THE OM COMPOSER'S BOOK

Volume 3

Edited by
Jean Bresson, Carlos Agon, Gérard Assayag

Preface by
Roger Dannenberg

Collection *Musique/Sciences*

Communication
Marine Nicodeau

Editions
Claire Marquet

Cover design
BelleVille

Volume editor
Jean Bresson

Editorial coordination and support
Aaron Einbond

ISBN 978-2-7521-0283-6

www.editions-delatour.com
www.ircam.fr

Contents

Musique instrumentale concrète: Timbral transcription in *What the Blind See* and *Without Words*

Aaron Einbond

Transcription is an increasingly influential compositional model in the 21st century. Bridging techniques of *musique concrète* and *musique concrète instrumentale*, my work since 2007 has focused on using timbral descriptors to transcribe audio recordings for live instrumental ensemble and electronics. The sources and results vary, including transformation of noise-rich playing techniques, transcription of improvised material produced by performer-collaborators, and fusion of instrumental textures with ambient field recordings. However the technical implementation employs a shared toolkit: sample databases are recorded, analysed, and organised into an audio mosaic with the CataRT package for corpus-based concatenative synthesis. Then OpenMusic is used to produce a corresponding instrumental transcription to be incorporated into the finished score. This chapter presents the approach in two works for ensemble and electronics, *What the Blind See* (2009) and *Without Words* (2012), as well as complementary real-time technologies including close miking and live audio mosaicking. In the process transcription is considered as a renewed expressive resource for the extended lexicon of electronically augmented instrumental sound.

Reproduction

Mimesis in music may date to the beginning of music itself, and its reach extends both historically and geographically, from Plato's *Republic* to Tuvan throat singing ([13], p. 58). It is also central to common-practice music, whether the imitation of imitative counterpoint, the programme of programme music, or the affect of the Baroque doctrine of the affections. In the 20th century mimesis has been reframed by "technological reproducibility", as in Walter Benjamin's landmark essay. In technology's wake, Pierre Schaeffer treats reproduction by juxtaposing "abstract" and "concrete" listening modes; it is with their synthesis that he defines "reduced listening" ([2], p. 37). Composers who combine instrumental and electronic technologies of reproduction include François-Bernard Mâche and his "phonography", Clarence Barlow's "synthrumentation", Gérard Grisey's "instrumental synthesis", Trevor Wishart's sonic "transformation", and Peter Ablinger's "phonorealism" [15, 5]. Surveying these perspectives as well as those of the early 21st century, Nicolas Donin writes: "'composition' [...] now includes as well the navigation between different 'reproducibilities' through operations of translation such as transcription, transcoding, or transformations" [5]. Transcription has been a particularly influential trend in what Donin terms "instrumental resynthesis".

In the purely instrumental domain, Helmut Lachenmann offers a historical counterpoint with his *musique concrète instrumentale*:

> The idea of "instrumental *musique concrète*"—i.e. sound as a message conveyed from its own mechanical origin, and so sound as experience of energy [...] signifies an extensive defamiliarisation of instrumental technique. [16]

Or as Rainer Nonnenmann clarifies, "instead of relating sounds to extra-musical causes, listeners were now called upon to relate sounds back to the genuinely intra-musical preconditions of their concrete instrumental production" [14]. Now, over 40 years after Lachenmann's initial explorations of *musique concrète instrumentale*, one could argue that the sound world of Lachenmann's instrumentarium has itself become familiar, and is no longer available to be "freshly illuminated" [16]. Yet judging by its ongoing popularity to a younger generation of composers, it is far from in danger of being abandoned. How can the composer of instrumental music employ these now familiar techniques and still hope for them to speak vividly?

The solution is to turn *musique concrète instrumentale* on its head: instead of focusing the listener on the mechanical origins of the sounds, to free the instrumental sources to suggest timbral details of other concrete origins. Instrumental sounds, stripped of their historical connotation and reduced to their timbral essentials by Lachenmann and the generation that followed, are now available for renewed connotation. Rather than *musique concrète* of instrumental sound, it is instrumental music that is perceived concretely: *musique instrumentale concrète*.

Behind the playful turn of phrase lies a wealth of evocative potentials, and associated technical questions, that have occupied my compositional work since *Beside Oneself* for viola and electronics in 2007. How can sonic *expression* be renewed without falling back on historically laden *expressivity*?[1] Technology has proven an indispensible means toward this goal, where concrete reduced listening is effected through processes including close miking, amplification, sampling, and audio mosaicking. This chapter will focus in particular in how OpenMusic is used in these processes for a range of compositional applications.

Transcription

A potential definition of musical transcription is the projection of material from one musical space to another—a space of frequencies or pitches, timings or rhythms, or other acoustic parameters. Especially when working with noise-based instrumental sounds, a suggestive metaphor is *timbre space*, a model of timbre as a multidimensional percept [21, 8]. Such timbral dimensions have been associated with audio features, or *descriptors*: "characteristics extracted from the source sounds, or higher level descriptors attributed to them" [17]. The notion of the descriptor is the perfect technological correlate of reduced listening: like Schaeffer's phenomenological *époché* [2], the computer takes no account

[1]This contrast is neatly denoted by two distinct German translations of "expression": *Ausdruck*/*Expression*.

of the sound's physical origins as it performs its analysis. Of course this does not apply to the human listener, who must set the parameters of the algorithm beforehand and interpret the results after, in both cases based on his or her own listening experience. So rather than transparent measures of perception, timbral descriptors are contextually sensitive materials upon which the composer necessarily exercises a subjective influence. Decisions include what samples to use, how to segment them, what descriptors to analyse, among many others. Far from automatic, timbral transcription can be a rich expressive resource.

In *What the Blind See* for viola, bass clarinet, harp, piano, percussion, and electronics (2009), sources for transcription include instrumental materials, physical gestures performed with an electronic controller, and field recordings of rain, ice, and snow. The process is termed *Corpus-Based Transcription* [7] as it takes advantage of corpus-based concatenative synthesis (CBCS) with the CataRT package for Max, and transcription with OpenMusic. While the details have been described elsewhere [7, 5], an example can be outlined as follows: a large database of instrumental samples (the *corpus*, in this case made up of samples of the five performers of the ensemble) is compared to a field recording (the *target*, in this case an arctic glacier) using CataRT and a purpose-built Max patch for analysis. After segmenting both corpus and target into short grains, and matching those segments with the most similar descriptor values, the glacier recording is reproduced with an audio mosaic of instrumental samples.

This mosaic is then used as a model for an instrumental score. The timings, filenames, and associated data are stored in an SDIF file, which can be loaded by the OpenMusic *SDIFfile* object. The data are extracted, organised, and rhythmically quantified in OpenMusic (Figure 1), then the contents of the resulting *poly* object are exported as a MusicXML file. This is imported into Finale and edited, with the aid of the descriptors and metadata stored in the SDIF file, by annotating playing techniques, dynamics, and articulations, as well as adjusting or simplifying rhythms and techniques to facilitate playability. The resulting transcription is incorporated into the work as the compositional process continues (Figure 2), and finally the ensemble performs the score live, reinterpreting the transcribed field recording in concert.

In the finished work, the original field recording of a glacier is not performed, but its shadow remains through the fusion of the ensemble into a colourful imprint. Even without an explicit concrete referent, the individual timbral details and identities of each instrument are subsumed into an "instrumental ambient"[2] soundscape. At the same time an association with other field recordings of rain and snow that are directly revealed later in the work may retrospectively conjure a secondary denotation of the texture's origin. In other examples transcription sources are made more or less explicit, opening up a powerful expressive resource with a decisive effect on perception.

One of the keys to this multi-valence is the simulation of reduced listening through close miking, sampling, and amplification. In *What the Blind See*, contact microphones are placed on the piano frame, piano soundboard, bass drum skin, and tam-tam, and miniature lavalier microphones are used on the bridge of the viola, near the embouchure of the bass clarinet, and beneath the soundboard of the harp. These microphones permit the amplification of sounds distorted out-of-scale to their mode of production, flattening

[2]Bernhard Günther, personal communication, 25 July 2010.

their concrete instrumental sources to the surface of a loudspeaker. So unfamiliar playing techniques are "misheard"—from the perspective of Lachenmann—as they cannot be identified with their productive energies. But on the contrary, through Schaeffer's reduced listening, they are freed to represent other sonic experiences that may be coaxed out through transcription.

Figure 1. OpenMusic patch to transcribe an audio mosaic of a field recording.

Figure 2. *What the Blind See* mm. 176-79. © Edition Gravis Verlag, Brühl, Germany. Printed with kind permission.

The same microphones are used for recording the pre-compositional sample databases as well as amplifying the instruments in concert, assuring the closest possible reproduction of the audio mosaic live. More important, the same musicians are called upon to record the database as will premiere the work, in this case Ensemble L'Instant Donné, and the imprint of their personalities and instruments becomes a collaborative part of the compositional process. This process is not only "radically idiomatic"[3] but radically personalised. Like the audio descriptors themselves, the performers and instruments too become sites of individual expressive variation, privileging the specific, concrete, and personal over general, abstract, or universal.

While using OpenMusic to transcribe samples of a performer is not new,[4] a significant difference here is the representation of noise-rich audio in its full detail without recourse to the time-frequency grid of a spectrogram. To reproduce this sonic timbral palette in symbolic music notation is a significant challenge. During sampling sessions, performers' playing techniques are documented with verbal explanations, photographs, and video. When the sessions are edited into sample databases, the filename of each sample is annotated by its playing technique in as much detail as possible, which is then used to edit the score in Finale. Nonetheless the noise-rich timbres push the limits of music notation: for many works the score is supplemented by sending the performers a selection of their own audio samples. The goal is to trigger their sonic and physical memories, so the symbolic notation serves as a mnemonic for personal oral history going back to the sampling session.

Notably, for many playing techniques, pitch may not be a salient feature of either the sound or its notational representation. Regardless, CataRT estimates the pitch of each grain using the *yin~* object in Max. However the *yin~* algorithm may not give a reliable estimate, or independently, the user may decide not to weight pitch in constructing the audio mosaic. When the mosaic is imported to OpenMusic, the estimated pitch is used in the *poly* display merely as a placeholder for each grain. As the score is edited into its final form this pitch information may be retained, adjusted, or eliminated entirely according to the playing technique needed. Comparing Figures 1 and 2, the viola pitches are replaced by x-shaped note heads (*pizzicato* on dampened strings) and harp pitches by rectangular note heads (fingernails *sur la table*), while those of the breathy low bass clarinet and marimba with Superball mallets are retained with adjustments and elaborations.

At the heart of the OpenMusic patch *omquantify* quantises the durations of grains in the audio mosaic, expressed in milliseconds, into notated rhythms (in Figure 1, *omquantify* is included inside the *make-voice* sub-patch). Significantly, this use of detailed rhythmic notation does not have a metrical function in itself, despite its "complex" appearance. As such it could be analogised to an elaborated form of graphic notation where the metrical grid serves a reference purpose only. Precedents include scores of György Ligeti, Gérard Grisey, and Tristan Murail.[5] But while the notated metre is not intended to be heard directly, the quantisation parameters are chosen finely enough to

[3]Richard Barrett and Daryl Buckley, *Dark Matter* programme text (2003): `http://richardbarrettmusic.com/DARKMATTERinterview.html`

[4]See for instance [20].

[5]For example: György Ligeti, *Études pour piano, premier* (Schott, 1985), p. 20, p. 28.

facilitate the performers' interpretation of the "expressive microtiming" [9] underlying the transcription target: whether the "feel" of a recorded improvisation, or the "groove" of the interconnected actors (animal, vegetal, mineral) in a recorded soundscape.

Gesture

A physical gesture rather than a field recording may be used as a target: in another example from *What the Blind See*, a Wacom tablet was used to "improvise" a short gesture mapped directly to the Max *lcd* object that CataRT employs as a two-dimensional representation of the sample corpus. As in the previous example, the exported SDIF file—including sample timings, textual metadata, and descriptor values—was imported and transcribed in OpenMusic (Figure 3) and a MusicXML file was exported to Finale for further editing (Figure 4). The gesture was drawn to connect the approximate spectral region of a preceding passage, based on an audio mosaic of rain, to other material that follows, with the goal of a smooth timbral transition. The interlocking timbral associations that result—including nails and fingers on the harp *sur la table*, plectrum *glissando* along the low piano strings, and wire brushes on the vibraphone—would have been difficult to conceive with pencil and paper alone. Indeed the smooth trajectory in timbral parameters does not necessarily match a symbolic pitch trajectory: although short segments comprise rising scales or arpeggios, they do not connect the instrumental registers of the ensemble in the most predictable way. Hearkening back to Lachenmann, the re-contextualisation of extended instrumental techniques in a timbre-driven composite gesture stimulates the composer's—and listener's—imagination unconventionally.

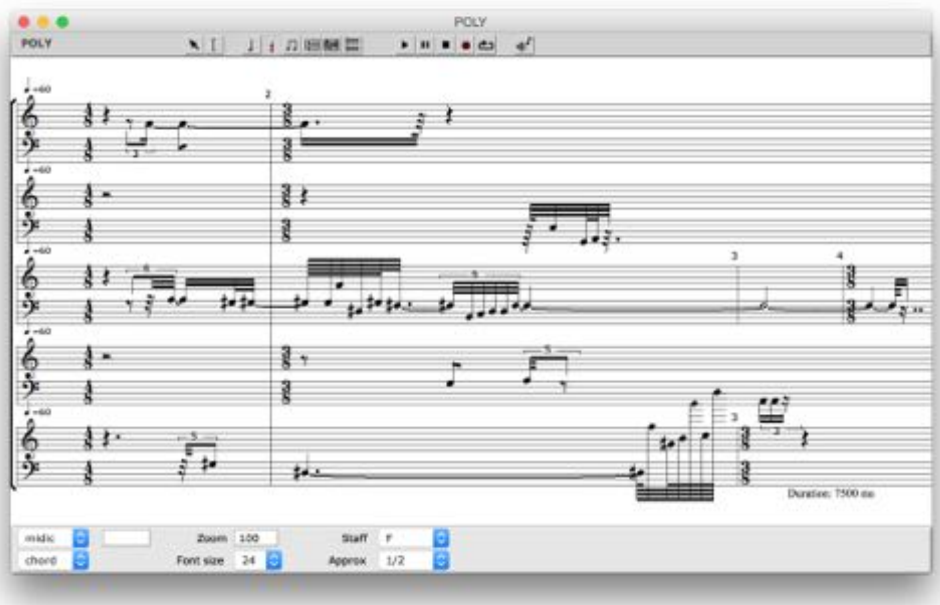

Figure 3. OpenMusic *poly* transcribed from a gesture improvised with a Wacom tablet.

Figure 4. *What the Blind See* mm. 157-159. © Edition Gravis Verlag, Brühl, Germany. Printed with kind permission.

Transformation

A third example is taken from the opening of *What the Blind See*, where transcription is applied to instrumental sample targets themselves instead of "extra-musical" sources. By taking a short instrumental figure from one instrument and transcribing it in varied form for another, the technique extends a classical form of musical imitation: motivic transformation. In this case, the audio recording of one instrument, the viola, is taken as the target for an audio mosaic using the samples of another, either the harp or vibraphone (Figure 5). This is closer to Lachenmann's *musique concrète instrumentale* than *musique instrumentale concrète*, as there is no "extra-musical" referent. However by transforming an acoustic target, rather than symbolic music notation, associations between materials including rhythms and playing techniques may suggest themselves beyond the composers' notational habits, conditioned and distorted by personal and historical practice.

Even more than in the previous examples, this one introduces details of the performers' collaborative interpretations on several levels. The viola part is derived from the earlier work *Beside Onself*, performed and recorded from the notated score by the violist of the ensemble. So the performer's interpretation of the written material, complete with shades of timbral variation and expressive microtiming, becomes the target of the audio mosaic of harp or vibraphone samples. The transcription then takes into account these variations, as well as the variations of attack time and responsiveness of the viola playing techniques. The resulting transcriptions differ in significant detail from the original notated viola part, but preserve its rhythmic and timbral flavour. This transcription takes advantage of a further layer of rhythmic choice: the OpenMusic patch is used to filter the audio mosaic output by CataRT according to a range of successive loudness thresholds, producing transcriptions of different rhythmic densities. The final harp and vibraphone parts in Figure 5 are drawn freely from these different rhythmic layers, permitting greater control over the result.

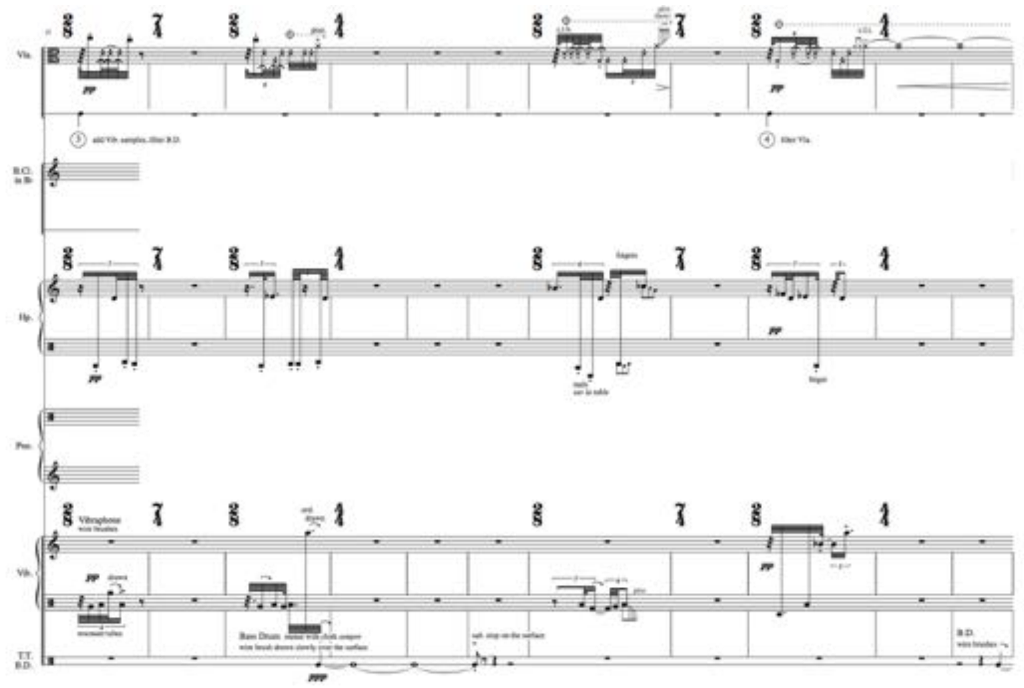

Figure 5. *What the Blind See* mm. 11-21, showing the viola part transcribed for harp and vibraphone. © Edition Gravis Verlag, Brühl, Germany. Printed with kind permission.

Of course all three parts, the viola target and harp and vibraphone mosaics, are subject to similar variations in timbre and microtiming when reinterpreted in live performance. Performed simultaneously in the finished work, the result is a homophonic doubling of the motivic material that fuses in terms of timbre and density. The live electronics (indicated by the pedal marking for cue 3 in Figure 5) respond through a similar process, complementing the viola with a real-time audio mosaic drawn from the harp and vibraphone sample databases. In effect this is another version of the notated harp and vibraphone parts, re-synthesised in parallel with the performance by the live players, complementing with a successive layer of heterophony.

In a related passage later in the work, a dense ensemble texture is transcribed from solo instrumental material. In this case, rather than responding to a notated score, the bass clarinetist improvised on a repeated low B♭ tongued as fast as possible: the natural

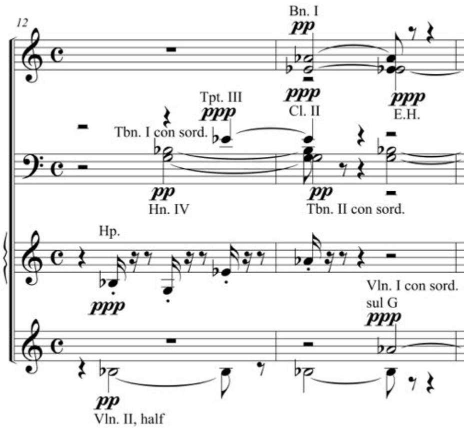

Figure 6. Arnold Schoenberg, *Fünf Orchesterstücke* Op. 16, No. 3. (reduction with C♭ pedal tone omitted for clarity).

variations in speed and response of the instrument produced a subtly shifting rhythm. A mosaic was then transcribed for the rest of the ensemble using sample banks combined from all four instruments excluding bass clarinet. In the finished score the bass clarinet accompanies with a similar figure, ideally producing a timbral fusion between the soloist and the ensemble texture.

Instead of the multiple-stranded heterophony of the opening viola solo, this passage unfolds a virtual timbral polyphony. The mosaic made with CataRT is conceptually monophonic, matching one sample at a time of the corpus to the target. However, when the notes of this line are distributed to the instruments of the ensemble each note is sustained or let ring longer than its monophonic value. The effect could be likened to the verticalisation of a melodic line into a sustained harmony, for example in Schoenberg's *Fünf Orchesterstücke* Op. 16, No. 3 (Figure 6). The result suggests a "timbral pedal" accompanying the soloist (Figure 7).

The rhythmic details of the bass clarinet part were derived from a different process. An audio recording of an improvisation with CataRT, using the mouse as controller, was segmented in OpenMusic using the *transient-detection* object from the OM-SuperVP library (Figure 8). By applying incremental transient detection threshold values, rhythmic patterns of decreasing density are produced. By using OM-SuperVP instead of AudioSculpt[6] the parameters can be flexibly adjusted directly within the OpenMusic patch, and successive analyses can be applied iteratively. The use of a list of threshold values to produce a rhythmic matrix resembles the varying loudness thresholds used to compose Figure 5, however now the results are used formally: each level of the matrix is composed out successively, to produce a carefully-controlled decrease in rhythmic density across a long-term section of the work. The bass clarinet rhythm in Figure 7 is drawn from the fourth line in Figure 8, partway through the process.

[6]The OM-SuperVP library provides the means to run most of the AudioSculpt sound analysis and processing features as part of OpenMusic visual programs. This has the advantage of connecting them directly to the compositional workflow and iterative procedures.

Figure 7. *What the Blind See* mm. 235-238 “virtual polyphony”. © Edition Gravis Verlag, Brühl, Germany. Printed with kind permission.

This rhythmic matrix is hierarchical,[7] as each transient detected at a given threshold is also detected at a lower threshold. However the notated rhythms output by *omquantify* are not strictly hierarchical, as at different levels the same rhythmic position in milliseconds may be approximated to different notated values depending on its context. These differing approximations were retained as-is, motivated both by readability for the performer, and providing an extra level of rhythmic micro-variation for potential expressive use. Such a rhythmic matrix, which has functioned in many of my recent works, could be heard as an extension of common-practice metre. It is intriguingly parallel to the hierarchical rhythmic structure defined by Lerdahl and Jackendoff in their *Generative Theory of Tonal Music* (1983) [12].[8] A compositional realisation can freely draw upon the matrix, crossing rhythmic levels for expressive purposes and underlining

[7] In the sense of Lerdahl and Jackendoff ([12], p. 13); see below.

[8] The rhythmic structure shown in Figure 8 satisfies “Metrical Well-Formedness Rules” 1 and 2, which Lerdahl and Jackendoff state are “defining conditions for metrical structures and are universal”, but not rules 3 and 4, which “define the metrical regularities possible within a given musical idiom”. This makes the structure comparable to a phonological “stress grid” (Fred Lerdahl, personal communication, 4 September 2015).

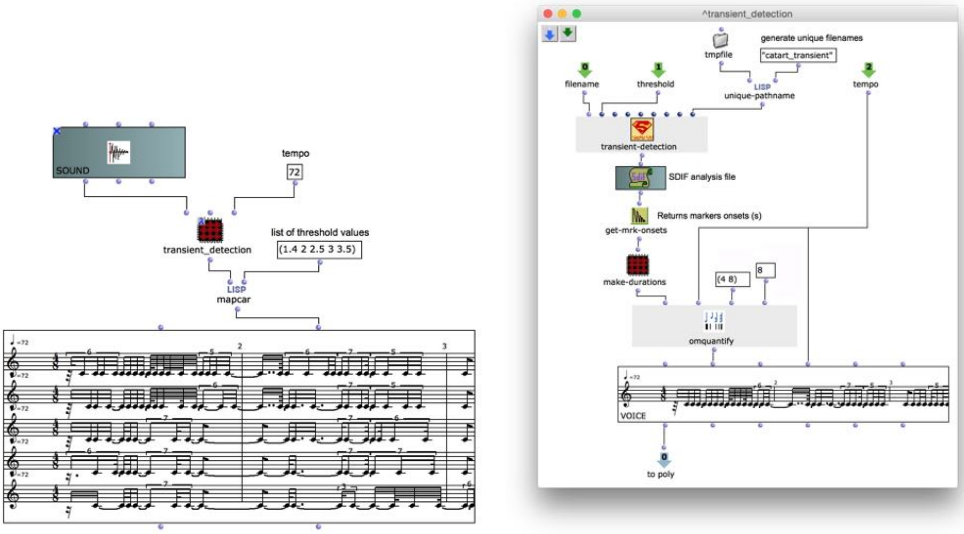

Figure 8. OpenMusic patch to apply *transient-detection* to a sound file at successive threshold values (left) with sub-patch *transient_detection* (right).

metrically strong attacks-points for emphasis. Of course unlike common-practice music, this rhythmic matrix is superposed on a contrasting regular metrical notation, but as mentioned above the metre serves a mere coordinating function.

Imaginary Gardens

In *Without Words* for soprano, 11 instruments, and electronics (2012), the sample databases and field recordings that have informed my work since *What the Blind See* are augmented by a database of vocal texts. The database was drawn from fragments, translations, and paraphrases of writers including Marianne Moore, Wallace Stevens, Matsuo Bashō, Douglas Huebler, Kenneth Goldsmith, and many other prose and poetic texts—chosen because of their connection to transcription, mimesis, metaphor, and place. They were recorded in collaboration with soprano Amanda DeBoer Bartlett who improvised with the texts as well as with various vocal "preparations"—objects to be sung through, including whistle, kazoo, coffee mug, cardboard tube, and spring drum.[9] In analogy to the examples above, here audio recordings of the texts are privileged over their symbolic printed versions.

This vocal database is treated as a mobile resource, ranging from isolated phonemes to full words, phrases, or occasional sentences. Accordingly semantic references may be entirely lost or more or less intelligible depending on vocal technique and whether or not a preparation is present. This continuum of semanticity is exploited in the transcription process, with smaller units available for audio mosaics suggesting other references, and longer units conveying intact linguistic denotations of their own. An example of the

[9] For a detailed account see [6].

former is the opening mosaic, performed by the live soprano, based on a target field recording of frogs. The work opens with the field recording alone, then gradually crossfades it with a recorded audio mosaic of soprano samples, and eventually integrates the live soprano herself in a seamless continuum. The tissue of whispered phonemes (square note heads in Figure 9), principally fricatives, sometimes sung through a police whistle (upper staff), imitates the high spectral range of the frogs, with breathy sung high D5 (diamond note heads) pointing to a prominent frequency of their call. The same field recording receives a contrasting transcription for the full instrumental ensemble a few minutes later in the work, by a similar technique to Figures 1 and 2.

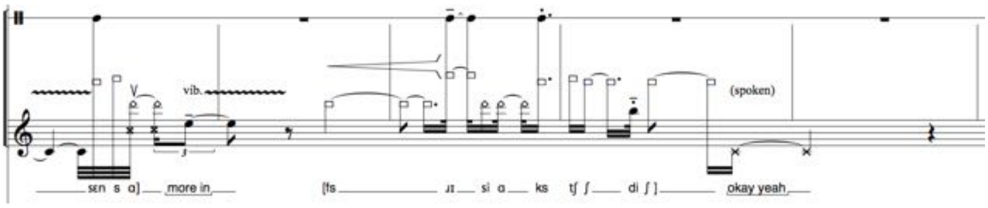

Figure 9. *Without Words* mm. 65-69 soprano part. © Edition Gravis Verlag, Brühl, Germany. Printed with kind permission.

In the vocal transcription shown in Figure 9, most of the phonemes are far too short for their source texts to be identifiable, while a few open out into words or phrases ("more in", "okay yeah"), hinting at the semantic origins of the utterances. Instead the connotation of the underlying field recording is more salient. Its compositional sense is later revealed in a soprano citation of Marianne Moore: "imaginary gardens with real toads in them".

Over the course of the 18-minute work, these textual fragments are hesitantly augmented, like folding fans that expand and contract. Eventually some vocal samples are exposed in their entirety, like the Marianne Moore text above. For another text, a haiku by Matsuo Bashō translated by Robert Hass, the simplest possible transcription process is applied: the original soprano sample is re-notated as faithfully as possible with the aid of AudioSculpt and OpenMusic. A *chord-sequence* analysis is exported as an SDIF file, imported to OpenMusic, and the most salient monophonic pitch is chosen from each segment. Nevertheless the "accuracy" of the result is belied by the subjective decisions throughout the process, essential to capture the fine shades of rhythm and pitch (Figure 10).

Figure 10. *Without Words* mm. 266-69, soprano part.

When the same soprano who created the original sample reinterprets the transcription, she reproduces the historical/autobiographical moment of the sampling session—the sample becomes a field recording. Yet as the soprano relearns to perform "herself" she underlines the tension between performance, transcription, and referent. As Peter Ablinger says of his *Weiss/Weisslich 36*, "the same is not the same. There is a difference. At least the difference between just being here and: listening".[10] Perhaps the same difference lies behind Basho's riddle-like haiku.

Traces

Textual fragments are taken as targets for audio mosaics as well as corpora. Brief excerpts of Wallace Stevens's own recorded reading of *Credences of Summer* are cited, placing historical recordings of the poet in dialogue with the live soprano. One recording was taken as the basis of an audio mosaic using a database of bass flute samples. The patch in Figure 11 adds another feature to the CBCS workflow: textual metadata output by CataRT are recorded in the SDIF file in name-value tables. Then using OpenMusic objects *getnvtlist* and *find-in-nvtlist* the values can be read and used to inform the manual editing of the score in Finale, as seen in Figures 11-12. For example, evaluating the *mapcar* box labelled "SoundFile names" (Figure 11) yields the filename "bfl-closed_o_harmonics-C3" for the second sample (in both the *poly* object in Figure 11 and the score in Figure 12). The pitch calculated by *yin~* is discarded and replaced with fingered pitch C3 documented in the filename. To indicate the flutist's closed embouchure (lips sealed around the mouthpiece), a black circle is added over a hollow rectangular note head. Finally, the flutist is directed to shape her mouth into the phoneme *[o]*. This process is repeated for each sample to build the bass flute part.

Reminiscent of the technique of Figure 8, this audio mosaic was made in CataRT with a range of different segmentations based on loudness threshold to produce mosaics of different densities. Beyond the contextual variations in rhythmic transcription produced by *omquantify*, there are also variations in the units chosen by CataRT according to the segmentation. As CataRT calculates an average of descriptor values over each segment, segments of different lengths corresponding to the same position in the target will have slightly different average values, and as a result the best match in the database might be different. So successive rhythms and playing techniques in the final matrix do not align precisely, even though derived from the same mosaic. This is a case in point of how slight variations in parameters can produce significantly different musical outcomes. The durations of the bass flute samples themselves also come into play, for example the 4-second tongued overtone *glissando* that concludes the third line of Figure 12. During transcription a subjective compromise is made between the durations of the target segments and the durations of the samples segments in the database.

These bass flute lines are reinterpreted in the finished work accompanied by repetitions of the Wallace Stevens target sample itself: "Trace the gold sun about the whitened sky/Without evasion by a single metaphor". As the sample loops, the varying bass flute lines are deployed in a gradual densification, in contrast to a gradual *descrescendo* by the voice of Stevens. Instead of a syllable-by-syllable transcription, the bass flute traces a more subtle mosaic privileging the weight and density of the grainy historical recording, and emphasising the timbre of the poet reading over the sense of the printed text.

[10]See `http://ablinger.mur.at/docu1515.html`

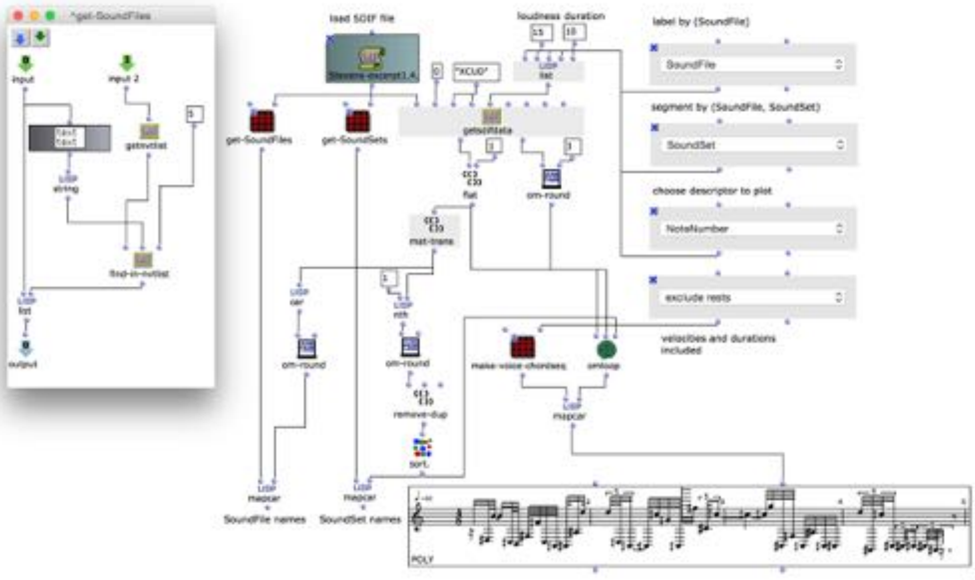

Figure 11. Patch for *Without Words* with sub-patch *get-SoundFiles* on the left.

Figure 12. Transcriptions of the same Wallace Stevens sample for bass flute at increasing loudness segmentation thresholds, subjectively edited in Finale.

At the end of the work, the field recording, the Stevens quotation, and their varied transcriptions return in a collage-like coda. While the recapitulation is perceptible on a semantic level, it also suggests an abstract musical return. The recapitulation of the same constellation of timbres, playing techniques, and spectral descriptors defined by a concrete recording and/or its transcription can be heard in analogy to the tonal return of common-practice music. In fact, given the statistical interpretation of tonality by scholars like Carol Krumhansl [10] and David Temperley [19], as well as the spatial tonal models advanced by Fred Lerdahl [11] and others, a direct connection could be made between the pitch distribution of a tonal centre and the statistical distribution of descriptor weights in a soundscape. Tonality as a place: a topic that invites further exploration.

Translation

Since 2012, transcription with CataRT and OpenMusic has advanced in several directions: the bach package for Max has been integrated into the CataRT workflow to facilitate real-time interaction with musical notation. Bach can also expedite exchange of data with OpenMusic, which remains useful for high-level computer-aided composition tasks. CataRT itself is available in a new version incorporating the MuBu package of Max externals, improving its portability, clarity, and especially permitting access to larger audio databases by taking advantage of 64-bit Max.

Posing the question of how mimesis can be generalised beyond strict transcription, I integrated CataRT with a machine-learning algorithm that permits the association of audio not only by descriptor similarity, but by shared context. Inspired by the existing program OMax [1] and the PyOracle package for Python [18], the resulting tool CatOracle is available in the MuBuForMax-1.8.5 distribution and later. It combines the powerful *Audio Oracle* algorithm for musical pattern matching with CBCS for applications including computer improvisation, high-level control of synthesis, computer-assisted composition, and musicological analysis. CatOracle extends the notion of transcription beyond a reproduction coordinated linearly in time, permitting an “improvisation” based on a target that can backtrack, repeat, or skip. Yet the shared contextual relationships compel the result to bear a degree of resemblance of the original. This could be described as behavioural transcription or “style imitation” [3]—or more generally “translation”,[11] a term chosen to suggest displacement of elements along the time axis, as well as expression through a new improvisational language. Applied to machine improvisation, it is especially complementary to other sites of spontaneity in the transcription process: collaborative improvisation during sampling sessions, gestural control with the mouse or Wacom tablet, and live reinterpretation of expressive micro-timing.

A first compositional application in *Xylography* for violoncello and electronics (2015) begins with an acoustic improvisation by cellist Pierre Morlet, transcribed in detail in the score and reinterpreted in performance, translated live into multiple strands of computer improvisation, the courses of which are further guided by ongoing transcription of the live performer. This layered process of mimetic and creative feedback points to promising territory at the intersection of composition, transcription, improvisation, and interpretation.

[11] An homage to the “central dogma” of molecular biology: “replication–transcription–translation” [4].

References

[1] Gérard Assayag, Georges Bloch, Marc Chemillier, Arshia Cont, Schlomo Dubnov. "OMax Brothers: A Dynamic Topology of Agents for Improvisation Learning". In *Proceedings of the ACM Multimedia Workshop on Audio and Music Computing for Multimedia*. Santa Barbara, 2006.

[2] Michel Chion. *Guide des Objets Sonores: Pierre Schaeffer et la recherche musicale*. Paris: Institut National de l'Audiovisuel & Editions Buchet-Chastel, 1983-1995. English translation: *Guide To Sound Objects. Pierre Schaeffer and Musical Research* by John Dack and Christine North, 2009.

[3] Arshia Cont, Schlomo Dubnov, Gérard Assayag. "A Framework for Anticipatory Machine Improvisation and Style Imitation". In *Anticipatory Behavior in Adaptive Learning Systems*. Rome, 2006.

[4] Francis Crick. "Central Dogma of Molecular Biology". *Nature*, 227, 1970.

[5] Nicolas Donin. "Sonic Imprints: Instrumental Resynthesis in Contemporary Composition". In Gianmario Borio (ed.) *Musical Listening in the Age of Technological Reproducibility*. Farnham: Ashgate, 2015.

[6] Aaron Einbond. "Subtractive Synthesis: noise and digital (un)creativity". In Aaron Cassidy, Aaron Einbond (eds.) *Noise In And As Music*. University of Huddersfield Press, 2013.

[7] Aaron Einbond, Diemo Schwarz, Jean Bresson. "Corpus-based transcription as an approach to the compositional control of timbre". In *Proceedings of the International Computer Music Conference*. Montreal, 2009.

[8] John M. Grey. "Multidimensional perceptual scaling of musical timbres". *Journal of the Acoustical Society of America*, 61(5), 1977.

[9] Vijay Iyer. "Embodied Mind, Situated Cognition, and Expressive Microtiming in African-American Music". *Music Perception*, 19(3), 2002.

[10] Carol L. Krumhansl. *Cognitive Foundations of Musical Pitch*. Oxford University Press, 1990.

[11] Fred Lerdahl. *Tonal Pitch Space*. Oxford University Press, 2001.

[12] Fred Lerdahl, Ray Jackendoff. *Generative Theory of Tonal Music*. Cambridge, MA: MIT Press, 1983.

[13] Theodore Levin. *Where Rivers and Mountains Sing*. Bloomington: Indiana University Press, 2010.

[14] Rainer Nonnenmann. "Music with Images—The Development of Helmut Lachenmann's Sound Composition Between Concretion and Transcendence". *Contemporary Music Review*, 24(1), 2005.

[15] James O'Callaghan. "Mimetic Instrumental Resynthesis". *Organised Sound*, 20(2), 2015.

[16] David Ryan, Helmut Lachenmann. "Composer in Interview: Helmut Lachenmann". *Tempo*, 210, 1999.

[17] Diemo Schwarz, Grégory Beller, Bruno Verbrugghe, Sam Britton. "Real-Time Corpus-Based Concatenative Synthesis with CataRT". In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*. Montreal, 2006.

[18] Greg Surges, Shlomo Dubnov. "Feature Selection and Composition using PyOracle". In *AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*. Boston, 2013.

[19] David Temperley. *The Cognition of Basic Musical Structures*. Cambridge, MA: MIT Press, 2001.

[20] Tolga Tüzün. "Maquette As Data Structure and Synthesis Agent in *Metathesis*". In Jean Bresson, Carlos Agon, Gérard Assayag (eds.) *The OM Composer's Book 2*. Editions Delatour France/IRCAM-Centre Pompidou, 2008.

[21] David Wessel. "Timbre Space as a Musical Control Structure". *Computer Music Journal*, 3(2), 1979.

Acknowledgements: I thank the musicians of Ensemble L'Instant Donné with whom I developed the instrumental samples and techniques for *What the Blind See* and who premiered the work at the Centquatre in Paris on June 12th, 2009. I thank soprano Amanda DeBoer Bartlett and the musicians of Ensemble Dal Niente with whom I developed the samples and techniques for *Without Words* and who premiered and recorded the work in Chicago on my portrait album released by Carrier Records in 2012 (`http://carrierrecords.com/index.php?album=einbond`). I gratefully acknowledge researchers at IRCAM including Diemo Schwarz and members of the ISMM team for developing and supporting the CataRT package, Eric Daubresse for production support on *What the Blind See*, Jean Bresson, Yan Maresz, Alexis Baskind, Emmanuel Jourdan, Mikhaïl Malt, Jean Lochard, and Cyril Beros. My participation in two years of the Cursus in Composition and Music Technology was made possible by the Fulbright Scholarship and Georges Ladd Prix de Paris.