

Una Visita al "Center for Computer Research in Music and Acoustics"

Ricardo Dal Farra

"Los parlantes controlados por computadora forman el medio más general de producción sonora que existe, pero hay sin embargo enormes dificultades que deben ser superadas para que este medio sea musicalmente provechoso. La música no surge fácilmente de las máquinas. Esto es verdad tanto si la máquina es un instrumento musical de tipo tradicional, o una computadora programada para producir sonido musical. En el caso de los instrumentos musicales tradicionales, años de entrenamiento son requeridos del ejecutante y el constructor de instrumentos; y en el caso de la computadora, un conocimiento sustancial acerca de procesamiento de señales, acústica y psicoacústica, es requerido del compositor/músico. Es con este nuevo instrumento, la computadora, que defrontamos nuevos problemas cuyas soluciones nos han llevado a trascender el medio, aumentando nuestro conocimiento, y enriqueciendo nuestra experiencia en el más amplio sentido."

John Chowning (de "Music from Machines", STAN-M-64, 1990)

Durante una reciente estadía en el "Center for Computer Research in Music and Acoustics" - CCRMA - de la Universidad de Stanford (al sur de la ciudad de San Francisco, en California, U.S.A.), tuve la oportunidad de conocer más de cerca los estudios y desarrollos que allí se están llevando adelante.

A partir del trabajo interdisciplinario de compositores y científicos que emplean a las computadoras como un nuevo medio artístico-musical y como una herramienta para sus investigaciones,

se desarrollan en CCRMA teorías y experiencias en campos tan diversos como la síntesis de sonido a través de modelos físicos, el reconocimiento óptico de una partitura, la producción de instrumentos musicales y controladores alternativos como el Celletto, el Radio Drum y el Biomuse, la confección de un archivo musical con grabaciones digitales de las primeras obras electroacústicas realizadas en todo el mundo, la creación de lenguajes de programación aplicados a la composición musical, la implementación de un sistema que permita controlar en tiempo real el movimiento espacial del sonido, el estudio del proceso por el cual el sistema auditivo separa los sonidos provenientes de distintas fuentes, la transformación interactiva por medio de computadoras del sonido de instrumentos musicales acústicos, y el análisis de factores psicoacústicos que influyen sobre la entonación musical, entre muchos otros.

El personal del CCRMA está integrado por John Chowning como director, Chris Chafe, coordinador técnico, Patte Wood, directora administrativa, Glen Diener, programador científico, Jay Kadis, ingeniero de audio, y compositores, docentes, e investigadores como John Pierce, Perry Cook, Leland Smith, Max Mathews, Marcia Bauman, Earl Schubert, William Schottstaedt, Bernard Mont-Reynaud, y Julius Smith. Completan este cuadro un considerable número de estudiantes avanzados y compositores invitados.

En cuanto al equipamiento de la institución, se utilizan fundamentalmente computadoras NeXT, y también varias Macintosh y PC AT&T. Cada NeXT viene ya preparada para permitir el procesamiento digital de señales de audio (Motorola 56001 DSP) con alta calidad, aunque algunas de ellas tienen un número adicional de plaquetas (Ariel Quint Processor) con varios procesadores en cada una. Algunas computadoras Macintosh han sido también equipadas con sistemas que permiten la grabación y edición digital de señales de audio (Digidesign Sound Accelerator y Spectral Innovations). La lista es extensa y continúa con: impresoras laser, sistemas analógicos y digitales para el almacenamiento de señales

de audio, diversos sintetizadores, procesadores y controladores de origen comercial (Yamaha, E-mu, Korg, Roland, J. L. Cooper, Buchla, IVL, etc.) conectados a las computadoras a través de MIDI, un estudio con posibilidades de grabar y mezclar en forma totalmente digital,

Algunos de los múltiples proyectos que están siendo desarrollados por músicos e investigadores en CCRMA son comentados brevemente a continuación.

MODELOS FISICOS

Según Chris Chafe, "la década pasada ha visto un rápido crecimiento en la aplicación de la tecnología digital a la música, y consecuentemente ha ido aumentando el contacto de los músicos con las computadoras. Los sintetizadores digitales se han convertido en un instrumento común, las estaciones de trabajo que utilizan audio digitalizado están reemplazando velozmente a los dispositivos analógicos de grabación, el compact disc es el medio más popular de reproducción, y tanto la composición como la impresión musical está asistida con una amplia variedad de programas.

Las mismas limitaciones que persisten en los sintetizadores digitales también aparecen en las aplicaciones de las computadoras en lo que hace a la voz hablada y la animación. Cuando se requiere alta calidad, nada reemplaza aún la cosa real, así nosotros aún utilizamos mensajes grabados para asistencia telefónica, y artistas humanos para lograr mejores animaciones en las películas. Solo pequeños progresos han sido hechos hacia la síntesis de ambientes totalmente automatizados, continuos, que puedan convencer en la imitación del habla natural o el movimiento de un carácter animado simulando uno vivo.

Los algoritmos empleados actualmente en los sintetizadores digitales caen dentro de una de estas tres categorías: aquellos basados en modelos espectrales de instrumentos (como la FM), muestras ("samples" - reproducción de sonidos grabados, cortos y aislados), o combinaciones híbridas de ambos. Cada aproximación

tiene sus propias ventajas particulares, y en consecuencia cada estudio tiene una combinación de estos instrumentos (interconectados via MIDI).

Una nueva clase de algoritmo de síntesis ha surgido de investigaciones recientes, que muestra una muy mejorada conexión en lo que hace a la articulación tanto en instrumentos musicales como en la voz humana. Las simulaciones están basadas en descripciones de la mecánica de un instrumento y son clasificadas, en un sentido amplio, como 'modelos físicos'. Interesantes resultados han sido presentados en instrumentos de viento, cuerdas frotadas y pulsadas, percusión, y voz cantada.

Aunque la idea de emplear modelos físicos para sintetizar sonidos no es nueva (los primeros modelos fueron desarrollados en los Laboratorios Bell en la década del 60), no había sido posible hasta ahora llevarla a la práctica. Así, la física de distintas familias instrumentales o fuentes sonoras es simplificada para crear modelos matemáticos que puedan ser repetidos en una computadora.

En los modelos físicos desarrollados los transientes de ataque, la transición entre notas ligadas, y los detalles expresivos que hacen a la interpretación de una frase musical son extremadamente buenos. El control del sonido es intuitivo, ya que el modelo responde como el instrumento a las manipulaciones del ejecutante. Los parámetros de control están directamente relacionados con los controles físicos con los que el músico está familiarizado (ej.: embocadura, presión del aire, ...).

La síntesis a partir de modelos físicos no ha sido implementada aún en ningún instrumento comercial."

El **T Bone** (ver fig. 1) es un modelo desarrollado por Perry Cook que permite sorprendentes resultados en la simulación de sonidos normalmente asociados con los llamados instrumentos de bronce. El instrumento está conformado por una boquilla de forma variable (las figuras 2a, 2b, y 2c muestran algunas variaciones posibles), cuatro secciones cilíndricas que se conectan mediante pistones, una sección cilíndrica de longitud variable (como la vara del trombón), y finalmente, una campana también variable.

Hay controles que permiten manipular individualmente la longitud de los cuatro tubos, la posición de la vara (en el caso del trombón), y la posición de los pistones, determinando así el largo total de la parte cilíndrica del instrumento. La forma de la boquilla puede alterarse totalmente a través de un editor. También puede variarse la campana, la posición de la vara y los pistones, la presión del aire de entrada, los diversos factores que hacen a la oscilación del labio, y otras características tanto del instrumento como del intérprete.

SPASM es otro modelo implementado por Perry Cook, y en este caso permite sintetizar el sonido de la voz cantada a través de la simulación del tracto vocal del ser humano (ver fig. 3). La fuente glotal que utiliza este sistema copia el desarrollo temporal y las características espectrales de la onda que produce la glotis humana. Las turbulencias de las fricativas y otras consonantes se simulan mediante una fuente de ruido filtrado que puede ubicarse en cualquier punto del tracto vocal. El mismo puede ser sensiblemente alterado en su forma (las figuras 4a, 4b, y 4c muestran algunas variaciones posibles). También puede modificarse la posición del velo del paladar y la forma de la cavidad nasal.

Los diptóngos se obtienen especificando un conjunto de parámetros iniciales y otros finales, así como el tiempo y la curva de interpolación para la transición. Un grupo de controles se relaciona directamente con la interpretación, y maneja parámetros como la frecuencia y amplitud del vibrato, y sus variaciones aperiódicas, entre otros factores.

Uno de los primeros modelos físicos desarrollados en Stanford partió del algoritmo ideado por Kevin Karplus y Alex Strong para sintetizar el sonido producido por instrumentos con cuerdas punteadas. Aquella idea básica fué notablemente mejorada y ampliada por David Jaffe y Julius Smith. Y actualmente es posible utilizar en las computadoras NeXT un instrumento llamado "Pluck", resultado directo de las mencionadas investigaciones. Rick Vander Kam ha implementado poco tiempo atrás, también en las NeXT, un modelo llamado "RokNRole" (ver fig. 5) que simula

el comportamiento de una guitarra eléctrica con distorsión y realimentación ("feedback").

También se encuentran en distintas etapas de desarrollo, modelos físicos que permiten sintetizar, por ejemplo, el sonido de un clarinete ("ClariNeXT"), ver figs. 6a, y 6b), o una flauta.

IDEAMA

El Center for Computer Research in Music and Acoustics junto al Zentrum Für Kunst und Medientechnologie (ZKM) en Karlsruhe, Alemania, llevan adelante el proyecto "IDEAMA", a través del cual se está creando un archivo digital que incluye las primeras composiciones electroacústicas creadas en el mundo. Marcia Bauman y Max Mathews son algunos de los investigadores que se ocupan, en Stanford, de llevar adelante esta importante propuesta.

Las obras comupestas en diferentes lugares del mundo durante los primeros años de trabajo con medios electroacústicos necesitan ser salvaguardadas del deterioro que está sufriendo el material empleado para su creación, reproducción y almacenamiento. Para ello, las grabaciones analógicas originales son transferidas a un formato digital. Durante este proceso es posible "limpiar" parcialmente la grabación primitiva de aquellos ruidos indeseados por el compositor.

Dado que se trata de un archivo totalmente digital, aún los materiales complementarios como partituras y notas de programa, serán solamente conservados en el citado formato (digitalizados a través de un "scanner").

Para 1994 se planea realizar un simposio en la sede del ZKM en ocasión de la apertura oficial de IDEAMA, donde compositores pioneros en el campo de la música electroacústica, técnicos y musicólogos, discutirán aspectos estéticos y tecnológicos relacionados con el tema: ¿cuál es la versión "original" de una temprana obra electroacústica en cinta?

COMMOM MUSIC

El **Common Music** es un lenguaje de programación desarrollado por Heinrich Taube para su aplicación específica en el campo de la composición musical. La idea de que un compositor escriba y produzca su obra utilizando una computadora es cada vez más frecuente, y este lenguaje de alto nivel construido en Lisp es una de las posibilidades para llevarlo a cabo.

Concretamente, el Common Music permite generar una "partitura" con la descripción de todos o parte de los eventos (notas) de una composición. Luego un programa distinto se encarga de la síntesis sonora, convirtiendo el contenido de la "partitura" en material audible (ver figura 7). Para la generación de sonido se emplean programas como el Music Kit, el Common Lisp Music (CLM), o el Csound, aunque también es posible codificar los eventos como mensajes MIDI. El Music Kit, creado por David Jaffe, es un recurso implementado en todas las computadoras NeXT que permite, entre otras cosas, el procesamiento y la síntesis digital de sonidos en tiempo real. En cambio el Common Lisp Music, desarrollado por William Schottstaedt, no permite la interacción en tiempo real, pero es un sofisticado lenguaje para el diseño de instrumentos (entendiendo por "instrumento" una configuración de elementos básicos generadores y procesadores que definen una estrategia de síntesis en particular, como por ejemplo: FM, aditiva, granular, u otras).

En el Commom Music existen conceptos de directa asociación con la organización musical tradicional, como: partitura, parte, escala, nota, ritmo, y dinámica; pero también implementa herramientas tales como los generadores, las funciones, y las agrupaciones de items, que permiten complejas operaciones del material compositivo.

Son múltiples las posibilidades de organización de eventos que pueden generarse mediante las llamadas "agrupaciones de items". Con ellas es posible estructurar en forma cíclica, secuencial, aleatoria, o de otros muchos modos, datos como: notas,

alturas, ritmos, intervalos, grados, pasos, amplitudes, frecuencias, y ubicación espacial.

Veamos algunos ejemplos sumples que nos permitan vislumbrar el potencial de las "agrupaciones de items". Habiendo previamente delimitado la generación total de eventos a un total de 9, una "agrupación de items" definida como:

(items 1 2 3 4 in sequence)

produce como respuesta:

1 2 3 4 4 4 4 4 4

dado que la lectura secuencial hace que se seleccionen los datos de la lista uno a uno hasta llegar al último, y luego continúa repitiendo este elemento hasta agotar la cantidad de eventos que se haya determinado en la "partitura", en este caso = 9.

Una "agrupación de items" definida como:

(items c d e f g in random)

puede producir muy diferentes resultados cada vez, dado que la lectura de datos es en este caso aleatoria. De este modo, si la generación de eventos a producir estuviese nuevamente limitada a un total de 9, podríamos obtener algunas de las siguientes respuestas, o tal vez otras distintas:

**c e g c e d f g e
f g f c f e e c g
e g d c f d e d g
d c g f c e d g f**

Otra "agrupación de items", definida en este caso como:

(items a b c in cycle for 8)

nos da como respuesta:

a b c a b c a b

ya que la lectura de items es definida aquí como cíclica, lo que significa que los datos son leídos de modo secuencial hasta llegar al último elemento de la lista, para entonces retomar ésta desde su comienzo hasta agotar el número de eventos a generar, limitados en este ejemplo a un total de 8.

Otra de las muchas posibilidades de generación y organización de eventos empleando el recurso de "agrupación de items" es:

(items 1 2 3 in palindrome)

que produce, limitando ahora el número de generación de eventos a un total de 12:

1 2 3 3 2 1 1 2 3 3 2 1

una variación de la lectura cíclica, llamada "en palíndromo", donde la selección de items es secuencial pero cambia de dirección cada vez que alcanza uno de los extremos del listado. Incluyendo en la construcción de este último ejemplo la opción **elided** (elidir = suprimir la vocal final de una palabra, cuando la que sigue empieza también con vocal), se consiguen alternativas como:

elided yes = 1 2 3 2 1 2 3 2 1 2 3 2

elided start = 1 2 3 3 2 1 2 3 3 2 1 2

elided end = 1 2 3 2 1 1 2 3 2 1 1 2

Uniendo los conceptos de lectura secuencial y cíclica, vemos en el próximo ejemplo que la complejidad del sistema puede incrementarse considerablemente cuando el último item de una "agrupación de items" es a su vez una nueva "agrupación de items":

(items 1 2 3 4 (items 5 6 7 in cycle) in sequence)

siendo el resultado esta vez:

1 2 3 4 5 6 7 5 6 7 5

si limitamos la generación de eventos, en este caso, a un número total igual a 11.

La figura 8 nos muestra un fragmento de la partitura de *Memorias* (Ricardo Dal Farra, 1992), composición producida utilizando el lenguaje Common Music y el programa de síntesis de sonido del Music Kit.

BIOMUSE

El **Biomuse**, desarrollado por Hugh Lusted y Benjamin Knapp, es un dispositivo que permite crear sonido en tiempo real a partir de la actividad bioeléctrica de aquel que lo esté usando, como así también utilizar señales volitivas para controlar las funciones de un sintetizador u otro instrumento que responda a mensajes MIDI.

Está compuesto por la interfase bioeléctrica y la unidad de procesamiento de señales. Pequeños sensores colocados sobre la piel son utilizados para tomar señales eléctricas de bajo nivel, las que son entonces amplificadas, filtradas y luego digitalizadas para alimentar el procesador. Señales provenientes de diversos músculos (electromiogramas), ondas cerebrales (electroencefalogramas), e incluso del movimiento de los ojos (electrooculogramas), pueden utilizarse como entradas de información al procesador, quien analiza estas bioseñales para finalmente producir como resultado los comandos MIDI. Una serie de bandas ajustables mantienen en posición a los electrodos que pueden colocarse sobre los bíceps, tríceps, pronadores, supinadores, cuádriceps, gemelos, u otros músculos, permitiendo así que ellos se conviertan en los biocontroladores del sistema.

Entre las aplicaciones del Biomuse aparecen las múltiples posibilidades musicales de emplearlo como controlador de instru-

mentos digitales durante un concierto, la relación con el movimiento a partir de una danza o una actuación teatral, su implementación en ambientes de realidad virtual, y también su utilización en la rehabilitación de personas afectadas por parálisis u otras enfermedades degenerativas del sistema nervioso.

El 16 de julio de 1992, en el anfiteatro de la Universidad de Stanford, el compositor Atau Tanaka estrenó su obra *Kagami* para Biomuse e instrumentos MIDI. El mismo compositor interpretó el Biomuse durante el concierto, y empleó sus músculos para generar la información básica con la que alimentó a una computadora, la cual a su vez procesaba dichos datos (con un programa escrito usando el lenguaje MAX) y transmitía finalmente los comandos MIDI que llegaban hasta los instrumentos musicales digitales.

(Agosto de 1992)

AGRADECIMIENTOS

Deseo agradecer al "Center for Computer Research in Music and Acoustics" por invitarme a compartir un ambiente de estudio y trabajo donde colaboración, confianza, experiencia, y creatividad, se conjugan de un modo muy poco frecuente. A la Fundación Rockefeller por su indispensable y decisivo apoyo económico. Y muy especialmente a John Chowning, Chris Chafe, Patte Wood, Rick Taube, Perry Cook, Jay Kadis, Bill Schottstaedt, Marcia Bauman, Glen Diener, y Heidi Kugler, por toda la ayuda que me brindaron.

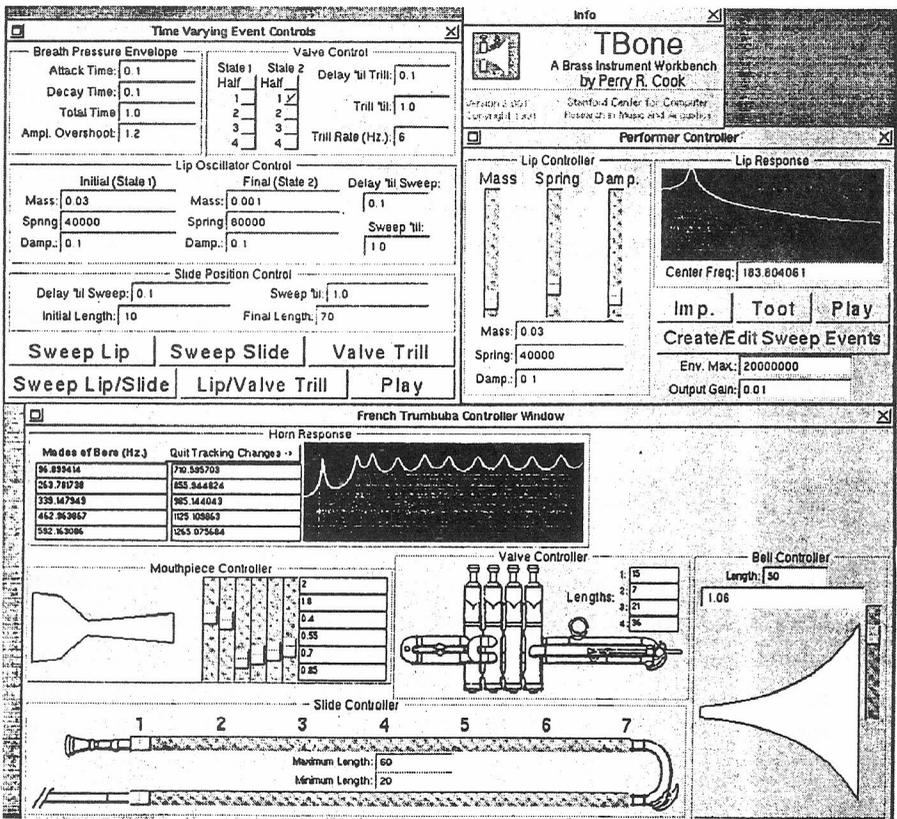


Figura 1. El TBone es un modelo desarrollado por Perry Cook que permite sintetizar sonidos normalmente generados por los instrumentos de bronce.

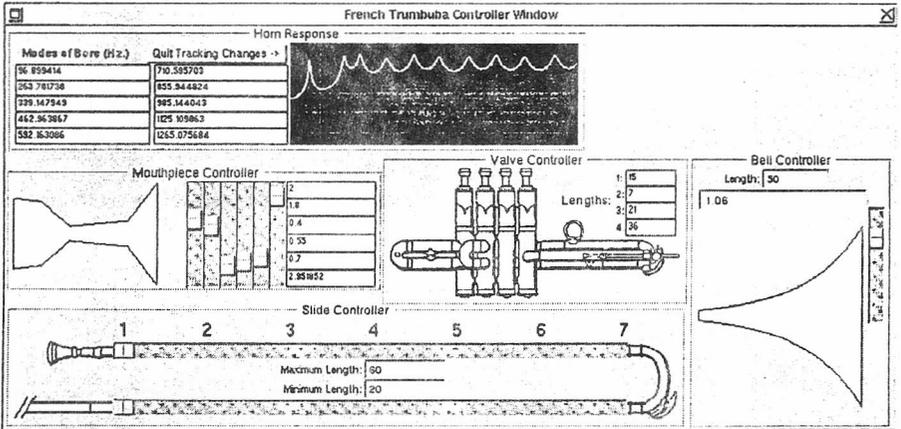


Figura 2a. En el TBone la forma de la boquilla puede alterarse totalmente a través de la interacción con un editor gráfico.

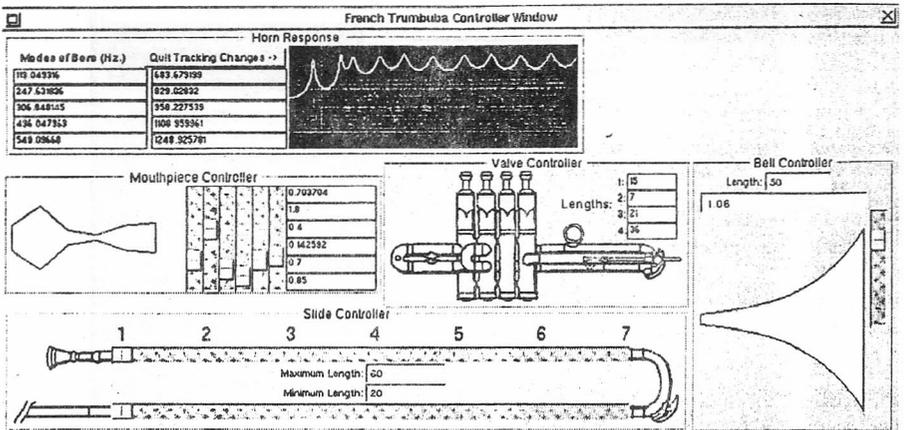


Figura 2b

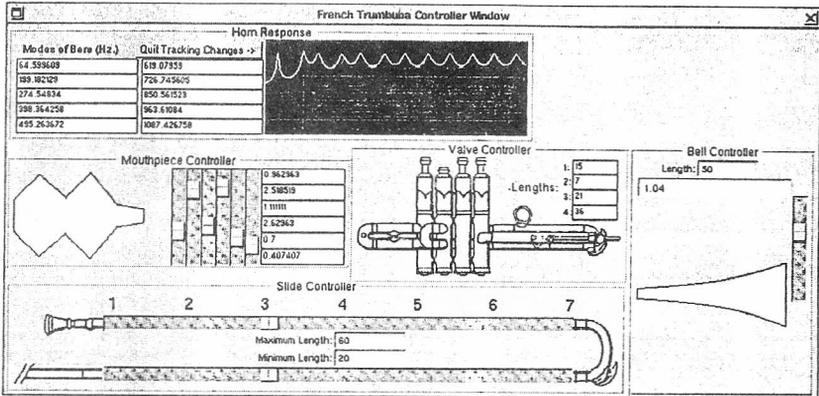


Figura 2c

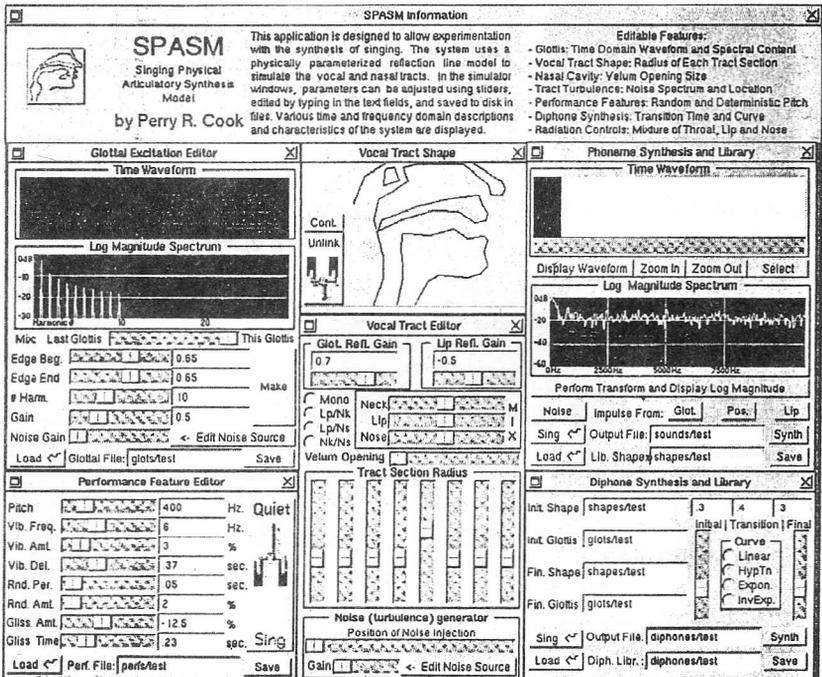


Figura 3. El modelo SPASM, desarrollado por Perry Cook, permite sintetizar el sonido de la voz cantada simulando las articulaciones propias del tracto vocal de un ser humano.

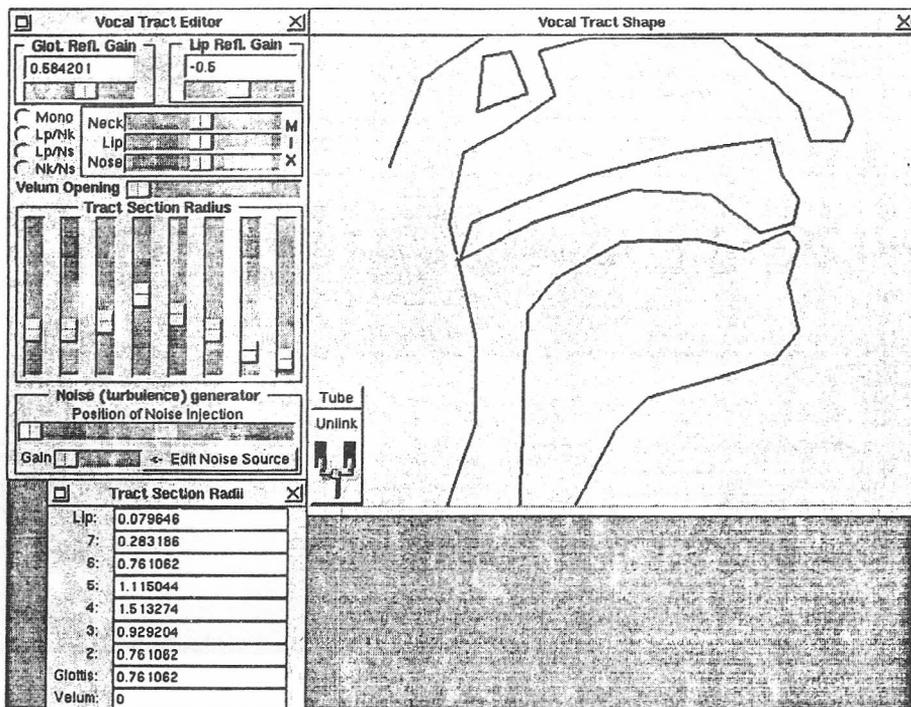


Figura 4a. En SPASM la forma del tracto vocal puede ser sensiblemente alterada. También puede modificarse la posición del velo del paladar y la forma de la cavidad nasal.

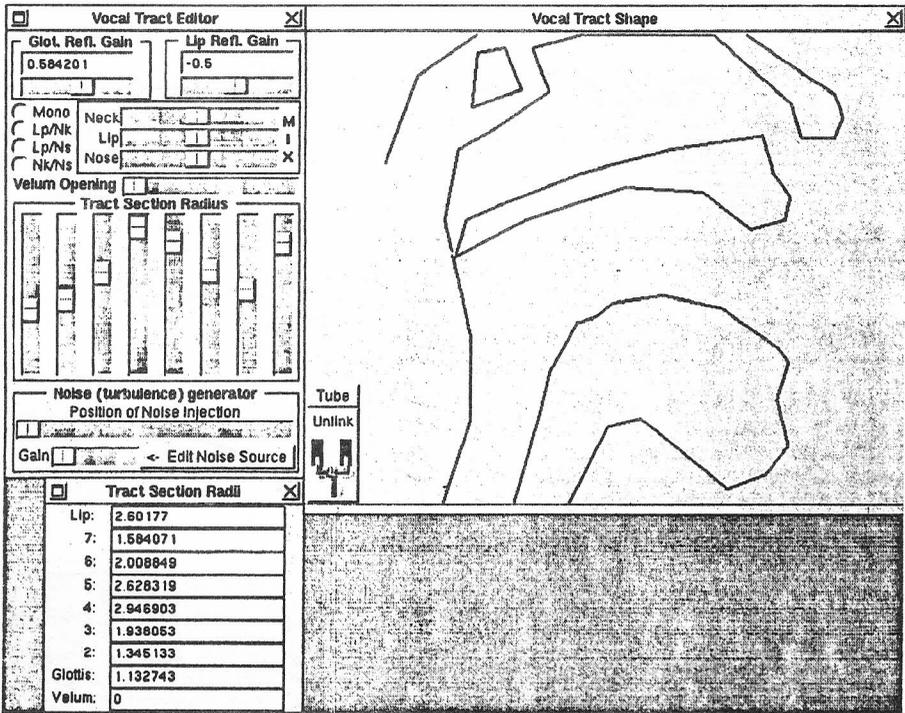


Figura 4b

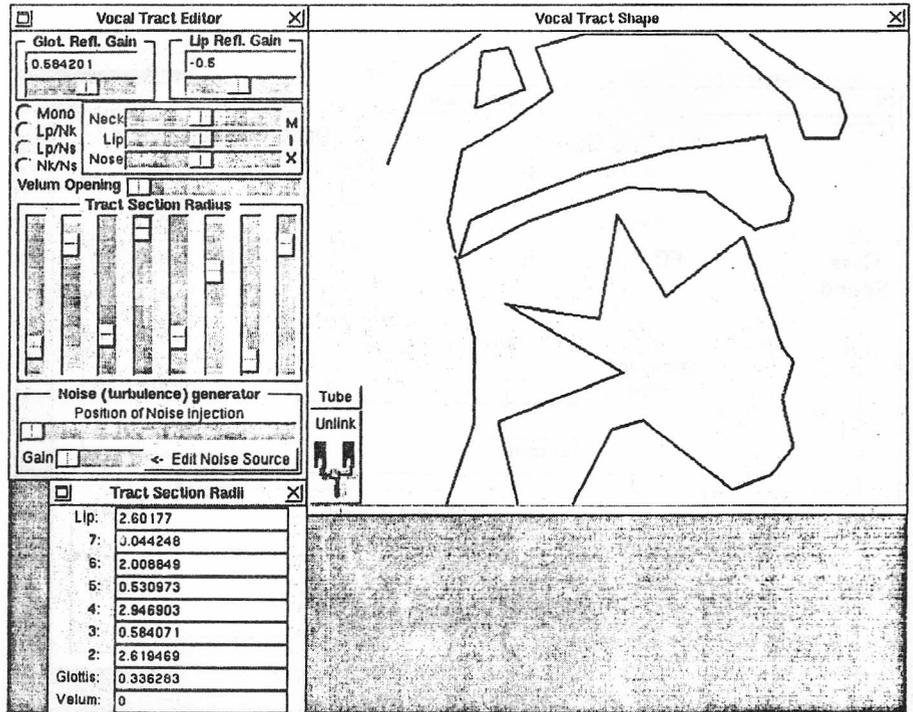


Figura 4c

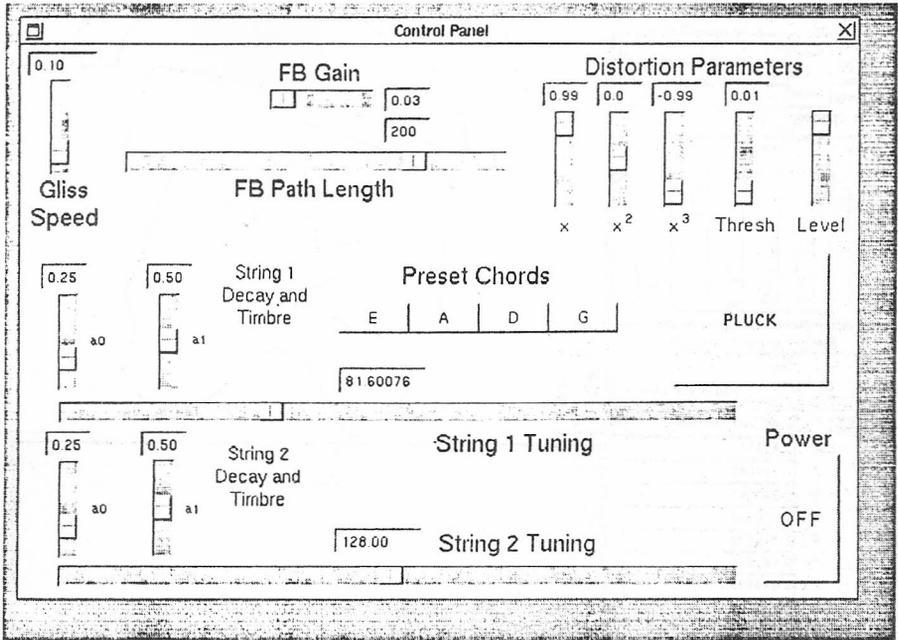


Figura 5. RokNRole, de Rick Vander Kam, es un modelo que simula el comportamiento de una guitarra eléctrica con distorsión y realimentación.

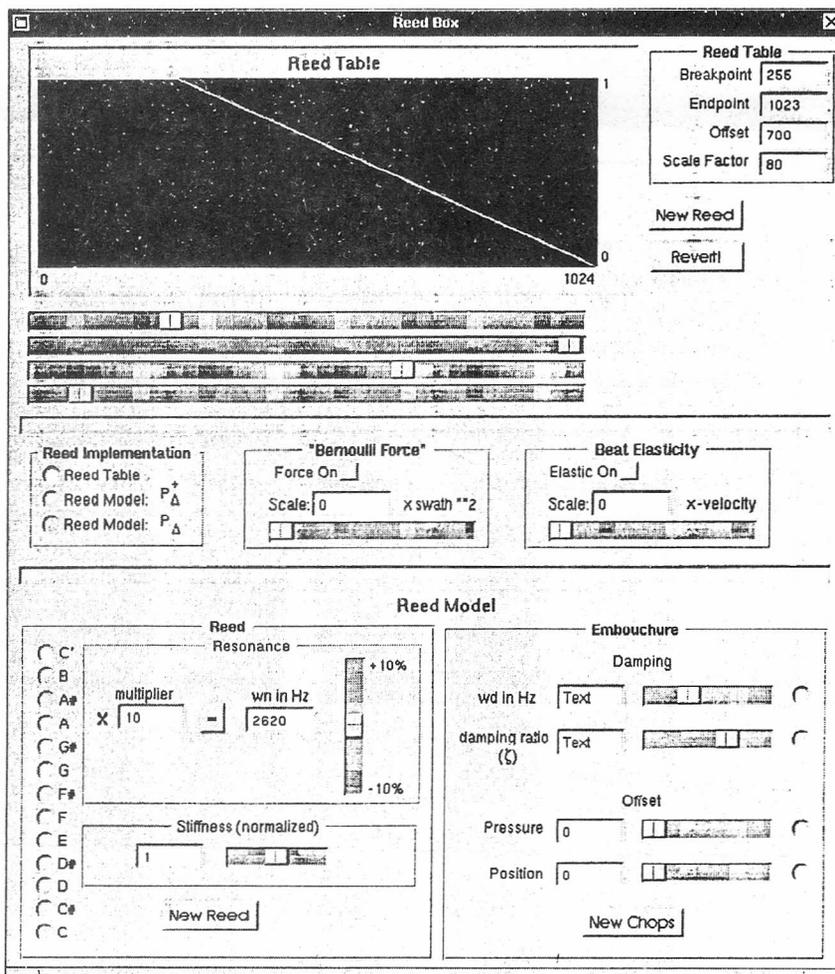


Figura 6a. El ClariNeXT, implementado por Suzanne Hirschman como un ambiente gráfico interactivo en la computadora NeXT, se basa en un modelo de clarinete creado por Julius Smith.

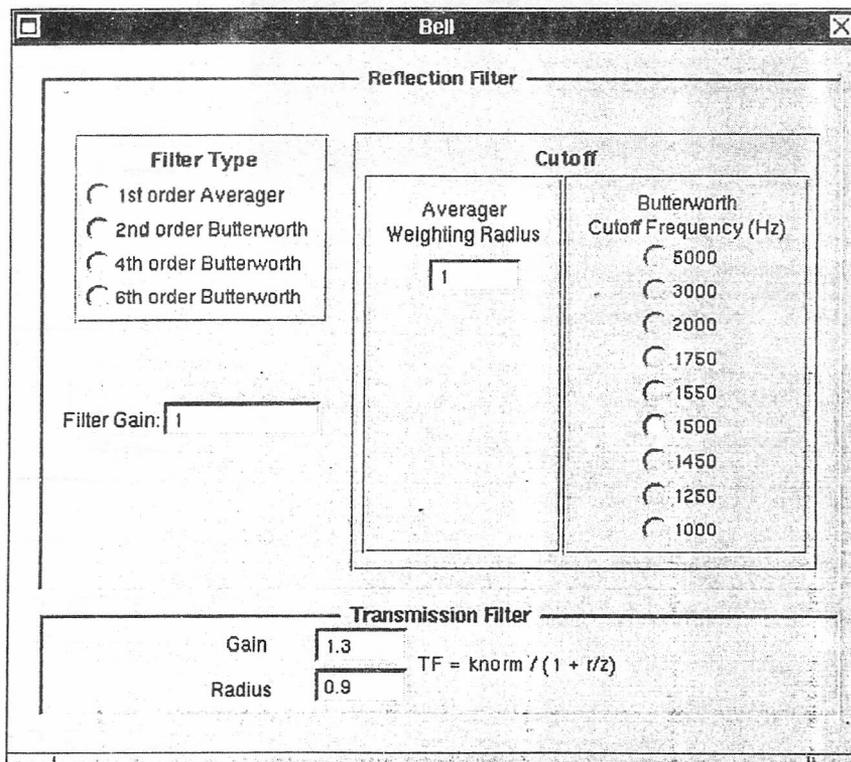


Figura 6b

COMPOSITOR -> Common Music -> "partitura"
"partitura" -> Music Kit ó CLM ó Csound ó MIDI -> SONIDO

Figura 7. El lenguaje Common Music permite generar una "partitura" con la descripción de todos los eventos (notas) de una composición. Luego un programa distinto se encarga de la síntesis sonora, convirtiendo el contenido de la "partitura" en material audible.

```
(with-part Fm1viPoly
(time 127.5 ; tiempo desde el inicio
 ampEnv "pepe" ; envolvente (amplitud)
 m1IndAtt .91 ; tiempo de ataque (env. modulador)
 m1IndRel 1.2 ; tiempo extinción (env. modulador)
 cRatio 3.76 ; portadora
 m1Ind1 8 ; índice de modulación cuando env.=1
 m1Ind0 2 ; índice de modulación cuando env.=0
 m1IndEnv "pipo" ; envolvente (índice de modulación)
 freqEnv "pepa" ; envolvente (frecuencia)
 sVibAmp .01 ; porcentaje de vibrato sinusoidal
 sVibFreq .2 ; frecuencia del vibrato sinusoidal
 rVibAmp .17 ; porcentaje de vibrato aleatorio
 events 37 ; número de eventos )

(setf ampAtt (item (items 4.2 .2 .57 1.75 1.2 .7 in palindrome)))
(setf ampRel (item (items 1.3 .6 3.2 (1.6 .3 .8 in random) in cycle)))
(setf amp (item (crescendo from ppp to mp in 37)))

(setf m1Ratio (item (items 1.4 6.9 2.3 1.17 3.15 8.2 3.29 4.2 7.125
2.14 8.6 7.98 3.11 4.24 2.0 8.61 5.33 4.28 1.07 3.96 in sequence)))
(setf bearing (item (items 20.0 0.0 -20.0 42.0 -37.0 -5.0 in cycle)))

(setf freq (item (items [77.1 228.3 67.1 925.9] 43.2 r 987.6 394.6
122.7 1057.1 [1007.8 47.3 667.3 721.2] 1092.1 52.0 212.4 in heap)))
(setf rhythm (item (items 7.3 9.5 4.3 7.4 8.7 6.3 9.4 in random)))
```

Figura 8. Fragmento de la partitura de *Memorias* (Ricardo Dal Farra, 1992), composición producida utilizando el lenguaje Common Music y el programa de síntesis de sonido del Music Kit.

BIBLIOGRAFIA

- CHAFE, Chris. "Pulsed Noise and Microtransients in Physical Models of Musical Instruments". Stanford University, Department of Music Technical Report STAN-M-65, 1990.
- CHOWNING, John M. "Music from Machines: Perceptual Fusion & Auditory Perspective - for Ligeti". Stanford University, Department of Music Technical Report STAN-M-64, 1990.
- COOK, Perry R. "Synthesis of the Singing Voice Using a Physically Parameterized Model of the Human Vocal Tract". *Proceedings of the International Computer Music Conference*. Ohio, 1989.
- _____. 1989. "TBone: An Interactive Waveguide Brass Instrument Synthesis Workbench for the NeXT Machine". *Proceedings of the International Computer Music Conference*. Montreal, 1991.
- KNAPP, Benjamin R. y LUSTED, Hugh, S. "A Bioelectric Controller for Computer Music Applications". *Computer Music Journal* 14(1):42-47, 1990.
- SCHOTTSTAEDT, William. "Pla: A Composer's Idea of a Language". *Computer Music Journal*, 7(1):11-20, 1983.
- SMITH, Julius O. "Viewpoints on the History of Digital Synthesis". *Proceedings of the International Computer Music Conference*. Montreal, 1991.
- _____. "Waveguide Synthesis Tutorial". Center for Computer Research in Music and Acoustics, 1992.
- TAUBE, Heinrich. "Common Music: A Music Composition Language in Common Lisp and CLOS". *Computer Music Journal*, 15(2):21-32.
- WOOD, Patte. "Research Overview". Center for Computer Research in Music and Acoustics, 1992.

Ricardo Dal Farra é compositor de música eletro-acústica e pesquisador, radicado em Buenos Aires.