specifies the number of steps to talk per beam search iteration (default 1).
- –bundle file.

Primers

In Magenta, primer files are an interesting way of inspiring model generation. As a parameter, a priming MIDI file can be set before generation. When set, the model will attempt to create something similar to the primer with the only tools it has, its training data. This is the primary way of creating model based cross-synthesis. For example, if a model has been trained on MIDI that represents human speech and a primer file of a musical example is set, the speech model will attempt to create something similar to the musical example through the lens of its training data, speech. Conversely, a musical model primed with a speech MIDI file will result in a generation attempting to imitate speech though a musical language.

Additionally, two other parameters can be set to control how the model uses the primer MIDI file. ‘Condition on primer’ and ‘inject primer’ during generation are both Boolean statements. If ‘condition on primer’ is set to ‘true’, the primer sequence will be used by the RNN for generation. ‘Inject primer’ during generation on the other hand tells the network to begin the generation with the MIDI primer in the actual generation output. If a primer is set for a generation and ‘condition on primer’ and ‘inject primer’ during generation are both set to ‘false’, the primer file will appear in the generated MIDI file before any newly generated material (Dubreuil 2020).

Training

Training machine learning models is comprised of a few steps. First, datasets of MIDI files must be created that contain features that the model can learn from. Once a dataset is collected, MIDI files must be converted into notesequences. This is done in a command line action, ‘convert dir to notesequnces.’ The result of this command is the creation of the file, notesequences.tfrecord. Next, the sequence examples can be launched by another command line action that points to the newly created notesequences.tfrecord document. This action creates a sequence example folder that contains a training set and an evaluation set. Finally, the actual training can begin. The following command line prompt tells Magenta to use the Polyphony RNN architecture, sets the output directory with run dir, points to sequence examples, asks that the network for training be a three layer with 128 units per layer and an overall batch size of 64. In addition this command is calling for the network to train for 30,000 steps (Dubreuil 2020).

```
polyphony_rnn_train \
–run_dir=tmp/polyphony_rnn/
logdir/run1  \
–sequence_example_file=tmp/
polyphony_rnn/sequence_examples/
training_poly_tracks.tfrecord  \
–hparams="batch_size=64,
rnn_layer_sizes= [128,128,128]" \
-num_training_steps=30000
```

Models Created Representing Traditional Musical Scores

These musical models are based on MIDI transcriptions of recorded piano music. Collections of these transcriptions form corpora that when trained into machine learning models, emulate musical styles of select composers and improvisers. This compositional process comprises four models of different piano styles based on these musical artists or works: D(u)0 by Mara Gibson, the compositions and improvisations of Thelonious Monk, the piano playing of James Carroll Booker III, and piano improvisations by the author.

Preparing Music Corpora for Training

Each of the four musical models are composed of a corpus of MIDI files which are created as a representation of audio recordings comprising each model respectively. Because several of the models are created from improvisations, a method of audio to MIDI transcription is necessary since commercially produced MIDI files and scores are not available for these examples. Additionally, because a rather large corpus of MIDI examples are needed for training the RNN, transcription by ear is not realistic. In order to computationally accomplish this task Magenta's dual objective piano transcriber, Onsets and Frames was used. (Hawthorne, Elsen, Song, Roberts, Simon, Raffel, Engel, Oore, Eck 2017)

The preparation for training each model is a three-part process. First, audio recordings that represent each model are collected. Next, the collected audio files are transformed into proper lengths for training. Finally, these audio files are converted into MIDI files as a corpus for model training.

The audio datasets collected for corpus creation are wave recordings of solo piano music. Considerations for dataset requirements include: high fidelity, limited audience noise in live recordings, and perceived validity of dataset uniformity. In some cases audio files that meet the first two criteria are disregarded because they seem out of character for the desired corpus. Each model's success is better measured when members of the training corpus contain similar features.

Before the conversion from audio file to MIDI file can be executed, file length must be considered. MIDI file length is an important factor for training models. Some MIDI files are regarded by Magenta as too long to create sequence examples for training while others are too short. Through the process of debugging and attempting to get as many sequence examples for training purposes an audio file length of 60 seconds is found to be most fruitful. Because of this, the audio command line processing library FFMPEG (Lei, Jiang, Wang 2013) is used to first concatenate all audio files in a directory. Next, FFMPEG is used to cut the now concatenated audio file into many 60 second audio files.

Once the audio files of each dataset are the appropriate length, it is time to convert audio into MIDI. With Onsets and Frames, polyphonic piano transcription is possible by using a deep convolutional and recurrent neural network which is trained to predict onsets and frames. When this method detects an onset audio event in a recording it will only activate if a pitch relating

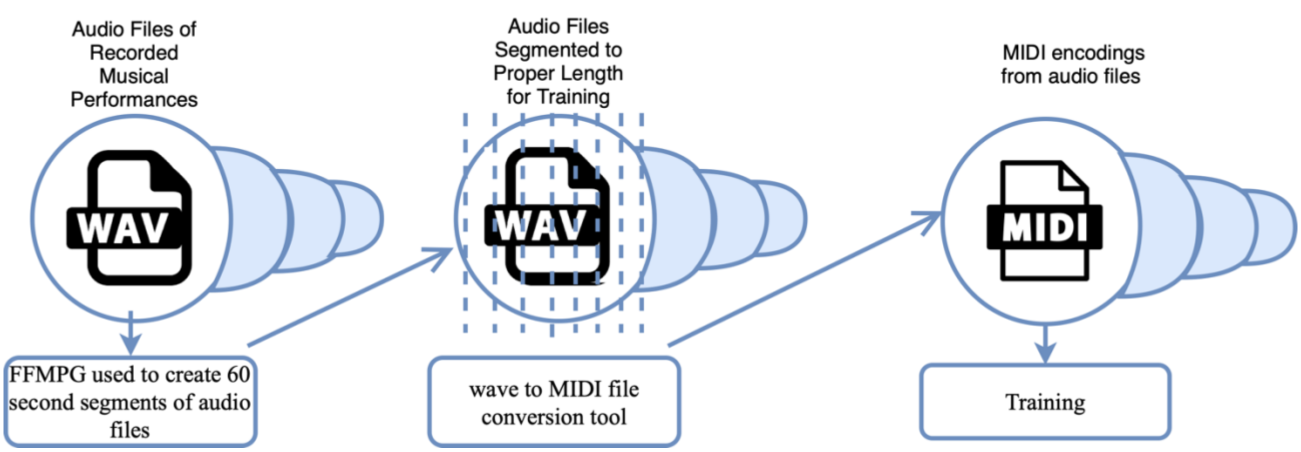


Figure 2. The preparation for training each model is a three-part process. First, audio recordings that represent each model are collected. Next, the collected audio files are transformed into proper lengths for training. Finally, these audio files are converted into MIDI files as a corpus for model training.

to the same frame is also detected. This method of transcription is much more accurate than prior automated piano transcription attempts by the Magenta team and others. Additionally, the model can predict relative velocities in audio files, resulting in more natural-sounding MIDI transcriptions (Hawthorne 2017).

Models Created Representing Characteristics of Speech

Machine learning models based on characteristics of speech are created through a similar process as models based on music. These speech models are made up of MIDI transcriptions of audio recordings of human speech. Collections of these transcriptions form corpora that when trained into machine learning models, emulate the sonic characteristics of specific emotional states of speech. This compositional process consists of three emotional states of speech: neutral speech, angry speech and amused speech. Each of these three emotional states is used to create a model intended for human pianist performance and a model intended for Yamaha Disklavier machine performance. The Yamaha Disklavier piano is a modern player piano that can read MIDI files (Kapur 2005). A total of six speech models are created to account for human and machine performance options representing each of the three selected emotional states of speech.

Preparing Speech Corpora For Training

Each of the six speech models are created from a corpus of MIDI files. These MIDI files are automatically transcribed from audio recordings of voice actors reading text spoken with the intent of emulating specific emotional moods of speech. The source of these recordings is the The Emotional Voices Database: Towards Controlling the Emotional Expressiveness in Voice Generation Systems, or EmoV-DB for short. EmoV-DB includes recordings of two male speakers and two female speakers. Each recorded speaker seeks to replicate the emotional styles of a neutral, sleepy, angry, disgusted or amused state. The original purpose of EmoV-DB is to aid in the field of emotional speech synthesis. The emotional styles of neutral, amused and angry are included in this compositional process (Adigwe 2018). The neutral, amused, and angry speech datasets of audio files are converted to MIDI files for model training through a similar process as the music models. Due to the complexity of sonic characteristics of recorded speech and the necessity of a large corpus for training models, transcription by ear is not practical. In order to computationally accomplish this task two separate transcription methods are used with two distinct goals. One transcription method is used to create MIDI corpora intended for machine performance, and another transcription method is used to create MIDI corpora that are performable by a human pianist.

To computationally transcribe audio files into a MIDI file corpora intended for machine performance, the command line tool, WaoN (Ichiki) is used to extract spectra from recorded speech audio files into MIDI files that contain polyphonic pitch information. The result is often a MIDI file that can contain very many notes based on an audio file's spectral content. When applied to speech recordings, WaoN can detect enough pitch information to make MIDI representations of speech recordings that are almost, if not actually, intelligible as specific words by the listener. Listening to these MIDI files is very similar to hearing the ”speaking piano” works of Peter Ablinger (Barrett 2009). This transcription method is ideal for machine performance by a player piano since these machine performances are not restricted to any range of possible simultaneously sounding pitches. However, because of the abundance of pitch information, these MIDI speech realizations are in most cases not performable by a human pianist.

To form the human performable MIDI corpora, Magenta's dual objective piano transcriber, Onsets and Frames is used. This transcription method is intended for piano transcription. However, when applied to human speech recordings, Onsets and Frames creates MIDI files that contain one to six simultaneous pitches on average, which are usually practical for human pianist performances.

Much like processes of preparation for training music models, the speech models for both human and machine performance are prepared in a three part process. First, audio

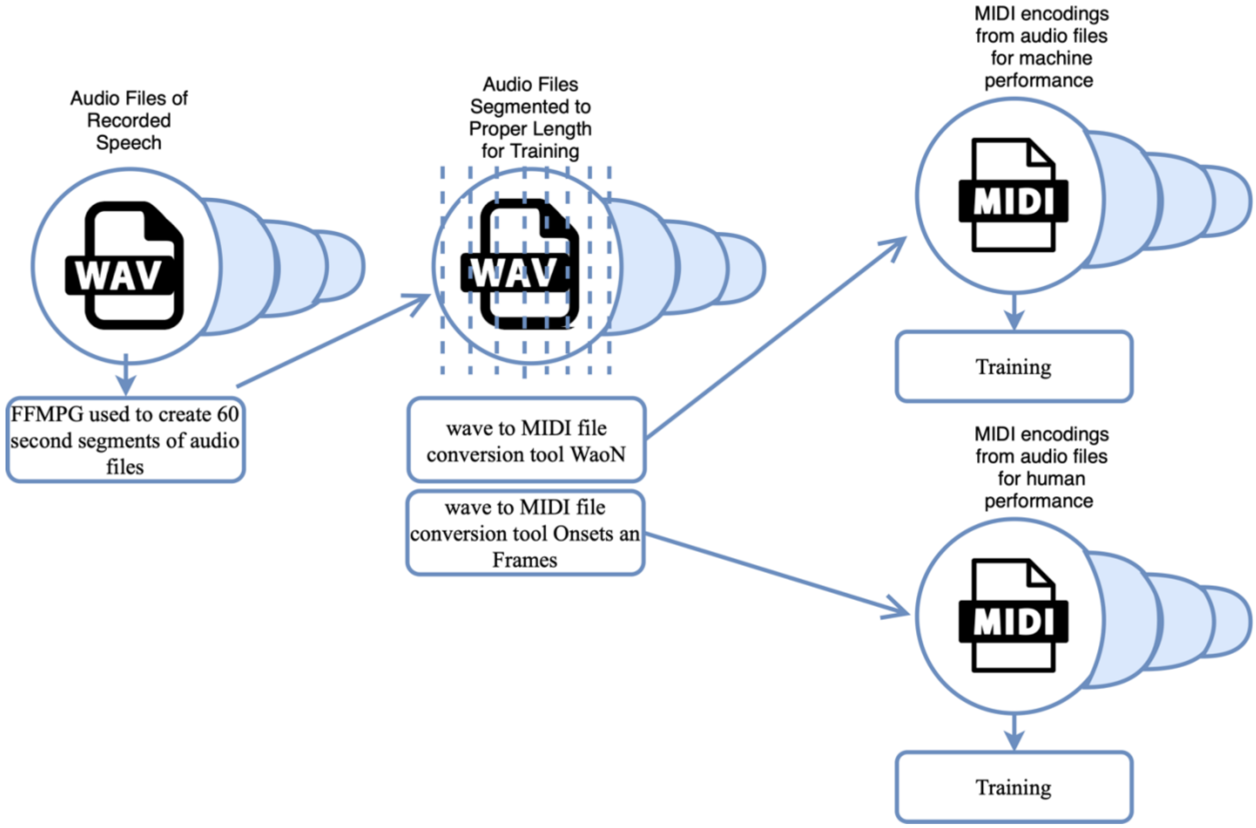


Figure 3. Speech models for both human and machine performance are prepared in a three process. First, audio recordings that represent each model are collected. Next, the collected audio files are transformed into proper lengths for training. Finally, these audio files are converted into MIDI files as a corpus for model training.

recordings that represent each model are collected. Next, the collected audio files are transformed into proper lengths for training. A speech audio file length of 60 seconds is found to be most fruitful. Because of this, the audio command line processing library FFMPEG is used to first concatenate all audio files in a directory. Next, FFMPEG is used to cut the now concatenated audio file into many 60 second audio files. Finally, these audio files are converted into MIDI files as a corpus for model training.

Cross Synthesizing Models

Computational steering is the interactive control over a computational process through adjustment of input parameters. (Mulder 1999) Each composition is created by adjusting parameters in the python script prior to generation. The resulting generation of MIDI data is realized as the score to be performed. This compositional process allows for model based cross synthesis through computational steering. Model based cross synthesis is a term used to describe the process of having one machine learning model generate material based on material that a separate model has either been trained on or has produced in generation.

Magenta is capable of setting a primer MIDI file before each generation. When set, the model will attempt to create something similar to the primer. Normative Magenta score generation using primer MIDI files can result in the creation of new MIDI files that continue and elaborate homogeneously on primer MIDI data. The pretrained Polyphony RNN model is trained on a corpus of Bach chorales. Using this model, a primer file that contains four-part counterpoint can easily result in a homogeneous generation that includes the injected primer and the following newly generated MIDI data. Magenta's pre-trained model is very good at this. The model configures itself around features of the primer and creates new MIDI data based on what it has learned from its training process. However, Models in this compositional process are asked to generate new MIDI data based on MIDI primers that often do not resemble the MIDI corpus upon which they have trained. In this process, musical models are asked to create music based on speech primers and speech models are asked to create speech based on musical primers.

Music Inspired by Speech and Speech Inspired by Music

One of the most important techniques employed in this compositional process generates speech like material inspired by music. Additionally, this process can generate musical material that is inspired by speech. This is an example of model based cross synthesis in which a model is generating material based on the training data for a separate model. This is possible because of the way that Magenta primers behave during generation. Examination of an example of this process is beneficial. In the following example only one sequence is generated for the sake of clarity. This generation is produced by a musical model that has been primed by a MIDI file that approximates human speech. This generation calls on the bundle file, booker.mag which contains checkpoint information from the model trained on the piano playing of James Booker intended for machine performance. The primer used in this generation is taken from the corpus upon which the neutral speech model was trained. As a result, it contains MIDI data that resembles spectra of speech in a neutral mood. The generation is set to condition on primer but it is not set to inject the primer. As a result, the score shows the primer file in the generation before the newly generated material and not at the same time.

```
generate(
      bundle_name="booker.mag",
      sequence_generator=
      polyphony_sequence_generat
      or, generator_id=
      "polyphony", midi_
      filename="neutral-
      booker_4bar.mid",
      total_length_steps=64,
      condition_on_primer=True,
      inject_primer_during_
      generation=False,
      temperature=1.0,
      primer_filename="bea_
      neutral_1bar.mid")
```

The generated score in figure 4 shows a MIDI approximation of speech spectra in the first measure. In the second measure there is a lot of MIDI data that looks and sounds very similar to the speech spectra. However, simultaneously, musical elements begin to appear since the Booker model is creating all new MIDI data. The most noticeable musical features consistent with the Booker models training appear in the bass clef. Rhythmically, measures two, three and four have the characteristics of a shuffle or likely ”boogie-woogie” pattern. Measure two seems to be centered around A in the left hand while in measure three F appears several times. Melodically there seems to be a bass line like structure with some often repeated pitches. These features are consistent with the Booker model. Harmonic structures also appear in the final three measures as thirds and fourths are much more common than in the speech spectra of the first measure. In some instances entire triads are present which are very consistent with expected Booker-like content. As this example progresses the presence of functional harmony is more clear. The final measure contains the following progression: E minor 7, E7 sharp 9, Ab augmented, E minor, E minor/B, and G major 7. Through the examination of this example it seems that the primer file, which contrasts the model, is most potent in influence at the beginning of the generation. Steps generated further from the site of primer injection seem to resemble model features more than those inspired by primer files.

In this next generation an example of a speech model primed by a musical MIDI file is observed. Here the bundle file, angry OnF.mag is called on to generate new MIDI data. This file contains checkpoint information that is the result

Figure 4. Speech Primer with Musical Model

of training on MIDI representations of angry speech for human performance. The primer used is a MIDI transcription of three measures of Thelonious Monk playing “Nice Work If You Can Get It.” Like the previous example, this generation is conditioned on the primer. In

Figure 5. Musical Primer with Speech Model

addition, the primer is injected into the new MIDI generation. MIDI data seems to more resemble characteristics of the model as time progresses in the generation.

```
generate(
      bundle_name="angry_OnF.
      mag",
      sequence_generator=polypho
      ny_sequence_generator,
      generator_id="polyphony",
      midi_filename="monk2angry_
      8bar.mid",
      total_length_steps=96,
      condition_on_primer=True,
      inject_primer_during_gener
      ation=True,
      temperature=1.0,
      primer_filename="monk_2bar
      .mid")
```

As can be seen in the score displayed in figure 5, the first three bars are made up by the primer. In the following three measures the primer is again present but surrounded by new MIDI data generated by the angry speech model. In the final measure the primer injection is no longer present. It’s much more difficult to understand how speech models react to music primers since speech spectra are not normally analyzed in musical notation.

In addition to conditioning and injecting music and speech onto one another, new generations can be created by one model and used as primers for other models. This cross model synthesis is explored through a new process referred to as successive primers.

Successive Primers

The use of successive primers is a new novel method of composing music with multiple primers and is an example of model based cross synthesis in which a model generates material based on MIDI data generated by another model. This is the major computational steering technique used in this compositional process. Successive primers work by sequentially looping multiple MIDI file generations and resulting in one fully realized composition. First, an initial primer file is assigned for the first model generation.

The first model generation then becomes the primer file for the second generation. The second generated MIDI file then becomes the primer for the third generation. This process is repeated as a for loop in the python script until the desired number of successive generations are completed resulting in one final composition. Each new generation is influenced by previous generations either by primer injection, primer conditioning or both. The use of successive primers allows for the use of multiple models that influence the outcome of one final generation. This process is referred to as model based cross synthesis because each generation, except the first generation, is reinterpreting data produced by another model. Additionally, the use of successive primers allows for the creation of one score that is the product of many different parameters that change throughout the generation. In normal single file generations parameters do not interpolate and are set to one condition before generation occurs. Each link in the successive generation chain can be set to a different number of measures, a different number of steps per measure, different models and different temperatures of randomness. Also, each link can be set to true or false for both conditioning on its primer and the injection of its primer. Examination of a specific example in the form of Etude III. 1712 Steam Engine (for human performer) can illustrate these principles well since its programming is more simple than other etudes.

This code generates only three sequences with one initial primer and two successive primers. The final output is one MIDI file that includes all previous steps as realized by the final model generation named ”III.human musicPrimer wativ slow18bars3.mid”.

```
   bundles = ("neutral_onF.mag",
   "booker.mag", "wativ.mag")
   filenames = ["primer.mid",
   "primer1.mid",
   "III.human_musicPrimer_wativ_slo
w18bars3.mid"]
   steps = []
   bars = [20, 8, 8]
   s_p_q = [1, 2, 2]
   totalBars  =  0
   for idx, x in enumerate(bars):
      totalBars = totalBars + x
      steps.append(totalBars *
            s_p_q[idx] * 4)
      print(totalBars)
   conditions = [True, True, True]
   injections = [True, True, False]
```

```
temperatures = [1.0, 1.0, 1.0,
   1.0, 1.0, 1.0]
primers = ["wativ_slow18bars.
   mid", "primer.mid",
   "primer1.mid"]
```

The initial generation in this short three generation composition uses the neutral speech for human performance model, is twenty measures in length, and is set to one step per quarter. The setting of one step per quarter means that the smallest possible rhythmic unit in generation is a quarter note. This is represented in the code as s p q. Because this is the first of three successive primers, this generation uses a primer file that is not generated in this sequence. The primer file, “wativ slow18bars.mid”, is an eighteen-measure piano improvisation and is one MIDI file included in the MIDI corpus used to train the William Thompson model. This generation is conditioned on the primer and injected with the primer. This is clearly seen in the score as the first eighteen measures include the primer file realized by the neutral speech for human performance model as seen in figure 6. The remaining two measures of the generation are completed by the model attempting to continue the injection with what it has learned from its training on neutral speech.

The second generation and first successive primer makes use of the James Booker model. Its primer input is the generated MIDI output from the first generation. In this second generation eight new steps are added making the total length of the generation now twenty-eight measures. The setting of two step per quarter means that the smallest possible rhythmic unit in generation is an eighth note. Again, this generation is conditioned on and injected with the primer. In measure 21 of Etude III. the material from the second primer is seen and is now reinterpreted with 8th subdivisions and James Booker model inspired embellishment. The third and final generation for Etude III. uses the William Thompson model and takes the twenty-eight measures generated in the second generation as its primer. Here eight more measures are added to generate a total output of 36 measures. This generation is conditioned on its primer but the primer is not injected. Because of this, the final eight measures of the etude are inspired by previous generations. However, previously generated material is not recounted exactly.

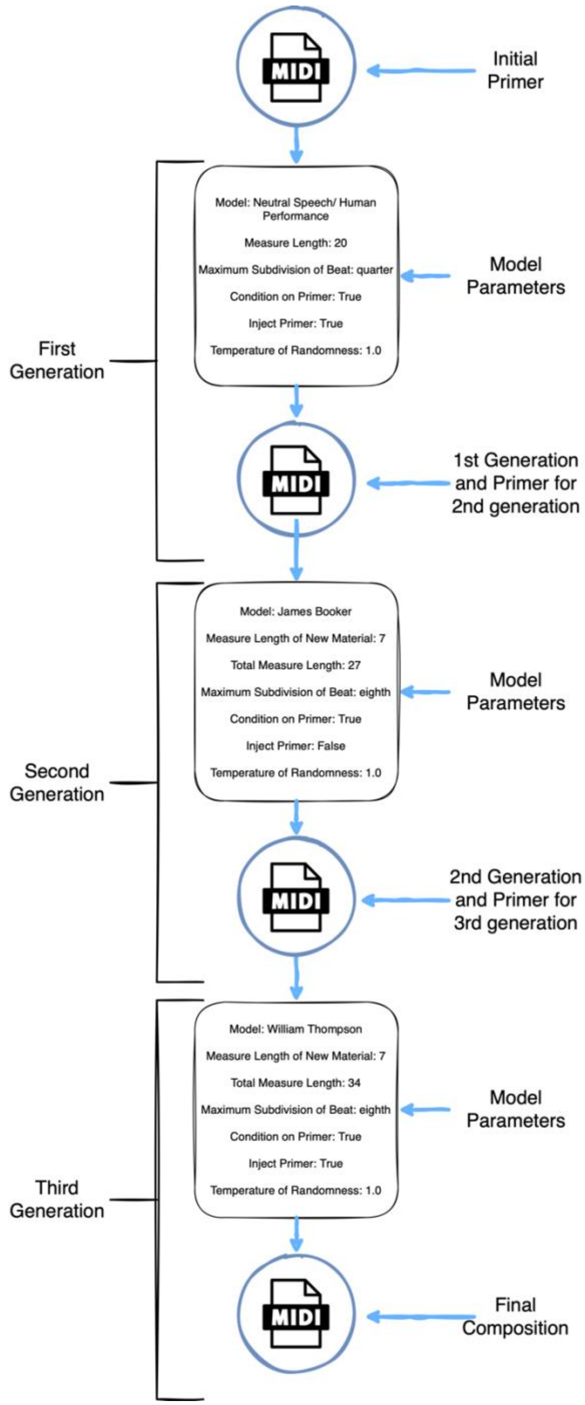


Figure 6. Successive Primers Process for Etude III.

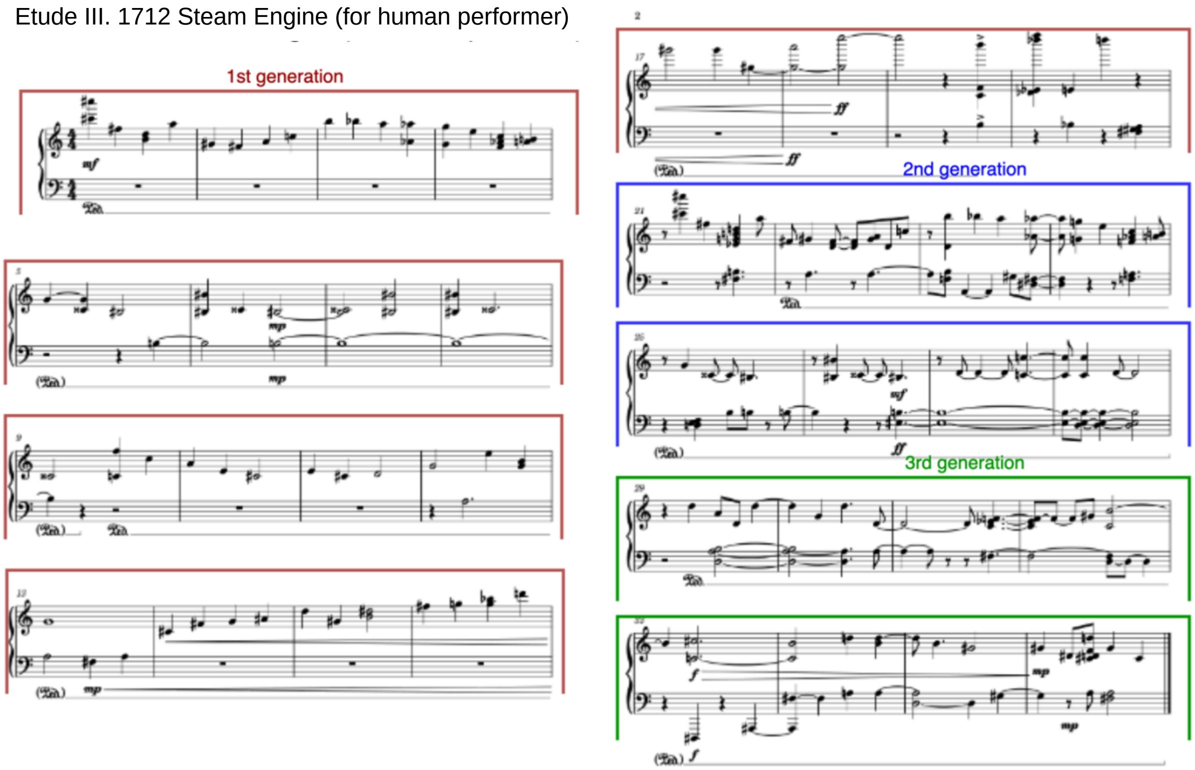


Figure 7. Etude III. Annotated Score

Results and Discussion

Seven Piano Etudes Speaks the Moody Machine, Compositional Observations

Seven Piano Etudes Speaks the Moody Machine was created via machine learning with the goal of hybridizing the musical qualities of three different emotional states of human speech with four different styles of piano music. Through this process, amused, angry or neutral speech have been melded with piano music created by Thelonious Monk, James Booker, Mara Gibson and William Thompson. These etudes are intended to be performed on one Yamaha Disklavier piano with two performers, one human pianist and one machine player piano. The Etudes that are intended to be performed by a human pianist are titled after notable events from the industrial revolution while those intended for machine performance are titled after possible future events regarding the rise of artificial intelligence as predicted by the author Ray Kurzweil. In the final etude both human and machine perform together representing the singularity. The contrast between events in the industrial revolution and future events is intended to entice listeners to reconsider each from different perspectives of time.

Seven Piano Etudes Speaks the Moody Machine

I. 1440 Type (Human Performer)
II. 2029 Claim of Consciousness (Machine Performer)
III. 1712 Steam Engine (Human Performer)
IV. 2030 Mind Upload (Machine Performer)
V. 1844 Telegraph (Human Performer)
VI. 2040 Full Immersion (Machine Performer)
VII. 2045 The Singularity (Human and Machine Performer)

Analysis

Each of these etudes are created with one generation that is the result of the successive primer process. Etudes intended for machine performance are not edited with the exception of added dynamic markings. Scores intended for human performance are edited in an effort to make them legible and physically performable by a human pianist. Dynamic markings are added by hand to both machine and human performed etudes. The entirety of the generated score can be found in appendix A. Additionally, all python scripts used to create these scores are available for review in appendix B.

Three distinct categories of generations are present in these etudes: those created for human performance, those created for machine performance and those created for simultaneous machine and human performance. A detailed examination of one etude from each of these three categories is discussed in an effort to better understand the compositional process and the musical interest created.

Etude V. 1844 Telegraph (Human Performer) is an interesting piece intended for human performance. The python script parameters that create this piece are listed here.

```
bundles = ("wativ.mag",
"neutral_OnF.mag", "duo.mag",
"wativ.mag", "monk_60sec.mag",
"amused_OnF.mag",
"monk_60sec.mag") filenames =
["primer.mid", "primer1.mid",
"primer2.mid", "primer3.mid",
"primer4.mid", "primer5.mid",
"V.human_music_wativ_drone_12bar
7.mid"]  steps = []
bars = [16, 2, 8, 12, 2, 4, 4]
s_p_q = [4, 1, 4, 8, 2, 4, 1]
totalBars  =  0
for idx, x in enumerate(bars):
totalBars = totalBars + x
steps.append(totalBars *
s_p_q[idx] * 4) print(idx)
conditions = [False, False,
False, True, False, False,
False]
injections = [False, False,
True, True, True, True, True]
temperatures = [1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0]
primers =
["wativ_drone_12bar.mid",
"primer.mid", "primer1.mid",
"primer2.mid", "primer3.mid",
"primer4.mid", “primer5.mid"]
```

The musical interest in etude V is in its use of theme and variation. Through the sequential primer process musical material is presented and then reintroduced several times in the piece. Each time material is sequentially present in the score it is altered through the manipulation of length, rhythmic scale and model interpretation. These manipulations are cumulative in the sequential primers process. If a series of sequential primers is injected five times its contents can be altered drastically since its final generation has undergone five different transformations. Additionally, each generation’s newly added material becomes part of the injection and conditioning for the next generation. Etude V. has some great examples of this. As indicated in the code there are five total primer injections in this piece. The initial generation is the injection of the first primer, wativ drone 12bar.mid. This primer is a piano improvisation by William Thompson. The next injection occurs at measure nineteen. Here the model being primed is one trained on neutral speech. Because the model has been asked to produce MIDI data with the largest subdivision of sixteenth notes (s p q = 4), the reinterpretation of the initial generation is injected with a new rhythmic scale with double the subdivision level. The original primer content is realized with sixteenth note possibilities. Here the model is not conditioning the generation on its primer. Therefore, it is generating new material around the primer. Things start to get interesting in the next four generations. Each subsequent injection uses newly generated material which was produced in past generations as its primer. As a result the fifth and final generation is linked to the first injection. However, it does not resemble it. Instead it is the product of distant variation. An examination of injection tracing can be seen in figure 8. Parts of the score circled indicate material from one generation to another. Yellow indicates material that first appeared in the first injection. Blue indicates the second injection while green and purple represent four and five. It’s interesting to examine how material changes from one generation to the next.

Etude IV. 2030 Mind Upload (Machine Performer) takes its title from author Kurzweil’s book The Singularity Is Near. (Kurzweil 2005) According to Kurzweil’s prediction in the year 2030 humans will perfect mind uploading which will make human minds software based. The composition is meant to capture this idea by shifting through the human experience of speech at inhuman speeds. Etude IV. is diverse in density while homogeneous in texture through its use of successive primer generation. The following are the parameters that create this score.

```
bundles = ("wativ.mag",
"angry.mag", "booker.mag",
"wativ.mag",
"duo.mag","wativ.mag",
"duo.mag")
filenames = ["primer.mid",
"primer1.mid", "primer2.mid",
"primer3.mid", "primer4.mid",
"primer5.mid",
"IV.machine_speechPrimer_sam_ang
ry_7bar_200bpm2.mid"] steps = []
bars = [16, 4, 4, 4, 4, 4, 4]
s_p_q = [4, 32, 1, 16, 2, 32, 1]
totalBars  =  0
   for idx, x in
enumerate(bars): totalBars =
totalBars + x
steps.append(totalBars *
s_p_q[idx] * 4) print(idx)
   conditions = [False, False,
False, False, False, False,
False]
   injections = [True, False,
True, False, True, False, False]
   temperatures = [1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0]
   primers =
["sam_angry_7bar_200bpm.mid",
"primer.mid", "primer1.mid",
"primer2.mid", "primer3.mid",
"primer4.mid",
   "primer5.mid"]
```

Etude IV. makes good use of the self playing piano since its performance would not be feasible by a human pianist. This is due to its extremely dense sections and the presence of 128th notes at a tempo of two hundred beats per minute. Etude IV. does not condition any generations based on a primer. Instead, this etude creates musical interest by contrasting sections that result from injection points and rhythmic density choices set

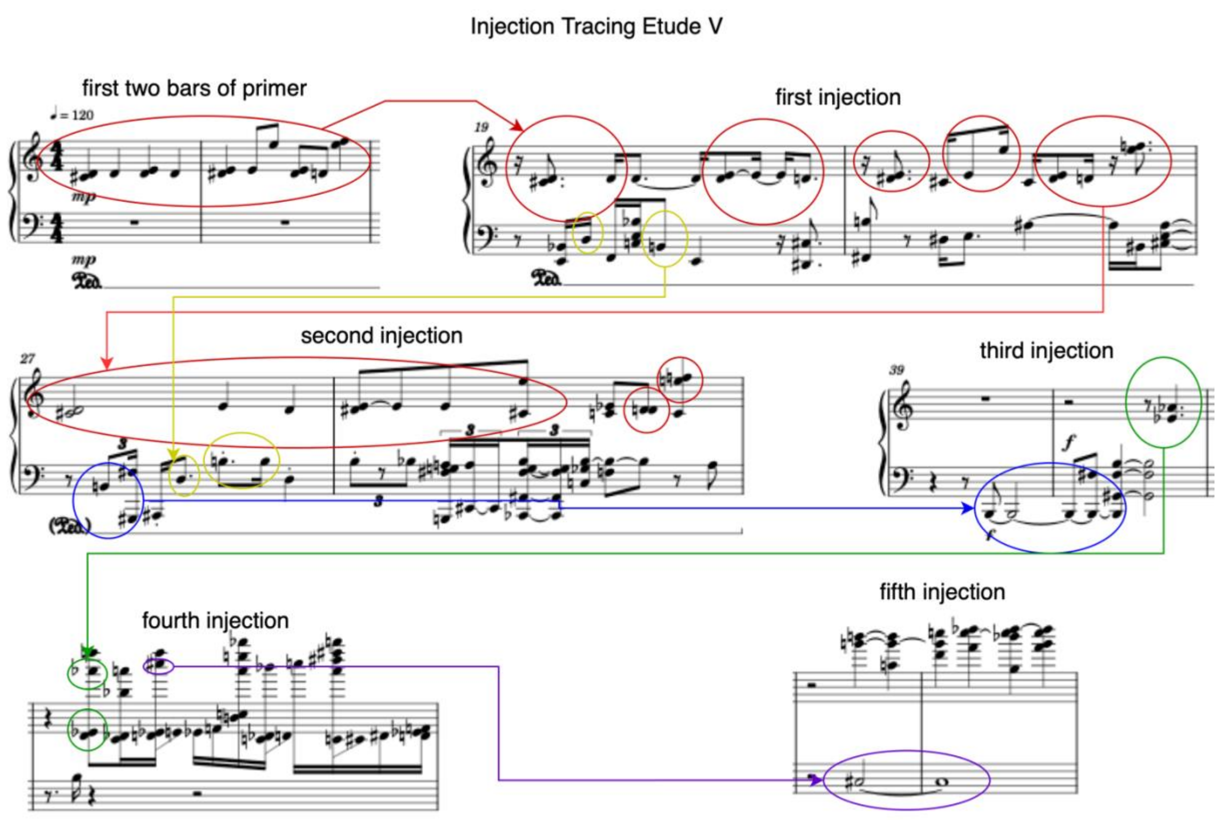


Figure 8. Parts of the score circled indicate material generation to the next.

before generation. Etude IV. is much more spectral in its imitation of speech than other etudes in this collection. As a result it has very little harmonically identifiable moments. Even though only one successive primer generation uses an angry speech model. Most of the piece sounds like angry speech. This is accomplished through a series of short generations and injections of speech like material. The first generation is sixteen measures in length. The remaining six generations are all four measures in length. This regularity in sequential primer length allows for greater clarity in the articulation of parameters for each generation. The first generation is injected with a MIDI file that is a part of the angry speech model. It is employed because its spectral MIDI representation is an effective example of angry speech. Additionally, the first sixteen measure generation is created

Figure 9. Etude IV. Angry Speech Spectra Primer

using the William Thompson model. However, it is difficult to distinguish this model's presence aurally or visually since the primer injection takes up all but two measures of the sequence. Figure 9 shows the speech primer in score representation.

The initial injection differs from the primer in pitch content only by addition. In measure eight the same pitch structure is present but surrounded by new pitches. As a result the injection sounds like angry speech and even more dense. The second generation does not inject the primer. However, it does retain the spectral qualities of angry speech by calling on the angry speech model to generate. Here things get interesting from a rhythmic standpoint. The steps per quarter note is set to 32, meaning that in this generation subdivisions as small as a 128th are possible. For this reason, several dectuplets are present. Some measures, such as measure nineteen, are so dense that one measure takes four systems to be realized in the score. This section is effective in created a fast and angry speech imitation. To contrast this, the third successive primer generation is set so that its smallest possible rhythmic subdivision is a quarter note. Interestingly, this generation also injects the former generation. In this sequence the initially injected primer is seen now in dense cluster chords at the rate of four to a measure. This does not sound like speech. However, it is easy to see and hear that it uses speech spectra as a source to create musical interest. The fact that the James Booker model is used is almost lost with the combination of rhythmic restriction and primer injection. The fourth generation at measure 25 is the first segment that sound as if it is based on any type of tonal harmony. This is because the William Thompson model is generating without injection. However, four measures later the speech primer is injected again using the D(u)o model at the eight-note rate. Finally, in the sixth generation the finale is presented as the William Thompson model generates a very dense sequence that allows for 32 steps per quarter note. To compose an ending the final generation uses the D(u)o model at the quarter note rate. One of the more interesting aspects to observe in the score of Etude IV. is the injection of the same primer material with different models and differing rhythmic density.

Figure 10 compares each injection in Etude IV. Etude VII. 2045 The Singularity (Human and Machine Performer) is unique among these pieces. Programmatically, it represents the

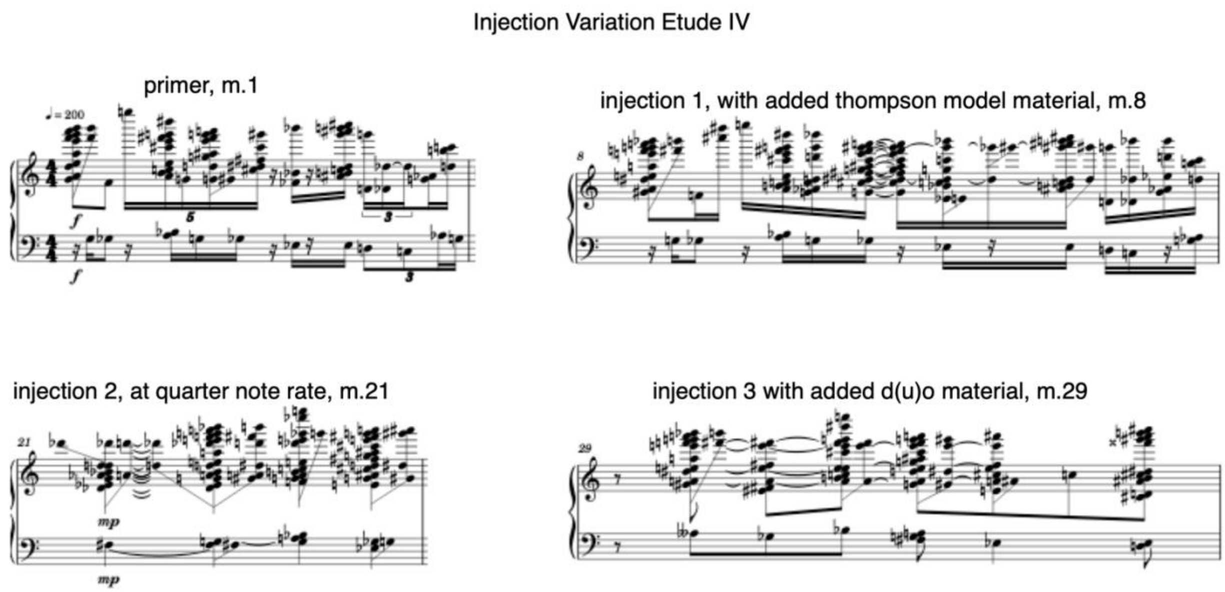


Figure 10. Etude IV. Injection Comparison

Singularity: a predicted moment in history in which technological growth becomes uncontrollable resulting in unimaginable changes for humankind. It is created to be performed by a Disklavier Piano and a human pianist simultaneously on the same piano. Because of this there could be no registrar overlap between the keys being pressed by the pianist and the keys activated by the machine. This possibility was realized by injecting the primer. In performance the pianist only plays these injected parts while the Disklavier performs only the model generated parts. Both parts were created from a group of sequential primers. The machine performed parts never overlap the injection. Because of the repetitive nature of the pianist's part, all the ferocity of this final etude is machine driven. The density of the machine part is intended to represent the predicted lack of human control in the idea of the singularity. The following is the code that creates this final etude.

```
bundles = ("angry.mag",
"angry.mag", "angry,mag",
"wativ.mag", "wativ.mag",
"neutral.mag", "wativ.mag",
"booker.mag" "duo.mag",
"wativ.mag")
filenames = ["primer.mid",
"primer1.mid", "primer2.mid",
"primer3.mid", "primer4.mid",
"primer5.mid", "primer6.mid",
"primer7.mid", "primer8.mid"
"VIII.machine_music_wativ_img_24
40_9.mid"]
steps = []
bars = [7, 5, 7, 1, 8, 4, 1, 4,
2, 3]
s_p_q = [4, 12, 32, 32, 32, 6,
6, 2, 1, 4]
totalBars  =  0
for idx, x in enumerate(bars):
totalBars = totalBars + x
steps.append(totalBars *
s_p_q[idx] * 4) print(totalBars)
conditions = [False, False,
False, True, True, False, True,
False, True, False, True]
injections = [True, True, True,
True, False, True, True, True,
False, True, False]
temperatures = [1.0, 1.0, 1.0,
1.0, 1.0, 1.4, 1.0, 1.0,
1.0, 1.0]
primers = ["wativ_img_2440.mid",
"primer.mid", "primer1.mid",
"primer2.mid", "primer3.mid",
"primer4.mid",
"primer5.mid", "primer6.mid",
"primer7.mid", "primer8.mid"]
```

In total, Etude VII makes use of ten generations as sequential primers and one final phrase which was not composed through machine learning. The generated phrases in this etude change swiftly and are odd lengths. As a result, listening to this etude or reading the score is easiest by following the original primer's injection. In this case the primer is a piano improvisation by William A Thompson which continues to surface in the composition for various lengths and with varying generated content. The first generation injects the primer. Because the primer's length is the same length as the generation, the initial generation is simply the primer with no additional material. Therefore, the first seven measures of the piece are performed by the pianist alone.

Figure 11. The primer is a piano improvisation by William A. Thompson which continues to surface in the composition for various lengths and with varying generated content.

As the piece continues the machine performer adds more and more cacophonous material which at times effectively hides the human performed theme. The second generation once again injects the primer. However, this time the pianist's hands are surrounded by MIDI data generated by the William Thompson model at a sixteenth note rate. Because this sequence is only five measures the theme is left incomplete before it is once again injected in the third sequence. This time the theme is complete. However, it is now surrounded by a lot of new material generated by the angry speech model.

Sequence four begins the theme again but for only one measure when it is disrupted by the presence of sequence five. This generation is the only significant sequence that does not inject the primer. As a result, the William Thompson model generates freely while the human pianist rests for eight bars. This sequence is particularly interesting since the tonal nature of the William Thompson model is at times very recognizable. Additionally, this sequence is very fast with potential subdivisions of 128th notes allowed by the code setting of 32 steps per quarter.

Generations six, seven and eight once again inject the primer. In generation six the theme is present in its entirety with obvious speech generated material from the neutral speech model. However, this speech is not easily aurally analyzed as neutral because of the speed and micro subdivisions. In sequence seven only one bar of the theme appears before sequence eight presents the theme for one final time with the James Booker model. The Booker model's presence is recognizable by the left hand broken boogie-woogie pattern.

The ninth generation does not inject the primer and is performed by the player piano. Its contents resemble the D(u)o model at the quarter note rate. The final seven measures of the piece

Figure 12. The beginning of the theme seen in Figure 11 surrounded by angry speech model material.

Figure 13. The Thompson model sounds tonal at times here and very fast.

Figure 14. The Booker model is recognizable by the left-hand broken boogie-woogie pattern.

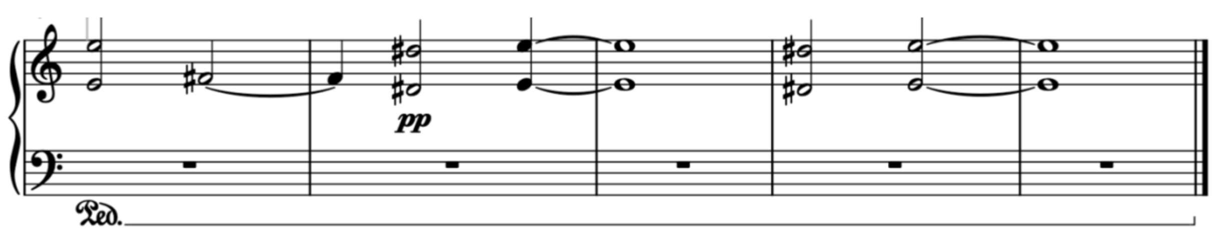


Figure 15. The final four measures of the piece are composed by hand and without the aid of machine learning. The right hand is performed by the pianist and the left hand is performed by the machine performer.

and the etudes as a whole was composed by hand and without the aid of machine learning. The right hand is performed by the pianist and the left hand is performed by the machine performer. It is a portion of the original theme.

Conclusion

This union of speech and music is justification for new compositional practices that exploit this idea. The compositional goal presented in this research attempts to make meaningful connections regarding human expression by finding a sonically interesting common ground between speech and music.

The creation of the piano etudes that resulted from this research do make for an interesting merging of music and speech. Speech-like symbolic data in the form of MIDI files does make sense as musical content because it can be perceived as and inspired by music. Additionally, music that is inspired by speech does sound musical while retaining speech-like qualities. In both cases the term “inspiration” is meant to reference the compositional processes of model based cross synthesis. In this process primers “inspire” a model to elaborate on foreign MIDI data through the lens of its own training data. MIDI data is not intended to imitate speech. In spite of this, MIDI does effectively capture speech spectra in a manner that is practical for musical score representation.

In many situations, composers draw inspiration from unknown resources. This is not entirely different from the processes outlined in this compositional method. Through machine learning composers can employ models that replicate the musical features desired by the composer. In spite of this, the training process occurs with little to no knowledge of what a specific machine learning model values and is learning from. On the other hand, what is unknown to the composer does not necessarily hinder good musical composition. Machine learning models are capable of learning and creating new musical grammar that a human might never perceive. This is difficult to discuss since we cannot discuss what we don’t know. However, complete understanding is not needed if machine learning results are interesting from a musical perspective. These techniques are a means and not an end.

Compositional tools that can access illusive musical features are especially advantageous when attempting to imitate the idiosyncratic emotional expression found in speech. Machine learning models created in this compositional process are effective in producing MIDI generations that resemble various emotional states of speech. Emotional states of speech such as anger, amusement and neutrality can be very complex in description. However difficult they may be to define, they are often easily discernible by the human ear. Human listening is by far the most effective way to evaluate models imitating emotional speech.

The challenge facing the artist utilizing this type of creative process is in the evaluation of generational output. There is a sweet spot that exists somewhere between models that generate new material that is almost identical to its training data and models that generate new material that does not resemble its training data at all. This is where interesting art can be found. Machines allow humans to create art that is not possible without their aid. However, human perception is the most effective measure of machine success.

Supplementary information

Audio Video of composition Seven Piano Etudes Speaks the Moody Machine: https://youtu.be/JCt6ugGv2SQ Included in the video is the score. A copy of the score in other formats is available by request.

References

Adigwe, Adaeze, Noé Tits, Kevin El Haddad, Sarah Ostadabbas, and Thierry Dutoit. 2018. "The emotional voices database: Towards controlling the emotion dimension in voice generation systems." *arXiv preprint arXiv:1806.09514.*

Barrett, G. Douglas. 2009. "Between noise and language: The sound installations and music of Peter Ablinger." *Mosaic: A journal for the interdisciplinary study of literature*: 147-164.

DuBreuil, A. 2020. "Hands-on music generation with Magenta (E-Book)." *Explore the Role of Deep Learning in Music Generation and Assisted Music Composition. Packt, Birmingham*.

Farrar, Frederic William. 1865. *Chapters on language*. Longmans, Green.

Hawthorne, Curtis, Erich Elsen, Jialin Song, Adam Roberts, Ian Simon, Colin Raffel, Jesse Engel, Sageev Oore, and Douglas Eck. 2017. "Onsets and frames: Dual-objective piano transcription." *arXiv preprint arXiv:1710.11153*.

Ichiki, Kengo. *n.d*. “WaoN, a Wave-to-Notes Transcriber.” Last modified October 6, 2018. https://sourceforge.net/projects/waon/.

Kapur, Ajay. 2005. "A history of robotic musical instruments." In *ICMC*, vol. 10, no. 1.88: 4599.

Kurzweil, Ray. 2005. "The singularity is near." In *Ethics and emerging technologies*: 393-406. London: Palgrave Macmillan UK.

Lei, Xiaohua, Xiuhua Jiang, and Caihong Wang. 2013. "Design and implementation of a real-time video stream analysis system based on FFMPEG." In *2013 Fourth World Congress on Software Engineering*: 212-216. IEEE.

Levman, Bryan G. 2000. "Western theories of music origin, historical and modern." *Musicae scientiae* 4, no. 2: 185-211.

Mulder, Jurriaan D., Jarke J. Van Wijk, and Robert Van Liere. 1999. "A survey of computational steering environments." *Future generation computer systems* 15, no. 1: 119-129.

Patel, Aniruddh D., and Joseph R. Daniele. 2003. "An empirical comparison of rhythm in language and music." *Cognition* 87, no. 1: B35-B45.

Patel, Aniruddh D., John R. Iversen, and Jason C. Rosenberg. 2006. "Comparing the rhythm and melody of speech and music: The case of British English and French." *The Journal of the Acoustical Society of America* 119, no. 5: 3034-3047.

Electroacoustic Practices of Radical Play and Resonance: Stuart Dempster's Use of Laughter, Illusion, and Environmental Homage

Elizabeth Hoffman

Department of Music, Arts, and Science
New York University
Elizabeth.Hoffman@nyu.edu

Elizabeth Frickey

Department of Music, Arts, and Science
New York University
Elizabeth.Frickey@nyu.edu

Abstract

Trombonist and educator Stuart Dempster, founding member of *Deep Listening Band* and a lifelong collaborator with Pauline Oliveros and others, has for more than fifty years now cultivated unique and profoundly valuable interventions to music making, to musical values, and to a humanistic approach to technology and education. This article focuses on his use of computer processing beginning in the 1980s to enable improvisation, composition, and new ways of listening to oneself and to one's environment. His use of the *EIS* (Expanded Instrument System) developed by Peter Ward and David Gamper, was distinctive towards this end. His *State of Washington as a Musical Instrument* (SWAMI) project was also remarkable— in itself and in its impact on Dempster's electroacoustic practices, and vice versa. Attention is paid here to details in the digital production process manifest on the album *In the Great Abbey of Clement VI* (1979). Notions of democracy, an embrace of the environment, and spiritual sound are explored through recent conversation with Dempster. As a tool that fostered the cultivation of illusion and augmented reality, electroacoustic design has been a natural part of Dempster's radical play throughout his career.

Introduction

Existing roughly contemporaneously with other, larger post-war radical musicking movements of the mid- to late-1960s (like the *AACM*, for example), assorted U.S. musicians whose artistic radicalisms were less institutionally articulated nevertheless also contributed to the shifting educational values of the time. They were smaller cohorts or even individuals whose activities, manifesting provocative attitudes towards institutionalized education or existing musical practices, may be harder to categorize and contextualize historically. Yet, they still have been heavily contributory in their own ways toward turning traditional ways of confining musical practice on its head. Of particular interest for this article is the connection of non-traditional practices and electroacoustic sound.

Stuart Dempster, trombonist, self-described "sound gatherer," and extended electronic instrument explorer has been one such practitioner. He has contributed in idiosyncratic ways to disrupting normative ways of listening, composing, performing, practicing, and teaching. His earliest experimental musicking includes recording sessions with rich digital transformations of his own acoustic playing. While he may have been "dragged kicking (not too much screaming) into [his relationship with computers]," he quickly became an intuitive user of real-time computer systems (Stuart Dempster, Personal correspondence with authors, August 26, 2023).

Dempster is the recipient of numerous experimental landmark pieces written for him by composers including Ernst Krenek [*Five Pieces for Trombone and Piano* (1967)], Luciano Berio

[*Sequenza V* (1966)], and Robert Erickson [*General Speech* (1998)]. But his experimentation has involved far more than shunned notes or extended techniques. Crucial to his pushing the confines of music has been his infusion of joy, humor, play, improvisation, risk, confusion into situations that typically run from these qualities. And he used any technologies at hand—whether it be Meyer Sound's ahead-of-its-time parametric *SIM* (*Source Independent Measurement*) system, or, alternatively, a massive underground cistern in Washington state. In terms of methodology, rubber chickens and fart whistles were in Dempster's back pocket wherever he went as he challenged propriety and seriousness almost as if he could not help himself, and quite often with disarming glee.

However, Dempster is also as serious and seriously skilled a musician as they come. Dempster knew and respected the classical literature; in fact, he played briefly in the Oakland Symphony, under Gerhard Samuel. He founded *Deep Listening Band* in 1988 with Pauline Oliveros and Peter Ward, a.k.a. Panaiotis; the *Band* welcomed David Gamper two years later as an assistant to Panaiotis, and about 16 months later Gamper became a performer as well. This article strives to demonstrate that this collective activity was only the tip of the iceberg of Dempster's radical musical activity. Dempster has prompted generations of students and audiences to make music that is primarily drastic, rather than gnostic.[1] Even more uniquely, he has carried out musicking that promotes the notion of "laughter as the world's most powerful change agent" (Wilson 2004, 137). This article asserts that his use of illusion and obfuscation of the real has been intimately tied to his experimentation with the electroacoustic medium, both through augmented reality (namely, his compositional techniques of using recorded sound in a merged past-present-future performer experience flow) and through repeated listening to the past (new listening practices, especially of focusing on study of environmental sounds, including sounds of the city and sounds of remote and undisturbed natural locations).

Of all *Deep Listening Band*'s members' activities, perhaps Oliveros' *Deep Listening Institute* has been the most widely known radical educational initiative. Dempster's own personal and delightfully quirky educational pathway evolved against this backdrop, simultaneously over several decades. Through his teaching position at the University of Washington, starting in 1968, and through his recordings, he, too, celebrated and taught the practice of listening but with an ethos all his own. Most uniquely, Dempster has cultivated an element of "play"—an extension of his tendency to want to catalyze irrepressible laughter. His radical approach to education and performance through play included a constant pull toward greater "democratization," whether or not he named it as such. Audience involvement, for example, was a pinnacle of satisfaction for him.

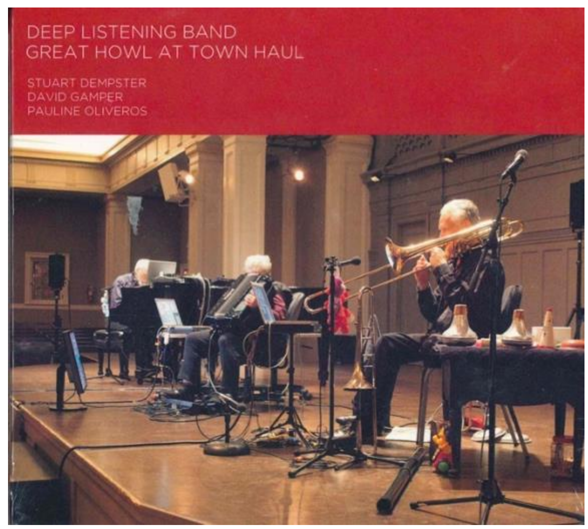


Figure 1. Deep Listening Band, *Great Howl at Town Haul*, IMPREC PO80, 2012

Electroacoustic music studies interleave with Dempster's body of work and practice in complex and important ways. The evolution of computer music (and digital arts) at the University of Washington in Seattle was itself positively impacted by the decades-long collaborative outreach of Dempster and William

[1] We allude here to Carolyn Abbate's (2004) seminal text on this topic.

O. Smith. Both of these performers were a constant presence in the School of Music Computer Center, and later, in the innovative interdisciplinary department created by Richard Karpen in 2001, DXARTS. They pushed envelopes constantly with their inventiveness and fearlessness. Here, we consider in more detail the pieces that have contributed to Dempster's radical stance on art, life, and teaching. The composite puzzle is far more than the sum of its parts. Those who have been lucky enough to study and work with Dempster know that firsthand. His vital contributions to the history of electroacoustic music convey the power of resonance as a fully holistic experience: sounds resonating across space and time, resonating amongst communities, and sometimes resonating with uproarious laughter through our individual and collectively assembled bodies.

Technology: Play and Delay, Illusion and Confusion

A key focus here is the role that electroacoustic sound has had in developing Dempster's methods of listening and composing, and, ultimately, teaching. In prior interviews, Dempster recollects the allure for him of studying sound recordings, simply as a listener—apart from using recording as a tool for creative or performance ends. He was drawn to recorded sound simply in a phenomenological sense, as a listening exercise to discover details he'd missed in real time. This predilection to study sonic detail was instilled in him at a very young age, through family outings to the Berkeley, California train station to watch and listen to trains (Dempster 2012). By the time he encountered the possibility of using technological applications to process recorded sound, he was naturally receptive. His penchant for emphasizing listening is thereby evidenced in use of extended pauses, rests, and delays in his compositions. Play with remarkably long natural delays (heard dramatically in the album *Underground Overlays from the Cistern Chapel* (released on New Albion in 1995)) was foreshadowed in his earlier project with fabricated delays, namely the third track ("Standing Waves – 78/87*")* that was added to the CD re-release of *In the Great Abbey*. It was thus a bit of life (1995, natural 45" delays in the cistern) imitating art (1978 and 1987, fabricated 40" computer created delays on fixed media at the Center for Computer Research and Acoustics at Stanford University and later recorded at Mills College) in an unplanned and casually coincident backwards progression.

The spontaneous nature of both albums attests further to Dempster's Zen-like openness. A trip to Avignon, France in 1976 with David Behrman and David Tudor, touring with Merce Cunningham, had several serendipitous and fortuitous points of inflection. The fact that he was traveling with like-minded musicians, always attuned to sound collecting, was a key element in the chance circumstances that led to the important early album, *In the Great Abbey of Clement VI* (released by 1750 Arch Records, on vinyl, in 1979 (S-1775)).

Virtual Spatialization and Reverberation and New Ways of Listening

> "This is where you have been forever and will always be forever." -- Stuart Dempster, speaking about what it feels like to be in a cistern where time seems to stop (Dempster 1994, 1).

Audio Example 1. "Standing Waves – 78/87*"*
Track 3 from *In the Great Abbey of Clement VI*, 1987 re-release.

*"*Standing Waves – 78/87" is, in many respects, the most remarkable demonstration of Dempster's interest in creating unusual spaces using electroacoustic manipulation. It is also an astounding demonstration of how he used digital sound practices to cajole himself into exploring new ways of listening, in particular, to what he has called in many interviews, "the past, the present, and the future, all at once" (Dempster 1994, 1).

In this performance, Dempster used the Lexicon 480L digital processor to create live reverberation for the solo trombone that matched the 40-second reverberation on a taped trombone choir that was being played back in the hall. The electroacoustic and digital sound developments were not trivial here. Meyer's SIM System(s) (of which many instantiations were released over many years, through to the present) are tools for acoustical analysis. The tool provides a clear, detailed picture of a sound system's performance in a space and can rapidly identify acoustic anomalies. It uses parametric equalization, and hence helps an astute engineer in real time to monitor a sensitive non-studio recording situation. The reverb is a key part of the live performance. It cannot be altered after the fact.

Performances in Avignon took place "in the Pope's Palace courtyard" (Dempster 1979, 1). By setting up equipment on another level, Behrman and Tudor discovered the unique resonance of the 14th-century abbey and brought it to Dempster's attention. As he writes in the album notes: "I played one note into that space. For fourteen seconds the note was floating suspended...It was the most amazing sound sensation I had ever experienced up to that time" (Dempster 1979, 1). The tour technician, John Fulleman, moved his portable recorder into the abbey and thus improvisationally initiated a days-long side project–Dempster recording, and then taking the tapes back to the hotel for a group listening. As such, the materials that became "the piece" were improvisatory in nature.

Whether the extended reverb and the equalization in the space was "natural" (as in the 1976 Abbey recording) or, whether it was computer generated (as in the 1987 Mills College recording), made little difference in terms of the listening challenges. The experience of listening to distorted time and space was something that Dempster learned to do quickly, and thereby spurred a unique approach to immersive, resonant spatialization seen through many of his later works, and even in his own teaching methodology. Dempster's statement in the album notes regarding the fluidity between artistic creation and teaching insights bears witness to this close symbiosis between experimentation and teaching:

> There is no score per se but the material is based on my 15 years of experience with resonance. Indeed, I occasionally use resonance as a teaching device with my trombone students to teach intonation and tonal center; a resonant space can point to faults in these areas quickly and clearly (Dempster 1979, 1).

As author Jennie Gottschalk says:

> The quality of sound is very distinct from both the Niblock and DeMarinis examples [Niblock's *The Movement of People Working* (2003) and DeMarinis' *Rain Dance* (1998)], but it is another way of enveloping the listener through the use of a reverberant acoustic. Dempster multiplies the forces in the 1978/87 version of this piece through layering of multiple tape parts, creating beatings and layering with the live performance that offer a different, more fluid form of submergence in the sound. This time the performance is within a concert hall, and is not as [naturally] reverberant as the chapel, but the five-channel sound system increases the sound mass significantly (Gottschalk 2016, 114-115).

The fact that "Standing Waves – 78/87" was recorded (with omnidirectional mics, mixed to a stereo format) in a space with an installed a five-channel system is worth noting. It is a landmark example of early multi-channel experimentation—for a project in which the multi-channel design was likely in service of greatest spectral clarity for the multiple sound layers through locational separation, rather than in service of maximally detailed gesture simulation. The nature of this project, and the specifics of its realization all demonstrate an acute sensitivity to sound itself in interaction with an acoustic space, and to sound as represented on a recording. It is also very significant, therefore, that the *Expanded Instrument System* (discussed in more detail below) stipulated the presence, always, of at least four speakers.

Before exploring the *EIS* more fully, it is also worth noting some approaches to education implicit in the practices of the often-touring *Deep Listening Band*. The "re-listening" *to-discover-more* process that Dempster (and his production team) discovered in the abbey, and that he relished as a younger musician, is not a normal prong in our formulaic music study in the United States. More frequently—despite postmodernism's post-authorial turn—we are urged to identify what is allegedly there. This distinction between our inner worlds and our outer world additionally continues to be immensely important in educational philosophies: Do we want to promote a shared and standardized hearing of a work of music? Or, do we wish to find ways to acknowledge and validate an ever-shifting allusiveness and elusiveness of the sounding object? The nature of Dempster's musical projects, especially the ones using computer collaboration, are rich in textural and timbral resonances. They effect a meeting of the acoustic and the electroacoustic in a seamless and elegant conversation. Akin to La Monte Young's *Dream House* installation, "standing waves" (which theoretically occurred as an acoustic phenomenon during the 1976 Great Abbey recordings) create many perceptual ambiguities, as well as bodily interactive phenomena. Turning one's head, for example, will simultaneously eclipse and reveal certain frequencies. "'There *is* no perceptual object!"' one might imagine Dempster exclaiming.

The Expanded Instrument System

The *Expanded Instrument System* was profoundly important to Dempster's music making for decades. It was, essentially, an active, participatory approach to using tape delays via a small but tightly knit network of expert musicians developing a community-oriented approach to improvisational experimentation with sound. The focus of their experimentation was on timbre, and on a sort of sculptural invocation of sound in virtual space(s). All creators were acutely tuned into distortions of time and space and place. While they never used the term 'futurism' to describe their aesthetic and their conceptual alliances, it is a label that seems apt. In fact, from a musicological stance, Dempster's cultivation of a floating ambient world on his "Standing Waves – 1976*"* preceded Brian Eno's release of *Ambient 1: Music for Airports* by two years.

The *Expanded Instrument System* was originally developed by Peter Ward in 1983; it was then continued in development by David Gamper roughly two years later. The *EIS* in its initial instantiation (in the early 1980s even before the formation of the Deep Listening Band) precluded the players' control over their own signal processing functions. These new 'limitations' became integrated into Dempster's psyche from the beginning, perhaps transformed into repurposed liberatory qualities:

> The *EIS* is an extended instrument that is manipulated by several people simultaneously, the notion of community performance is heightened by the fact that a performer has only partial control over the outcome of what she/he produces with an instrument…With extended delays a performer is constantly listening to something that was created in the past while performing in the present something that will come back in the future (Oliveros, Panaiotis 1991, 407).

Working with such a lack of authorial control goes beyond improvisatory tropes. Or, at least, it is something notably different. Embracing this lack of control in a social communication context (which is what playing chamber music arguably is) extends into realms of anthropology and philosophy. While improvisation moves within existing tropes to allow unpredictable events to happen, the *EIS* exacerbated and celebrated a more decided liberation from control. It also promoted an acceptance of an impossible to parse self—whether a musician's sonic identity might be sought perceptibly by a listener or by a group member.

In this sense, the *EIS* exemplifies not only Dempster's experimentation with sonic resonance, but, in perhaps a more metaphorical sense, a social resonance between ensemble members, and audience members by extension. The blending of forces and mechanisms to manipulate the routing and processing ensure that a player's or listener's ability to find the fringes of a precise sonic object was, in other words, greatly attenuated by the *EIS*. The players were truly part of a "system" that was a vibrantly intertwined sound mass in which a mirror held up reflected only a composite whole. No one could actually see her/him/themselves. As space itself became a "dynamic parameter," shifting through the course of a performance, it became even harder to maintain a background grid reference that might clarify the location of objects.

Vibrant and Agential Matter

As Dempster describes in his notes to the *Underground Overlays from the Cistern Chapel*, 'The reverberation time of 45 seconds is so great that it is nearly impossible to communicate unless you are grouped together (Dempster 1994a, 1). The overlays in the cistern (which grew from a masterclass that Dempster held there for his students) were arguably on an interesting continuum with the techniques he used in playing with the *EIS*. It is notable that the merger of individual layers (in the cistern it was ten layers, i.e., ten players) contributed to qualities that transcend the ambient. According to Dempster's liner notes for *Underground Overlays,* it was the agential act of the cistern, "acting as its own audience in sustaining [the sounds]," along with the extreme slowness of events (the timing of which was conducted by Dempster, improvisatorially) that created a spiritual aura. It was, in other words, something more than ambient in its more typical definitions. The merger of the players was a large part of the process and the sound result. This is audible if one listens carefully to the texture. Perhaps the term "texxture"—especially as invoked by Caroline Miller in a musical context—is an apt descriptor here (Miller 2018, 5). "It was nearly impossible to distinguish direct from reflected sound" (Oliveros 1995, 22).

Dempster's use of electroacoustic tools early on seem to have cemented his appreciation for such vast and ethereal potentials for disembodiment, since spatial illusions enabled by the EIS and (conversely) its initial technical limitations made reality blurry and made imaginings real. As Pauline Oliveros reflected further on this life-long passion that she shared with Stuart Dempster:

> Stu and I have both worked with reverb and echo effects a lot, in my case mostly with electronic delays. I started twenty years ago working with Echoplex-type delay loops... And, in recent years, MXR… It gives you up to 3.5 seconds of delay, and it's a completely clean sound. And, so, working with the live performers actually for me was very much like working with my digital delay… quite a few of my friends and students thought [as much]. (Meyer 1986, 197)

The Resonant Performer-Educator

To piece together other parts of the puzzle that is Dempster's multi-threaded radicalism, a glance at least at his trombone scholarship is vital. Dempster's 1979 text, republished in 1994, *The Modern Trombone: A Definition of Its Idioms,* is another marker of his care and concern for education. Being a trombonist in particular—and having to find partials as a primary playing technique—was not an incidental contribution to his development as a radical listener. Dempster, as a composer and improvisor, sought to create musical textures and harmonies which more closely aligned with acoustic principles, e.g., just intonation. He talks, in fact, about the vibrancy of "uplifting" musics:

> [In equal temperament,] resonance has gone missing for the most part. Equal temperament is not especially 'resonant' but is its own thing. There is some value in equal temperament, but orchestra musicians have to make significant compromises. Generally, the orchestra overall plays in equal temperament while each section of the orchestra plays in just tuning. Could it be that this is why orchestral players are a bit crazy? In the trombone section, the first and third players tune easily in a given chord, but the second player has to figure out what tuning to use to have the chord be resonant, and that is not easy (I speak from experience). (Stuart Dempster, email correspondence with authors post interview, March 9, 2023)

For Dempster, an understanding of resonance is rooted first and foremost in this practice of performing as an ensemble musician. To create a resonant sound, therefore, requires careful listening and collaboration with fellow players. In his playful and therapeutic *Sound Massage Parlour* (1986), Dempster takes this practice even a step further in order to soothe physical and psychic aches and pains of his listeners. However, these somatic benefits of resonance do not come without careful consideration:

> Oh, my, indeed, resonance does provide healing and therapeutic elements. There are some caveats, however. One has to play absolutely in tune. It is easy to begin the pitch in an uncentered way and that would *not* be therapeutic. There are so many ingredients to keep track of; however, the sonic delights come quickly when one realizes that (a) the tonal center has to be super clean when starting a pitch, and (b) the overall tuning has to remain true throughout. Once that is grounded, the performer has to honor the entire sonic spectrum, particularly the 7th partial that few people are used to using. However, once a performer tastes that particular partial the ones further up (and down) will begin to settle in…One good thing about trombone is that there is no excuse to play out of tune with perfect sonic resonance. (Stuart Dempster, email correspondence with authors, March 9, 2023)

Dempster's process of commissioning and collaborating with a range of out-of-the-box creators included teacher and friend, composer Robert Erickson. Erickson, in particular, enabled Dempster's engagement with Aboriginal didjeridu playing (and its ethics) which flowed naturally into Dempster's already heightened attention to tuning's alleged impact on musical affect, including experiential optimism and spirituality.[2] As Dempster says about this instrument's unique tuning and timbre, and the possibly still radical idea (in the 1980s) of music's healing capabilities:

> Generally, the therapeutic output comes from players of both [trombone and didjeridu] who know what to do about that. It is matter of playing *to* your audience and not *at* your audience. The best on either instrument is for the performer to love the audience and bring them into the experience. (Stuart Dempster, in interview with authors, March 9, 2023)

This quote is a testament, too, to the reasons that electroacoustic performances which create illusions of "who," "what," "where," "when," and "how," have always appealed to Dempster. Enticement to active engagement by posing a perceptual and cognitive question, as well as by enveloping a listener in literal full-bodied sensorial vibration, often serves as the performance substrate. The result is an audience listening with much more acute sensitivity; their anticipation for clues to these answers is on a pleasurably heightened alert.

Listening to Time[less]-Space, and Acoustical Accidents

As discussed above, Dempster's experimentation with resonance led him easily towards an environmental (and geographically expansive) compositional methodology; his ears were always open, always working, always listening. His attention was never oriented towards or restricted by culturally defined listening spaces. But, responding to the difference between mundane resonance, and the resonance of hidden or unusual spaces was a serendipitous turning point for Dempster that merged these two conceptual foci of play with computer processing and play in and with environmental spaces. Technological experimentation through the Expanded Instrument System ultimately addressed the stretching and manipulation of time-space, which is, in effect, resonance. Delay effects are a de facto playing with time—an exploration of past, present, and future simultaneously. Dempster's deliberate exploration of acoustic resonance also ripples continuously across time and space, from the natural acoustic environment of The Great Abbey in 1976, to the pre-designed and fabricated reverb generated eleven years later, used in the "Standing Waves – 78/87" track. Time and space distortion pulls Stuart Dempster back, this time to a truly cavernous space of an underground cistern

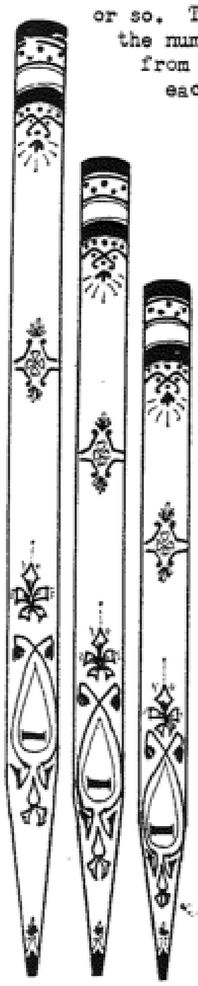
[2] In 1973, Dempster spent one month in the Northern Territory of Australia visiting two camps with Aborginal didjeridu players supported by a Fulbright-Hays Scholar Award.

Stuart Dempster SWAMI page 6

PORT TOWNSEND - on the Quimper Peninsula in Jefferson County

Port Townsend contains a major acoustic site complex for the state of Washington in FORT WORDEN STATE PARK, discussed elsewhere in this document. In the Fort there is the cistern, the "crown jewel" of acoustic phenomena. Besides Fort Worden there are two other locations in Port Townsend that warrant special mention: The Clock Shop in the Port Townsend Antique Mall, and the 1889 Whalley and Genung Pipe Organ in the First Presbyterian Church.

Clock Shop - in the Port Townsend Antique Mall at 802 Washington Street, open 10:00 to 5:30 daily. The clock shop is located about one third of the way down the "mall" on the right, in a separate room. If you are patient you can be in there by yourself or with just a few understanding friends. It is essential that you be quiet while in there and just listen to the random ticking, and even random chiming, of probably 100 clocks or so. The "or so" means that it may only seem like 100 clocks, but the number is unimportant. What is important is the randomness that from time to time seems to put a great many of them in sync with each other. Treat the experience as a meditation...



1889 Whalley and Genung Pipe Organ - in the First Presbyterian Church at Franklin and Polk. One hundred years old as of this writing, this pipe organ is the oldest such in the state in its original intended setting. There are a number of things that are unique about this instrument besides its original setting. For one thing, there are few Whalley and Genung organs remaining in the country. For another, the first organist, H. Ambrose Kiehl, specified that a bass stop be added--not usually done in organs of this size. The organ has a fine, if somewhat mellow, sound due to the use of lots of wood pipes. It is very appropriate to the space. In order to hear the organ you should be aware of the organ concert series which takes place in approximately October, February, and April of each season. The resident organists have been known to feature music from the church archives, which are extensive. The church also is carrying out a major restoration project to bring the organ back to as near to new as possible. Any donatinns would be most welcome. If you wish any information about the organ, the series, or wish to make a donation to the Organ Preservation Fund, please write to the church at 1111 Franklin Street, Port Townsend, WA 98368.

Figure 2. Page 6 from *State of Washington as a Musical Instrument*, Dempster 1989.

and a limestone quarry in the two years following.

Stretching time, after all, means responding to sound in an alluring and playful manner. That Dempster returned to the cistern—a chance of collaborating with one's environment, exploring it as such, and even learning from it as a newly found teacher—became amplified in his later project *SWAMI*, the *State of Washington as a Musical Instrument* (1989), a radically attentive set of "acoustical journeys" (Dempster 1989, 1). The piece takes the form of a "preliminary self-guiding booklet," indeed, directly inspired by the cistern at Fort Worden. Throughout, Dempster gives the reader glimpses into his own field notes from his acoustic explorations across the state of Washington. In approaching Palouse Falls, for example, Dempster remarks:

> The Concert starts almost immediately! [...] The road is so much a washboard that it is like range road grating, which was right here, too. In looking in the rearview mirror I would see the back end of the car shaking so much that it was all out of focus. Slower was no help; indeed, it was often worse! I was amused by the sign(s) saying 'slow to 35 mph' (for curves). I could barely make 20 mph without tearing the car to pieces (Dempster 1989, 1).

Alongside careful considerations written to other fellow acoustic explorers ("Watch for Rattlesnakes," "Take water if a warm day," etc.), Dempster clearly illustrates these sites of sonic intrigue. As he clarifies, these spaces are either human creations or natural phenomena. Yet what they have in common is that they are 'likely to be acoustical accidents' (Dempster 1989, 1). Indeed, the sounds of nature unto themselves are dealt with rather minimally. If the state of Washington is to be treated as a musical instrument, therefore, Dempster does so in regard to his own relationship *to it*. In this sense, *SWAMI* is very much akin to *EIS* – an expanded methodology for exploring resonance, chamber music performance, and communal play. The project is not only collaborative in that the reader (or perhaps, the audience?) is invited into the compositional procedure, but also as we imagine how the sonic environment might exercise its own improvisatory forms of past, present, and future.

Conclusion: Resonant Play as Radical Education and De-Centering

As we consider the role Stuart Dempster has played in slow, layered pulls away from modernist, and then post-modernist music education models, towards something more radical, these shifts in relationality are particularly illustrative. Amidst the arising of The SoundScape Project, SoundMap projects, and Biomusics which often exhorted preservationist orientations, the relational stance of Dempster's electroacoustic performance projects were skewed slightly differently. The augmented realities of his sonic realizations are not in opposition to a reverence for our natural environment. Dempster's uses of technology in various projects were sometimes but not always a direct means to an environmental end (of appreciation or conservation); but they have always been a means to a philosophically prior phenomenological project that is equally activist. Stuart Dempster's teaching inculcates a listener and practitioner/listener to develop refined self-awareness and ever more conscious interpersonal experience, and, yes, empathy. Dempster's work on this path encourages us not only to resonate with the environment as our tool and collaborator, but to decentralize the human performer quite emphatically. Within a traditional educational paradigm, this may be seen as a radical act. To what extent does turning the environment into a symbiotic part of the creative process shift our role within it?

Part of the process of decentering the human performer for Dempster involves elevating everything else around us. He playfully and naturally yet quite thoughtfully personifies everything—the walls of the cistern, the mistake a player might make (which to him is actually "a mistake visit[ing]" the situation) and so on. "We are all in this together" (Dempster, email correspondence with authors, June 28, 2023). The radicality of inclusion. Do we cede control or are we just exerting a new kind of control? How can we play with those boundaries in a productive, mindful way? Dempster's work and teaching asks all of these questions. His use of live computer processing, in environmental contexts which further blur the end of the human sound and the beginning of the artificial or natural facsimile of it, is another facet of ceding control, since once a drone starts it must find its way to conclusion, including its resonance. A single event might be upwards of a minute in length typically, in the contexts we are referencing and that are discussed earlier in this article.

In sum, whether using dog toys, fart whistles, coffee percolators, an abandoned nuclear reactor site, or natural outdoor and underground spaces as instruments, Dempster's interest in radical play, alongside serious provocation, has allowed him to embrace digital techniques as an essential part of a continuum of influential performance and teaching methods. His love of resonance as a generative force, both intimately embodied (as he experienced it through his trombone and didjeridu playing) and disembodied (as he experienced it through diverse performances with *Deep Listening Band*), seeped into the conceptual realm. In other words, Dempster was practicing 'sound study' before there were departments of such. And, for him, this also bled into the domain of ethnomusicology, or, into the domain of early anthropology-leaning courses such as *Music and Society*. Dempster's radicalism certainly included an embrace of interdisciplinarity, long before it became popular.

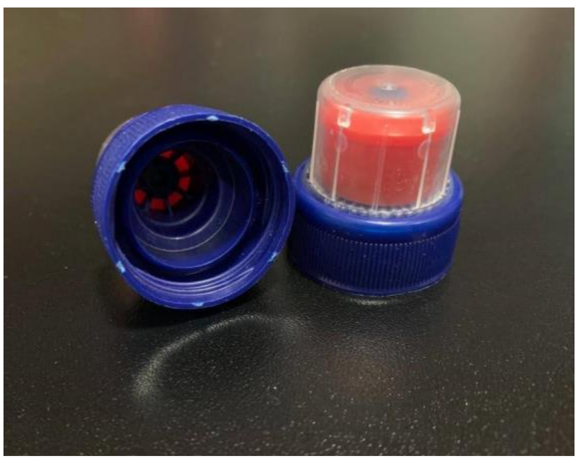


Figure 3. Bottle caps, part of Dempster's arsenal of accidentally manufactured instruments

Teaching through lived practice, always, Dempster pursued his sonic and theatrical provocations with a sense of pleasure and genuine politeness—which, perhaps, was part of the message. Experimentalism has been, for him, more a way of life than a chosen artistic direction. As mentioned, it has always gone beyond welcoming unexpected happenings in improvisation. It has, more pointedly, involved an insouciance about social propriety, and a passionate belief in the value of releasing creativity in untrained non-specialists. It has involved a consciousness of the allure of 'limitless' [non-stylistic] improvisation (Dempster, in correspondence with authors, March 9, 2023). It has involved an almost Artaudian attraction to an edgy discomfort that might lead to some sort of 'wow' moment for anyone involved. 'Better to ask forgiveness than permission' seems an apt descriptor for so many documented Dempster happenings. That he has always done everything he has in such a meditative, flow-like way, is testament to the fact that he has been guided, always, by a real joy of giving people "something to chew on," (Dempster, in correspondence with authors, March 9, 2023) and by the visceral (and cerebral) impact of unusual or particular sonic phenomena.

Most of the phenomena discussed above require a live playing *with* the given space and with the sounds produced, in a feedback process. The role of a practiced rehearsal thus lessened in importance or changed in nature. Whether through environmental resonance or through use of electroacoustic delay systems, feedback for Dempster has always been a form of communal process. That community might be his students, a random audience, or the limestone walls of a chasm in the earth. Is that not a unique educational philosophy?

Whether Dempster would have developed similar practices without the oppressively traditional system of cultural containment is an apt question for theorists, perhaps. But, to sum all of this up in his own eloquent words: he has been hell bent on "putting the 'play' back into playing [and teaching] music" (Dempster, in correspondence with authors, March 9, 2023). But why is it so important, in such a serious pursuit, that the audience (whether listener or reader) leave with a smile? It is undeniably healthy to laugh. Producing laughter is also an uninhibited behavior, and, as such, an invaluable addition to a repertoire of pleasurably resistant actions in response to our normative constraining behavioral codes.

Acknowledgements

Huge thank you to Stuart Dempster for his generous time and availability to the authors as we prepared this article.

References

Abbate, C. 2004. “Drastic or Gnostic?” *Critical Inquiry* 30, no. 3: 505–35.

Dempster, Stuart. 2024. [1989.] *State of Washington as a Musical Instrument*. Seattle: CENTRUM Foundation.

Dempster, Stuart. 2012. “Training for Listening: A Lifelong Practice,” in *Anthology of Essays on Deep Listening*, edited by Monique Buzzarté and Tom Bickley, 13–28. Kingston, NY: Deep Listening Publications.

Dempster, Stuart. 1994a. “Underground Overlays from the Cistern Chapel” New Albion Records NA076CD. [Original album notes.]

Dempster, Stuart. 1994b [1979]. *The Modern Trombone: A Definition of its Idioms*. Accura Music Publishing.

Dempster, Stuart. 1979. *In the Great Abbey of Clement VI*. 1750 Arch Records S-1775. [Original album notes.]

Gottschalk, Jennie. 2016. *Experimental Music Since 1970*. New York: Bloomsbury Academic.

Meyer, Neal. 1986. "An Interview with Stuart Dempster and William O. Smith." *Perspectives of New Music* 24, no. 2: 190–223.

Miller, Caroline Louise. 2018. “Texture, Materiality, and Sensation in the Digital Production of Electronic Dance Music Subgenres.” *Journal SEAMUS* 29, nos. 1–2: 3-19.

Oliveros, Pauline, and Panaiotis. 1991. "The Expanded Instrument System (EIS)." In *Proceedings of the International Computer Music Conference*, International Computer Music Association, 404–7

Oliveros, Pauline. 1995. “Acoustic and virtual space as a dynamic element of music.” *Leonardo Music Journal* 5: 19–22.

Wilson, Marie C. 2004 [1984]. *Closing the Leadership Gap: Why Women Can and Must Help Run the World*. New York: Penguin Books.

Synesthetic Perceptions: An Interview with Composer Yi-De Chen

Yi-De Chen

yidechenmusic@gmail.com

Introduction

Yi-De Chen is a composer from Taipei, Taiwan who experiences both music and life in a unique way. As Yi-De explains, "I am a composer with retinitis pigmentosa, which adds a unique perspective to my musical expression." This perspective, along with Yi-De's experience of synesthesia, has influenced Yi-De's creative process, as well as his outlook on music, sound, and life more generally.

The following interview was conducted over email in late 2023 and early 2024 with Journal SEAMUS Editor-in-Chief Drake Andersen, and has been lightly edited for clarity.

For more information on Yi-De's music, please visit: https://yidechen.com/

Journal SEAMUS: What is the role of sound in your everyday life?

Yi-De Chen: Sound has played an integral role in my daily life, even from a young age. Interestingly, as my mother observed, my visual development started relatively late, but I demonstrated an early fascination with sound. When I was just a baby, I exhibited behaviors like tapping the floor to create rhythms or crinkling newspapers to listen to the sounds they produced. These were the early signs of my responsiveness to the auditory world around me. My hearing is empathetic, allowing me to sit by the window on a morning and absorb the sounds of pedestrians walking down the street, cars passing by, birdsong, the rustling of leaves in the wind, and even the sounds of my neighbors cooking.

When did you first realize that you wanted to become a musician?

While studying at Taipei National University of the Arts, I studied with several renowned Taiwanese composers. I received support from my closest music fellows and alums, who provided invaluable guidance and support, deepening my understanding of classical and contemporary music. They generously shared their own experiences in composition, inspiring me to contemplate the evolution of music and the diversity inherent in contemporary classical music. These interactions enriched my musical perspective and ignited a stronger desire to delve deeper into this field. I began experimenting with various musical styles and techniques to find and develop my unique voice. However, the more apparent desire to become a composer first sparked when I could truly appreciate the beauty of the works of Bach, Mozart, Chopin, Debussy, and Messiaen.

What other influences have been significant for you?

Since childhood, I was exposed to various music styles and genres, such as Western classical, contemporary classical, pop, rock, dance, Jazz, world music, and video game music; simultaneously, the musical experience I obtained has again transformed my musical perspectives. As I've gained more understanding and knowledge, my ability to absorb, digest, and reuse these influences in my compositions has grown. My musical education and the cultivation of my listening sensations have deepened my comprehension of music. In addition, my synesthetic and intuitive perceptions uniquely influence my music comprehension, as I often translate sounds into visual imagery. These various influences and my personal experiences have contributed to my development as a well-rounded musician. Here, I want to express my profound gratitude to all my composition teachers in Taiwan and the United States who have worked with me over the years.

During my time at the Indiana University Jacobs School of Music, I had the privilege of studying under Prof. John Gibson and Prof. Chi Wang, who, with their infinite patience and passion, guided me through the challenges of computer music, enhancing my electroacoustic music composition skills.

Tell me more about your experience of synesthesia.

I possess a unique ability to translate sounds into visual images, a phenomenon known as synesthesia. This ability allows me to vividly perceive both the sounds and the visual representations they evoke, providing me with a multi-sensory perspective of the world around me. I rely heavily on my acute sense of hearing to navigate safely. It enables me to avoid collisions with vehicles and pedestrians, particularly in the city's bustling streets. When I walk through urban environments, my hearing becomes my primary tool for detecting changes in my surroundings, such as construction work on buildings or moving large vehicles. This heightened reliance on sound adds an extra layer of importance to auditory experiences and underscores sound's profound impact on my ability to interact with and understand the world.

A childhood memory lingers in my mind: sleeping with my parents and brothers, my bed was close to the road, and sometimes, the sound of car brakes or motorcycles passing by would often wake me up in the middle of the night. In those moments, I would see a serrated, sawtooth-shaped image in my mind's eye, with the serrations blending into shades of green and orange, and yellow pen strokes accompanied this visual. These sensory experiences were so intense that they would occasionally cause me to cry out loud, leaving my parents and brothers puzzled and perceiving me as a sensitive child.

Other sounds also trigger unique sensations for me. For instance, the ticking of a clock evokes the sensation of the shape of fingernails. The sound of water in the bathtub creates a visual sensation that I can only describe as the surface of a coin. Additionally, when I hear birds singing, it generates a stroke of yellow color in my vision, making it seem as if yellow hues surround me. These experiences demonstrate how my synesthesia influences my perception of everyday sounds, creating a multi-sensory world that is both fascinating and, at times, overwhelming.

But it also has an impact on your music, right?

Synesthesia allows me to emotionally connect with and understand the feelings or sensations conveyed by music, especially in terms of how the music evokes visual imagery or other sensory experiences. For instance, the violin timbre makes me perceive red, the flute is yellow, the horn is green, the celesta is blue, and the cello is brown (or purple). The connection between my visual sensations and the tone of different instruments is a crucial part of my creative inspiration.

My synesthesia and associations also cause different keys to evoke specific colors: C major is white, D major is light gray, E major is yellow, F major is green, G major is red, A major is purple, and B major is blue.

Are there any particular musical works that give you a strong synesthetic impression?

As I listen to J.S. Bach's The Art of Fugue, BWV 1080, my synesthesia is also triggered by the instruments used in the recording and the key of D minor. For instance, the piano brings me a white color, while the harpsichord induces a yellow hue. Moreover, I associate the music atmosphere with the intricate structure of tree branches, forming an organic unity that conveys a sense of sorrow and a quest for truth. The dissonance and resonance between voices create tension, drawing me into the piece's complexity.

W.A. Mozart's Clarinet Quintet in A Major KV 581 evokes images of beige colors and wooden textures, along with a sense of sweetness. When the clarinet leads with the melody, I envision contrasting purple and yellow colors. This piece differs from the Bach work mentioned earlier, with its clear harmonies and square shapes that remind me of chocolate.

Have you ever met another musician or composer with synesthesia? Did you find any similarities between your respective visual imagery?

I haven't personally met anyone with synesthesia, but I've encountered descriptions of synesthetic experiences from others through YouTube videos and news articles. Many people associate specific letters with particular colors; for example, the letter A might be perceived as red and B as blue in English. However, I don't have synesthesia with English letters or vocabulary. My synesthetic experiences apply to the sequence of numbers. For example, one is red, two is blue, three is yellow, and four is green.

When it comes to numbers beyond ten, the colors become less distinct. Typically, I don't experience synesthetic reactions until I encounter more significant numbers, often linked to my elementary school exam scores—it's amusing! For instance, I associate 100 with red, 90 with blue, 80 with green, and 70 with yellow. However, it's worth mentioning that I don't have any synesthetic reactions to numbers like 99, 98, 97, 86, 75, or similar, except for 70, 80, 90, and 100. Synesthesia is genuinely fascinating, and it's interesting to discover how our brains connect different sensory experiences.

What other factors influence your music?

Improvisation plays an essential role in my writing. This does not mean I don't rely on written composition to complete my piece. For me, the written composition I use most of the time can be considered a kind of improvisation in the relatively long term; conversely, the characteristics of inspiration, musical ideas, and motivation originate from my improvisation based on my inner thoughts. By selecting the desired materials, developing specific ideas, and abandoning unnecessary elements in various ways based on a severe writing plan, my creative vocabulary, melodic passages, and whole pieces can gradually form to some degree.

Also, I want to elaborate on the "transformation" process in music composition. I want to use the language learning analogy to explain that music composition has a similar approach, just as learning a language requires stimuli, comprehension, imitation, and creation. Looking back to my previous works, I have used various themes in music composition, including impressions from the material world, emotions, and abstract concepts. As a composer, through continuous listening, reflection, and creation, my musical experiences continue to evolve, allowing me to express my musical ideas in diverse ways.

Does electronic media/sound contribute an additional dimension to your synesthesia in any way? For example, if you are working with a synthesizer, are you able to change the color you perceive by changing the timbre of the synth?

The colors I perceive change based on the frequency and timbre of the sounds I'm working with. For instance, when I experiment with a synthesizer, adjusting the timbre can alter the colors I see in my mind's eye. As the frequency increases, the edges of these shapes become more defined. When I use a synthesizer with increased high frequencies, it's like a kaleidoscope of light, making the colors lighter and more radiant.

Let me give you an example using sine and triangle waves within the 23 to 3000 Hz range. In my synesthetic experience, a sine wave might evoke a radiant, light green color, while a triangle wave takes on a teal-green hue. Now, when I say "shape," I'm referring to the visual representation of the sound in my mind. It's not a physical shape but a mental image accompanying the sound. If you push both beyond 3000 Hz, they start to sound more alike, and their colors lose their distinctions, gradually merging towards a whitish appearance.

When various shapes become similar due to increased high frequencies, they all emit a bright sensation, and the colors shift towards white or yellow. This is what I refer to as "visual blending." It's like mixing different colors, and they create a new color. In my case, it's a visual blending that occurs when the sonic characteristics align in a certain way. This fascinating interplay between sound and color never ceases to amaze me, and it adds a whole new level of creativity to my work as a composer.

How would you describe your artistic philosophy?

Though intangible, music possesses a profound spirituality that can profoundly impact the human soul and culture. As a composer, I draw inspiration from the tapestry of everyday life, engage in philosophical contemplations, and seek spiritual insights from various sources, including religion. This exploration often gives me an intuitive sense of the mood and themes that will shape my future compositions.

Furthermore, I am deeply committed to unlocking the rich layers of music. I strive to infuse my compositions with a narrative quality through the meticulous refinement of harmony and timbre and the artful manipulation of tension and release within musical sections. Above all, my primary focus remains on the emotions conveyed by my compositions. My ultimate goal is to establish a profound connection with the audience, allowing them to grasp the essence of my work.

Can you give me an example of how this plays out in a particular work?

In my recent orchestral work, *Spirituality,* I blend the elements of traditional Beiguan music and folk beliefs with the vibrant essence of Taiwan's fruits, creating a symphony that encapsulates the country's rich cultural and historical heritage. Fruits are an integral part of Taiwanese cuisine and traditions, featuring prominently in dishes like ice cream, tea, and cake and playing a role in religious practices and even Feng Shui. I chose seven of my favorite fruits—kiwi, pineapple, grape, papaya, watermelon, mulberry, and blueberry—to weave into the orchestral fabric of the piece. These fruits, with their diverse tastes and smells, go beyond mere sweetness or sourness; I aimed to create music that embodied each fruit's complex flavors and fragrances.

My synesthesia played a crucial role in the early stages of the composition process. For instance, the sweetness of watermelon reminded me of the harmonic series, so I used this sonority to represent that aspect of the fruit. However, since synesthesia can produce vague visual graphics, I also relied on associations, such as the fruit's large size and how its flavor changes as you eat it. After the piece was performed, I could enjoy the synesthesia again through the music recording. Although some aspects of my perception might be unique to my synesthesia, many friends have given positive feedback, sharing that they could feel the various colors and imagine the qualities of the seven fruits after listening to the recording.

Each fruit's color and taste were represented through carefully chosen musical elements based on my synesthesia and associations. I utilized chords with third intervals to capture the sweetness of all seven fruits, altering pitches to express their unique tastes. While my synesthesia for smell isn't as prominent as for vision, my orchestration still partially mirrors the complex colors, flavors, and fragrances, creating a symphony that resonates with my senses.

You can immerse yourself in the fantastical world of *Spirituality,* where the notes bring to life the vibrant colors, tastes, and textures of these seven beloved fruits, at the following link: https://yidechen.com/works-by-instrumentation/

What are you working on right now?

I'm currently immersed in an exciting project for the upcoming Alba Music Festival in Italy, scheduled for Summer 2024. It's a piece for piano trio, and the talented SOLI Chamber Ensemble will perform it during their residency at the festival. I've been placing my creative energy into crafting this composition by employing particular timbre, sonority, and texture. I can't wait to see it come to life on stage in such a beautiful and culturally rich setting like Alba.

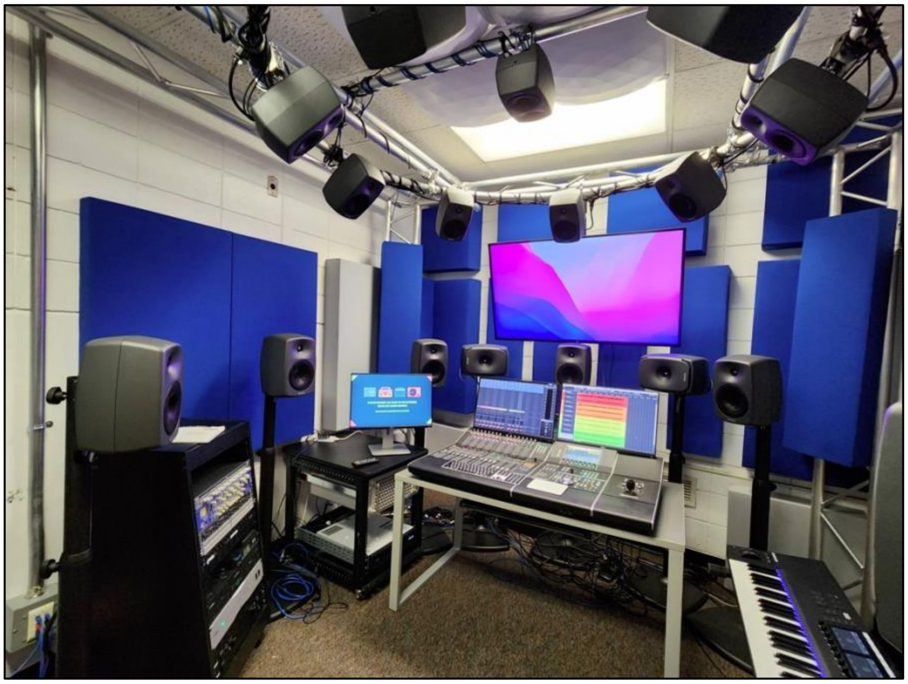
Where We Create (Part 1)

Introduction

Where We Create is an ongoing series in which we feature SEAMUS members' working spaces, past and present. This inaugural edition features a wide dispersal of such spaces in both time and space—from a Buchla 100 from 1968 to a Pacarana in Hawaii in the present day. We look forward to featuring more spaces in subsequent issues. If you would like to be included in a future edition, please submit your photograph by email with the subject line "Where We Create" to journal@seamusonline.org.

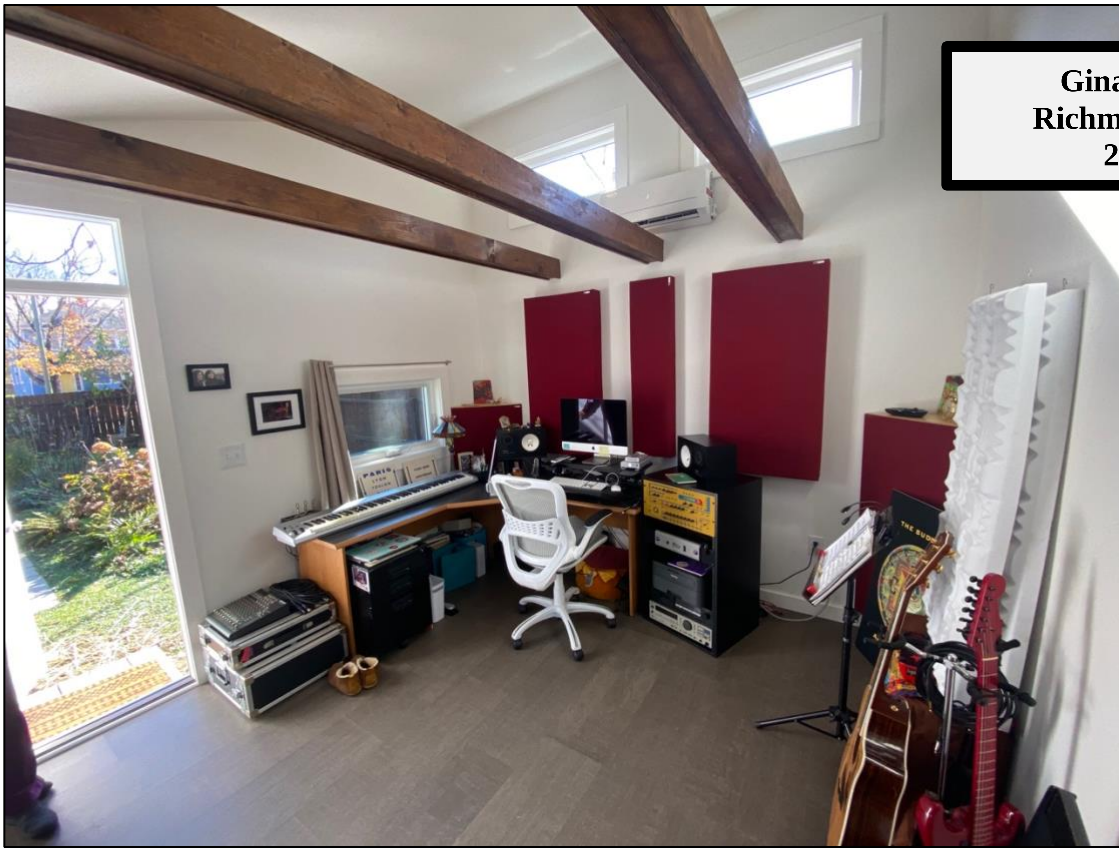
Brian Belet
Clark University
Worcester, MA
1992

Jeff Kaiser
Warrensburg, MO
2023



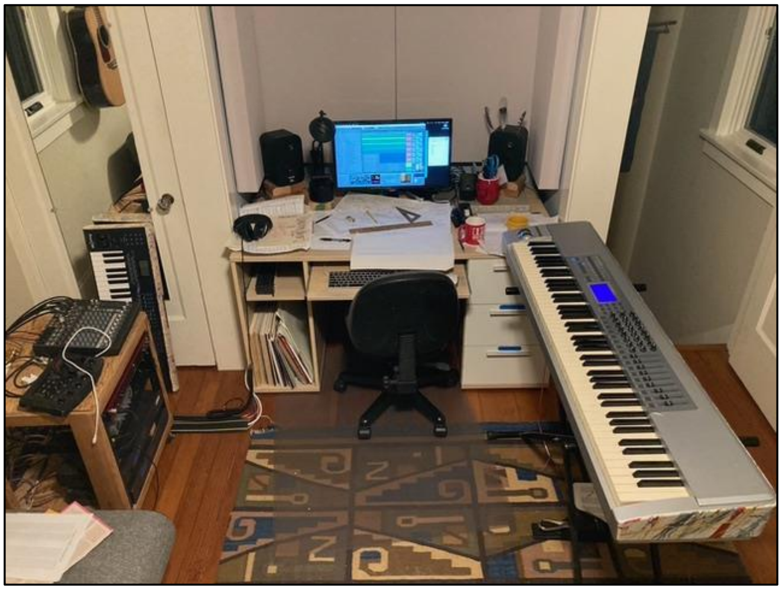
Jason Bolte & Linda Antas
Montana State University
Bozeman, MT

Elizabeth Hinkle-Turner
WDR Studios
Köln, Germany
1989
(Photo: Volker Müller)



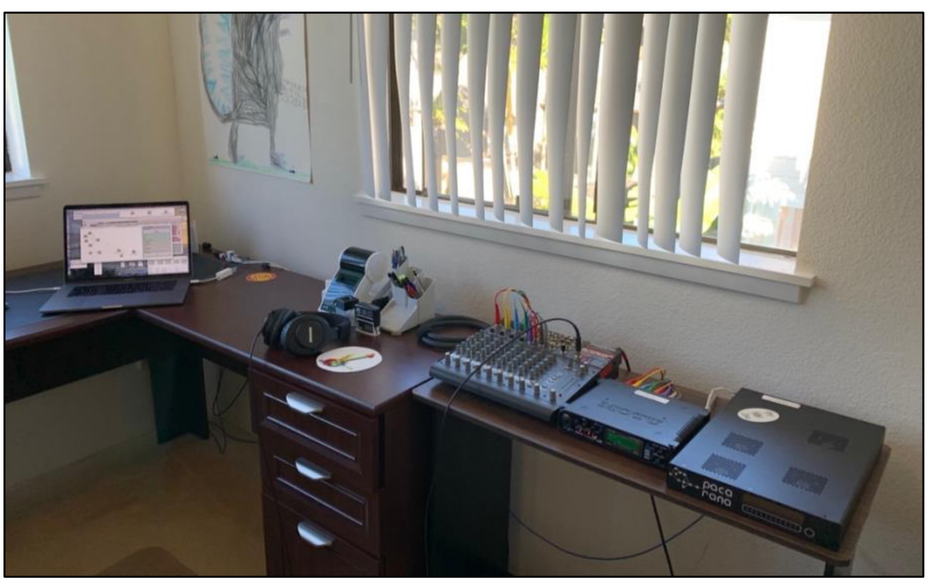
Gina Biver
Richmond, VA
2023

Eric Chasalow
Bowdoin College
Brunswick, ME
1976

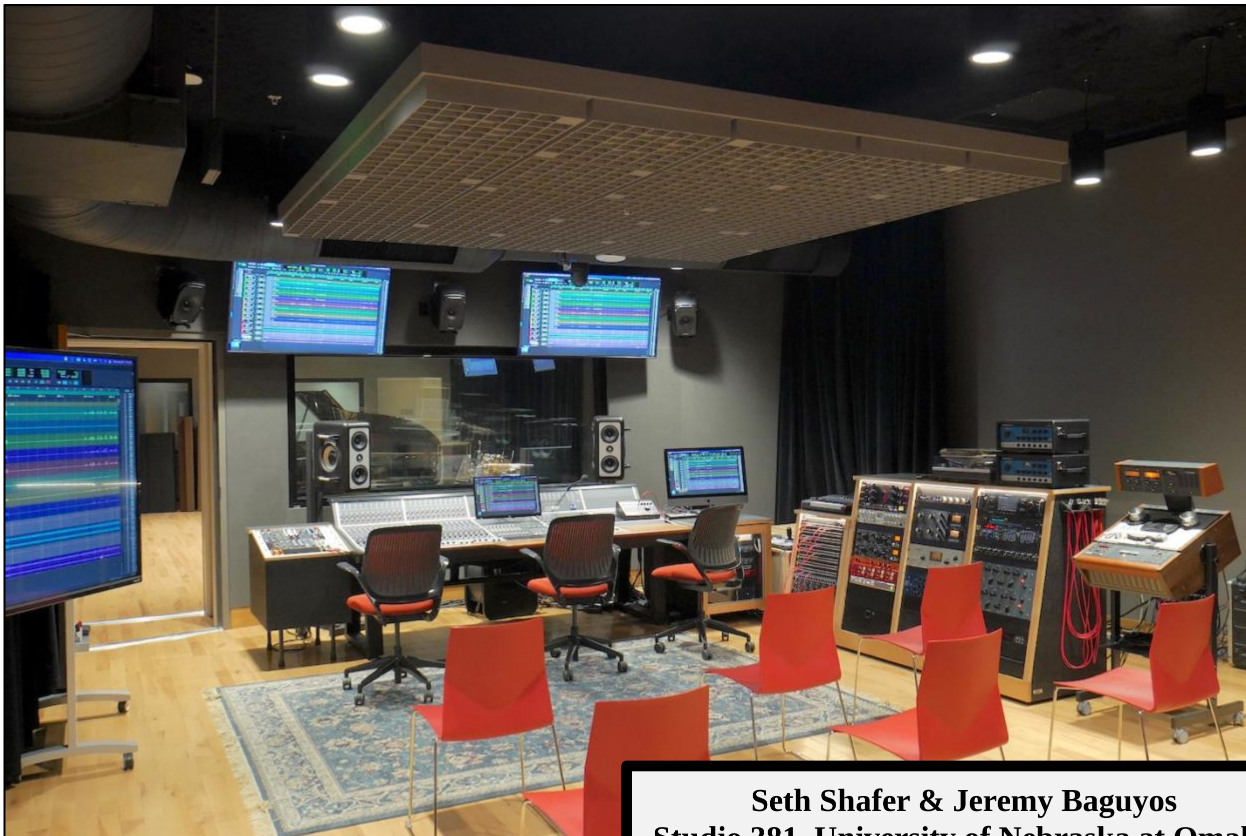


Ted Allen
Oakland, CA
2023

Fabio Fabbri
Santa Margherita
Ligure, Italy
2023

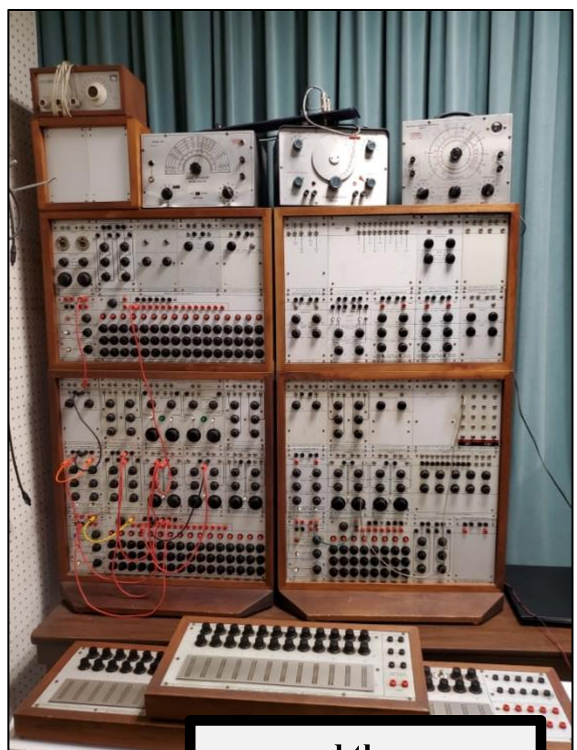


Brian Belet
Mauka Studio
Mililani, HI
2023

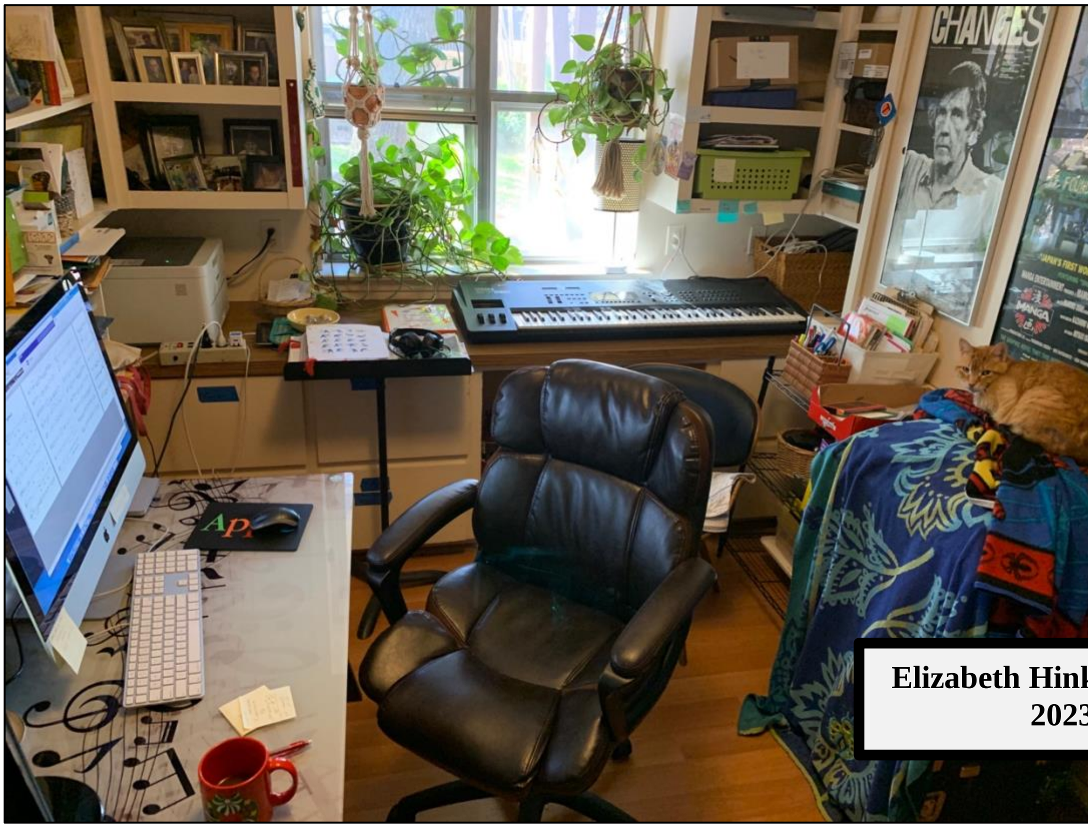


Seth Shafer & Jeremy Baguyos
Studio 381, University of Nebraska at Omaha
2022

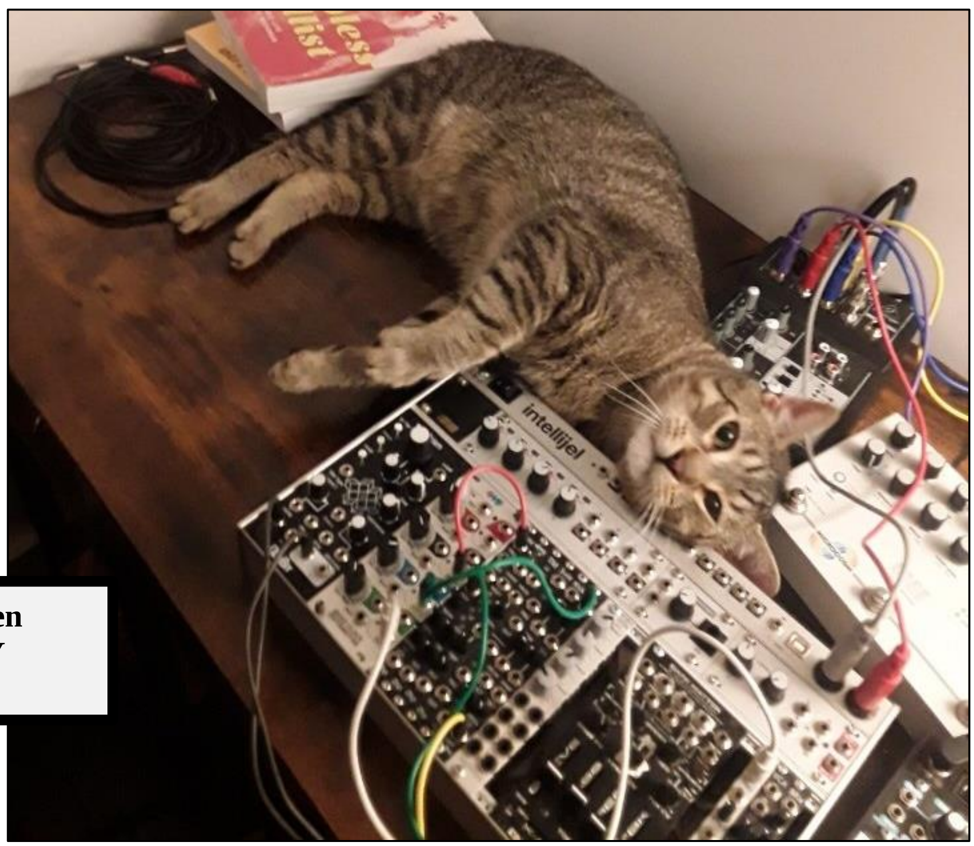
Eric Chasalow at BEAMS
(Brandeis University) in 1993…



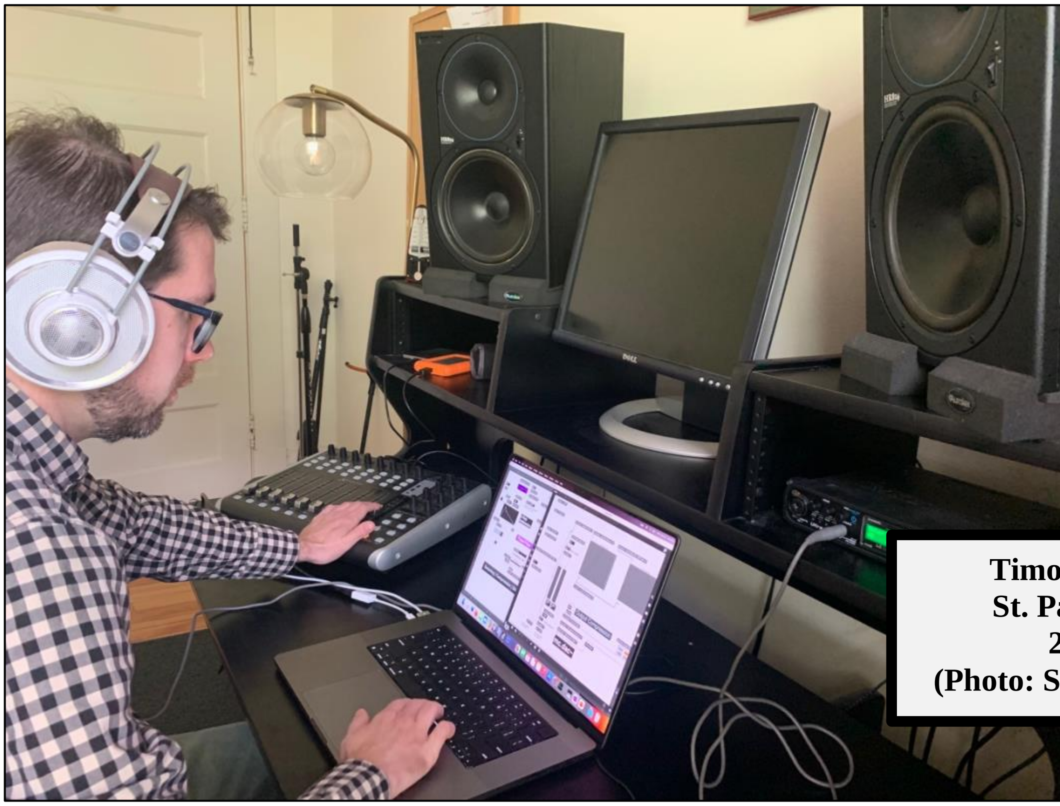
…and the same
Buchla in 2022!



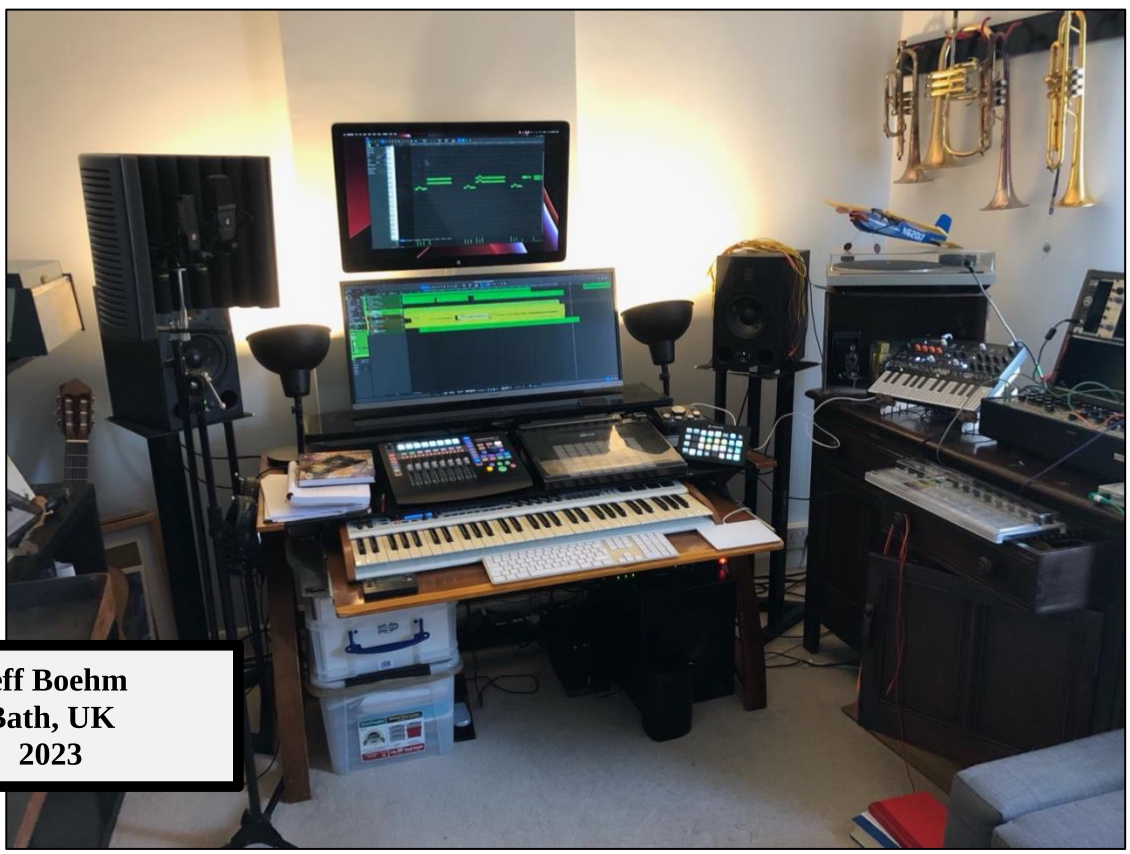
Elizabeth Hinkle-Turner
2023



Drake Andersen
Brooklyn, NY
2022



Timothy Roy
St. Paul, MN
2022
(Photo: Sarah Bauer)



Jeff Boehm
Bath, UK
2023

Drake Andersen
Editor-in-Chief, Journal SEAMUS
journal@seamusonline.org

Department of Music
Vassar College
124 Raymond Ave.
Poughkeepsie, NY 12604

www.seamusonline.org