Evaluation of the NeXT Computer System for Psychoacoustic and Music Perception Research
Author(s): Michael P. Lamoureux
Source: *Computer Music Journal*, Vol. 17, No. 3 (Autumn, 1993), pp. 78-84
Published by: The MIT Press
Stable URL: http://www.jstor.org/stable/3680950
Accessed: 05-06-2018 18:00 UTC

Products of Interest

User's Reports

Evaluation of the NeXT Computer System for Psychoacoustic and Music Perception Research

Reviewed by
Michael P. Lamoureux
Calgary, Alberta T2N 1N4 Canada
and Annabel J. Cohen
Wolfville, Nova Scotia B0P 1X0 Canada

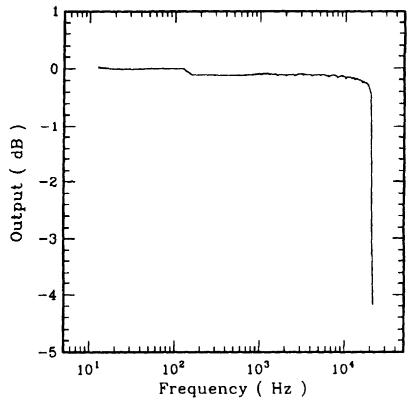
[Editor's Note: When *Computer Music Journal* originally accepted this article for publication, the NeXT machine it evaluates was growing in popularity for use as a computer music workstation, used either alone or in combination with coprocessors in the form of the IRCAM Signal Processing Workstation or the Ariel_ Corp. QuintProcessor. Since that time, NeXT has announced that it is discontinuing production of its hardware and will instead deliver the NeXTSTEP software environment on Intel-based platforms (for which no widely standardized sound I/O and DSP solutions exist). The editors feel, however, that this article serves the community in presenting a set of criteria and an excellent discussion of the requirements of psychoacoustics and music research.]

With the discontinuation of the NeXT hardware production, musicians and acousticians have lost a useful, integrated platform for sound and music work. While the software tools that came with the NeXT will live on in NeXTSTEP, one must still consider what audio hardware is available for computer-assisted sound generation. Three years ago, the NeXT was the only workstation-class machine that offered "16-bit, CD-quality" audio as a standard. Today, there are several other "16-bit" options, including the various SPARCstations from Sun Microsystems, Inc., and the IRIS Indigo manufactured by Silicon Graphics, Inc. There are also several sources for 16-bit sound cards for expanding the audio capabilities of both Intel-based and Macintosh PCs, and rumours of forthcoming multimedia PCs with integrated 16-bit audio. Of course, as revealed in the analysis of the NeXT audio hardware, having 16 bits in the computer specifications does not necessarily imply true 16-bit performance, and a careful characterization of the hardware is necessary to verify that it meets the audio requirements of the discerning user. An examination of the sound software tools available on a given platform is also important. The procedure described in the report below suggest one way of establishing a benchmark.

Defining the limitations and capabilities of human listeners is one of the goals of research in auditory perception. Such research can interest those composers and performers who wish to apply knowledge of the auditory capabilities of listeners in their music. For example, how many different tones in a new musical scale can people keep track of, or what is the smallest difference in frequency that can be discriminated? The field of psychoacoustics is that part of auditory research that aims to define precisely the relation between the physical characteristics of sound and its subjective representation in the mind of the listener. In psychoacoustic experiments, carefully controlled sounds are presented to listeners who may be asked to provide any of a number of responses: for example, to rate the similarity of each pair of sounds, to state whether a pair of sounds are the same or different, or to state whether a tone is higher or lower than the tone preceding it. Mathematical descriptions of the data from such studies define the relationship between the physical and subjective representations and reveal what physical differences in sounds make psychological differences. Because psychoacoustics research relies on accurate measurement and quantification, it benefits from advanced computer technology for stimulus presentation, response collection, and data analysis. If the same computer that is used in psychoacoustic research also supported composition and performance, the stage would be set for collaborative work in psychoacoustics and music. For example, it would be a simple matter to employ the composer's musical materials in psychoacoustic studies. To date, only major computer installations such as those at IRCAM or CCRMA have provided opportunities for such interaction. Suppose, however, there were a desktop computer that served both the musician and the psychoacoustician. The NeXT computer system has been well received by composers (e.g., Lansky 1989, 1990). Can it serve psychoacousticians? In this report, we evaluate the NeXT for use in applications of psychoacoustics and music perception research.

The NeXT computer combines a powerful Unix graphics workstation with high-quality sound capabilities (Jaffe and Boynton 1989; Webster 1989). At the time of its introduction, it was the only available workstation-class computer with 16-bit sound output, digital synthesis capabilities, and large storage capacity as standard features. The result is a single system that allows for stimulus generation, response collection, and data analysis. These features suggest that NeXT would be an attractive option for researchers in psychoacoustics and music perception. For psychoacoustic research, in particular, it is important that acoustic signals can be very precisely defined.

Fig. 1. Frequency response (no filter).



To verify that the system was appropriate for use in psychoacoustics and auditory and music research, we ran a complete set of audio measurements to evaluate the sound quality of the machine, and also made extensive tests of the software tools provided with the system for algorithmically generating auditory signals. This is a report of our measured specifications for the NeXT computer system and its applicability to various experiments, along with some caveats and limitations of which the experimenter must be aware.

Audio Measurements

The published audio specifications of the NeXT can be summarized as indicating telephone-quality sound input (8-bit, 8-kHz CODEC-format ADC) and CD-quality output, with stereo 16-bit resolution at 44.1-kHz sampling (NeXT, Inc. 1989). As the input is clearly not high-quality, it is appropriate to restrict our evaluation to the digital audio output. We will discuss below how digital sounds can be created for output, but first simply report our measurements of the sound output quality. To perform a complete yet reasonably standard verification of the performance, we followed the AES Standard method for digital audio engineering measurement of digital audio equipment (AES 1991). This standard provides for measurements of a variety of parameters under controlled conditions, including output characteristics, linearity, frequency response, and noise. All measurements were carried out on the 68030-based NeXT "cube" running version 1.0 of the NeXTStep software environment. Upgrades of the NeXT to the 68040 processor have not changed any of the audio components, as these are housed separately in the video display unit. There are minor improvements in the documentation provided with software upgrades 2.0 and 2.1, which we note, but no functional changes to the sound software. The list of instrumentation used for calibration is found in Appendix 1.

The results of these measurements are summarized in Table 1. The various noise and separation figures of roughly 90 dB are typical of a 16-bit system. Figure 1 shows excellent frequency linearity, essentially flat to 20 kHz, with the 4 dB drop at the Nyquist rate indicating the lack of sin(x)/x correction. Several precautions should be noted, however. First, and most significantly, a low-pass filter is provided with the system, which can be switched in and out of

Table 1. NeXT Audio Measurement

Output Characteristics	
Bandwidth	1 Hz to 20 kHz
Out-of-band noise	–90 dB of full-scale (FS) or less
Full Scale Output	2.0 V RMS
Stability	0.2 dB variation or less
Digital output	DSP port standard
	AES/EBU available as from third parties
Linear Response	
Frequency response,	+0 dB, –0.5 dB in band
with filter	Flat to 1 kHz, –10dB at 20 kHz
Phase response	10°-variation or less
Interchannel phase delay	6 μsec, uniformly
Polarity	Positive
Amplitude Nonlinearity	
Gain	+0 dB to –1 dB, on a 75 dB range
Intermodulation	Not measurable
Signal-to-Noise	
Idle channel noise	–90 dB FS or less
Noise with signal	–88 dB FS
Power line noise	–90 dB FS (video monitor horiz. scan)
Crosstalk	
Channel separation	90 dB
Input to output leakage	Not measurable

Fig. 2. Frequency response (with filter).

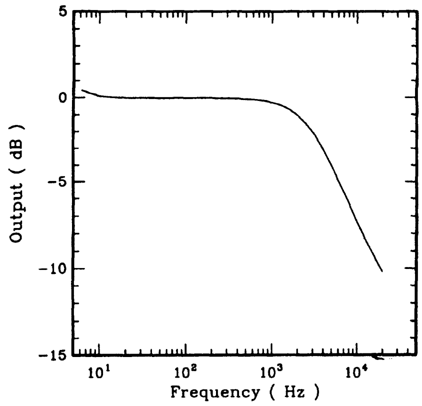


Table 2. NeXT as a Tone Generator

Frequency Accuracy	
Crystal controlled	±0.002 percent or less
Harmonic Distortion	
First harmonic	–75 dB below fundamental
Second harmonic	–80 dB below fundamental
Amplitude Accuracy	
Gain	+0 dB to -1 dB, on a 75 dB range
Frequency response	Flat to 1 kHz, -10 dB at 20 kHz
Signal-to-Noise	
Noise with signal	–88 dB FS
SNR, full signal	72 dB
Summary	
16-bit dynamic range	13-bit linearity

circuit by an unusual keyboard sequence (command–loudspeaker key). This introduces a 10 dB drop in the high-frequency range, as indicated in Fig. 2, leading to very poor frequency response. This deemphasis filter is provided as a feature of the system for playback of raw analog-to-digital data, but as it was undocumented in System 1.0, its hidden presence caused considerable frustration in our measurements of the computer as we switched unknowingly from one mode to the other. Presumably, it would also cause problems in an experiment if the human subject accidentally switched modes via a keypress during a test. System 2.0 does document how the programmer can disable the filter in software; however, as the above-mentioned keystrokes could reenable it at any time, the programmer should be sure to disable it after every keyboard event. Other precautions to note are that bandwidth does not extend to direct current, although for audio work this is not significant, and that there is a fixed left-to-right channel delay of over 6 μsec due to a time multiplexing of the digital-to-analog converter (DAC) output. At frequencies near the Nyquist rate, this would cause a phase difference between channels of close to 60 degrees. Finally, a spectral analysis of the noise indicates that much of it originates from the video monitor, which houses the sound hardware. For instance, the "power line noise" measured is actually a 68-Hz signal corresponding to the video refresh rate, while an out-of-band signal measured at 57 kHz is the horizontal scan rate. Note that the new color systems house the audio components separately from the video monitor and should be free of most of this noise.

Since the NeXT machine would often be used in experiments as a general-purpose signal generator, it is appropriate to measure the system in the generation of sine waves. Table 2 summarizes the characteristics that are relevant here. Notably, the frequency is crystal controlled and is easily set with software to an accuracy of 0.002 percent, while the amplitude may be accurately set over a range of 75 dB. Harmonics appear at 75 and 80 dB below the fundamental, as shown in Fig. 3, giving a signal-to-noise ratio of 72 dB for a full-amplitude signal. This response is typical of a 16-bit system with (at most) 13 bits of linearity.

This points out an important limitation of the NeXT's sound output components. Although the DAC has 16 bits of resolution, it does not have 16-bit linearity. Thus, harmonics will always appear around the -75 dB level, and while a single tone may be presented at amplitudes ranging over 90 dB, two tones cannot be separated by more than 70 dB. Note that single-bit dithering does not improve the situation, unless added in at the 14th bit, which would effectively reduce the dynamic range. Nevertheless, this 13-bit linearity represents relatively good performance.

A more significant limitation appears when using the alternative 22-kHz sampling rate that is available on the NeXT. The system does not compensate for the undersampling, and significant out-of-band imaging occurs, giving a false spectral image

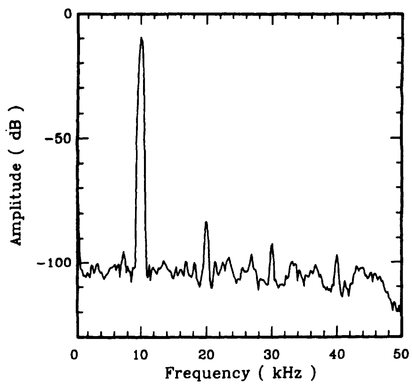


Fig. 3. A 10 kHz tone (showing harmonics).

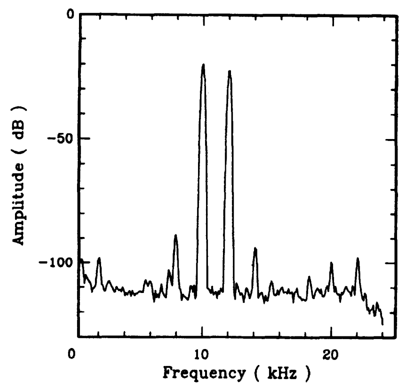


Fig. 4. A 10 kHz tone at 22 kHz sampling (showing aliasing).

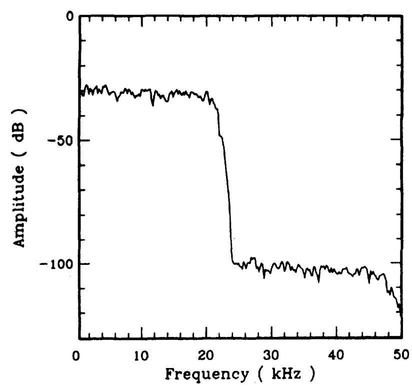


Fig. 5. U-Noise generator (showing flat white spectrum).

within the normal human auditory range. For instance, in Fig. 4, the spectrum of a 10-kHz signal is illustrated, which shows a false image peak at 12 kHz, that is only 3 dB down from the true fundamental. The optional deemphasis filter mentioned above helps to reduce the out-of-band signal, but an external low-pass filter would have to be added to eliminate the error. Significantly, this same out-of-band imaging occurs when the input CODEC sounds are sent to the output. Although the lower sampling rate saves disk space and computation time, it should not be used as the source for the highest quality sound.

Software Toolkits

Besides the hardware to produce sounds, NeXT provides two software libraries for the creation of digital sounds. These have been well described by their originators, David Jaffe and Lee Boynton (1989). We focus here on features of particular relevance in psychoacoustics research.

First, the Sound Kit, written by Lee Boynton, is a collection of software modules that allows the programmer to create, display, and play back digital sounds, providing a container to hold sampled sounds that were either digitally recorded or algorithmically produced. These recordings can be organized into a library of sounds that can be stored on disk, edited, and recalled by name. By managing this collection of sounds in the form of, for example, notes or words, the programmer can create melodies or sentences to be played back in an experiment. The concept is to use the computer as a digital tape recorder, under software control, with tight control over each segment of sound.

The Music Kit, written by David Jaffe, is a more complex set of modules for the representation and performance of music on the NeXT, with many similarities to Music V–style languages. Using the built-in DSP, a number of instruments can be synthesized, which act under the control of a so-called orchestra of software "performers" which play through a "score" containing many "parts," which are collections of notes to be played with specific frequencies, amplitudes, envelopes, and even position in stereo space. As this description suggests, this is a rich and complex collection of objects that can be used to simulate a performance in a multitude of ways.

As a simple experiment, the Music Kit may be used to create a single performer and instrument that will generate tones or a melody, either presented immediately to the subject, or saved to disk for later use by the Sound Kit. The instruments are capable of producing high-quality digital signals. For instance, the interpolating sine generator (`Waveli`) produces a digital signal with a total harmonic distortion (THD) of 1.4×10^{-9}, which is much higher resolution than the NeXT DAC hardware can reproduce. The noise generator (`Unoise`) produces white noise with flat spectrum from DC to 20 kHz, as shown in Fig. 5. By combining "unit generators," one can produce a new DSP instrument with specific characteristics that may be used as any other instrument in the orchestra. We also created a Shepard tone instrument using the equations of Shepard (1964) consisting of 10 independent sine oscillators and a number of scalers and adders, as described in the signal flow diagram in Fig. 6.

The stereo imaging in the Music Kit was also measured. It operates by varying the left and right amplitudes of an instrument's sound, proportional to the cosine of the bearing, giving a scaling contour as shown in Fig. 7. The stereo placement of an instrument has no effect on its phase.

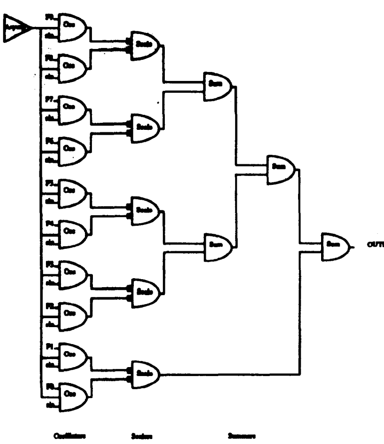


Fig. 6. Shepard tones "synthpatch" block diagram.

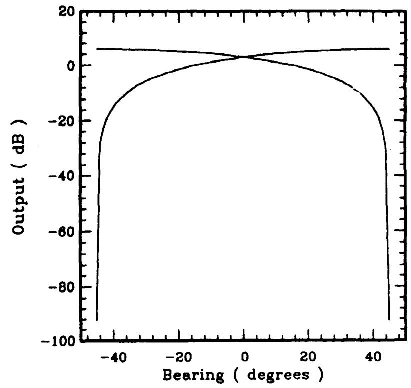


Fig. 7. Stereo Sound Bearing.

To be sure, in real-time performances and experiments, the Music Kit has a finite limit on the number and complexity of instruments or performers allowed at any one time. Typically, we found only four to six relatively simple instruments could be played at once, depending on their complexity (e.g., six sine wave generators each with amplitude- and pitch-envelope control). There is also a limit on the complexity of instruments; our Shepard tone instrument, with its 10 interpolating wave generators and numerous summers, appeared to be close to that limit. And while both the Sound Kit and Music Kit can run at the lower 22-kHz sampling rate, for saving space and computation, they also provide only half the audio bandwidth and require an external low-pass filter to remove the out-of-band signals mentioned above. Finally, a "leak" in the virtual memory system, due to a bug in both System 1.0 and 2.0, will cause the swap file on disk to grow to overflowing when performing extensive sound editing with the Sound Kit, which has occasionally led to system crashes during our experiments.

NeXT also supplies tools to program the DSP chip directly (in DSP56000 assembly language due to the lack of a compiler for any higher-level language). This is, however, an order of magnitude more difficult than using the other tools, perhaps more so for real-time sound generation, which we did not explore extensively. However, we did use the DSP directly in the acquisition and analysis of analog data for the measurements made above. Because third-party analog spectrum analyzers available to us allowed measurements accurate to only an 80 dB signal-to-noise ratio, we constructed a digital spectrum analyzer using a single Motorola DSP56ADC16S analog-to-digital converter, which had the required noise characteristics. Initially it was used in conjunction with the DSP56001 on an IBMPC version of Motorola's Application Development System; however, we found it more convenient to connect directly to the DSP56001 on a second NeXT. The IBM and NeXT gave essentially the same measurements, although the Unix tools on the NeXT resulted in a much more flexible measurement device.

General Software and System Management

Besides the facilities for sound processing described above, the NeXT machine has a number of general software development tools. Most programming is done in C, or Objective C, though other languages such as Lisp and Smalltalk are available from third parties. Objective C is a hybrid object-oriented programming language that is essentially a mix of C and Smalltalk; it comes with a large library of useful modules to work with. NeXTStep System 2.0 also provides the C++ language. Most useful is the *Interface Builder* user interface construction tool, which allows the building and testing of a graphics interface to an experiment without any text-oriented programming. Figure 8, for example, shows an interface that has been used in a variety of studies to explore the role of scale structure and sequential structure on memory for strings of sounds (Cohen et al. 1990, 1991). This window contains buttons for the subject to control the progression of the experiment, sliders for the subject to record the interpretation of a melody (as a two-dimensional representation), and information boxes that provide instructions and feedback to the subject. This interface was created without any text programming using a "drawing" paradigm whereby user interface components such as buttons, sliders, and text labels are composed graphically. The programmer then took the core routines of the experiment (sound generation, response recording—created as program libraries written as text) and linked them to the graphical interface, as shown in Fig. 9. The advantage is that not only can the experimenter design the experiment without any complicated programming, but also a completely new experiment can be created just

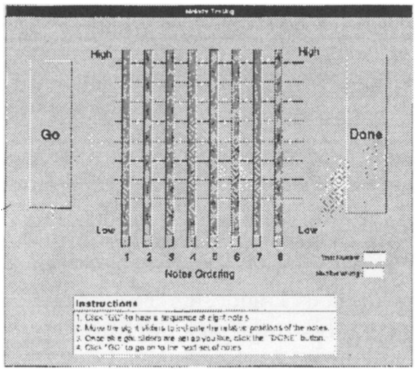


Fig. 8. Graphical user interface to Sequence Tracker.

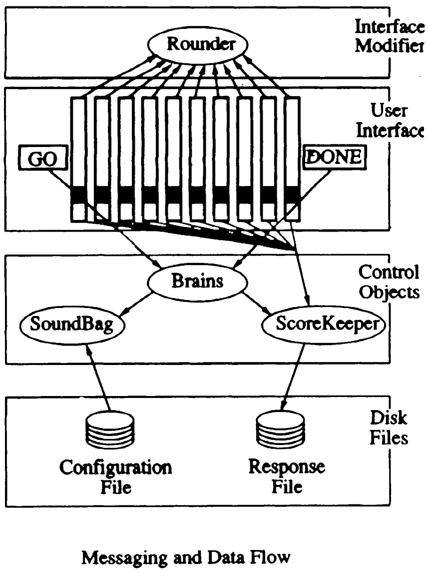


Fig. 9. Sequence Tracker objects.

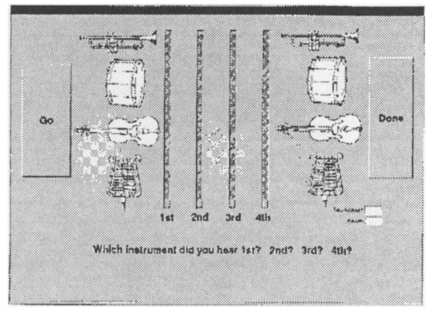


Fig. 10. Modified interface for testing memory for order of instruments.

by changing the interface and the sounds. For instance, Fig. 10 shows an interface that is based on the one shown in Fig. 9 but works with different musical instrument timbres instead of different musical pitches. Of course, there is no need to confine these choices to traditional instruments or symbols. We have also developed a similar, simple interface with animal sounds, which is appropriate for testing very young listeners (see Cohen et al. 1991). It took little time to create this new experiment, with no programming. In principle, a library of tools could be created in which a complete experiment can be produced simply by graphically connecting a variety of objects. It should be pointed out that although some modifications of experiments are rather straightforward, others, such as changing from mouse to keyboard input, may require considerable reworking of a program. Once created, however, such an adaptation can generally be transported to other programs with ease.

Collecting the subject's responses is also a critical issue when running an experiment. On the NeXT, a user enters a response by the keyboard, by the mouse that controls the graphical objects on the screen, or by voice input, although we have not fully implemented the latter. In addition to the particular response made by a subject, the time to respond is often a useful measure in psychological and psychoacoustic research. The NeXT system has a microsecond clock that can be used for general timing, but the keyboard and mouse are essentially timed by the video refresh clock running at 68 Hz, which may be a limitation in experiments requiring millisecond-accurate timing.

Since the NeXT runs Unix, it should be straightforward to port other programs to the machine, although, to us, it seemed such useful applications as X Windows were rather long in coming to System 2.0. In particular, a number of sound research software tools for the NeXT are now available both commercially and in the public domain. The NeXT also connects via Ethernet or a serial line to other computers, allowing the sharing of data, as well as via other network standards including Novell and NSF. This allowed us to dump our data to a more powerful SunMicrosystems workstation for data analysis, in order to utilize a program that had not yet been ported to the NeXT.

Conclusions

In general, the sound-generation abilities of the NeXT make it well suited for psychoacoustical and music perception experiments in the 20-Hz to 20-kHz audio band. The 16-bit resolution hardware, with 13-bit linearity, provides ample dynamic range and signal-to-noise ratio for all but the most demanding experiments. The software tools provided with the system reduce the burden of the creation and management of digital sounds. The experimenter must remain alert to a number of deficiencies, in particular, the presence of a deemphasis filter that may be inadvertently activated from the keyboard, out-of-band imaging errors at low sampling rates, software limitations on the complexity of sounds that can be generated in real time, and possible timing problems resulting from multitasking. Careful programming and experiment design can avoid these trouble spots, once one is aware of the difficulties. Overall the NeXT meets many of the technical specifications required for psychoacoustic instrumentation and also provides the tools to create an accessible computation platform for composers as well as auditory researchers. The combination promises much for the acceleration of the application of psychoacoustics to musical issues.

Acknowledgments

Fred Cotaras, Wayne Higgins, Stephen Hughes, Geoff Szabo (Defence Research Establishment of the Atlantic), and G. Stroink (Department of Physics, Dalhousie University) provided invaluable access to certain instrumentation. Gary Burns of Bruel & Kjaer, Canada, provided consultation. We thank Debora Dunphy for assistance and acknowledge Bob Bergeler and the Motorola University Support Program. This work was supported by an NSERC operating grant and NSERC equipment grant made to A. J. Cohen. A preliminary version of this report was presented at the 120th Meeting of the Acoustical Society of America, November 1990 (Lamoureux and Cohen 1990).

References

Cohen, A. J., B. Frankland, M. Lamoureux, and D. Dunphy. 1990. "Effects of Tone Sets and Serial Order on Melodic Tracking." Abst. 5MU7. *Journal of the Acoustical Society of America* 88:S91.

Cohen, A. J., M. Lamoureux, and D. Dunphy. 1991. "NeXT in the Psychology Laboratory: An Example of an Auditory Tracking Task." *Behavior Research Methods, Instruments and Computers* 23 (4):523–536.

Draft AES Standard Method for Digital Audio Engineering Measurement of Digital Audio Equipment. Draft AES17-1991, ANSI S4.51-1991. New York: Audio Engineering Society.

Jaffe, D., and L. Boynton. 1989. "An Overview of the Sound and Music Kits for the NeXT Computer." *Computer Music Journal* 13 (2): 48–55. Also in S. T. Pope, ed. 1991. *The Well-Tempered Object: Musical Applications of Object-Oriented Software Technology*. Cambridge, Massachusetts: MIT Press.

Lamoureux, M., and A. J. Cohen. 1990. "Evaluation of the NeXT Computer System for Auditory Research." Abst. 5MU10. *Journal of the Acoustical Society of America* 88:S92.

Lansky, P. 1989. "It's About Time: Some NeXT Perspectives (Part One)." *Perspectives of New Music* 27(2):270–281.

Lansky, P. 1990. "It's About Time: Some NeXT Perspectives (Part Two). *Perspectives of New Music* 28(1):170–191.

NeXT, Inc. 1989. *The NeXT System Reference Manual*. Redwood City, California: NeXT, Inc.

Shepard, R. N. 1964. "Circularity in Judgments of Relative Pitch." *Journal of the Acoustical Society of America* 36: 2346–2353.

Webster, B.F. 1989. "What's Next?" *Byte Magazine* January: 108–117.

Appendix: Measurement Instruments

HP3562A Dynamic Signal Analyzer, Dual Channel FFT

HP5315A Universal Frequency Counter

HP400E Average Responding Voltmeter

HP True RMS Voltmeter

Bruel and Kjaer Dual Channel Realtime Frequency Analyzer, Type 2133

Motorola 56ADC16S True 16-bit Digitizer, interfaced to NeXT DSP port.

Motorola 56000 Application Development system

Product Announcements

OSC Deck for Apple Macintosh Computers

Deck 2.0 is a multitrack digital audio recording and mixing system for Apple Macintosh computers with Digidesign sound cards. Formerly distributed by Digidesign, Deck 2.0 offers 16-bit, four-track nondestructive hard-disk recording, real-time onscreen fader automation, timeline-style multitrack waveform editing with track-slip, SMPTE synchronization, unlimited track bouncing, and synchronous QuickTime picture and audio playback. Deck also works simultaneously with MIDI sequencer programs, including OSC's Metro program. The retail price of Deck 2.0 is about $300. Contact OSC, 480 Potrero, San Francisco, California 94110 USA; telephone (415) 826-1121; fax (415) 826-2292.

AT&T DSP 3210 Digital Signal Processor

The AT&T DSP 3210 digital signal processor chip, which has been bundled into recent Apple Macintosh computers, is a 32-bit floating-point circuit with four memory accesses per instruction cycle. All instructions are single cycle. The 32-bit byte-addressable memory space of the 3210 allows it to share the address space of a host microprocessor.

AT&T offers a number of development tools for the 3210, including the VCOS operating system, a C-language compiler, assembler, link editor, make utility, and interactive simulator. The application software library includes a variety of routines for floating-point arithmetic, matrix pro-