TOUCH'N'GO: ECOLOGICAL MODELS IN COMPOSITION

by

Damián Keller

Maestro, Esc. Nac. de Música J.P. Esnaola, 1988

Bacharel, Universidade de Brasília, 1995

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF FINE ARTS

in

the School for the Contemporary Arts

SIMON FRASER UNIVERSITY

May 1999

Approval

Name: Damián Keller

Degree: Master of Fine Arts

Title of thesis: **TOUCH'N'GO: ECOLOGICAL MODELS IN COMPOSITION**

Supervisory Committee:

____

Barry D. Truax, Professor

Senior Supervisor

Professor

School of Communications

School for the Contemporary Arts

____

Martin S. Gotfrit, Assistant Professor

Assistant Professor

School for the Contemporary Arts

Date approved: ____

Abstract

touch'n'go is a fifty-minute tape and text piece, published as an enhanced compact disc, and as a custom CD-ROM. The CD-ROM includes the eleven sections of the piece, short Spanish and English texts - in HyperText Language Markup format - for each section, Csound code used to produce the sound material with most of the necessary samples. **touch'n'go** is the first attempt to apply an ecological approach in composition. This approach draws from soundscape techniques, physical models, and ecological models, expanding the existing palette of tools for synthesis and transformation of everyday sounds.

Acknowledgements

I would like to thank the people that helped me plough through my work during my stay in Vancouver: (in chronological order) Hildegard Westerkamp, Barry Truax, Tanya Petreman, Chris Rolfe and Trude Heift, Martin Gotfrit, my big Argentinean and Brazilian family, Ana Lúcia and Nahuel.

This project was partially funded by the School for the Contemporary Arts, Simon Fraser University. Sections of this work were presented at the *International Computer Music Conference* (Ann Arbor, 1998) and at the *XII Colloquium on Musical Informatics* (Gorizia, 1998).

Contents

Tables

Figures

Introduction

Unlike science, art, and specifically music, does not center its attention on the observation of natural phenomena. On the contrary, the main trend in European and North American music composition and theory seems to be the production of art objects completely unrelated to everyday sounds. The reproduction or description of environmental sounds falls broadly outside the field of music composition and analysis. Nevertheless, during the last thirty years there has been a growing interest on mundane sound as source material for composition and even as an object of study by itself. Based on the work of R. M. Schafer (1977), two new areas of music research opened: acoustic ecology and soundscape composition (Truax, 1996).

Another aspect that has been noticeably neglected in compositional and theoretical practices is the social context in which a given musical piece is produced. How this context influences musical thought has been generally approached from the anthropological and/or sociological perspective. Nevertheless, with few exceptions, little attention has been paid to the structural aspects of music while focusing on the relation between sound environment, cultural context and music structure.

Just as natural sound does not generally have a place in compositional techniques, perceptual processes in music listening have only recently been regarded as relevant for music composition. In this respect, J. Tenney's and P. Oliveros' work stand as an example of an alternative approach. A healthy exchange between music composition and music psychology has timidly started, but a full-blown interaction is yet to come. In fact, a great corpus of research in music perceptual phenomena has been gathered during the last ten years, creating a firm ground for the development of music psychology. This knowledge could eventually be applied to music creation processes and music theory. Similarly, new developments in music composition and theory could raise well-focused questions in experimental research. Our work is intented as a humble contribution to bridge this gap.

We need to make several non-trivial assumptions to advance in our argument. As put forth by Kelso (1995,

92) and tested in various experimental settings, sound production mechanisms and perceptual processes are inextricably interrelated. Keeping in mind the innovative profile of his theory, it would be useful to provide specific examples which pinpoint its implications in environmental sound compositional techniques. We will not deal directly with these issues here. The reader should refer to (Keller, 1998b) for an in-depth discussion. We propose that musical processes cannot be isolated from sound structures. Thus, morphology, syntax and sound (Keller & Silva, 1995) are just levels of musical structure which can be used for analytical purposes. Nevertheless, no hierarchical relationships can be implied since all time levels interact influencing the final musical percept. This is shown in several synthesis examples included with this paper, and is documented in experimental research (Tróccoli & Keller, 1996). This interaction among levels provides the key to a dynamic-pattern formation process which underlies time-varying percepts (Kelso, 1995, 224).

Finally, we define sound source recognition as the axis that links sound production mechanisms, music perception processes, and natural and cultural context. These processes can be observed at various simultaneous time levels. The interaction of these processes results in musical phenomena which we can manipulate compositionally. Therefore, understanding the processes of sound source recognition becomes a key aspect of composing with environmental sound.

Given the breadth of issues relevant to an ecologically-informed compositional approach, we will summarize in a few paragraphs our two-year research work in ecological models. After discussing the theoretical foundation of our work, we will briefly describe the technical issues involved in the use of ecological models in composition. As a case study we will present the organizational procedures, the underlying compositional concepts, the socio-contextual references, and a few sound examples of our tape piece, **touch'n'go**.

The four domains

The problem of sound source recognition can be studied from four perspectives: (1) the formal domain: structure of systems and models of sound production (theoretical and algorithmic), (2) the acoustic domain: sound parameters, organization in time, frequency, and space (3) the perceptual domain: processing of sound stimuli and correlated mental processes, (4) the social domain: the natural and cultural sound environment.

Table 1. Perspectives in sound modeling.

Perspectives in sound modeling			
formal domain	**acoustic domain**	**perceptual domain**	**social domain**
structure of systems and models of sound production (theoretical and algorithmic).	sound parameters, organization in time, frequency, and space.	processing of sound stimuli and their mental representation.	the natural and cultural sound environment.

Although the formal domain belongs mainly to epistemology, as exemplified in Kampis (1991), music theory and sound synthesis methods can deeply affect this domain (Truax, 1992). System theory has already been applied to musicological studies (Georgescu & Georgescu, 1990; Boon, 1995), unveiling a wide range of questions that need to be confronted. How do the microproperties of acoustic events influence the perception of macroproperties? Could we hypothesize the existence of a meso-level representation? Is there a signal space defined by social, perceptual and physical properties of the sound environment? Are sound events categorized in relation to a “perceptual grain?” Are dynamical models more suitable than fixed categories? Is musical time irreversible?

There is a wealth of research in acoustics relevant to synthesis and sound organization techniques (Smith, 1992). Physical modeling generally places emphasis on the study of resonant structures that respond dynamically to different types of excitation. As J. O. Smith (1997, 221) puts it, “a mathematical model of a musical instrument becomes the instrument itself.” Simplified resonant systems are strings, tubes, membranes, either elastic or stiff. Sound-producing systems are the source of energy fed into the resonant system (Fletcher, 1992, 6). Another useful technique for modeling sound-producing systems is granular synthesis (Roads, 1996; Truax, 1988). Since granular synthesis can be applied to produce complex excitation patterns, bringing together both approaches provides a comprehensive way to deal with dynamic excitation of existing resonant structures.

During the last few decades, auditory perception has become a flourishing area with well-grounded and, often, conflicting theoretical approaches. The oldest paradigm belongs to psychoacoustics, a branch of psychophysics that focuses on sound. Psychophysics attempts to relate physical descriptions of stimuli with measurable behaviors to sensation (McAdams, 1987, 5). Its methods and underlying philosophy are tightly linked to the behaviorist approach. In other words, psychoacoustics concentrates on the problem of finding models that relate subjects' sensations with physical representations of stimuli (Parncutt, 1989). Since an exact definition of the sound object is needed to implement a computational model, this acoustic representation may be useful when mapping sound parameters onto perceptual results - a fancy way to say 'composing.'

A body of theories that has gained weight during the last fifteen years is what has been called 'the cognitive approach' to music. Many disparate methods with dissimilar objects of study share some basic assumptions which entitle them to be labeled as 'cognitive.' Krumhansl (1990, 5) describes cognitive psychology as "a subarea of experimental psychology concerned with describing human mental activity." "One of the major contentions of the cognitive approach to psychology is that all mental activity is mediated by internal (or mental) representations." (McAdams, 1987, 18). As we discussed in (Keller & Silva, 1995), this mental representation can take two forms, symbolic or subsymbolic. The idea of a mental representation as separate from the actual perceptual process is the main difference between cognitive psychology and ecological psychology (Kelso, 1995, 194; Keller, 1998b). Ecological psychology studies acoustic phenomena by observing the physical characteristics of a sound event, the high-order configuration of variables, and the listener's ability to detect the information provided by the event (Gibson, 1966; Kelso, 1995; Michaels & Carello, 1981).

Within the ecological approach, we have discussed the relationship between music and its social environment (Keller, 1998a). These are the conclusions of this study: (1) music production and perception are dependent upon the social structure where they take place; (2) social structures are not only based on economical and political dynamics but also on their cultural representation. Music is a key factor in the shaping of this representation.

Ecological models: basic concepts

Sources and lack of resources

Environmental sounds have received attention from various fields in auditory research. They can be treated as an instance of the problem of source recognition. Let us outline this process in simple terms. When a human being, in his[1] everyday environment, perceives a sound coming from an object, several listening strategies are used simultaneously. Reflections produced by the surrounding space are combined with waves coming directly from the object. The auditory system parses and abstracts time-varying frequency patterns as belonging to a unique resonant body. Finally the characteristics of the excitation process are identified by their effect on the resonant body.

The mechanisms employed in auditory source recognition have been studied from different perspectives. Grantham (1995) reviews the literature related to spatial cues in the context of sound multisource determination. Bregman (1990) focuses on the problem of grouping and parsing sound events - i.e., 'streaming' - within the approach of auditory scene analysis. Handel (1995) uses the terms "auditory object identification," and addresses the issues involved in identifying everyday sounds.

From an ecological perspective, the concept of 'source' should be brought into question. Even if we overlook background sounds, moving sound sources, multiple excitations, short-term and long-term memory, emotional state, cultural and musical context, and so forth, we still have to ask, what is a "sound source?" Far from getting a definite answer, we are bound by several practical limitations. Ellis (1996) deals with the whole auditory scene instead of treating source and background as unrelated acoustic phenomena. In other words, given that the signal characteristics are modified by the space where they occur, and that in an ecological context they interact with other sources, we need to consider all processes simultaneously, not just a single isolated source. Bregman (1990, 488) puts it this way, "timbre [of a sound source] is not a result of a certain acoustic input. Timbre is to some degree created by our processes

1 'His' stands for 'his' and 'her', and 'he' stands for 'he' and 'she' throughout the paper.

of auditory scene analysis."

These observations suggest that real-world sounds are still outside the reach of current paradigms of theoretical and empirical research. For the time being, we have to settle for simplified sound models and strive not to lose the relevant characteristics of complex real-world interactions. Of course, these limitations must be acknowledged: our synthetic sounds are just toy examples of environmental sound complexes!

Invariants, where are they?

"The ecological approach combines a physical analysis of the source event, the identification of higher order acoustic properties specific to that event, and empirical tests of the listener's ability to detect such information, in an attempt to avoid the introduction of ad hoc processing principles to account for perception. (. . .) The information that specifies the kind of object and its properties under change is known as the *structural invariant* of an event; reciprocally, the information that specifies the style of change is the *transformation invariant*." (Warren & Verbrugge, 1984, 705-706). A single exciting source acting on different resonant bodies and a single resonant source excited with varying strength have been described as *acoustic invariants*. (Warren et al., 1987).

"Invariant patterns over time refer to constant patterns of change - that is, manners or style of change." Michaels and Carello (1981) go on to give the example of a melody as a typical time-pattern which keeps invariant relationships among its elements. The structural properties are the same, even if the pitch material is transposed. By now, we know that this is quite a naive example. Is a melody played without dynamics, agogics, and timbre inflections still a melody? There are several examples of the perception of pitch being modified by interactions with other parameters (Tróccoli & Keller, 1996). So we would question the idea that a single parameter could be perceived as constant across different transformations in varying contexts. More likely, as we will discuss later, the relationship among several parameters, or collective variables as Scott Kelso (1995) likes to call them, is a better candidate for structural invariance, if such a thing exists at all.

Going back to the Gibsonian approach, “if we define events as changes in objects or collections of objects, structural invariants are those properties that specify the object or collection participating in the event.” (Michaels & Carello, 1981, 26). What remains to be accounted for are the changes that occur to the objects themselves. Michaels and Carello use the term ‘transformational invariance.’ “A transformational invariant is the style of change in the proximal stimulus that specifies the change occurring in or to the object.” From this perspective we recognize that an object is breaking apart by some type of invariant ‘breaking’ characteristic. We argue that we will not find anything invariant in the process itself. The place to look for ‘regularities’ (Bregman, 1993) is one level higher than the physical variables themselves. It is in the dynamical coordination among the parameters that define the sound event. And although these processes are tightly constrained, they are random or highly complex in their micro-level details.

Rate of change

Bregman (1990, 71) discusses amplitude modulation (AM) of a tone as an instance of auditory grouping formation. He observes that when “AM is slow, the tone is simply heard as a single tone whose loudness alternatively rises and falls. As the AM speeds up, a series of separate tone bursts is perceived, the bursts corresponding to the loud phase of the AM cycle and the inter-unit gaps to the attenuated phase. (. . .) Presumably this happens because the rate of amplitude exceeds a critical value and a perceptual boundary is formed. (. . .) It is likely that it is really the rate of rise in intensity that controls the formation of units.” From a multifunctional perspective, we would be reluctant to agree that a single variable can produce any form of meaningful percept. Correlations with other variables should be considered. Alternatively, higher order variables such as the range of change in AM rate can greatly affect the perceptual effect. This is clearly exemplified by sounds processed by asynchronous granular algorithms (Roads, 1996; Truax, 1994).

Bregman's onset hypothesis predicts that a sudden rise in amplitude should serve as a clue to indicate the beginning of a sound. We claim that this depends on the context. If this onset is embedded in a granular texture with an average amplitude distribution that approximates the isolated sound, it will be integrated as part of a high-order percept.

Bregman (1990, 72, 708) criticizes the view that temporal units are tied to periodicities in the signal, as was suggested by Mari Jones (1976). He defines a unit as having uniform properties representing distinct events in the environment, whenever a fast change of properties occurs a perceptual boundary is formed. "Units can occur at different time scales and smaller units can be embedded in larger ones. When a sequence is speeded up, the changes that signal the smaller units may be missed by the auditory system, and other changes, too gradual to form units at the slower speed, may now be sudden enough to control the formation of units." (Bregman, 1990, 644). This hints at clear range limitations in the production of ecologically feasible sound patterns. Physical parameters in the acoustic environment do not change independently. Thus, models of sound production need to correlate changes in amplitude and spectral modifications within narrow time constraints. Failure to do so will produce sounds that fall outside identifiable environmental sound classes.

Although rate of change is a major factor in the organization of the auditory scene, correlation among variables also seems to play an important role. The use of fast changing, widely varying parameters in granular sounds hints at organizational *strategies* that can distinguish different elements even within the 'mess' of several simultaneous granular sources. Phase coherence among streams, similar rate and type of change in different variables, correlated changes in delayed signals, all point to a common sound source. Moreover, it is not unlikely that the auditory system applies similar comparison strategies to completely different functional tasks, such as the separation of source and reflected sound, the extraction of a vibrating body's resonant characteristics, and the estimation of the number of similar sources.

Units, who are they?

Digital signal processing techniques provide us with a reliable method to represent sound signals at a sample level (Moore, 1990; Orphanidis, 1996). Although these techniques are well-suited for time-sensitive models such as sound localization cues or spectral filtering, it is difficult to find percepts that could be directly mapped onto a single variable at the sample level. Interactions among several acoustic mechanisms, such as those discussed in physical modeling (Smith, 1992), provide a useful prediction of higher level properties from locally defined characteristics. Nevertheless, computationally efficient implementations are generally done by lumping, i.e., simplifying, descriptions of the sound behavior to provide an output that approximates a generic acoustic model (cf. Smith, 1997). In spite of the fact that some of these models are perceptually convincing, this approach does not start from perceptual processes but from the physical sources that produce the sounds. Although there are some exceptions (Chafe, 1989; Cook, 1997), research in this area has mainly concentrated on modeling the spectral behavior of resonant bodies, leaving aside descriptions of time-related excitation patterns. This lack of research in time processes has been pointed out by Dannenberg (1996) among others.

The next higher level of signal description falls approximately in the range of grain durations. A grain, i.e., a very short sound, is simply a windowed group of samples. Its duration goes from a few samples, one to ten milliseconds, to a few hundred milliseconds. It has been popularized as the sound unit in granular synthesis (Truax, 1994), though from a broader perspective it can be defined as the window of observation (Lynn & Fuerst, 1994) in several analysis and synthesis methods (short-time Fourier transform, Wavelet transform, FOF, pitch-synchronous granular synthesis, etc.) (Cavaliere & Piccialli, 1997). The granular description of sound shares some properties with sample-based techniques, such as the possibility of shaping the spectrum from the time domain, or controlling the micro-temporal structure of sound. But it also permits the use of ecologically meaningful sound events and time-patterns that are hard to tackle within a sample-based approach (Keller & Truax, 1998).

Granular sounds require high densities of short events to produce aurally convincing sound textures.

Therefore, computer music composers have adopted statistically-controlled distributions of grains limited by tendency masks, averages, deviations, probability densities, and similar methods (Xenakis, 1971; Truax, 1988). Besides the use of quasi-synchronous (periodic) grain streams in FOF (Rodet, 1984) and pitch-synchronous granular synthesis (De Poli & Piccialli, 1991), some composers have recently proposed deterministic control methods. Roads (1997) suggests a traditional note-based approach for long grain durations which can be extended to fast grain rates in order to produce micro-temporal and spectral effects. He calls this traditional compositional technique "pulsar synthesis." Di Scipio (1994) and Truax (1990) have explored the possibilities of controlling granular streams from the output of nonlinear functions. This technique offers good possibilities for the generation of macro-temporal patterns, though up to now only arbitrary mappings of isolated acoustic parameters have been used, e.g., grain frequency and grain duration. The common trend in all these approaches is to take a time line, isomorphous to absolute time, as the underlying space where the events are placed. In other words, it is in the hands of the composer to make all decisions regarding the duration, density, distribution and organization of the grains.

To constrain the space of possible organizations would mean to throw away the idea of "limitless possibilities," "the vast soundscape that synthesizers produce" (Paradiso, 1997, 18), "any sound that could ever come to a loudspeaker" (Smith, 1991, 1), and the usual rhetoric that we have heard in computer music for thirty years. As Smith (1991, 9) says, "most sounds are simply uninteresting." The ecological approach suggests that time be parsed into informationally relevant events. The perceptual system is constantly searching for new patterns of information. Thus attention-based processes are triggered by organized transformation, not by redundancy or randomness. As we have discussed, this transformation can be tracked by co-dependent variables, at several levels, changing at strictly constrained rates. Thus, to define ecologically meaningful sound events, the grain distributions and sample-based processes have to be controlled from parameters defined by a higher level transformation. This transformation needs to be constrained to a finite event which is feasible, at least in theory, within our day-to-day environment. In other words, we are not working on an abstract time line, but from a representation which parses time into ecologically-constrained events.

Events

From a Gibsonian perspective, information is structure that specifies an environment to an animal. Thus, it is carried by a high-order organization that occurs over time (Michaels & Carello, 1981, 9). As the animal gathers information from the environment, it exerts changes on its surroundings. This activity is clearly goal-oriented.

Traditional theories of information processing consider the stimulus to be a discrete time-slice (cf. Massaro & Cowan, 1993). Michaels and Carello (1981) argue that time should be directly related to the informational structure of the stimulus. "Time is not chopped into an arbitrary succession of nows, but organized into naturally occurring events of varying duration (. . .). If time is viewed as an abstraction from change we might as well question the value of that abstraction. After all, change itself (events in space-time) is of interest to a behaving animal, not absolute time. (. . .) The notion of absolute time is given up in favor of space-time on the belief that perceivers do not perceive space and time, but events in space-time." (Michaels & Carello, 1981, 13).

Instead of working from the assumption of an absolute time which is detached from actual occurring events, we propose a model where time is parsed in event-dependent chunks. This implies that the perceptual system gets reconfigured whenever it finds new information. Change acquires a new meaning. It is not simply the variation of observed variables, but defines how these variables should be observed. That is, the significant unit of observation is the event defined by ecologically meaningful boundaries. These boundaries can be tracked by monitoring incoming information compatible with the behavior of sources existing in the environment. Sources that are not compatible with the current environment-individual state have fewer probabilities of being processed by the perceptual system. Nevertheless, when new information is found it modifies the state of all the subsystems, triggering the perception of a new event.

Patterns of change

Sound events can be described by the interaction between two systems: excitation and resonance. The excitation establishes a temporal pattern of energy input. The resonance produces a pattern of energy dissipation. When a resonant system is excited its losses are unevenly distributed, thus some frequencies are less damped than others. Generally, objects react linearly to excitations. Their response lasts a finite amount of time after the energy source stops.

Resonant systems reach a final stable state because energy is not generated within the system but it is received from an external source - an exciting system. Exciting systems may exhibit unstable states, e.g., fire, rain, dripping water. If the exciting system behaves heterogeneously, we can safely infer that there is more than one source of energy. A resonant system is heterogeneous when it comprises various subsystems excited by a single energy source.

An excitation pattern is perceived as continuous because it forms a perceptual unit at a higher level. By observing the system at several time levels we see why events that are discrete at a low level form a fused percept at a higher level. Organization at one level influences the others. Therefore, this is neither a top-down nor a bottom-up process, but a pattern-formation one. For example, the micro level characteristics of a sound grain influence the meso and macro properties of the sound event.

<u>Ecological hypotheses:</u>

1. Patterns of change (PCs) should be recognizable.
2. To be ecologically valid, PCs should form higher-order percepts. In other words, the interaction of low-level elements show emergent properties at higher levels. This interaction occurs among all levels.
3. PCs should be perceptually distinguishable from other PCs to be classified as different.
4. There is no unique PC for an ecologically valid percept. There is no class of PCs that has only one item.

5. Previous stimuli modify the current state of the perceptual system.

6. The perceptual system is biased toward percepts with which it has previously interacted. These percepts are more stable than unfamiliar percepts.

7. Sound models form a continuum of stable models and unstable ones. Stable models are perceived as ecologically valid.

8. Basins of attraction are defined by a memory trace of interactions between the perceptual system and the sound environment. Short-term memory shows a fast decaying trace, long-term memory a slowly decaying trace.

These simple hypotheses provide a development methodology and a validity test for ecological models. By generalizing ecological event concept to multi-level perceptual and physical patterns of change, we can establish a model of timbre perception consistent with micro, meso, and macro sound phenomena. This model is documented at length in (Keller, 1998b).

Acoustic environs

Everyday sounds occur in various surroundings which modify their temporal and spectral characteristics. "The environment destroys any simple invariance between an acoustic event and the waveform that arrives at the ear. (. . .) The perceptual system thus faces a general problem of separating properties of the source from the properties of the [environment]." (Darwin, 1990, 220-221).

Sounds reflected from surfaces are either perceived as part of the sources or are heard as separate from them. Darwin (1990) offers a speech-oriented account of the interactions between source and environment. He mentions reverberation, echo, static and dynamic spectral effects. The temporal effects that he discusses can be grouped as synchronous or asynchronous change, such as onset and offset differences, or varying modulation rates. Grantham (1995) provides a thorough survey of recent advances in spatial hearing, covering several issues which are further discussed in (Keller, 1998b): reflections, reverberation, localization, and lateralization.

In his (1993) paper, Bregman discusses the cues used by the auditory system to deal with the complexity of natural environments. He calls them regularities. These are his guidelines for understanding the perceptual processes involved in sequential and simultaneous integration of environmental sounds.

Regularities:

1. Unrelated sounds seldom start or stop at the same time.
2. Gradualness of change: (a) a single sound tends to change its properties smoothly and slowly, (b) a sequence of sounds from the same source tends to change its properties slowly.
3. When a body vibrates with a repetitive period, its vibrations give rise to an acoustic pattern in which the frequency components are multiples of a common fundamental.
4. Many changes that take place in an acoustic event will affect all the components of the resulting sound in the same way and at the same time.

A deeper understanding of the source separation process, what psychoacoustic researchers call 'the cocktail party effect,' allows us to apply ecologically-constrained sound processes to design a compositionally effective sound space. Some tools used for this purpose, such as convolution and phase-controlled granulation (Keller & Rolfe, 1998) are discussed in the next section.

Summary

We have outlined the basic ideas of the ecological approach to auditory source recognition. This approach is defined by the interaction of environmental constraints with the individual's goal-oriented activity. This activity not only takes place at the level of auditory processing, but it is also related to interactions among other sensory modalities. We address these issues in (Keller, 1998b). Listening takes place within a specific cultural context. The individual, through his sound-producing and sound-listening activities, modifies and is modified by his environment. This process was described as structural coupling by Varela et al. (1989). Its relevance to music composition is discussed in (Keller, 1998a).

The ecological approach can be characterized by a few theoretical assumptions. A system cannot be studied by isolating its parts. All models should be constrained to ecologically feasible events. Ecological validity is defined by the observation of complex interactions actually occurring in the environment. The action of the individual on the environment and the influence of the environment on the individual determine a process of pattern formation. This process can be approximately modeled by algorithmic tools. Both spectral and temporal characteristics of sound events need to be accounted for in the modeling process.

An ecological approach to composition

The limitations of the early Gibsonian approach were acknowledged by Michaels and Carello (1981, 168). "Inattention to algorithmic concerns is a problem of resources rather than a systematic bias against that particular class of scientific questions. Ecological psychologists recognize that the identification of the algorithms that are embodied in living tissue is a necessary part of a full theory of knowing." It took nearly twenty years to bring together well-known digital signal processing (DSP) techniques with the relevant theoretical ideas. But it seems the time is ripe for an algorithmic implementation of the concepts put forth by ecological psychologists.

Surprisingly, a few simple techniques account for varied and flexible sound synthesis and processing methods. The ecological methods rest on two pillars: generic physical models and meso-pattern control of granular sample pools. The basic techniques are further extended by placing synthesized sounds within the context of soundscape recordings, and by parametric transformation and convolution of granular samples. Probably, the biggest advantage of the ecological approach over other processing methods is the ability to independently work at several temporal levels on a single sound source. This is outside the scope of the DSP techniques that use a linear time line as the basis of their algorithms (Lynn & Fuerst, 1994; Orphanidis, 1996).

We have discussed the advantages and limitations of ecological models and their implementation details in (Keller & Truax, 1998), (Keller & Rolfe, 1998), and (Keller, 1998c). Thus, we will concentrate our discussion on compositional methods and bring up technical issues only when they are directly related to the focus of our discussion. So let us dive into the secrets of the ecological approach to composition.

Our compositional method relies on using recognizable sound sources, keeping a consistent spatial placement of the sound material and applying ecologically-feasible transformations on the sources. These transformations provide the basic compositional strategies of our method. Thus, the minimal compositional element is the sound event, an ecologically meaningful unit. The macro-organization of the material is a result of interactions of events at a meso and micro level. Finally, meso and micro level transformations consist of a single ecologically-based process: model interaction.

Within a constrained parameter space, excitation patterns interact with resonant systems. This process provides material which is consistent with real-world sound production. The temporal organization of the events is defined by time-patterns that occur, or at least might occur, in real-world situations, such as scraping, pouring, etc. The spectral transformations are done by means of resonant structures that fall within common broad classes: woods, metals, glasses, strings.

Sound sources

Defining the object of our study has proven to be a tricky task. McAdams (1993) talks about 'nonverbal' sounds, including 'natural' ones. Handel (1995) talks about 'objects' and 'events' interchangeably. Bregman (1990, 10) defines 'stream' as the perceptual representation of an 'acoustic event' or 'sound.'

There has been some confusing use of the terms 'source' and 'cause.' Young (1996) and Smalley (1993) identify complex sound-objects as responses of a source, a physical vibrating object, to causal energy, or excitation. We prefer to avoid this usage because it goes against the standard subtractive synthesis approach, "a sound source feeding a resonating system" (Moore, 1990), and against the widely-used model of speech production.

Computer interface work (Gaver, 1993; Darvishi et al., 1995) treats environmental or everyday sounds as 'auditory icons' or 'earcons.' Presumably these icons are useful to convey information in software interfaces by reproducing the sound behavior of objects being excited in different ways.

All these definitions are too general to characterize specific classes of sounds which can be linked to perceptually relevant parameters or algorithmic models. An early experiment by Lass et al. (1982) uses four types of environmental sounds: (1) human, (2) musical, (3) inanimate, and (4) animal. They hypothesize a relationship between exposure to the sound and subjects' accuracy in identification tasks. Gaver (1993) suggests three classes: (1) vibration of solids, (2) motions of gases, (3) impacts of liquids. Ballas (1993) picks up Lass's concept to include stimulus properties within his sound categorization scheme, coining the term 'ecological frequency.' This is a measure of how often a sound occurs in a given subject's daily life. Within the limitations of his study, he obtains four classes: (1) water sounds, (2) signaling sounds, (3) door sounds and modulated noise sounds [sic] - he probably means amplitude-enveloped white noise - and (4) sounds with two or three transient components.

From a musical point of view, Wishart (1996, 181) proposes two dimensions for classification of complex sounds: (1) gestural morphology and (2) intrinsic morphology. Since these correspond to the source-filter model, where (1) is the excitation and (2) is the resonance, we are again within a standard approach. Wishart (1996, 178) divides the types of sounds generated by gestural - imposed - morphology in: (1) continuous, (2) iterative, and (3) discrete; but when it comes to the intrinsic morphology there is an explosion in the variety of classes.

To be able to link the sounds to the actual objects that generated them we have proposed a loose classification scheme (Fig. 1) which includes Wishart's types and takes account of the difference between excitation and resonance mechanisms in sound production (Keller, 1998c). This classification, being based on the sound phenomena and not the causes, suits the needs of a synthesis guideline but would need to be refined to account for auditory constraints.

Classes of environmental sounds

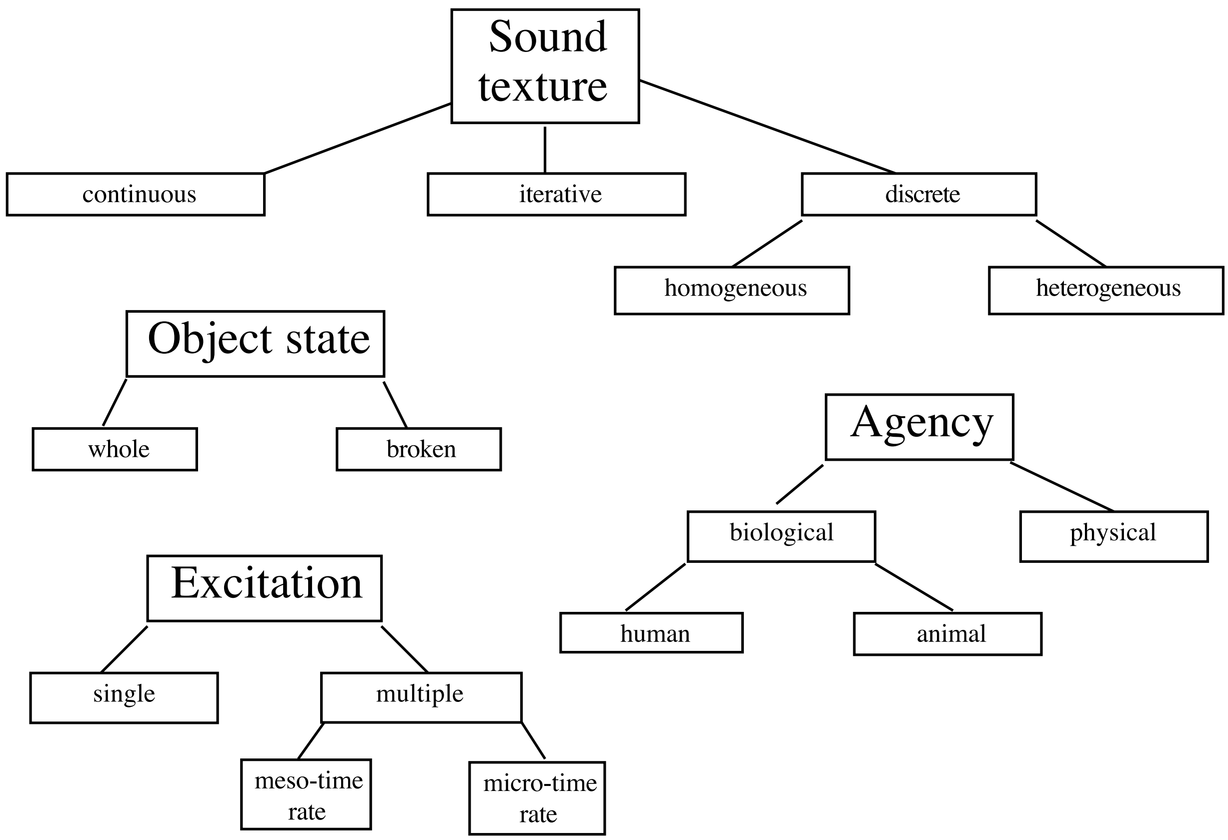

Figure 1. A broad classification scheme for environmental sounds.

By defining two classes of interaction between excitation sources and resonating bodies, i.e., single or multiple excitations, we establish a clear way to link physical models and granular synthesis with mundane sounds. All terrestrial objects suffer energy loss. After a resonating body is excited, it (usually) damps higher frequencies first, then lower frequencies until reaching zero amplitude. No acoustic object generates unchanging sinusoidal components until infinity, or increases in amplitude after the attack transients have ended. Physical modeling provides an appropriate framework to deal with single-excitation sounds and granular synthesis is best suited for multiple-excitation ones.

Multiple excitations are produced by physical agents, such as fire or running water, or by biological agents, such as a clapping hand. The cues provided by these types of sounds may be similar in their micro-temporal or spectral structure but will certainly be different in their macro-time structure. It is very likely that these cues have been decisive for survival, i.e., the sounds of a predator stepping on leaves have to be discriminated from the same type of sounds produced by a prey. When visual cues are not readily available this becomes a question of life or death.

The distinction between human and animal agency only makes sense if applied within a specific social context. In other words, human beings have become familiar with different systems of musical conventions which influence the way they perceive musical sounds (Shepherd, 1992). Playing an instrument would be an example of a human agent interacting with a whole resonant body.

Another possible cause for multiple excitations is the change of state of an object, as occurs in breaking. Breaking is a unique pattern because of its heterogeneous characteristics. The initial resonance is produced by the whole vibrating body. After breaking has occurred, the subsequent sounds have a higher spectral profile, corresponding to the smaller glass pieces resonating after each collision.

When an object is excited at a micro-time range there is no perceptible time gap between successive excitations. So the sound texture is perceived as continuous. In blown tubes or in the friction between two surfaces, the resonant body and the excitation source establish a pattern of interaction that generates nonlinearities. This interaction has been roughly modeled by dynamical systems with feedback structures (Schumacher & Woodhouse, 1995).

Time scales

As discussed in the previous section, within the ecological approach sound sources are characterized by invariants (Shaw et al., 1981). Invariants are common characteristics of a sound class that listeners use as cues for recognition of the source's physical structure. Contrastingly, Gestalt-oriented psychology has generally given importance to the physical sound parameters, such as frequency or intensity. This is reflected in the way sources are classified and how parameters are interpreted.

McAdams (1993, 179) proposes two classes of sound properties to define an acoustic source: (1) microproperties and (2) macroproperties. The first class corresponds to events with durations from ten to one hundred milliseconds. This class is further subdivided into: (1) spectral and (2) temporal microproperties. Macroproperties span longer periods of time, i.e., one hundred milliseconds to several seconds. These can be (1) temporal patterning macroproperties or (2) spectral variation ones. The former ones are related to the way an object is excited - the gesture as Wishart (1996) would put it - or the state of the object, that is, if it is whole or broken. According to McAdams (1993, 181), the spectral variation provides a cue to the nature of the material being stimulated and to the number of sources, i.e., one or many.

Suddenly we find ourselves into deep trouble! Which cues are used to identify the number of objects, the spectral or the temporal ones? If the excitation changes the spectral content of the signal, how can we be sure that the excitation changed and not the object itself? For temporal patterns that range from hundreds of milliseconds to less than five milliseconds - such as bouncing - how do we separate macro time from micro time?

Since properties that change over several seconds or more are not considered by McAdams, we find these categories inappropriate. Keller & Silva (1995) proposed to use 'macro level' for durations over several seconds. Events occurring at less than ten milliseconds are generally fused into a continuous sound. Thus, it seems reasonable to use the term 'micro level' for this time span. Most environmental time patterns fall within a range between ten milliseconds and several seconds. Following Kelso (1995), we adopt the term 'meso level' for this range.

Smalley (1993, 41) comments on the relationship between different time levels in compositional sound structure. He makes loose use of vocabulary, such as identity as equivalent to recognition, or evolution as meaning change over time. "There are types of spectromorphology [spectral patterns over time] whose existence can only be established after certain evolution time because the completion or partial completion of a pattern integral to identity. Motions based on rotation are examples of spectromorphologies whose timbre (if that is the right word) is embodied in the spectral changes over at least one rotation-cycle. (. . .) It becomes impossible to distinguish between its timbral matter [micro-time structure] on the one hand and its short-term evolution [meso-time structure] on the other."

Given that micro level granular synthesis parameters determine the global sound result, Clarke (1996) suggests the use of a frequency-time continuum to define compositional strategies. When dealing with granular sound, the boundary between micro and macro properties is blurred and the focus is placed on parameter interaction across levels instead of the usual independent parameters. We believe this approach should also be adopted in relation to environmental sounds. The idea of parameters interrelated at different time levels is consistent with an ecological perspective. We are usually exposed to sound events that present coherent patterns of spectral and temporal characteristics. There is no breaking water or bouncing wind in our world!

Table 2. Time scales applied to sound organization.

Time scales: do they interact?			
Micro level	Meso level	Macro level	Natural and cultural context
less than 10 ms.	from 10 ms. to several seconds.	from several seconds to hours.	music conventions and everyday sounds.

Sound space and soundscape

After twenty years of its original formulation (Schafer, 1977), a strong tradition in soundscape composition has already formed. Several active composers routinely use soundscape techniques in their works (Westerkamp, 1996; Truax, 1996). As Truax (1996) has discussed in his various writings, the great difference between soundscape composition and the acousmatic tradition is that the latter uses sound objects separated from their context while the former keeps the sounds as an integral part of their social, cultural and aural context. Following this line of thought, recorded sounds as used in most soundscape compositions provide an ideal and simple way to incorporate the untouched sound environment within tape composition.

By keeping a consistent mix of environmental ‘background’ sound with sound events generated by ecological models, we can give a real-life feel to algorithmically designed sound. Though, we should warn against the cheap use of the concept background. We mean, by no means, a drone or permanent room tone with no life of its own. In this context, we use the word background only because we need to convey a stable and believable sound space. But the events contained in this environmental sound should be as complex, lively, and meaningful as in any other purposeful recording. Thus, the events generated by ecological models are allowed to interact and fuse with the dynamics of the environment.

The other useful technique for placement of ecological sounds within a sound space is convolution. Convolution consists in applying the spectral dynamics of a source sound onto the spectral dynamics of a target sound. Although Roads (1997) proposes this technique as a special case of granulation, we have found that it is neither flexible nor computationally efficient to provide a handy tool for shaping time patterns - at least in its current implementations on personal computers. On the other hand, it provides an exquisite tool for shaping isolated grains, which can be later used in ecologically-defined time-patterns.

These convolution-designed grains may consist of an ecologically meaningful short sound, such as a water drop or a bubble, which is convolved with the impulse response of a cavern, or any other reverberant space. When distributing these convolved grains as a meso-level time-pattern, the result is a stream of events that occurs within the space defined by the impulse response used, for example, bubbles inside a cavern. Given that we can use several types of grain, the number of simultaneous spaces created depends on the limits of our auditory system in discriminating sounds coming from different reverberant spaces. This limit is undoubtedly low.

The last method that we have to mention is phase-controlled granulation. Given that we have not used it in our current compositional work, we leave a detailed discussion for a future paper. The idea behind this type of processing is to increase the volume of the source sound (as defined in Truax, 1994) by superimposing several granulated versions of the processed sound. If the phase-delay among these streams is kept constant, the result is an effect very akin to the reflections produced by a reverberant space. The number of 'reflections' is roughly proportional to the number of streams.

Summary

Ecologically-based composition applies ecological models, physical models, and soundscape techniques to create musical processes consistent with a given natural and cultural environment. The sound sources provide direct references to everyday sounds. The structural unit is the sound event. The processing techniques apply transformations at three levels: micro, meso, and macro. Excitation processes and resonant structures interact to produce events which take place or might take place in our everyday environment.

Since ecological models form the basis of ecologically-inspired composition, we will present some models that can be applied to generate and process sound material. The next sections presents an overview of three paradigmatic examples: bouncing, scraping, and breaking. Although the models are implemented in Csound, the descriptions of algorithms are not language-specific. Csound users can inspect the code included in the CD (Keller, 1999).

Implementing Ecological Models

When developing ecological models, the emphasis is not to obtain exact resynthesis of a given sound but to approximate the behavior of a class of sounds. Thus, statistical descriptions and perceptually relevant cues are more important than identical reproduction. In fact, literal resynthesis of a sound which serves as a validity test for most analysis / resynthesis techniques (Ellis, 1996; McAulay & Quartieri, 1986; Risset & Wessel, 1982), proves to be ecologically invalid. Ecological models produce sounds that can be identified as belonging to the same class. But each realization, or instance of given a model is never exactly the same as any other instance of the same model.

Granular synthesis

More than a synthesis technique, “granular synthesis is a way of realizing sound production models using locally defined waveforms” (De Poli & Piccialli, 1991). Authors generally refer to any windowing technique as a form of granular synthesis. From a wide perspective, short-time Fourier transforms, wavelets, filterbanks, etc. would be included in granular synthesis methods (Cavaliere & Piccialli, 1997).

Looking at the granular approach as a two-stage method, we can differentiate the control-function generation from the sound synthesis stage. First, we establish a time-frequency grid of grains (Roads, 1996, 172) by means of analysis (Short-Time Fourier Transform, Wavelet Transform) or algorithmic generation (screen, cloud, density). Then, we produce the sound by placing either synthesized grains (e.g., sine waves, filter parameters) or sampled-sound grains (from one or several sound files).

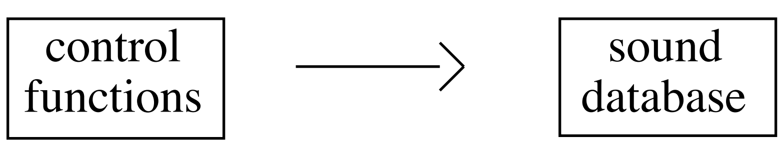

Figure 2. Granular synthesis as a two-stage method.

Table 3. Granular synthesis parameters.

parameter	grain rate	grain duration	grain amplitude	grain frequency	grain envelope	grain sound file	number of streams	grain overlap
level	local	local	local	local	local	local	global	global

Control functions

Whether the control functions are derived from analysis, or generated algorithmically, similarly to signals, they can be classified in two broad classes: (1) deterministic, and (2) stochastic. Paraphrasing Damper (1995, 258), a deterministic signal is one for which future values can be correctly predicted from a mathematical model of its generation process. Observed past values are used to find the parameters of the model. On the other hand, a stochastic signal is unpredictable because its generation process is too complex or poorly understood.

Deterministic processes can be produced by linear or nonlinear dynamical systems. A linear system is usually described by linear difference equations with constant coefficients (Damper, 1995, 36). Its output is a function of the input and the given coefficients (Bosch & Klauw, 1994, 9). These are some of the properties of linear systems: (a) the output is independent of previous inputs; (b) their impulse response is finite (FIR); (c) they are stable (Damper, 1995, 44). Examples of linear systems are the filters used in subtractive synthesis. By introducing feedback, the output of the system is made dependent on previous inputs. Thus, the impulse response becomes infinite and for some parameters the system may present instability and nonlinearity.

Based on these general classes of control functions, it is possible to group the synthesis methods in granular synthesis (as opposed to the analysis methods) in two rather simplified categories (Roads, 1997, 427): (1) synchronous, mostly based on deterministic functions; and (2) asynchronous, based on stochastic functions.

Synchronous methods are found in FOF synthesis (Rodet, 1984), VOSIM, quasi-synchronous GS, and pitch-synchronous granular synthesis (De Poli & Piccialli, 1991). Asynchronous methods have been used in synthesis by 'screens' (Xenakis, 1971), real-time granular synthesis (Truax, 1988), FOG synthesis (Clarke, 1996), and pulsar synthesis (Roads, 1997). In this context, the functions control the delay between grains for a single stream. Alternately, Clarke (1996) measures the time between grain onsets and uses this parameter to control grain rate.

There are some limitations in the traditional control method of independent grain streams, grain generators, or voices (Truax, 1988). As Clarke (1996) points out, these models do not take into account the difference between synchronized and independent grain generators. In ecologically-based granular synthesis, we use the term 'phase-synchronous' for several streams that share the same grain rate, and 'phase-asynchronous' for independent streams.

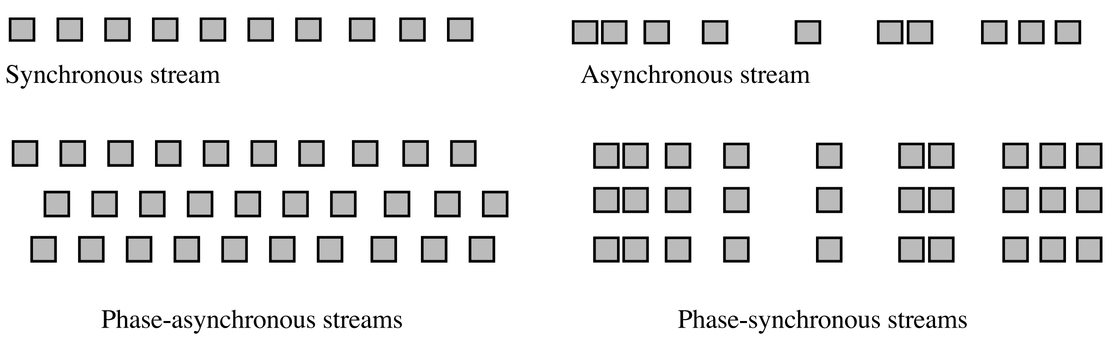

Figure 3. Classification of granular synthesis methods.

Sound database

As we mentioned previously, control functions define local parameters in granular synthesis. The relevance of each of these parameters depends on what GS approach is adopted. For example, envelope shape is important in FOF synthesis because this local parameter determines the bandwidth of the resulting formant. By contrast, the same parameter in asynchronous granular synthesis has little or no effect. Random sample-based processing causes spectral “blurring” and the sound is further modified by the complex interaction of overlapping spectrally rich grains. In part, this explains the gap between GS techniques that use simple synthetic grains to try to synthesize existing sounds, and the granular compositional approaches that start from more interesting and complex grains which produce less predictable results. “Tell me what grain waveform you choose and I’ll tell you who you are.”

A literature review has shown that GS techniques have used three types of local waveforms: (1) sine waves, in FOF synthesis (Rodet, 1984); (2) FIR filters derived by spectral analysis, in pitch-synchronous synthesis (Cavaliere & Piccialli, 1997; De Poli & Piccialli, 1991); and (3) arbitrary sampled sounds, in asynchronous granular synthesis (Truax, 1988), FOG (Clarke, 1996), and pulsar synthesis (Roads, 1997). Given that the local spectrum affects the global sound structure, we have used grain waveforms that can be parsed in short durations (20 to 200 ms) without altering the complex characteristics of the original sampled sound. Thus, we use water drops for stream-like sounds or pieces of bottles crashing for breaking-glass sounds.

Methods

The synthesis technique used in our study is implemented in Csound (Vercoe, 1993), and the grain events are generated with our own score generator, CMask (Bartetzki, 1997), and Algorithmic Composer Toolbox (Berg, 1998). The local parameters provided by the score determine the temporal structure of the resulting sounds. These parameters are processed by one or several instruments in the orchestra. The instruments function as grain stream generators. There are three possible configurations: (1) a single stream generator, (2) parallel phase-asynchronous stream generators, (3) parallel phase-synchronous stream generators.

The total spectral result is given by the interaction of the local waveforms with the meso-scale time patterns. Thus, the output is characterized by emergent properties, which are not present in either global or local parameters.

Table 4. Parameters of ecological models.

Grain	Sound sample defined by: (1) frequency, (2) amplitude, (3) duration, (4) envelope shape, (5) sound file.
Grain pool	Several short sound files. Ecological models use complete sound events, i.e., the attack and decay of the recorded sounds are kept intact. Micro-models allow to apply transformations to the original samples.
Grain rate	Delay between the onset of two consecutive grains in the same stream.
Grain duration	Time interval between the onset and the end of the grain.
Grain amplitude	Maximum amplitude over grain duration. The control function is usually normalized to 1.
Grain frequency	Sample rate, expressed as a transposition ratio. 1 is the original sound file frequency. 2 raises an octave and cuts the duration of the sample to half its original length.
Grain sample	Spectral and micro-temporal content of the grain. It depends on the pointer location within the sound file. Several sound files can be used simultaneously.
Grain envelope	Shape of attack and decay of the grain. Quasi-Gaussian in asynchronous granular synthesis. Ecological models generally do not need windowing because they use pre-defined sample pools.
Grain overlap	Time interval during which two or more grains are sounding simultaneously. The grain overlap is determined by the difference between grain rate and grain duration. There are three possible configurations: (1) positive, there is a gap between the end of a grain and the onset of the following grain; (2) zero, a grain starts when the previous ends; (3) negative, before the grain ends the next one starts. There can be as many overlapping grains as memory and patience allow.
Stream	Grains synthesized by a single grain generator, i.e., one oscillator in the Csound orchestra. There are three implemented instruments (which can be extended): (1) single stream generator, (2) multiple phase-asynchronous stream generators, (3) multiple phase-synchronous stream generators.
File pointer	Location in sound file. These are the four possible ways to access the file contents: (1) no reset, the file is read from beginning to end; (2) loop, the file is read from beginning to end repeatedly; (3) cycle, the file is read from beginning to end and backwards repeatedly; (4) random, the file is read at randomly picked locations. The first one is the standard one.

Procedure:

1. Collect several samples of everyday sounds produced by self-excited objects, such as running water or fire, and objects with an external source of energy, e.g., cracking wood, struck metal, etc.

2. Observe the temporal patterns and the spectral characteristics of the samples.

3. Extract grain samples to be used in the Csound synthesis language and define the meso-scale temporal behavior of the simulation.

4. Produce the synthetic sounds and compare results with the original samples.

Bounce

The bounce pattern can be approximated by an exponential curve or by a recursive equation. The former can only be used for one instance of the class of bounce sounds. On the other hand, the latter provides a general representation of all possible forms of bounce patterns. It can easily be adjusted just by changing the damping parameter. This function produces a family of exponential curves that we will use to control grain rate and grain amplitude.

Figure 4. Bounce control function.

This function provides an idealized bounce pattern in which damping is not affected by external factors such as surface irregularities or shape of the bouncing object (assumed to be perfectly round). Furthermore, we need to scale independently the damping factor controlling grain rate, from the one controlling grain amplitude. We have found that the range of damping values needs to be consistent with the object's elasticity. When simulating bouncing bottles, the damping factor for grain amplitude may be lower than the damping factor for grain rate. These settings allow the grain rate to reach sub-audio rate before the amplitude has been completely damped. The effect is a rising pitch that reproduces the phenomena heard in real bouncing bottles.

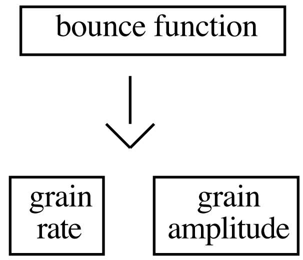

Figure 5. Simple bounce model.

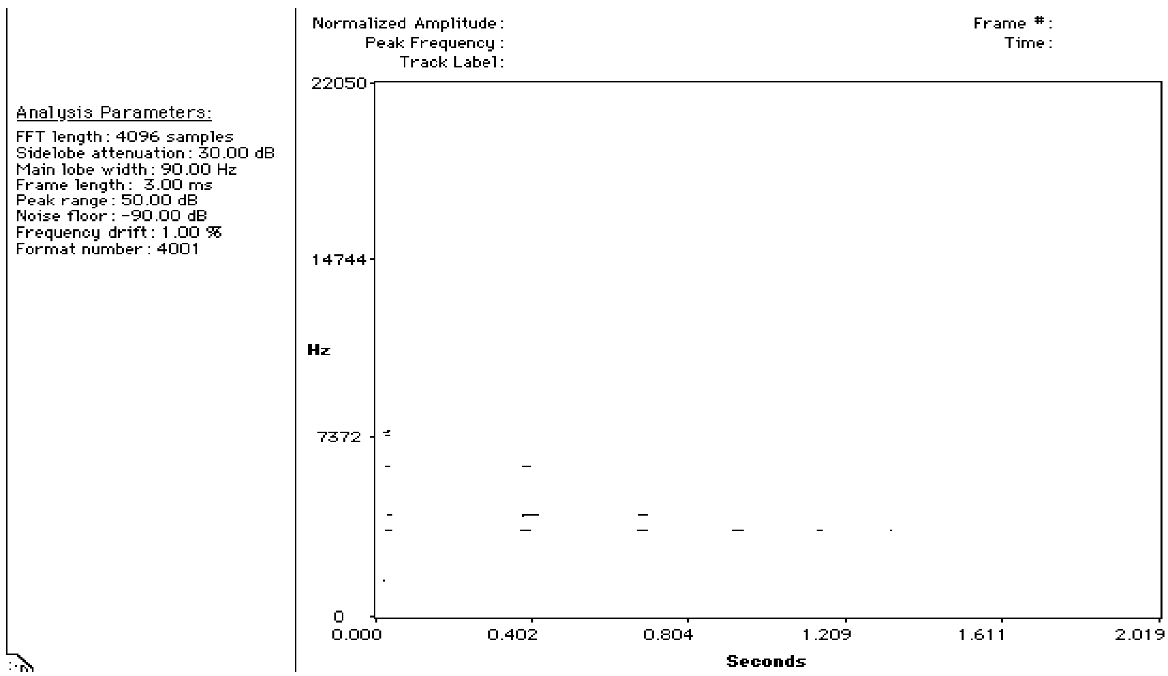

Figure 6. Spectrogram of bouncing bottles.

The initial grain rate is equivalent to the acoustic cue used to determine the distance between the bouncing object and the surface (in this context surface can also be understood as another object). Thus, if the first collision of object A occurs at 0 ms. and the second at 50 ms., and the first and second collisions of object B are at 0 ms. and 200 ms., we can safely infer that object B is initially more distant from the surface than

A is, and that this difference is not determined by different elasticities. Contrastingly, the change over time in grain rate which is controlled by the damping factor provides a cue to the elasticity, or 'bounciness,' of the object and the surface. Given the same surface, a very elastic object exhibits a slower acceleration in collision rate than a less 'bouncy' one.

No surface is perfectly smooth and no object is perfectly round. Therefore, random variations in bouncing behavior should be expected. Given the inability of the auditory system to identify the source of these random variations, i.e., is the surface uneven, or is the object not round? We can express these irregularities as a single random parameter added to the damping factor in the bounce function. The roughest surfaces and the most irregular objects get the highest random variations. Similarly, a pool of varied grain samples can be used to account for changes in spectral characteristics at a micro level. Therefore, a more refined bounce model should include random changes in grain rate, grain amplitude, and grain sample. Time-dependent changes in spectral profile, to account for progressive loss of energy, could also be included.

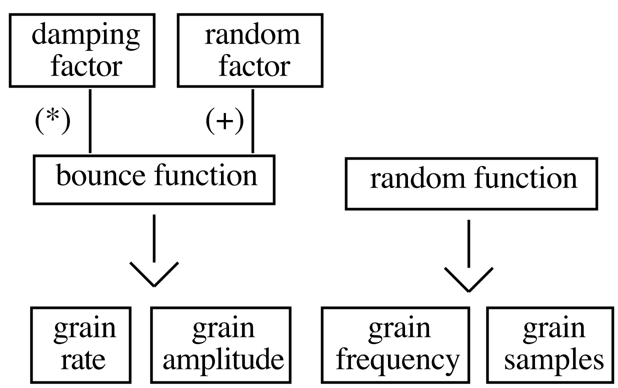

Figure 7. A more refined bounce model.[2]

No sound in nature happens exactly the same way twice. The unnatural quality of loops and repeated sample playback is caused by the lack of variation at the micro and meso-level sound organization. The auditory system can readily recognize sounds with micro-level repetitions as ecologically unfeasible events (Tróccoli & Keller, 1996). On the other hand, sudden changes in micro-level characteristics usually cue the

2 Arrows indicate data flow. Solid lines stand for parameter settings.

beginning of a new event. Thus, synthesized environmental sounds should exhibit dynamic micro-level characteristics within carefully constrained ranges.

In models that use sampled grains, grain frequency stands for change in sample rate. This rate is expressed as a ratio. Thus, 1.0 keeps the original sample rate and 0.5 drops the frequency one octave and doubles the length of the sampled sound. We have observed that variations in a range of 0.1% (.001) to 10% (.1) produce very subtle to dramatic effects. Of course, this depends on the interaction with other variables and on the characteristics of the samples used.

Scrape

The scrape model that will be discussed in this section is based on the assumption of a single point / surface interaction. This is clearly an idealized case but provides a basis for more complex models that could be controlled by high-level transformations. The results obtained with this model are perceptually more satisfying than the ones reported in (Gaver, 1993, 233), i.e., frequency of band-limited noise corresponding to dragging speed, and filter bandwidth correlated to roughness of the surface.

Scraping is usually linked to gestures produced by human agents. Therefore, scrape events should be constrained to finite durations feasible by human movement (unless the acoustic cues are intended to suggest a machine-generated process). The ecological event satisfies perfectly these requirements. Given this context, the scraping action cannot start suddenly from a high-energy level but needs to develop slowly from zero amplitude. To mimic this behavior, we use a tendency mask which allows for random variations in grain amplitude and constrains the initial and final ranges to zero. Similarly, increase or decrease in scrape speed happens at relatively slow rates. We simulate scraping action by controlling the grain rate with a fairly simple algorithm.

The scrape function consists of a random number generator that sends values uniformly distributed over zero and under zero to an accumulator. The requirement of a single absolute limit for positive and negative values is needed to avoid overflowing the accumulator or getting stuck at a boundary value. The accumulator is limited by a low and a high boundary which establishes the fastest and slowest possible rates. Depending on the number of positive or negative values produced by the number generator, the scraping speed increases or decreases accordingly. The absolute limit value given to the number generator defines a virtual time grid, 'grid factor.' If this value is high the delay among grains is usually long, if it is slow average grain rate will be fast.

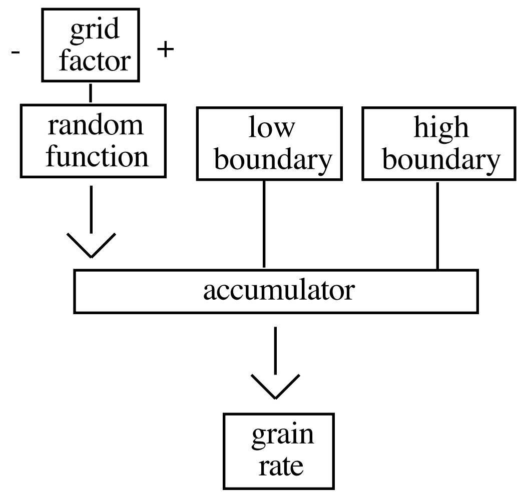

Figure 8. Scrape control function.

Aside from the gestural action produced by a human agent, scraping is also determined by the characteristics of the surface being scraped. Given that the gesture is assumed to be fairly constant, irregularities in the rate of interaction between point and surface can be attributed to surface roughness. By randomizing the grain rate control by a small percentage (up to 10%), we obtain various degrees of roughness without affecting the gestural feel of the model. The material of the surface being scraped is established by the timbral quality of recorded grains or by modifying the spectral result through the use of resonators.

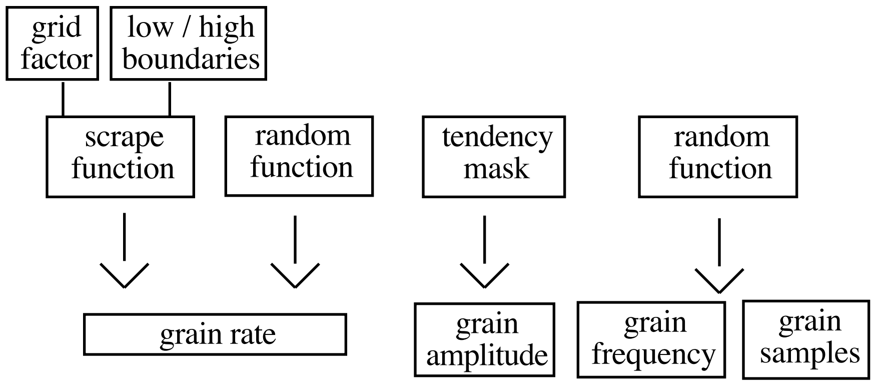

Figure 9. Simple scrape model.

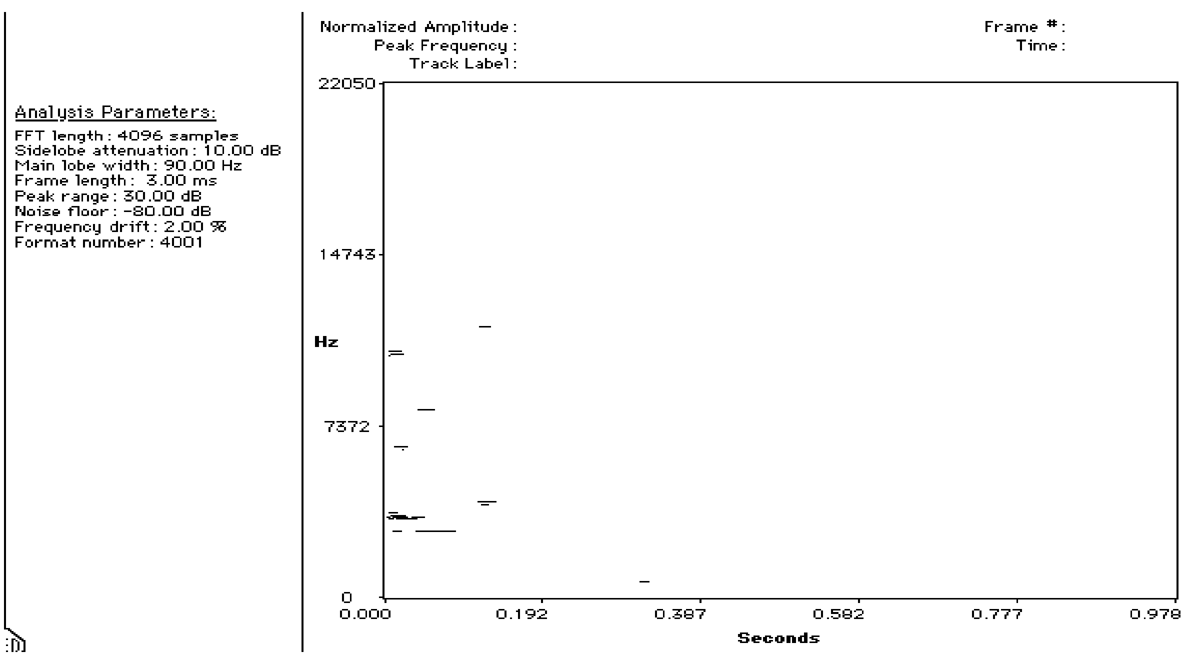

Figure 10. Spectrogram of breaking glass.

Break

Breaking patterns are time-varying, highly structured events. Straightforward random distributions or superimposed bounce patterns (Warren & Verbrugge, 1984) do not approximate the complex behavior of these sounds. A multi-level approach is needed. Careful observation of recorded breaking glass sounds provides some useful insights. The initial strike produces broadband resonances in the whole object. These resonances last for approximately five milliseconds and can be simulated with an enveloped burst of white noise. As explained in (Keller & Rolfe, 1998), the envelope shape provides a simple way to filter out high frequencies. Alternatively, a sampled grain can be used.

After the first five to ten milliseconds the object starts to break. Many small pieces hit a surface (possibly the floor) producing a dense cloud of impacts. Although these sounds share the spectral characteristics of the whole object, given the pieces' small size their spectral profile has a higher frequency content. A more noisy sound is produced by glass pieces hitting each other. Thus, it seems reasonable to use a short-impact grain pool with two spectral characteristics: glass hitting a surface, and glass hitting glass. The amplitude of these impacts can be approximated by an exponential decay.

A third complex sub-event in breaking consists of randomly varying bouncing patterns. These patterns are produced by pieces scattered around the initial impact spot. Depending on the surface elasticity and the shape of the broken pieces, these patterns range from highly random with high damping factor to slightly random with low damping factor. The last configuration produces bounce-like patterns. The algorithms used for bouncing can be used to produce these meso patterns.

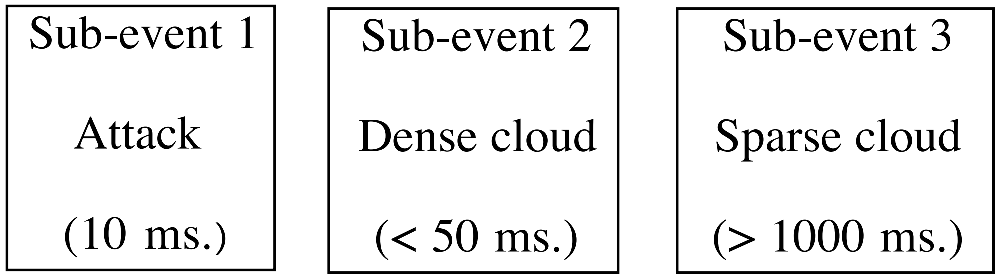

Figure 11. Breaking glass model.

Summarizing, breaking glass is composed of three sub-events: (1) A broadband attack, lasting five to ten milliseconds. (2) A dense random cloud of short impacts: glass onto surface and glass onto glass, lasting less than fifty milliseconds. (3) Several overlapped bouncing patterns with durations depending on the surface elasticity (one second or more).

Summary

This section discussed the implementation details of ecological models, focusing on three examples: bouncing, scraping, breaking. Ecological models use grain pools of isolated meaningful events taken from real-world sounds. The distribution of grains is controlled from meso-level functions, following patterns found in nature.

touch'n'go: A Case Study

touch'n'go, or **toco y me voy**, is a fifty-minute tape and text piece. The piece is divided into eleven sections of varying length, going from less than a minute to ten minutes. Each section explores simple, direct ideas expressed in short Spanish and English poetry or prose. Musically, each section uses a limited number of sources which are usually constrained to a single or very few sound classes. **touch'n'go** is one of the first attempts to apply a consistent ecological approach in the context of a tape composition.

This section will succinctly describe the main concepts which inform **touch'n'go**, its structural characteristics, the ideas presented in text format, and a few illustrative examples of the processes involved in the development of the sounds. The section concludes with a discussion of why **touch'n'go** should be labeled as an ecologically-based composition.

The text

All the text of the piece is laid out in HTML (HyperText Markup Language) format (Keller, 1999). Each section has its self-contained text but all the text forms a network of interconnected ideas physically linked by hypertext anchors. The text consists of excerpts of *Martín Fierro* (Hernández, 1872), *The Garden of the Forking Paths* (Borges, 1962), two verses of a poem by Julio Cortázar (1966), and poetry and prose written by myself (Keller, 1999).

To summarize the ideas contained in the piece would be pointless, so we will describe its structural and thematic axis. The piece's structure (or lack of it) is centered on the idea of the forking path (Borges, 1962). In other words, the piece is designed by the listener / reader at the moment he establishes a path through the material of the piece. The texts written for this piece make use of algorithm-inspired methods, such as progressive expansion (**Coin a name**) or progressive reduction of material (**sCRATch**), modular and randomly combined sections (**Pandemonio**), and a seven-element series (**toco y me voy**).

As discussed in (Keller, 1998a), **touch'n'go** makes reference to the key problem of cross-culturalism in Latin American and Argentinean intellectual practices. Jorge Luis Borges and Julio Cortázar have dealt with cross-cultural issues in all their works. Going back to one of the pillars of Argentine culture, the traditional epic poem *Martín Fierro* by José Hernández addresses the interaction between indian, black and white culture and how it shapes the emergence of a new socio-cultural identity: the gaucho. A paradigmatic example of this social identity is Atahualpa Yupanqui. This philosopher, poet and composer has created a body of work which is assimilated into popular culture not by market dynamics, but by the force of its internal coherence and its direct reference to everyday life. As we have said in (Keller, 1998a) it is clear that if a concept of Argentine identity can be used at all, it is determined by the dynamics between local and immigrant cultures. In this sense, Argentinean identities are not one homogeneous, abstract entity, but several coexisting and ever-changing approaches to the world. This dynamics ensures a broad and flexible understanding of cross-cultural issues, and explains the traditional Latin American and Argentinean skeptical rejection of fundamentalist and pseudo-"universal" theories.

The music

Although each of the sections of **touch'n'go** is a complete, self-contained piece, most of them share material and provide sound references to each other. Thus, the piece can also be divided into four sections of ten to eighteen minutes that comfortably fit a concert setting. Besides using common sound sources, most pieces were created with similar sound processing tools.

All the material for **touch'n'go** was produced with ecological models (Keller & Truax, 1998) implemented in Csound (Vercoe, 1993), CMask (Bartetzky, 1997), and the AC Toolbox (Berg, 1998), and with real-time asynchronous granular synthesis using MacPOD (Keller & Rolfe, 1998). MacPOD was developed by Chris Rolfe following the POD model documented in Truax (1988). MacPOD was mostly used for time-stretching and real-time gestural transformation of material. Its intuitive and responsive interface allows for transformations which are cumbersome or even impossible to do in batch mode. An example of the effectiveness of this sound processing software is provided in the voice material in **let me see . . . how can I word it?** The exaggerated inflections in Yupanqui's voice were created by changing the stretch ratio parameter at the exact beginning of consonants and vowels. Thus producing inflections such as 'muuuuuy, ' meaning 'very' in Spanish, or 'proffffffundo' which means 'deep' and lasts for approximately a minute.

Direct sound references (Shepherd, 1992) and gestures (Smalley, 1993) which are directly related to the meaning of each section are used throughout the piece. **sCRATch** uses scraping as its basic sound model. **spill, spiel, spoil** uses spilling and pouring. **Coin a name** employs splashing and water sounds. **Action to be taken in the event of a fire** uses the sound of a lit match as its only source material, and different models of fire. **Realpolitik** shares some elements with **Action**, and introduces cracking woods, shattered glass, and explosion-like sounds. **A waltz in a ball** uses only a sample of a rolling marble.

Two structural processes are exploited in **touch'n'go**: emergent macro-structural properties by interaction of lower-level elements, and isomorphism. **Farewell, welfare** provides an example of isomorphous processes at a conceptual, algorithmic, and perceptual level. The piece develops the idea of the infinite labyrinth: a straight line traveled by covering half its length every time (Borges, 1962; Kampis, 1991). The algorithm used for producing all sounds in this piece is a variation of the Karplus-Strong (1983) model and uses a single short sample as source material. This sample is 'walked' by two pointers at different sample rates producing an ever-rising or ever-falling sound, depending on the parameters fed to the model. The sound is akin to a combination of Shepard tones with self-similar events that occur at ever-expanding time-spans.

The model developed for **Vox Populi** uses a small sample pool of conch shell sounds. Being harmonic with some noise content, these samples when combined randomly approximate the behavior of choir-like formants (Keller & Rolfe, 1998; Rodet, 1984; Truax, 1992). These formants are the result of the interaction of granular samples at a meso level. So the effect at a macro level is qualitatively different from the source sound characteristics.

A similar phenomenon can be observed in the 'structured rain' material used in **least, but not last**. This sound was generated by a three-stage process. First, several different drop sounds were produced by convolution. Then these drops were organized as constrained random meso-level events. Finally, the events were combined in a slowly evolving macro-structure, which distributed the meso-events in a slightly irregular rhythmic pattern. The result is a sound that resembles wind and rain with metallic reverberations in the background. **least, but not last** also makes use of two contrasting spaces - metaphorically speaking, the space of the living and the space of the dead. The first space was produced by mixing a recording of a big, open building space. The other space was realized by applying convolution to the grain samples.

As can be inferred from our theoretical discussion, cultural references are not absent from **touch'n'go**. The most obvious and probably most meaningful one is the use of the voice of the great poet and philosopher Atahualpa Yupanqui (1980). Although his voice is kept recognizable, we have taken care of always processing it to heighten its presence and to give it a deeper timbral quality. Sound transformations are tightly related to the meaning of the text and its metaphoric implications. The first verses can be translated thus: *during those times such things happen that do not happen anymore*. The atmosphere of story-telling set by the verses is reinforced by an outdoor background sound which frames all sound events in the first section. As previously mentioned, the word 'deep' in the verses: *if the river is wide and deep, that one who swims well gets across*, is time-stretched to last approximately a minute. The same verses are used to produce a hybrid sound by means of Linear Predictive Coding. This technique lets us filter Yupanqui's voice through patterns produced by synthesized bubbles. The effect is akin to whispering inside water.

Pandemonium incorporates cultural references in its sample materials and in its algorithmic models. Brazilian percussive instruments, such as *cuica* and *pandeiro*, were transformed by ecological models to provide a continuum of instrumental and environmental sounds. The meso time-patterns employed closely approximate traditional Brazilian rhythms without losing an overall 'careless' and 'aimless' feel. The sound palette is further enhanced by introducing hybrid sounds such as the breaking-glass *cuica*, or the broken *pandeiro*. These sounds were generated by applying breaking time-patterns to instrumental sampled grains.

The format

touch'n'go proposes a system that requires a flexible and modular format. The ideology underlying the piece is expressed in its thematic axis: the gaucho or the traveler, in its structure: a path established by the listener / reader, and in its sound material: direct references to a social and environmental context. The ideas developed in the text are related to (and match one-to-one) the sections of the tape piece. Nevertheless, they are given to the reader as a network of short excerpts that are freely combined. This same process is mirrored in the music.

Once I had established a rough plan for **touch'n'go**, a pressing question had to be answered: what format could comprise sound and text in a modular, readily accessible layout? Both Enhanced CD and CD-ROM seemed to fulfill these requirements. I chose to present two versions of the piece: a data CD-ROM for the Macintosh platform and a commercially released Enhanced CD, playable on any CD player, Macintosh or IBM-PC computer. The text material of the piece, including the Csound code, is freely available on my website (Keller, 1999). I only refer to the data CD-ROM in this documentation.

The random-access structure of the data CD-ROM allows the user to freely peruse the text in his browser and simultaneously listen to any section of the music. This has two implications on the way the piece is developed and perceived. First, the material cannot be structured in a linear, one-dimensional fashion because in any case this layout can be overridden by an active user. Second, the piece requires a thoughtful involvement from the user: clicking randomly and arbitrarily cutting sections becomes boring and uninteresting. (Of course, following the spirit of the piece we establish no rules or restrictions whatsoever). Given this context, the user / listener / reader is asked either to leave his passive role and engage in an active, conscious, creative process, or not to engage at all.

Ecological echoes

After presenting the basic methods and concepts behind ecological modeling and composition and after providing an overview of the issues addressed by **touch'n'go**, it seems wise to restate the elements that characterize **touch'n'go** as an ecologically-inspired piece.

The material used in **touch'n'go** makes direct reference to everyday and environmental events. This material also incorporates cultural elements that place the piece within a very specific and distinct cultural context: Latin America. Most sound processing is done through models that apply ecologically feasible transformations on the material. Transformations that produce abstract material are used sparingly. They are usually introduced by means of simple compositional processes, e.g., expansion of the number of elements, time-stretching, hybridization, resonant filtering with uniform spectral characteristics. Sound spaces are kept consistent and stable, i.e., there are no spatial changes without appropriate cues and transformations in the related sound events.

The sound unit used throughout the piece is the event. In other words, there is not a single sound that has a suddenly cut decay, or an attack without transients. Most sounds were produced by the interaction of exciting systems - granular algorithms - with resonant systems - generic physical models. At a meso level, all algorithms follow ecologically feasible patterns, that is, they produce events that occur or may occur in our terrestrial environment. At a macro level, the piece explores the emergent properties resulting from the interaction of complex elements at a micro and meso level. The format chosen for the piece requires the active participation of a conscious user. In other words, the listener / reader engages in a process of mutual-determination (Varela et al., 1989; Keller, 1998a) with the material of the piece. Thus, the general model of environment-individual interaction is scaled down to the artwork-individual pattern formation process. Indeed, if **touch'n'go** is not an ecologically-based piece, it is nothing at all.

References

Ballas, J.A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 19(2), 250-267.

Bartetzki, A. (1997). CMask. *Software package*. Berlin: STEAM.

Berg, P. (1998). Algorithmic Composer Toolbox. *Software package*. The Hague.

Boon, J.P., Noullez, A., & Mommen, C. (1990). Complex dynamics and musical structure. *Interface*, 19, 3-14.

Bosch, P.P.J., & Klaw, A.C. (1994). *Modeling, Identification, and Simulation of Dynamical Systems*. Boca Ratón, FL: CRC.

Bregman, A.S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Bregman, A.S. (1993). Auditory scene analysis: hearing in complex environments, *Thinking in Sound: the Cognitive Psychology of Human Audition*, S. McAdams & E. Bigand (Eds.). Oxford: Oxford University Press, 10-36.

Cavaliere, S., & Piccialli, A. (1997). Granular synthesis of musical signals, *Musical Signal Processing*, C. Roads, S.T. Pope, A. Piccialli, & G. De Poli (Eds.). Lisse: Swets & Zeitlinger, 155-186.

Chafe, C. (1989). Simulating performance on a bowed instrument, *Current Directions in Computer Music Research*, M.V. Mathews & J.R. Pierce (Eds.). Cambridge, MA: MIT Press.

Clarke, J.M. (1996). Composing at the intersection of time and frequency. *Organised Sound*, 1(2), 107-117.

Cook, P.R. (1997). Physically informed sonic modeling (PhISM): synthesis of percussive sounds. *Computer Music Journal*, 21(3), 38-49.

Damper, R.I. (1995). *Introduction to Discrete-Time Signals and Systems*. London: Chapman & Hall.

Darvishi, A., Munteanu, E., Guggiana, V., Schauer, H., Montavalli, M., & Rauterberg., M. (1995). Designing environmental sounds based on the results of interaction between objects in the real world, *Human-Computer Interaction: Interact 1995*, N. Knut, P. Helmersen, D.J. Gilmore, S.A. Arnesen (Eds.). London: Chapman & Hall, 38-42.

Darwin, C.J. (1990). Environmental influences on speech perception, *Advances in Speech, Hearing and Language Processing*, Vol. 1, W.A. Ainsworth (Ed.). London: JAI Press, 219-241.

De Poli, G., & Piccialli, A. (1991). Pitch-synchronous granular synthesis, *Representations of Musical Signals*, G. De Poli, A. Piccialli, & C. Roads (Eds.). Cambridge, MA: MIT Press.

Di Scipio, A. (1994). Micro-time sonic design and timbre formation. *Contemporary Music Review*, 10(2), 135-148.

Ellis, D.P.W. (1996). Prediction-driven computational auditory scene analysis, *PhD. Thesis in Electrical Engineering*. Cambridge, MA: MIT Media Lab.

Fletcher, N.H. (1992). *Acoustic Systems in Biology*. New York, NY: Oxford University Press.

Gaver, W.W. (1993). Synthesizing auditory icons, *Proceedings of the INTERCHI 1993*. New York, NY: ACM, 24-29.

Georgescu, C., & Georgescu, M. (1990). A system approach to music. *Interface*, 19, 15-52.

Gibson, J.J. (1966). *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton Mifflin.

Gilkey, R.H., & Anderson, T.R. (Eds.) (1997). *Binaural and Spacial Hearing in Real and Virtual Environments*. Mahwah, NJ: Erlbaum.

Grantham, D.W. (1995). Spatial hearing and related phenomena, *Hearing*, B.J.C. Moore (Ed.). San Diego, CA: Academic Press, 297-345.

Handel, S. (1995). Timbre perception and auditory object identification, *Hearing*, B.C.J. Moore (Ed.). New York, NY: Academic Press.

Iverson, P., & Krumhansl, C.L. (1991). Measuring similarity of musical timbres. *Journal of the Acoustical Society of America*, 89(2), 1988.

Jones, M.R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323-355.

Kampis, G. (1991). *Self-Modifying Systems in Biology and Cognitive Science*. Oxford: Pergamon Press.

Karplus, K., & Strong, A. (1983). Digital synthesis of plucked-string and drum timbres. *Computer Music Journal*, 7(2), 43-55.

Keller, D. (1998a). The social domain. *Unpublished paper*. Burnaby, BC: Simon Fraser University.

Keller, D. (1998b). The perceptual domain. *Unpublished paper*. Burnaby, BC: Simon Fraser University.

Keller, D. (1998c). The acoustic domain. *Unpublished paper*. Burnaby, BC: Simon Fraser University.

Keller, D. (1999). *touch'n'go / toco y me voy*. Enhanced Compact Disc. Burnaby, BC: earsay productions. http://www.sfu.ca/~dkeller

Keller, D., & Silva, C. (1995). Theoretical outline of a hybrid musical system, *Proceedings of the II Brazilian Symposium on Computer Music*. Canela, RS: SBC.

Keller, D., & Truax, B. (1998). Ecologically-based granular synthesis, *Proceedings of the International Computer Music Conference*. Ann Arbor, MI: ICMA. http://www.sfu.ca/~dkeller

Kelso, J.A.S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, MA: MIT Press.

Krumhansl, C.L. (1990). *Cognitive Foundations of Musical Pitch.* New York: Oxford University Press.

Lass, N.J., Eastham, S.K., Parrish, W.C., Scherbick, K.A., & Ralph, D.M. (1982). Listeners' identification of environmental sounds. *Perceptual and Motor Skills*, 55, 75-78.

Lynn, P.A., & Fuerst, W. (1994). *Digital Signal Processing with Computer Applications.* Chichester: John Wiley.

Massaro, D.W., & Cowan, N. (1993). Information processing models: microscopes of the mind. *Annual Review of Psychology*, 44, 383-425.

McAdams, S. (1987). Music: a science of the mind? *Contemporary Music Review*, 2, 1-61.

McAdams, S. (1993). Recognition of sound sources and events, *Thinking in Sound*, S. McAdams and E. Bigand (Eds.). Oxford: Oxford University Press.

McAulay, R.J., & Quartieri, T.F. (1986). Speech analysis/synthesis based on a sinusoidal representation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(4), 744-754.

Michaels, C.F., & Carello, C. (1981). *Direct Perception*. Englewood Cliffs, NJ: Prentice-Hall.

Moore, R.F. (1990). *Elements of Computer Music*. Englewood Cliffs, NJ: Prentice-Hall.

Orphanidis, S.J. (1996). *Introduction to Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.

Parncutt, R. (1989). *Harmony: a Psychoacoustic Approach*. Berlin: Springer-Verlag.

Risset, J.C., & Wessel, D.L. (1982). Exploration of timbre by analysis and synthesis, *The Psychology of Music*, D. Deustch (Ed.). Orlando, FL: Academic Press, 26-58.

Roads, C. (1996). *The Computer Music Tutorial*. Cambridge, MA: MIT Press.

Roads, C. (1997). Sound transformation by convolution, *Musical Signal Processing*, C. Roads, S.T. Pope, A. Piccialli, & G. De Poli (Eds.). Lisse: Swets & Zeitlinger, 411-438.

Rodet, X. (1984). Time-domain formant wave-function synthesis. *Computer Music Journal*, 8(3), 9-14.

Schafer, R.M. (1977). *The Tuning of the World*. New York: Knopf.

Schumacher, R.T., & Woodhouse, J. (1995). The transient behaviour of models of bowed-string motion. *Chaos*, 5(3), 509-523.

Shaw, R.E., Turvey, M.T., & Mace, W.M. (1981). Ecological psychology: The consequence of a commitment to realism, *Cognition and the Symbolic Processes*, W. Weimer & D. Palermo (Eds.). Hillsdale, NJ: Erlbaum. Vol. 2, 159-226.

Shepherd, J. (1992). Musical as cultural text, *Companion to Contemporary Musical Thought*, Vol.1, J. Paynter, T. Howell, R. Orton, & P. Seymour (Eds.). New York, NY: Routledge, 128-155.

Smalley, D. (1993). Defining transformations. *Interface*, 22, 279-300.

Smith, J.O. (1991). Viewpoints on the history of digital synthesis, *Proceedings of the International Computer Conference*. San Francisco: ICMA, 1-10.

Smith, J.O. (1992). Physical modeling using digital waveguides. *Computer Music Journal*, 16(4), 74-87.

Smith, J.O. (1997). Acoustic modeling using digital waveguides, *Musical Signal Processing*, C.Roads, S.T. Pope, A. Piccialli, & G. De Poli (Eds.). Lisse: Swets & Zeitlinger, 221-263.

Tróccoli, B.T., & Keller, D. (1996). The function of familiarity in timbre recognition. *Technical Report* (in portuguese). Brasilia, DF: FAP-DF.

Truax, B. (1988). Real-time granular synthesis with a digital signal processor. *Computer Music Journal*, 12(2), 14-26.

Truax, B. (1990). Chaotic non-linear systems and digital synthesis: an exploratory study. *Proceedings of the International Computer Conference*. San Francisco: ICMA, 100-103.

Truax, B. (1992). Electroacoustic music and soundscape: the inner and outer world, *Companion to Contemporary Musical Thought*, Vol. 1, J. Paynter, T. Howell, R. Orton, & P. Seymour (Eds.). London: Routledge, 374-398.

Truax, B. (1994). Discovering inner complexity: time shifting and transposition with a real-time granulation technique. *Computer Music Journal*, 18(2), 38-48.

Truax, B. (1996). Soundscape, acoustic communication and environmental sound composition. *Contemporary Music Review*, 15(1), 47-63.

Varela, F.J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Vercoe, B. (1993). Csound. *Software package*. Cambridge, MA: MIT Media Lab.

Warren, W.H., Kim, E.E., & Husney, R. (1987). The way the ball bounces: visual and auditory perception of elasticity and control of the bounce pass. *Perception*, 16, 309-336.

Warren, W.H., & Verbrugge, R.R. (1984). Auditory perception of breaking and bouncing events: a case study in ecological acoustics. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 704-712.

Westerkamp, H. (1996). *Transformations*. Compact Disc. Montreal: Empreites Digitales.

Wishart, T. (1996). *On Sonic Art*. Amsterdam: Harwood Academic Publishers.

Xenakis, I. (1971). *Formalized Music*. Bloomington, IN: Indiana University Press.

Young, J. (1996). Imagining the source: The interplay of realism and abstraction in electroacoustic music. *Contemporary Music Review*, 15(1), 73-93.

Literary References

Borges, J.L. (1962). El jardín de los senderos que se bifurcan (The garden of forking paths), *Ficciones*. Grove Press. Translated by Helen Temple & Ruthven Todd.

Cortázar, J. (1966). Razones de la cólera, *La vuelta al día en ochenta mundos*. México: Siglo Veintiuno. (Fifth edition, p. 198).

Hernández, J. (1968). *Martín Fierro*. Buenos Aires: Centro Editor de América Latina. (First edition published in 1872). Translated by Ivan Roksandic.

Yupanqui, A. (1980). *El payador perseguido*. Compact Disc. Buenos Aires.