

Martin Riches

The Talking Machine

Das erste automatische Musikinstrument, das ich gebaut habe, war eine Flöte spielende Maschine. Danach habe ich eine Reihe von mechanischen Instrumenten gebaut, Orgelpfeifen, Percussions und Saiteninstrumente. Diese Maschinen werden von Musikwalzen oder vom Computer angetrieben. Sie werden auf Kunstaustellungen gezeigt, nehmen an Konzerten teil und haben viele Komponisten inspiriert, Musik für sie zu komponieren.

Als ich die Pfeifen für eine mechanische Orgel baute, bemerkte ich, daß diese, wenn sie falsch spielten, manchmal Laute erzeugten, die der menschlichen Sprache sehr ähnlich waren. Ich fragte mich, was wohl geschehen würde, wenn ich alle Orgelpfeifen durch Pfeifen ersetzen würde, die speziell für die Erzeugung von Sprachlauten konzipiert wurden, und ob diese neue Orgel sprechen lernen könnte. Das Ergebnis war die »Talking Machine«.

»Die TALKING MACHINE von Martin Riches ist Musikinstrument, Skulptur und Sprachkunstwerk in einem. Sie vereinigt das Prinzip der Orgel mit dem der sprachlichen Lauterzeugung, indem die Form der »Orgelpfeifen« Mund- und Zungenstellung imitieren, um den Tönen ihre jeweilige Vokalfärbung zu verleihen. Gleichzeitig ist für das Publikum deutlich sichtbar, wie und wo die Sprachlaute entstehen. Aber gerade die Offensichtlichkeit der Funktionsweise macht das Vertraute fremd. Statt eines Menschen ein künstlicher Mechanismus, statt eines Mechanismus eine Skulptur, statt einer Skulptur ein Musikinstrument, statt eines Musikinstruments ein Sprachwerkzeug – nichts ist mehr, was es scheint.«

Die »Talking Machine« ist ein mechanischer Sprachsynthesizer. Die Sprachlaute werden in Stimmkanälen und nicht durch einen Lautsprecher produziert. Man könnte fast sagen: ein Instrument für jeden Laut, aber einige Instrumente können mehr als einen Laut von sich geben.

Ein Beispiel: Wie in der natürlichen Sprache wird das T genauso wie das S gesprochen, nur sehr viel kürzer. Einige andere Instrumente haben ein zusätzliches Ventil, um Zunge oder Lippen zu simulieren, so daß sie mehr als einen Laut erzeugen können. M hat z. B. ein Ventil, das sich öffnet, um den Klang schlagartig zu modifizieren: mmm-ah.

Jedes Instrument besteht aus zwei verschiedenen Teilen. Erstens aus einem Geräuschmacher, einer Zungenpfeife oder einer Zischpfeife, und zweitens aus einem Resonator, der den gewünschten Laut aus dem Geräusch filtert. Die Resonatoren reproduzieren die Hohlräume, die in Mund, Nase und Hals während des Sprechens entstehen.

Für die Entwicklung der Resonatoren mußten eine Menge anatomischer, akustischer und phonetischer Informationen gesammelt und verarbeitet werden. Die folgenden Methoden wurden verwendet:

1. Das Thema nachlesen und Experten konsultieren;
2. Beobachtungen mit Hilfe eines Zahnarztspiegels und einer Schieblehre;
3. Röntgenbildern;

4. Berechnungen mit Hilfe von akustischen Formeln und Tabellen;
5. Experimente mit einem einstellbaren Resonator und einem Modell und;
6. Interpolation zwischen existierenden Resonatoren;
7. Glück. Einige Resonatoren, die nicht den erhofften Laut erzeugten, erwiesen sich als genau richtig für einen anderen!

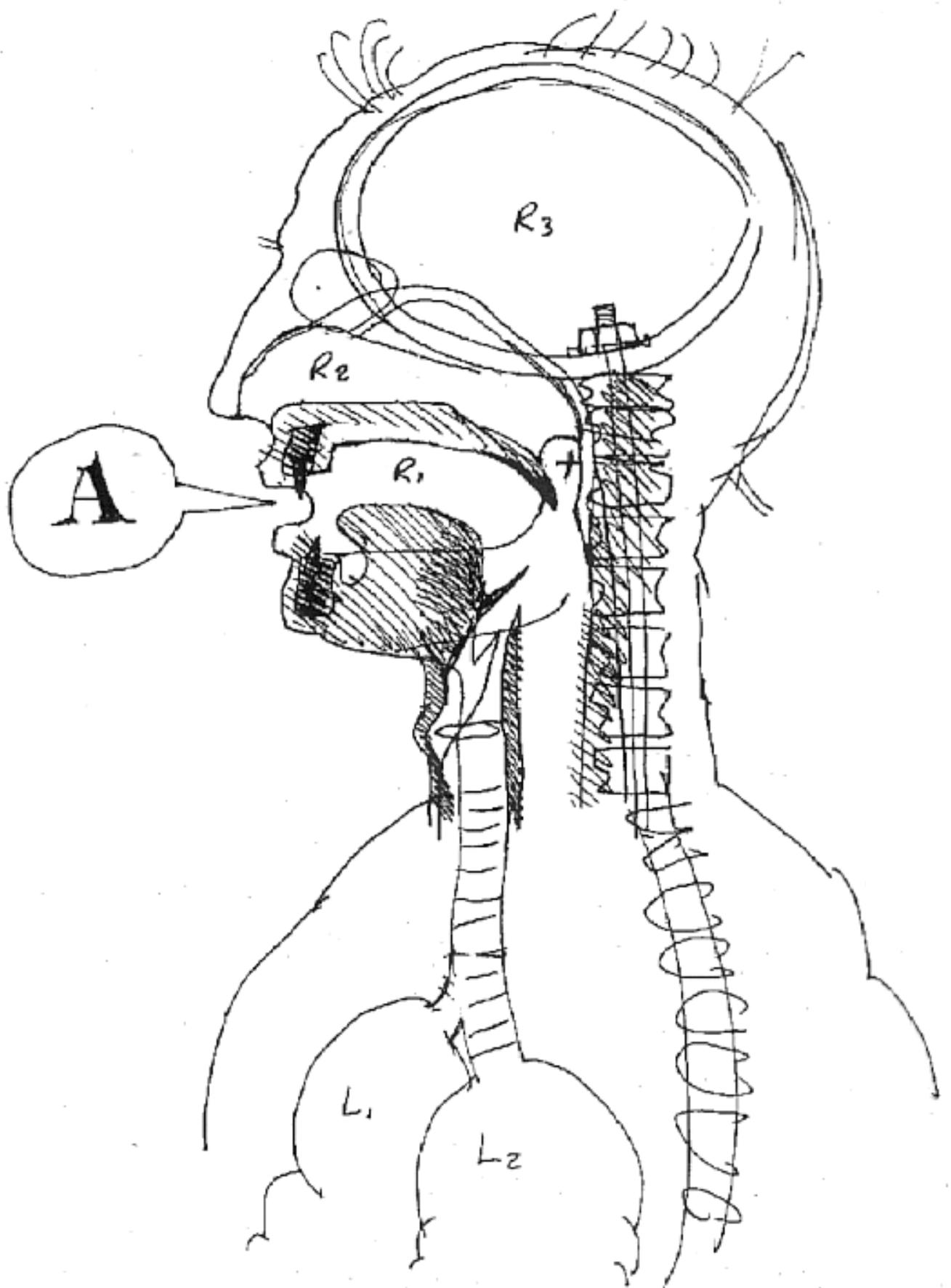
Die Entwicklung und Prüfung eines vollständigen Satzes von 32 Resonatoren war ziemlich ermüdend für die Ohren. Ich mußte die Hörsitzungen kurz halten und versuchen, eine positive Einstellung und Geduld zu wahren. Es war vielleicht deshalb so frustrierend, weil die geprüften Resonatoren zweifellos irgendwie menschlich klingende Sprachlaute erzeugten und man sich leicht vorstellen konnte, daß man mit einer höchst eigensinnigen menschlichen Persönlichkeit konfrontiert war.

Wenn diese Maschine verständlich ist, dann deshalb, weil unsere Ohren Sprünge von einem Laut zum anderen akzeptieren können, genau wie unsere Augen beim Betrachten eines Films Sprünge von einem Bild zum anderen akzeptieren. Eine Voraussetzung dafür ist die Einhaltung eines genauen Zeittaktes. Die aufeinanderfolgenden Laute dürfen sich weder überlagern, noch dürfen ungewollte Pausen zwischen den Lauten entstehen. Ein Problem ist, daß die Instrumente nicht sofort ansprechen, wenn sie angeblasen werden, und, was noch gravierender ist, jedes Instrument hat seine eigene Verzögerungszeit. Aus diesem Grund werden die Luftventile durch einen Computer gesteuert, der programmiert ist, diese Probleme zu lösen.

Englisch ist die Muttersprache dieser Maschine, sie spricht auch deutsch und japanisch. Weitere Sprachen sind theoretisch möglich, werden aber mit englischem Akzent gesprochen, weil bestimmte Laute fehlen. Sie könnte nicht un bon vin aussprechen, weil sie keine nasalen Vokale hat. Sie könnte auch keine chinesischen Dialekte sprechen, bei denen die Bedeutung von der Tonhöhe abhängig ist. Im Augenblick spricht die Maschine immer in einer Tonhöhe: 220 Hertz – also mit einer Tenorstimme. Bisher habe ich nicht versucht, Betonung oder Intonation zu simulieren. Intonation könnte man durch Modulation des Luftdrucks erreichen, und es ist auch relativ leicht, die Sprechgeschwindigkeit zu ändern.

(Ich möchte den folgenden Personen meinen herzlichen Dank aussprechen für Inspiration, Information und Hilfe verschiedener Art: Prof. Manfred Krause, Folkmar Hein und Klaus Hobohm, TU Berlin; Jasia Reichardt, London; Peter Richards und dem Bibliothekar des Exploratoriums, San Francisco; dem Bibliothekar des Instituut voor fonetische Wetenschaap, Amsterdam; Francisco Flores, Caracas.)

(Der Text ist zuerst in dem Katalog *INTERFACE II, Computergestützte Visionen*, hrsg. von Klaus Peter Dencker im Auftrag der Kulturbehörde Hamburg, 1995 erschienen)



L = Lunge, R = Resonanzkörper (R₃ = bei einem Hohlkopf)

© positionen, 25/1995, S. 17-18