Realizing Musical Gestures with the Computer: Paradigms and Problems

By Christopher Bailey

I have a gesture, a musical shape, in my head, and I want to translate it into actual sound. A simple task, it would seem; after all, this is my job, I am a composer. Why should this process be more difficult on a computer (where I have, supposedly, complete and direct control over sound), than with live musicians, where my ideas are (at the very least) twice-filtered—through the sieve of notation and through the sieve of the mind of the reader of that notation, with all of that mind's training(s), tradition(s), etc.? How do I approach the creation of a gestural language on this instrument that has so recently blossomed into a real compositional tool? To answer these questions, I will step back for a moment, and discuss briefly some aspects of the composition, notation, and performance of musical gestures for acoustic instruments.

When writing for the latter, I am heir to what might be thought of as a huge bias—a lens, prism, or filter—through which any gesture notated in the Western system of notation becomes, as it is played by Western-trained musicians, related or relatable to the vast repertoire of traditional Western musical gestures, built up over centuries. The action of this filter is reinforced through the traditional structure of Western (classical) music-making:

(1) By the composer, because of the fact that I rarely (relative to the number of gestures I compose) choose to write something that isn't part of this inherited repertoire in some way, and because my mind, aiming for efficiency, instinctively guides me, whenever possible, toward making use of what I already know; and through the fact that notation itself also filters my compositional output.

(2) By the performer, who interprets whatever events I notate, no matter how exotic, complex, or bewildering, into something at least dimly relatable to a traditional gesture.

(3) Finally, by the listener, who, when trying to make sense of a performance of my music, will, again, no matter how unfamiliar it might be, (try to) relate it to his or her own cultural experience (for most of my listeners, the Western concert tradition).

One can look on this filter with a kind of resignation, perhaps feeling that there is, and never will be, anything new under the sun. But let's take a look at some of the things that happen when a composer tries to resist

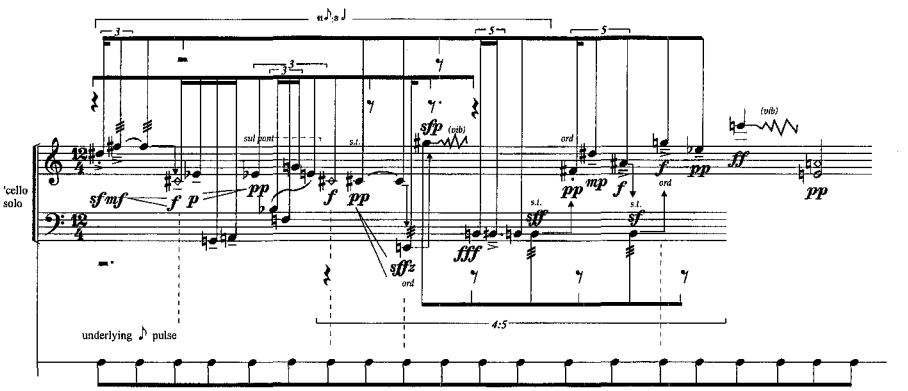
Current Musicology 67 & 68

the action of this filter, when this reliance by composer and performer on tradition is stretched nearly to its breaking point.

The score to *Megalomaniac,* for solo cello, contains a number of passages of rather nasty-looking notation. Part of my purpose in writing these passages was to see how the performer would come to terms with the notation, to see what, in fact, the performer would do to "traditionalize" these hideous "things" (pieces of notation). That is, I wanted to see how the performer would pass (or perhaps, *squeeze*) the musical work through the filter that I've been talking about—an action they must commit for the sake of sheer "survival," negotiating some way to wade through the complexities of the musical moment.

Here is an example of such a passage:

Figure 1: *Megalomaniac,* for solo cello, excerpt.



How might a performer approach this passage? I offer some advice in the performance instructions to the piece:

> The piece might be approached as follows: learned at first in a "lick-by-lick" manner, perfecting each individual gesture and giving the said gesture a maximum of expressive and dramatic content as suggested in the score. At this point in the learning process, the player should be concerned, more or less, only with the basic 8th-note pulse, and how the activity above relates to that pulse. In other words, at this beginning stage of learning the piece, it can be treated as a "graphic" score, rhythmically, with the stipulation that all gestures be learned with respect to the underlying 8th-note pulse. (The pulse is indicated underneath the score.)

> After the individual gestures are learned, the player then proceeds to string them together, into larger and larger formal units. At first, this process should still take the rhythmic point of view of "graphic score against basic pulse," but as the general flow of the work comes into fruition, the player should attempt to feel the larger-scale rhythmic strands that are interacting. For example, in a passage that is composed of a 7-tuplet and an 11-tuplet strand interacting, "feeling" the passage "in 7" or "in 11" will reveal different shades of meaning brought about by different weightings of rhythmic strands. It is this kind of interpretive exploration that I hope the piece's complexity, in terms of its rhythmic notation, will inspire. (Bailey 1997)

It is in the "perfecting [of] each individual gesture and giving the said gesture a maximum of expressive and dramatic content" that the filter of tradition will no doubt come into play in the strongest way. On the other hand, the "stringing together into larger and larger formal units" is where something new happens. I like to think of this piece (and others written in a similar vein) as a series of gestures, many with strong associations, musical or extramusical, but ripped out of context, and with those associations "left hanging," perhaps posing unanswered questions, often with these gestures toppling over one another, frequently denying or canceling one another's associational implications. This makes for a difficult musical experience, for performer and listener, one which taxes one's ability to concentrate and give each gesture the focus it deserves. To pose that kind of challenge was, in this piece, one of my goals.

With acoustic instrumentalists, this filter, this bank of assumptions is in operation. When you give them something different, something wild, outside of their experience and training, they attempt (assuming they approach it in good faith) to give it what is called a "musical interpretation," to render it as some warped form of (their) musical tradition. It is this interaction that I often seek when composing for performers.

It is important to note that many aspects of the filter are built into the physical characteristics of instruments themselves, their methods of sound-production, and the way a player moves to cause that sound-production. This is, in turn, passed on to notation. If I want to get a particular gesture out of an instrument or group of instruments, I know how to notate it, based not only on my musical aural training but also on my ability to imagine the gesture's physiological instantiation in performance.

With the computer, especially with synthesized (as opposed to sampled) sound, this is not the case. Everything must be done from scratch. When I first started to realize music on computer, the following would often occur: I would think and hear a gesture in my head—"oomph"—and attempts

to realize that gesture, through programming, sound synthesis, mixing of samples, and so on, would result in "aaamph" (so to speak)—not quite (or sometimes quite distant from) what I originally wanted.

At that time I had little knowledge of acoustics, so it was often difficult to ascertain why a gesture wouldn't come out the way I wanted it to. Now, armed with a greater knowledge of acoustics, I often know *why* a gesture doesn't come out the way I want it to, but it usually turns out that the *how* of correcting the problem would lead me into a complex, low-level web of research into software and acoustics—a path that (although I have some interest in it) I'm not really qualified to follow, and I'm not really interested in following: as a composer, I want to be given a set of reasonably flexible tools, and make use of them to create effective music. It is, of course, possible to make quite effective computer music without losing oneself on that full-fledged programmer/acoustician path. But in that case, the approach one takes towards the creation of gestures, and thence to complete compositions, is fundamentally different from the one taken when writing for acoustic instruments.

Put simply, I take less of an "I've got to get *this* gestural effect" kind of approach, and more of a "Let's experiment with what this machine does: generate some musical material consisting of gestures whose characters and effects I can't quite predict, and figure out how we can modify the musical *contexts* in which we place those gestures so that the gestures 'work' (musically and dramatically)" kind of approach.

The experimental process whereby I "generate some musical material" is partially an intentional one, partially an arbitrary one: I might begin by trying to get a certain gesture out of the machine; what comes out is something different from what I had in mind originally. I may then try to modify the gesture to get it closer to what I wanted originally, changing the parameters I gave the machine to create the material; but eventually I change paradigmatic gears entirely, and I begin to think about how to shape context A to accept or fit gesture X instead of how to achieve gesture X to fit context A—in other words, *creating* the context that will make things seem as if the accidentally created gesture was not accidental, but created *for* the context.

This is part of my composing process in works for acoustic instruments as well, but it has a special relevance for computer music, since the relationship between what I tell the computer to do and the resultant sound is far less well understood than the relationship between what I tell an acoustic instrumentalist to do, notationally, and the sound that is produced. The plethora of unexpected material coming out of the machine demands this approach.

Of course, as one continues with this looser, more experimental approach to sound and sound-gesture creation, something that starts to re-

semble a "tradition" is built up, amazingly fast. One learns quickly the kinds of effects that most often result from certain methods of sound synthesis on the computer. Frequency modulation, amplitude and ring modulation, physical models of instruments, different types of sound processing, and so on, all have their characteristic tone colors or families of tone colors. Extrapolating from there, typical resultant gestures or families of gestures, with corresponding emotive associations, arise from the use of a given computer music tool.

Computer music history itself is, in large part, the story of a series of discoveries of new sound-generation techniques, each followed by a flurry of excitement and new pieces that use the technique; then the gradual realization that, as Milton Babbitt put it, "Nothing gets old faster than a new sound." The technique is then absorbed into the community as simply another tool for making sound.

This lesson of history I have taken to heart; hence, I try to focus my compositional energies not (entirely) on the method of synthesis of a particular gesture, but on its harmonic, rhythmic, and timbral content/context. I deliberately do not seek out "new" sounds; rather, I try to use old ones, to combine and recombine them in a rapid kaleidoscopic fashion to produce event-complexes in which the interaction of different sound components, though they may each be individually familiar, yields a combined event-complex that, in a subtle yet striking way, is something new.

Thus, in my first mature computer music work, *Ow, My Head,* I decided from the start not to utilize any kind of synthesis or processing at all, but to deploy into the musical fabric only raw, unprocessed, recorded *musique concrète* sounds from the environment. In all of my pieces that use this type of material, the sounds are usually recorded in one *place* (in the case of this piece, the house where I grew up, in the 'burbs of Philadelphia). Although this certainly does not provide any source of immediate sonic or musical unity, it does provide for me, psychologically, a desired *spiritual* unity—a unity of spiritual source, so to speak.

I'll say a word about why I choose to use concrète sounds in particular (out of all the choices of material provided by computer music composition). What I find fascinating about the use of "found sounds" is the emotional effect of the *displacement* of a sound—a sound with a very clear origin —*from* its origin. Each sound brings with it an illusion of its original space or place, be it a kitchen, a washroom, a subway train, or whatever. Yet the sounds are brought together in a "musical" space. I find that the interaction between these spaces has a powerful emotional effect in the mind of the listener. It is similar to the effect one experiences while watching a movie with bright, sunny, images: sometimes the mind loses itself in the illusion of the sunniness, then it realizes that all of this is taking place in the

darkness of a movie theatre, possibly in the dead of winter around midnight. There is something almost frightening about this paradox (similar in character to the idea from physics that 99 percent of matter is actually empty space). This kind of spooky disjunction is the emotional basis of my attachment to concrète sounds.

On a more technical level, musique concrète is (still) a wide-open field of discovery, in terms of the idea, mentioned above, of combining sounds together rhythmically, harmonically, and timbrally to produce new event-complexes or meta-timbres. About *Ow, My Head,* I am often asked how I processed (filtered, reverbed, delayed, etc.) or synthesized certain sounds. In fact, there was no processing, no synthesis; instead, the particular *combinations* of raw sounds in the piece, achieved only through relative rhythmic and amplitude adjustment, produced the "new" meta-timbres.

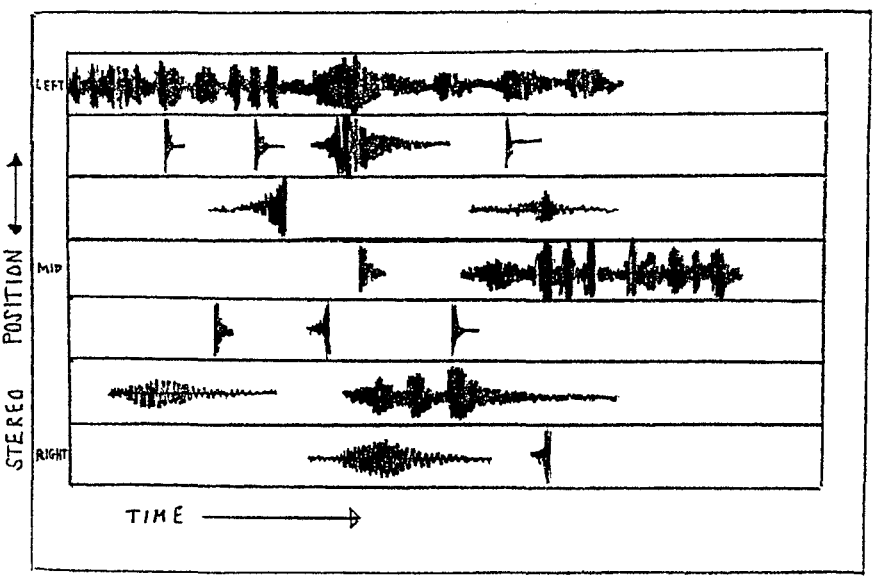
The piece was composed in small blocks, each consisting of a sequence of only a few gestures (something analogous to a phrase). Later, the blocks would be joined to form sentences and, eventually, a complete form for the work.

The process of composing a gesture, or a small sequence thereof, was, in general, something like this: "Spray" a random set of sounds (a subset of the complete set of about 200 sounds, varying in duration from 0.2" to 5" or so) into a mix. Adjust their rhythmic positioning, amplitude, etc. to make a viable musical gesture. With appropriate rhythmic and amplitude adjustment of the component sounds in a mix, a context will be created whereby every sound fits, and coherent line(s) of rhythm and pitch (arising from the sometimes obscure, sometimes quite clear, pitched qualities of found sounds) will be formed. Later, the gestures themselves are treated the same way, to make phrases and, eventually, the complete piece.

To explain exactly what I mean by "spraying" sounds randomly into a mix, it might help if I say a bit about sound-mixing programs. The principle of all of these programs is quite simple. A visual display is used to represent the sounds and their placement in time and (stereo) space. The *x*-axis represents time, the *y*-axis represents stereo position, from far left to far right. Each sound is represented by a shape, which represents its amplitude curve (see fig. 2).

Note that the tricky part about this is that pitch-content is not displayed in any way. With found sounds, pitch-content is often complex, so that a simple "score" representation (i.e., each sound having a single fundamental "pitch") would be problematic. In the case of this piece, I relied mostly on my ear and aural memory to keep track of what sounds were associated with what pitches, in order to be able to construct contours and harmonic combinations of the pitch-contents of different sounds.

Figure 2: Typical mixer image.



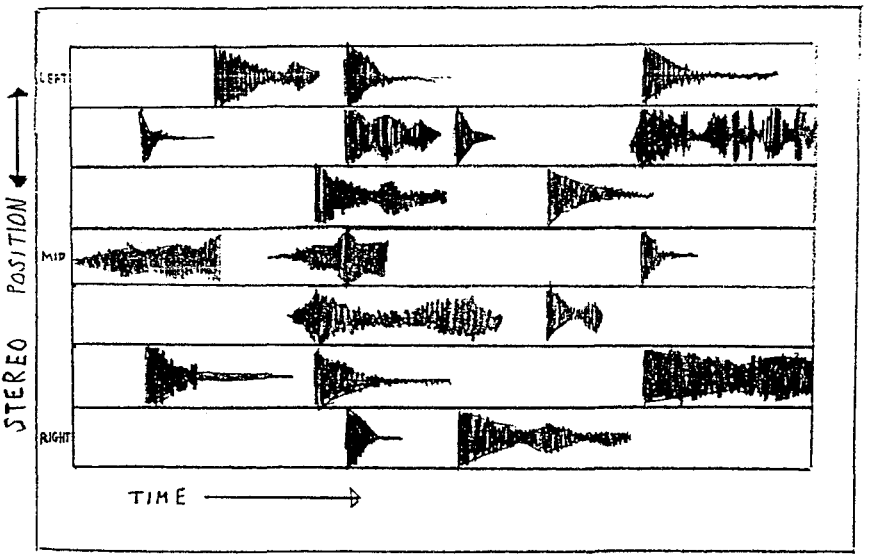
The next step is to hone the "sprayed" mix that appears. There are a few common methods I use to hone the randomly generated sound-sequences:

1) The most common technique is to line up attack-points (or points of high amplitude) between selected sound objects (see fig. 3). Since the ear will often hear several sounds with the same attack-time as a single, new, combined sound or timbre, these kinds of events probably account for people's questions as to what processing and/or synthesis techniques I use: the new events seem familiar, yet skewed in some way.

2) Frequently, I will use these simultaneous attacks as goals (or origins) of rhythmic activity for preceding (or succeeding) sound complexes. Then I will use increasing or decreasing density of sounds (i.e., *accelerando* or *decelerando*) as rhythmic patterns of approach to, or departure from, these goals.

Hence, the music (especially in this piece) often becomes a series of waves. I like to think of my use of waves as being analogous to Elliott Carter's use of wave-forms in his large ensemble pieces (*Concerto for*

Figure 3: Lining-up attacks 'twixt sounds.



Orchestra, Double Concerto, etc.). As in Carter's works, some of the waves in *Ow, My Head* are composed of simultaneous, overlapping tempi.

One of the ways I like to work as a composer is in a kind of dialectical manner, taking aspects of two seemingly unrelated compositional languages, and combining them to form something interesting or expressive in some way. In this piece, I was interested in taking the rhythmic language of the "uptown" New York composers (Babbitt, Carter, Davidovsky, etc.) and applying it to a sound-world not explored by these composers, that of concrète music.

Hence, Carter's waves of overlapping pulses in different tempi, Babbitt's rapid, unpulsed, and highly individuated rhythmic cells, and Davidovsky's play of different timbres on the same pitch are all elements to be found within this piece.

3) Returning to the subject of different methods of building musical gestures with concrète sound material: More difficult to explain (mostly because it depends very heavily on the particular sounds used in a particular context) is the use of timbral/harmonic characteristics of the sounds themselves to guide their placement in relation to one another. Often this amounts to something similar to common-tone modulation in tonal music; we might call it "common-partial modulation." Thus, two successive sounds may be very different in terms of features such as attack-hardness,

fundamental pitch, presence or absence of internal repetition or agitation, etc., but the ear will still hear these timbral/harmonic connections between them; or, a certain sound might "fade in" from another's timbre, entering in a smooth blend (having several common partials) with the first, thus forming a line begun by the first sound. A chain of such relationships can create a continuous line of timbral change (see fig. 4).

In either case, the continuity of certain partials allows the ear to hear the sequence as developmental, and is thus an important way of achieving the coherence of a gesture or phrase.

There are other, analogous ways of achieving continuity and coherence. For example, noisier sounds, with no strong individual partials, can be thought of as frequency bands of noise in a given register. Thus they can lead smoothly to other acoustically and spectrally similar sounds (see fig. 5).

Another kind of progression illustrates the exploitation of a psychological-analogy relationship: in *Ow, My Head* (4:07) the sound of a toilet flushing (essentially a band of noise, acoustically) and the sounds of vocal weeping (vocal tones with downward glissandi) are heard in counterpoint. To my ear, this meshing works particularly well, and the reason is not an acoustic one; rather, it is because both sounds communicate a sense of *down:* toilets flush *downwards,* and weeping involves *downward* motion (of musical pitch, spirits, tears, and so forth).

I would like to mention a few observations concerning large-scale form that I made while composing *Ow, My Head.* Many of the sound-objects in the piece return later, still unprocessed, but recombined in various ways. (These returns are usually at some distance from the original appearances, for I wanted to avoid the "sampler" effect of repeating a recorded sample immediately.) One of the most prominent of these returning *leitmotivs* is a set of long, vocal tones, often combined to create a choral harmony, tuned approximately (see fig. 6).

No doubt its perceptual prominence is due to the fact that it functions as a sort of signifier of traditional pitched-instrument composition. (This is another example of that "tradition filter" affecting the listener's end of things.)

The gesture-sequences of the work were composed independently, without any thought (at the time of their composition) of how they would eventually be strung together into a larger form. This lack of precompositional large-scale formal planning was intentional, for I enjoy, as part of the compositional process, watching large-scale relationships (echoes/flashbacks, premonitions, *leitmotivs*—like the vocal chord) appear spontaneously as I mold the gesture-sequences into an effective large-scale musical form.

Figure 4: Continuity through common partials.

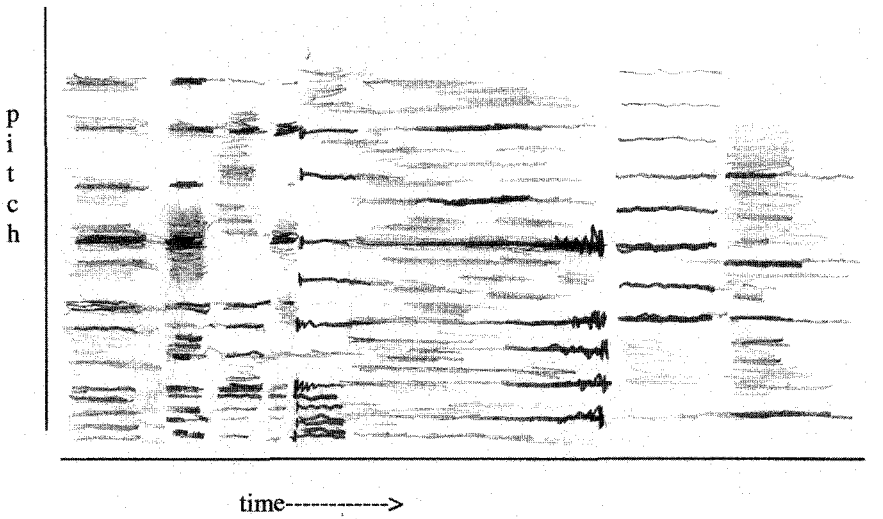


Figure 5: Continuity through similarly pitched bands of noise.

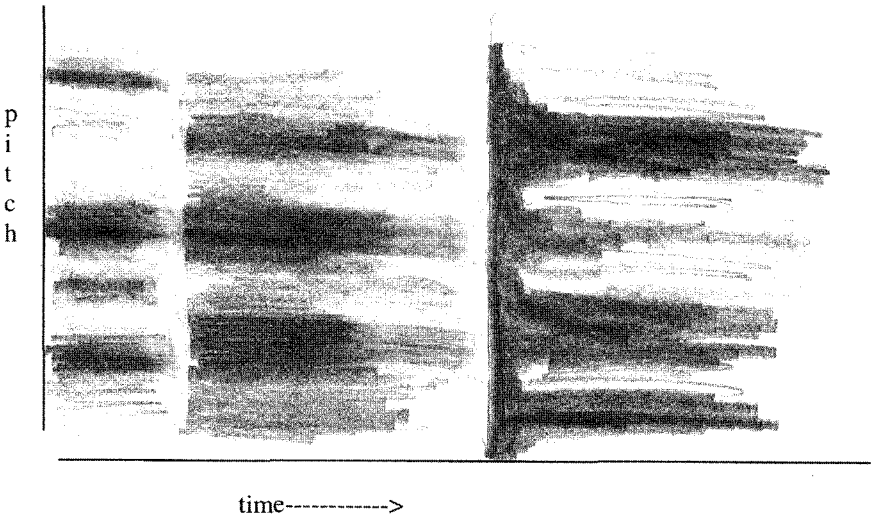
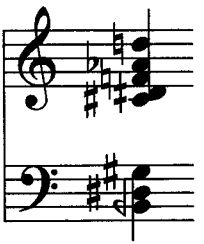


Figure 6: *Ow, My Head,* vocal chord "leitmotiv."



The ending of the work gave me some difficulty. Originally, I wanted the piece to lack any sort of obvious climax, to end abruptly, and thus to be a sort of window onto a sequence of gestures, a sequence that flowed nicely but did not necessarily feel the need to go anywhere. In the end, I felt I wasn't achieving this goal effectively, and thus the second half of the work became a more consciously kinetically formed event-sequence—a build to a climax.

Ow, therefore, ended with a fairly traditional kinetic build-up. *Duude,* my next computer music work, ended up relating to musical tradition by being, formally, a kind of rondo-like alternation between two textures.

One of these textures came into existence as the development of one sonic idea. Occasionally, a single sound suggests an entire sequence of gestures. In *Duude,* one of the sounds I found (a creaking door), when slowed down by a factor of about 20 (without changing the pitch), produced a sound that reminded me of some sort of blaring, "dirging," medieval, bass trumpet. I decided to make this the entire basis of certain sections of the piece. I deployed a single, long line of "door trumpet," counterpointed against smaller fragments and phrases of itself, to create an entire ensemble of door trumpets.

The second of the main ideas making up the quasi-rondo came from my desire to achieve maximum rhythmic density, for at least parts of the work. In *Ow, My Head,* I had been more concerned with the idea of individuated, clear, musical gestures, or small sequences of gestures. In *Duude,* I wanted to achieve a massive gestural density—one in which the individual sounds could still be more-or-less clearly made out, but where their toppling over one another would create a continuous, frenetic web of sounds.

To achieve the "frenetic web" texture, my working procedure went something like this:

To begin with, as in *Ow, My Head,* I had a collection of found sounds, about 300 of them, the basic material for the work. Most were very short—objects (bottles, plates, silverware, etc.) being scratched, hit, rubbed, etc. I decided that these high-density "wads" of concrète counterpoint would

be, at most, a minute long each, a minute into which I'd pack all 300 of my basic sounds. The procedure for making these wads thus became one of randomly spraying the 300 sounds into the first minute of the mix, then, as in *Ow, My Head,* adjusting the timing of the sounds in the wad so that each sound would flow, lead, or leap into the next one(s). The difference was that this time, high-density sound per unit time was a guiding desideratum.

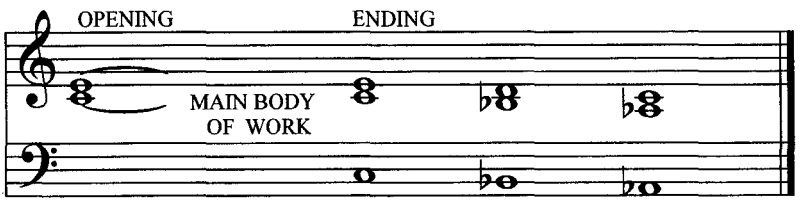
In this piece, I also began to worry just a small bit about issues of large-scale pitch structure. In *Ow, My Head,* I had wanted to leave the pitch domain in a "primitive" state. That is to say, only on the local level, where the harmonic/timbral content of the different sounds led me to sequence them intuitively in a certain way, was there a pitch structure of any kind. This had been an interesting departure for me, since in my acoustic-instrument pieces, I'm fond of using various types of algorithmic techniques (e.g., especially serial) to generate pitch-structures. The development of the latter is usually the first stage in the (pre)compositional process of these works. On the other hand, in these computer music works, any pitch-structure was more of a resultant, a by-product of the random sound-spraying and local rhythmic adjustment.

This was certainly the case with *Ow, My Head.* With *Duude,* I decided to introduce a very simple large-scale pitch structure. One of the sounds I collected, that of air being blown through a bottle, was pitched on a middle C (C4). I decided to build a major third on this pitch; this dyad is emphasized near the beginning of the piece. At the end of the work, in the first explicitly pitched and "synthesized" texture of this piece, this third returns, but this time followed by a slow descent through two other thirds, to make the whole-tone scale progression shown in figure 7. This structure, simple as it is, manages to quite effectively impart a sense of rest and ending to a work filled to the brim with density and activity.

In my next major computer-music work, *Ooogaaah: Dungeony Specimen Spaceship,* I dealt with pitch in more complex ways. I was also dealing with the idea of gesture on a new level, since the piece was written in collaboration with dancer/choreographer Ania Majewska.

This aspect I found to be particularly inspirational. As a composer, I'm very susceptible to "metaphorizing" visual images, structures, gestures, and so on, into sound. (Perhaps this is the reason for my concern with "gesture" to begin with: music as a sonic modeling of bodily—or other—movement.) The collaboration process behind this piece was interesting: at first, we agreed upon a large-scale kinetic (density/energy level) form for the work (see fig. 8). (It happened that, eventually, when all of the music had been composed and realized up to 66, we decided to end the piece.) This, then, was the gesture of the piece on the large scale: a gradual loss of

Figure 7: *Duude* "pitch structure."



energy. I also mapped this gesture onto other domains. For example, the piece gradually moves from humorous/silly to a more "serious" mode of expression: it begins with more noisy or percussive sounds, and moves towards being dominated more and more by pitched music. Registrally, the piece develops from activity localized in the middle register, to activity filling extremes of low and high; there is also a gradual process that reveals the harmony upon which the piece is based.

Our agreement at first was to divide the piece into small sections; I would compose music for the first small section, she would compose dance for the second, I'd compose music for the third, etc. Then we'd switch over and compose our respective other domains for the sections.

However, this process was not followed with any respectable degree of discipline, mostly because dance is normally composed in a much shorter time than music. Ania completed her assigned sections of choreography well before I completed my assigned sections of music; I ended up seeing many of her movements before I had finished the sections I was assigned to write music for. Because visual images and gestures are, as mentioned earlier, such a vivid inspiration for me, the piece was ultimately written more as music-to-choreography rather than the other way around, or as the balanced mixture we originally intended. This affected the gestural content of the work.

As I mentioned earlier, this work used pitch and pitch-structure much more consciously, including the use of synthesized (not just concrète) sounds as explicit carriers of structural pitch information (i.e., like instruments in most acoustic-instrument music); of course, adding pitch to the gamut of materials in use greatly increases one's ability to form musical gestures.

I'll say something about the derivation of pitch structure in this piece, since it involved another of my quasi-dialectic ideas mentioned earlier. I

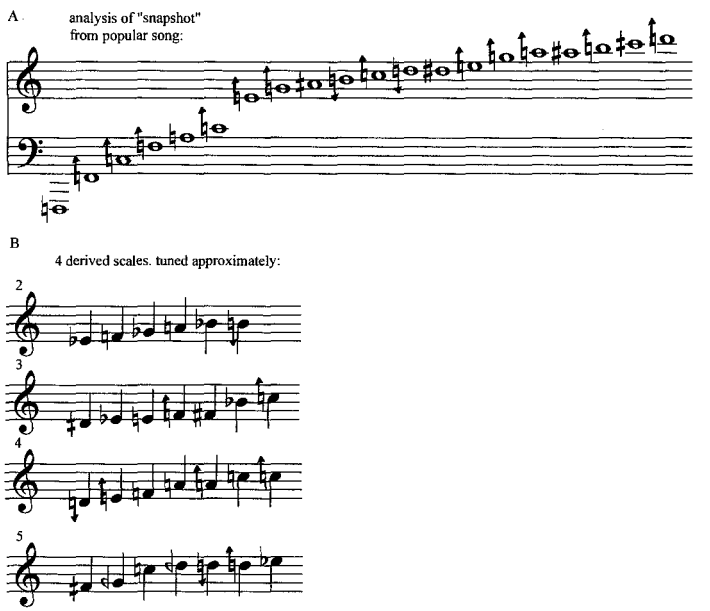
Figure 8: Kinetic- and work-plan for *Oogaaah: Dungeony Specimen Spaceship.*

Section numbers. Single digits were to be composed: music first, dance second; double digits: the reverse.

had recently come into contact with the music of the French spectral composers (Tristan Murail, Gérard Grisey, etc.), and also with music of their American cousins, the just-intonation composers (particularly La Monte Young, Harry Partch, Ben Johnston, etc.). Both schools are involved with microtones. Generally, the spectral composers derive vast harmonic complexes from analyses of real-life timbres; usually these complexes are variations (distortions) of the overtone series, although occasionally they experiment with inharmonic timbres (such as that of a cymbal). The just-intonation composers derive their pitch material directly from the *pure* overtone series, but unlike the spectralists, they transpose the tones, with octave duplications, to form scales. (Generally speaking, recent European composers tend to regard the idea of using a scale—especially over a large span of time—as an old idea; musical passages based on scales aren't found in very many current European scores.)

For this piece, then, I decided to combine these ideas: I would derive scales, but from analysis of a real-life sound, rather than from the pure overtone series. The real-life sound I chose was an instant from a popular song. The frequency analysis of the "snapshot" is shown in figure 9, along with four of the scales derived from it, which I used in the piece. (The chord derived from this moment also appears in the piece, explicitly as a harmony, towards the end of the work.)

Figure 9: "Ur" chord and derived scales from *Ooogaaah: Dungeony Specimen Spaceship.*



In this piece, sound material is created using the many different computer music techniques at my disposal. Concrète material is mostly limited to extremely scratched-up LP records from my childhood. Synthesized material was created with well-worn techniques: as I mentioned earlier, my intent was not to present "new" sounds, but to present old sounds in new, complex combinations. Hence, in this piece I used plucked-string imitations, bell-like timbres produced with frequency modulation, and many samples of pitched instruments, including my own voice.

Integrating the less obviously pitched concrète material with the overtly pitched material was done with several simple techniques. The first was to use the noisy concrète material in a traditional, adjunctive manner (i.e., as percussion, to mark off phrases or sections), to add "unsolicited commentary" in the name of humor (most of the concrète material consisted of text fragments, noises, and sound effects—all masked by a large sheen of noise arising from the scratched surface of the LPs), and occasionally to mark metrical rhythms.

The other method of integration was to process the concrète sounds to bring out inherent pitches within them, which could interface with the pitched elements of the rest of the musical context. Usually this processing involved fairly straightforward filtering—emphasizing partials of the sound that matched those of the reigning harmony at a given musical moment. For examples of this, as well as the idea of transferring gestures from the choreographer's ideas to music, I will now speak about some of the events in the first minute of the work, the most gesturally frenetic of the piece.

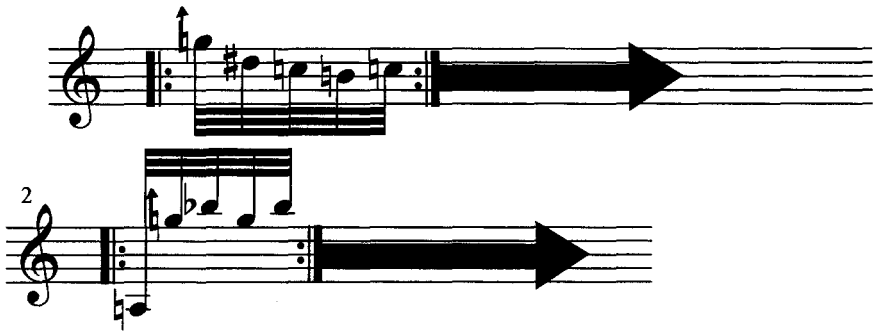
Ooogaaah begins with an introduction of several text fragments, followed by approximately three seconds of a machine-like texture, which consists of several noisy text fragments, looped in different tempi, in counterpoint with one another. This "machine texture" was in fact inspired by machine-like, repetitive, mechanistic gestures on the part of the dancer.

The next gesture in her choreography sequence consisted of a repeated leaping motion; I matched this musically, introducing the first bit of pitched material into the piece, derived from a small cut of the original harmony (see fig. 10).

I realized the pitches with samples of a piano. Since the instrument I recorded was already badly out of tune, I decided to take advantage of the computer's capabilities and have the tuning of the piano samples waver with each repetition of the gesture. The machine-like gesture/texture then appears again, but this time, as I hinted previously, filtered by the first of those scales listed above.

This alternation between the "machines" and the "leaping" continues, but the textures themselves begin to develop. For example, the leaping ideas become more complex in terms of contour. More complex contours

Figure 10: "Leaping" gesture derived from "Ur" chord.



were produced with the aid of a (very rudimentary) computer "improvisor" programmed in the computer language LISP. The process, or algorithm, followed by this improvisor was something like the following:

(1) A set of contours to choose from:
(numbers indicate # of scale-steps)
+1 +1 +1 +1 −1 −1
−1 −1 +2 −1
+1 +1 +1 −3 −4 +2 −1 +2 −1
etc.

(2) A set of rhythms to choose from:
(duration/attack-distance in units of pulse)
2 1 2 1 2 1 1 1 1 1 1 3 1
1 2 1 2 3 1 1 1 2 1 1 2 3 1 1 1.5
3 1 1 2 1.76 1.35 1 1 1
etc.
(decimal fractions add a "micro-rubato")

(3) Produce a line of pitch contour from strung-together random members of the set provided above.

(4) Do the same for rhythm.

(5) Match up rhythm and pitch, and have the contour "play through" the reigning scale that is being used in whatever portion of the piece we're in at the moment.

(6) The composer edits the results, removing unsuccessful, or drab, portions of improvisation, and places the excerpts in appropriate musical contexts (that is, in a rhythmic relation with other elements of the mix to produce the most interesting musical result).

Having finished *Ooogaaah,* I moved on to write several instrumental works. However, during this time, I continued to think about some of the questions raised by my computer music endeavors, especially by the more concrète-based *Duude* and *Ow.* I wanted to get back to "achieving gesture X to fit context A" rather than making gesture X from randomly selected materials and then "shaping context A to fit gesture X." I wanted to accomplish this with collections of raw, unprocessed found sounds.

Much of the computer music world is concerned with processing a sound until it becomes unrecognizable. Recognizability and association may result in an affect that is too sentimental or "cheesy." This happens when the most relevant thing (or even the only thing) that the listener hears in the individual sounds in a mix are their associations. In other words, the listener thinks only, "Huh . . . these are pots from Christopher Bailey's kitchen"—not, perhaps, the most "musical" reaction. This is a worst-case scenario, and because of even the shadow of this possibility, many computer-music composers are driven to "hide" their sounds behind a wall of processing and transformation. This "safety procedure" does not interest me: I do not want to rid the sounds of all recognizability and therefore all associations; instead, my goal is to produce music where overall gestural shapes and phrases take precedence over the autonomy of the individual sounds, where the individuality of the sounds is sacrificed to these greater musical wholes—and yet, those individual associations and references are still there. This leads to a multilevel musical experience: structural musical listening (in terms of how a phrase or sentence works) *and* associative listening ("this finely crafted phrase . . . just so happens to be made of pots from Christopher Bailey's kitchen").

In order to produce a "finely crafted phrase" out of found sounds, we first describe the phrase as a sequence of events, each of whose parameters can be specified exactly. We store information about all the sounds in our source collection in a database. We can then ask the computer to search the database, matching the specified parameters of an event against the parameters of sounds in the database, thereby ultimately retrieving an appropriate sound for each particular event. What would such a database look like? Figure 11 shows a portion of one that I'm using for a current work-in-progress.

Each sound is described in terms of 11 parameters. The first, *filename,* is simply information about where the sound-file lies on the computer disk. *Duration,* measured in seconds, is self-explanatory. *Pitch* indicates one or more strong pitches or partials in the sound—most often, the fundamental or first harmonic. (It is indicated here in MIDI notation, where middle C = 60, C♯ = 61, etc.) Of course, some sounds have no clear pitch, and

Figure 11: Portion of a found-sound database.

filename	duration	pitch in MIDI format	loud-ness	attack hard-ness	bangs list	noisiness/ harmonic		color		agitation	material/ category	tessitura/ register
("ds4.pan.rhythm"	1.619	(63)	5	6	(0.057	0.31	0.52	0.695	0.857	1.464)		
						4		5		7	("metal" "rhythm")	5)
("ds4.pan.scr.rhythm"	1.995	(63)	5	6	(0.066	0.499	0.938	1.677)				
						3		5		7	("metal" "scrape")	5)
("fs.pan.drum"	0.570	(66)	6	7	(0.0)	3		6		2	("metal")	4)
("fs4.cowbell.MONOIZE"	0.722	(66)	4	7	(0.0)	1		6		2	("metal" "glass")	4)
("fs4.jarscrape.RIGHTIZE"	0.737	(67)	6	4	(0.0)	5		4		5	("glass" "scrape")	5)
("h.creak.2"	2.32	()	2	2	()	6		5		6	("creak")	5)
("h.creak.3"	1.052	()	2	2	()	3		6		4	("creak")	6)
("h.cup.klink.4.RIGHTIZE"	0.66	(83)	5	7	(0.0)	3		7		2	("glass")	6)
("h.cymbal.pan.2.MONOIZE"	1.827	()	7	6	(0.009)	6		7		2	("metal"}	7)
("h.glass.bnk"	0.264	(84)	3	6	(0.0)	4		7		1	("glass")	6)
("h.hit.jiggle"	0.556	()	5	5	(0.1)	7		6		6	("blech" "metal")	5)
("h.jar.balls.shake"	4.169	()	4	3	()	7		7		7	("blech" "crunch")	6)
("h.klingk.2"	0.18	(95)	2	5	(0.0)	4		7		2	("metal")	6)
("h.klingk.gk.2"	0.295	(88)	3	5	(0.0)	3		7		4	("metal")	7)
("h.klingk.gk.complex"	0.643	(88)	4	5	(0.0 0.420)	5		5		5	("metal")	6)

therefore this parameter is left empty. *Loudness* is not about sheer amplitude or volume, but rather perceptual loudness—a light whisper is a qualitatively soft sound even when highly amplified. This parameter is measured from 1–7, as is *attack hardness,* describing the "violence" of the beginning of the sound—whether it fades in, enters with a bang, or something in between. The 1–7 range applies to many of the parameters.

The reader might recall, from the discussion of *Ow, My Head,* the idea of "lining up attack-points between different sound objects" (see fig. 3). The *bangs list* is a list of those attack-points. Later, we can use this information to have the machine line up those points automatically.

Noisiness/harmonic is also more or less self-explanatory: a voice or a bell would be a harmonic sound (value of 1); crumpling paper would be noisy (value of 7). Rubbing a washboard, producing both a pitch and a fair amount of noise, would be somewhere in between. *Color* describes whether the sound tends toward being "dark" (value of 1) or "bright" (value of 7). *Agitation* describes the internal state of the sound during its duration: is there much movement and change (for example, vigorous rubbing or scraping) (value of 7) or is there simply a decay (a bell rings) (value of 1 or 2), or something in between? *Tessitura/register* describes the general pitch register of the sound (even if it is too noisy to have an exact pitch), from low (1) to high (7).

Finally, *material/category* remains as a sort of catch-all "semiotic" parameter, describing associations, concepts or words that the sound brings to mind. Thus, often it is simply a matter of material (e.g., "metal," "glass") or action ("creak," "scrape"); sometimes it describes some important musical characteristic of the sound (e.g., "rhythm" if the sound is "rhythmic").

Measuring some of these parameters from 1–7 might seem very crude, but the crudeness is appropriately matched to the extreme heterogeneity of the materials. For example, what would be softer, a whisper or a recording of soft, distant ocean sounds? The question is a bit silly, yet we'd all agree that they are both soft sounds. Thus, 1–7 seems like a reasonable compromise.

You might imagine how this database would be used. As I began to describe above, we can essentially think of a musical gesture as a sequence of events, each event being described in terms of one or more of the above parameters.

Thus, a simple gesture might be: three short, high sounds, with hard attacks, descending in register, made of glass or metal; a couple of simultaneous, longish (two or three seconds), highly agitated mid-register sounds, slamming down into a low metallic sound, with a hard attack, not agitated but with a very long decay (see fig. 12 for a quasi-pictogram of the gesture). To the computer, we feed a quasi-spreadsheet of the same gesture (see fig. 13).

Figure 12: Quasi-pictogram of a simple gesture.

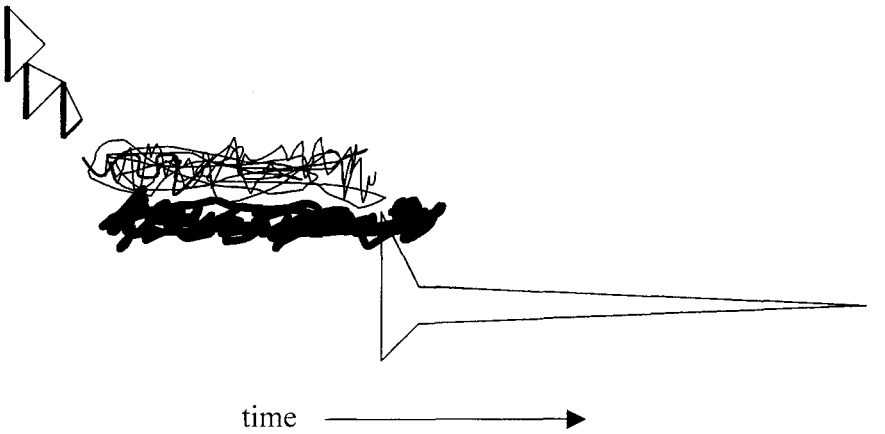


Figure 13: Quasi-spreadsheet of a simple gesture.

sound number	1	2	3	4	5	6
duration	0.0–0.2	0.0–0.2	0.0–0.2	2.3–2.7	2.3–2.7	4.0–6.0
perceptual loudness	6–7	6–7	6–7	6–7	6–7	6–7
hardness of attack(s)	6–7	6–7	6–7	any	any	6–7
how many bangs	any	any	any	0	0	1
noisy-harmonic	5–7	5–7	5–7	1–3	1–3	5–7
agitation	2	2	2	6–7	6–7	2
words	"metal"	"metal"	"metal"	"blech" "scrape"	"blech" "scrape"	"metal"
tessitura/register	7	6	5	3	3	1
start time for given sound	0.0	0.13	0.18	0.25	0.25	2.75

time ——————————→

The computer, when fed the "spreadsheet," looks at each column, figures out what parameters a sound would need to have to satisfy the criteria of that column, and grabs a random sound from the collection that satisfies those criteria. It then places the sound in a mix (as in fig. 2), in which the composer may modify the order of the sounds, delete sounds, etc. Furthermore, it is easy to generate ten or twenty versions of a given gesture—each a different attempt by the computer to realize the specified gesture with a different combination of sounds—so that eventually the composer can get, more or less, the gesture he or she had in mind.

(Another possibility, of course, is that the computer will come up with something pleasantly unexpected.)

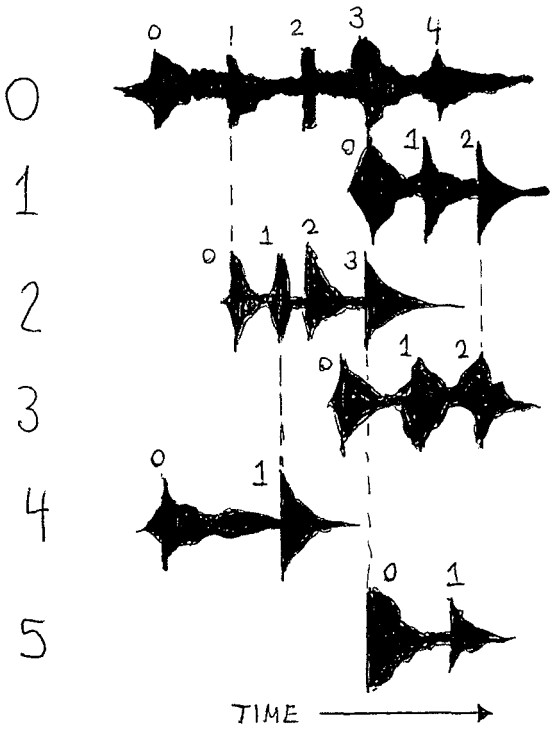
Finally, it is also possible to specify what I call a *bang tree.* This is a special rhythmic specification that arises from the *bangs list* parameter mentioned earlier. Let us begin with the pictogram shown in figure 14.

You can see that the idea is one of a gesture whose sounds relate rhythmically through their common peaks or attack-points—as discussed in *Ow, My Head.*

We can then feed to the computer a list of the qualities of these sounds (as in the example above) together with a bang tree: a list of how the bangs in the sounds relate in time. A bang tree takes the following form:

```
(mother-sound    (child-sound mother-bang child-bang)
                 (child-sound mother-bang child-bang))
(mother-sound    (child-sound mother-bang child-bang)
                 (child-sound mother-bang child-bang)). . . . . etc.
```

Figure 14: Sounds relating via "bangs."



Thus, in fig. 14, sound 0 is the "mother" of sounds 1 and 2. Then, in turn, sound 2 is the mother of 4 and 5, and so on. 1, a "child" of 0, attaches its bang #0 to sound 0's bang #3. Sound 3 attaches its bang #2 to 1's bang #2, and so on. Thus we get, as the whole tree:

(0 (1 3 0) (2 1 0))
(1 (3 2 2))
(2 (4 1 1) (5 3 0))

The computer's task is to find sounds with the appropriate number of bangs (as well as any other qualities we care to specify), and mix them as we request, placing them in time so that the appropriate bangs line up.

* * *

The idea of the gesture, its origination in the creative mind, and the way it shapes itself in the process of composition are, for me at least, very deeply intuitive processes, which seem at once too simple to even merit discussion ("you want it to go *oomph*? just write *oomph*!") and at the same time ultimately elusive. I hope I have made some tiny scratch on the surface of the understanding of how these things happen.[1]

Note

1. I am grateful to Professor Bradford Garton of Columbia University for suggesting the topic of this article, which was given initially as a talk in his Advanced Computer Music seminar.

Reference

Bailey, Christopher. 1997. *Megalomaniac* performance instructions.