Journal of Music Theory Pedagogy

Volume 1 Article 8

1-1-1987

# A Comparison of Synthesized and Acoustic Sound Sources in Lower-Division Theory Courses

W. Kenton Bales

Roger E. Foltz

# A COMPARISON OF SYNTHESIZED AND ACOUSTIC SOUND SOURCES IN LOWER-DIVISON THEORY COURSES

W. KENTON BALES

ROGER E. FOLTZ

## BACKGROUND

In 1981, the Department of Music at the University of Nebraska at Omaha initiated a new approach for teaching aural perception of music based on the Personalized System of Instruction (PSI), also known as the Keller Plan (Keller and Sherman, 1964). With this approach, students were required to pass a number of competency tests during the semester. The final grades were based on the number of exams successfully completed. Students were allowed to go on to a new test only after they had exhibited competency on the current test. *Competency* was defined as a score of 85% or better on each exam. Students were provided with a great deal of individualized instruction from the teacher, graduate assistants, and a computer-assisted instruction laboratory that included lessons on PLATO and Apple II computers. After two years of study, the results of this experiment suggested that PSI in aural-training classes did not significantly increase students' performance in the classes, but did increase student morale and had a positive effect on attrition, which is traditionally high in college-level classes in aural training (Bales and Foltz, 1985).

As an ongoing part of the research conducted with these classes, the effect of synthesized sound as a stimulus was examined, since this was a major sound source for student practice in the PSI project. Synthesized sounds have long been used as stimuli in music courses and for research in music, speech pathology, physical acoustics, foreign languages, and aural perception. Much research has been done in the area of auditory perception (Deutsch, 1982). The perception of pitch and timbre has been found to be mutually dependent (Plomp and Smoorenburg, 1970), and the dominant frequencies in complex auditory stimuli have also been described (Ritsma, 1967). The role of height, melodic contour, and timbre in the recognition of melodic patterns has been studied (Massaro, Kallman and Kelley, 1980), and dependence of timbre on the tonal loudness produced by musical instruments has been examined (Clark and Milner, 1964).

However, the research mentioned above deals primarily with audio and psychophysical perception of isolated musical events. We were unable to locate any studies that compared synthesized sound directly to acoustic sound when used as a stimulus in courses in aural training. Remez, Cutting, and Studdert (1980) compared response to synthesized violin tone with electronically gated speech in order to support the hypothesis that these sounds do not share the same feature detectors. Cope reviewed the mechanics of listening to synthesized music (1977). However, no conclusive research has been done to ascertain whether or not the timbre (color) and feature detectors of synthesized sounds is an acceptable stimulus for use in music courses, particularly in the area of aural-skills training. The perception of musical timbre has been linked to the dynamic processes by which sound is created (Balzano, 1986). These processes are encountered and initiated outside, as well as inside, musical contexts and never lead to a static spectrum analysis idealized by the use of Fourier Analysis, except in degenerate cases (297).

It occurred to us that this question of timbral appropriateness or "goodness of fit" (Risset and Wessel, 1982) is a concern of all who teach aural training and conduct research in perception using synthesized sound as a stimulus. Our speculation, based on personal observation and discussion with colleagues who taught similar classes at other institutions, was that there was no significant difference between sound sources with different timbres when used as stimuli in aural-training courses if those sources were equally familiar to students. However, also through personal observation, we noted that subjects seemed to respond differently to the sounds of a piano played in class than they did to the synthesized sounds in the computer-assisted instruction lab, which they were required to use. This observation was further substantiated when subjects expressed concern that the practice they were taking in the computer lab was given by an instrument whose timbre in no way resembled that of a piano, the instrument with which they were tested in the classroom. The same phenomenon was reported by colleagues in other similar situations, and led to the question "What is the difference between these two sound sources with regard to success in aural perception?" At first this research was undertaken in order to make sure that students were treated fairly, since they were practicing with one sound source, the computer, and being tested with an entirely different one, the piano. Since recent research supporting the effectiveness of computer-assisted instruction (Lee, 1975) has led to significant increases in the use of computer-assisted instruction and, consequently, synthesized sound, we were further motivated to compare the use of the piano with that of synthesized sound within the PSI environment.

After the problem outlined above was defined, a second question arose. It has not been proven that musicians who learn to respond to a synthesized sound source can respond in a similar way to any other sound source. Nor is there overwhelming data proving that musicians who respond to a piano can respond similarly to any other instrument. It is quite likely that the students who expressed concern about the sound source with which they were practicing out of class also could have complained if they had been listening to brass ensembles, string ensembles, an organ, woodwind ensembles, or any other traditional instrument or instrumental or vocal ensemble. With this in mind, we wished to determine whether or not subjects could respond in a similar manner to unlike sound sources. The problem is a practical one, since there is not conclusive evidence that a student who learns to identify a particular interval on a piano, for example, will be able to identify the same interval when it is played on a flute.

Musical sounds are intrinsically complex (Risset and Wessel, 1982), and musical instruments have a complex behavior (Benade, 1976). During most real musical experiences, such as listening to a piano, the tones are not generated by a standardized mechanical player. Music is produced by human musicians who introduce intricacies both intentionally and unintentionally (Risset and Wessel, 1982): "Even if a human player wanted to, he could not repeat a note as rigorously as a machine does. More often the performer will not want to play all notes the same way. All these considerations, which involve different disciplines—physics, physiology, psychology, esthetics—certainly make it difficult to isolate characteristic invariants in musical instrument sounds" (30). With this in mind, it was obvious that the subjects who received aural stimulus from a piano played by the teacher in the classroom were not hearing the same example each time. Further, these same students were likely to hear more accurate mechanical performances from the computer in the listening laboratory. Do the exercises presented by the computer, then, lack aural relevance? If synthesized versions of musical sounds are to be useful to musicians and music educators, they must have feature detectors similar to those found in traditional musical instruments; in other words, they must have aural relevance, or "goodness of fit."

The purpose of this study, then, was twofold. First, we wished to examine the aural relevance of synthesized sound, both in the classroom and in individual practice sessions, for students in aural-training classes. This was addressed by asking, "Is the timbre of the synthesizer used in the computer-assisted instruction laboratory good enough for students to use as a major resource in aural-training classes?" Second, the experiment was designed to determine what, if any, transfer of skills was identifiable between sound sources—in this case, between the synthesizer and the piano. In other words, can students who successfully complete exercises on the computer transfer their skills and complete the same or similar exercises when the sound source is a traditional piano?

We felt this work was necessary for several reasons. First, as stated above, recent expansion of computer-assisted instruction laboratories at all levels of instruction (elementary, secondary, post-secondary, and adult-continued education) is making extensive use of synthesized sound of relatively low quality with equally low fidelity. Almost none of these laboratories makes use of the high-quality sound reproduction currently available through the use of advanced digital synthesis. These new machines are capable of synthesizing all the necessary feature detectors to fool even practiced listeners with regard to identification of electronic or acoustic sounds. Second, it has never been conclusively proven or strongly suggested that skills gained by practice with these computer-synthesized stimuli can be transferred by students to any other sound source. Third, while a great deal of research has been done in the realm of audio acoustics, no work has been done that directly compares acoustic and synthesized sound sources. Finally, while a great deal has been done in the area of sound synthesis, very little has been done that examines the effect of these synthesized sounds on human behavior.

## PROCEDURES

Since the context of musical cues is important (Rasch and Plomp, 1982), the experiment was carried out during the course of two regular semesters of aural-training classes. Two classes totaling sixty-one (n=61) first-semester aural-training students were divided into four random groups as prescribed by the Randomized Solomon Four-Group Design (Campbell and Stanley, 1966). The groups were designated as C1 (n=13), C (n=18), E1 (n=15), and E2 (n=15). Groups C1 and E1 were given a pre-test based on the content of the course for first-semester aural skills. Groups C2 and E2 were not pre-tested. The four groups were then assigned to two sections of the course, with C1 and C2 in the control group and E1 and E2 in the experimental group. The control group (n=31) was taught using a piano, the traditional instrument for sonic stimulus in aural-perception classes. The experimental group (n=30) received all classroom instruction from a synthesized sound source, a Yamaha DX-7 with a Crate 40-Watt amplifier and two Crate loudspeakers. A sound was programmed and used during the course of the semester that closely matched that generated by the MMI DAC Board, the sound source in the Apple II computers. Both groups were subjected to the same material in the same order for the entire semester. All musical examples used during class were determined before class periods and were uniform for the two groups. The instructors and graduate students involved in teaching the class used identical class plans for each day of the classes during the semester-long study.

In order to minimize the effect of different teaching techniques on the experiment, the instructors of the two groups alternated every two weeks during the course of the sixteen-week semester, so both groups were exposed equally to both teachers. Graduate assistants periodically visited both classes and regularly met with the instructors to further insure uniform instruction. All subjects were required to spend two thirty-minute sessions of supervised practice each week outside class. The subjects in the control group practiced in groups of five or fewer at a piano with a graduate assistant playing examples similar or identical to those used by the experimental group. The students in the experimental group practiced at Apple computers equipped with digital synthesizers (MMI DAC boards). The Yamaha DX-7 used in the classroom for all examples and testing was programmed with a sound that emulated the sound of the synthesizers in the computers. Both groups practiced the same material each week.

At the end of the semester, all subjects were given a post-test with the piano as the sound source. The test was exactly the same as the pre-test given to groups E1 and C1. Test results were tabulated, and, using the post-test score as an arbitrary measure of success in the class, the following results were recorded.

## RESULTS

Group C1 scored a mean of 55.38% with a standard deviation of 15.385 on the pre-test. Group E1 scored a mean of 49.6% with a standard deviation of 15.52 on the pre-test. T-test results of these two groups give $t = -.952$ with $df = 26$. From this it was determined that there was no significant difference between the groups at the .05 level of confidence for a two-tailed test.

At the end of the semester, all subjects in all groups were given a post-test that was identical to the pre-test. The results of the post-test are shown in Figure 1. The two control groups, C1 and C2, and the two experimental groups, E1 and E2, were compared in order to ascertain whether or not pre-test bias was a factor in the result of the experiment. For groups C1 and C2, t = .749 with df = 29. For groups E1 and E2, t = .2547, with df = 28. In both cases the t values do not exceed the cutoff point for significance at the .05 level of confidence for a two-tailed test. Since the pre-test bias does not appear to be statistically significant, the post-test data for all subject groups was used for further analysis.

Figure 1.

| CONDITION | MEAN | STANDARD DEVIATION |
|---|---|---|
| C1 (n=13) | 82.154 | 8.184 |
| C2 (n=18) | 78.944 | 13.756 |
| E1 (n=15) | 77.933 | 6.453 |
| E2 (n=15) | 76.667 | 18.192 |

The total mean for the control group was 77.3% with standard deviation of 13.201 (n = 31). The mean for the experimental group was 80.29, standard deviation was 11.498 (n = 30). A t value of 1.3 with df = 58 was found not to be significant at .05 level of confidence for a two-tailed test.

With the sample sizes of n=31 and n=30 observed here, we can calculate a 95% confidence interval with interval bounds of ± 6.35%. This gives an estimate of the difference between the experimental and control group to be between +9.34% and -3.36%. In short, whatever differences may exist must be small.

Further processing of data included analysis of variance and analysis of covariance with pre-test scores used as a covariate. Results of the analysis of variance is shown in Figure 2. Results of the analysis of covariance with the pre-test scores for groups E1 and C1 used as co-variate is shown in Figure 3.

Figure 2. Analysis of Variance of the Four Groups.

| | | |
|---|---|---|
| Method: | Piano x Synthesizer | SS= 158.863<br>df=1<br>f = .98 |
| Treatment: | Pretested x Postested | SS= 75.374<br>df=1<br>f = .465 |
| | Method x Treatment | SS= 14.199<br>df=1<br>f = .088 |
| | Error | SS=9236.9<br>df=57<br>MS= 162.05 |

Figure 3. Analysis of Covariance with Pre-test Scores as a Covariate.

<u>Pre-Tested Groups (C1 and E1)</u>

| | | | Error |
|---|---|---|---|
| SS = | 60.304 | SS = | 1085.657 |
| DF = | 1 | DF = | 25 |
| MS = | 60.304 | MS = | 43.426 |
| F = | 1.389 | | |
| P = | .248 | | |

<u>Homogeniety of Regression</u>

SS between regression = 108.973
SS within regression = 976.684
F (1,24) = 2.678 P = .11

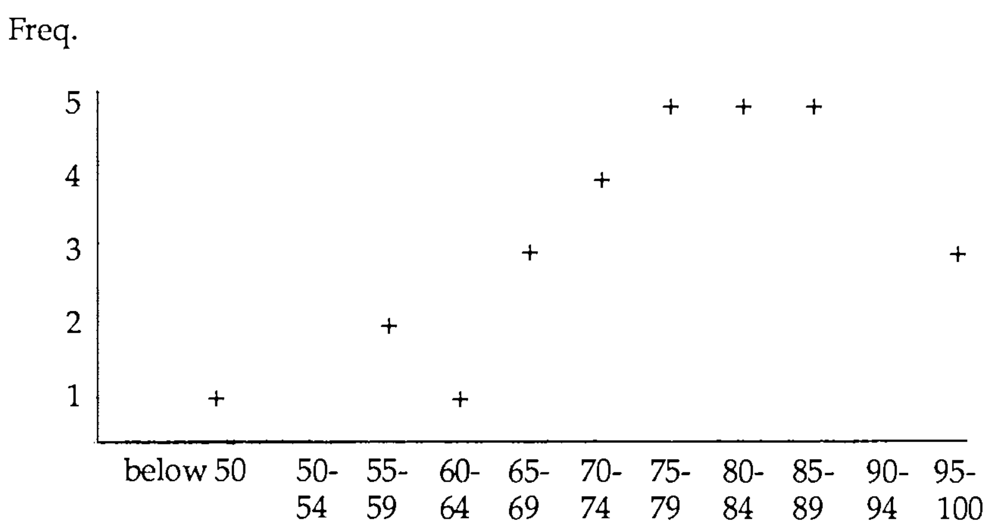
Within Group Correlation

R = .466

## DISCUSSION

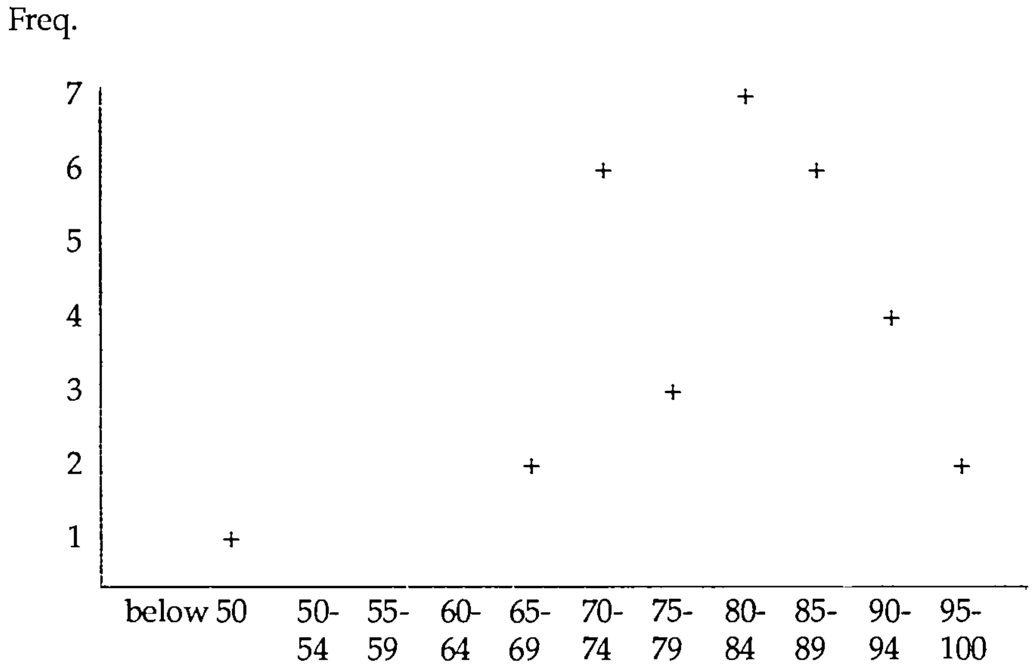
The excessive potential for error as indicated by the above results would lead to caution with regard to generalizations. This may be partially explained by the distribution of scores of the two groups (see Figures 4 and 5). The definite skew to the right as well as the relatively high standard deviations are to be expected, since the purpose of the class was to produce subjects who had more developed aural skills at the end of the term than at the beginning. However, the results of the experiment imply that there is little difference between the synthesizer and traditional instruments, such as a piano, as a choice for sonic stimuli in courses.

Figure 4. Dot-Array Diagram of Post-test Scores, Experimental Group




N = 30 Mean = 77.3 Standard Deviation = 13.201

Figure 5. Dot-Array Diagram of Post-test Scores, Control Group




N = 30 Mean = 80.29 Standard Deviation = 11.498

Students were given continuous reinforcement of desired behavior in both sections, and the results as reported above are posited to be similar in other situations where skills in aural perception are important. The course content and method of presentation were not intended as a variable in the experiment, since the purpose of the work was not to test a method of teaching, but to test the appropriateness of the synthesizer as a sound source in any course dealing with aural perception. In addition to the data discussed above, an attitudinal survey was administered. This measurement revealed that only 16% of the total population, with an opinion, disapproved of the use of the synthesizer in the classroom. No students who used the synthesized sound on the computer outside the classroom disapproved of that sound. This would suggest that any significant differences found in these experiments are based on listener prejudices and experience rather than on the chosen sound source. Further, this work suggests that transfer of skills gained in certain circumstances does occur when instruments with disparate timbres are used. Subjects who took the post-test on the piano, when they had had only a synthesizer as a sound source for the entire term of the experiment, did not score statistically higher than those who had used the piano for the entire term.

## SUGGESTIONS FOR FUTURE RESEARCH

From data collected and studied to date, it has become apparent to us that more research needs to be done in this area, since students and teachers are placing increased reliance on synthesized sound sources both in and out of the classroom. It seems that familiarity with the timbre of the piano as opposed to that of a synthesizer accounts for some of the perceived differences. However, a difference was indeed perceived by us and the subjects involved. What causes this perception? If indeed, there is no significant difference between the sounds produced by traditional acoustic instruments and electronic instruments, when used pedagogically, then the trend toward expanded use of electronic instruments in music training should continue. Otherwise, more work should be done to assure that subjects receive the best possible instruction on traditional instruments while continued research and development create electronic instruments that meet the needs of teachers and students alike.

We suggest that repeats of this study be carried out by others. It would be interesting for researchers at other institutions, with differences in class demographics and course content, to repeat this experiment, to see if similar results are produced. If the results of these repeats are similar to the results outlined above, it would seem that educators are justified in their move towards greater use of the synthesizer as a sound source.

## POSTSCRIPT

When this research was begun, high-quality digital synthesis and sampling technology were not readily available to most computer-assisted instruction laboratories in schools. The conclusion of little difference between sound sources, when used in a classroom situation, may seem to have little pertinence, given the fact that these newer technologies are now widely distributed and relatively inexpensive. However, the results of the study indicate that these advances in technology may not necessarily improve the quality and/or efficiency of ear training in lower-division theory classes. The results of the study also have important implications in the prioritizing of limited financial resources for equipment related to theory teaching. No doubt, students will continue to comment on what they perceive to be the poor quality of sound in the music laboratory. The results of this study are not intended to refute these students' feeling, but rather to serve as a reassurance of benefit from training in many circumstances—even those not approaching the optimum. Technological advancements will, of course, continue to have significant impact on the training of musicians, but they do not negate the values of older technologies and certainly do not replace practice and hard work.

## REFERENCES

Bales, W. K., and Foltz, R. E. "Competency-based Ear Training Using the Personalized System of Instruction." *Journal of Research and Performance in Music* 2 (1985): 107-120.

Balzano, Gerald J. "What are Musical Pitch and Timbre?" *Music Perception* 3/3 (1986): 297-314.

Benade, A. H. "The Physics of Brasses." *Scientific American* 229/1 (1976): 24.

Cadoz, C., Luciani, A., and Florens, J. "Responsive Input Devices and Sound Synthesis by Simulation of Instrumental Mechanisms: The Cordis System." *Computer Music Journal* 8/3 (1984): 60-73.

Campbell, D. T., and Stanley, J. C., *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally and Company, 1966.

Chowning, John. "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation." *Journal of the Audio Engineering Society* 21 (1973): 526-534.

Clark, M. and Milner, P. "Dependence of Timbre on the Tonal Loudness Produced by Musical Instruments." *Journal of the Audio Engineering Society* 12 (1964): 28-31.

Cope, David. "The Mechanics of Listening to Electronic Music." *Music Educators Journal* (October 1977): 47-53.

Deutsch, Diana. "Musical Recognition." *Psychological Review* 76 (1969): 300-307.

Deutsch, Diana, ed. *The Psychology of Music.* New York: Academic Press, 1982.

Keller, F. S., and Sherman, J. G., eds. *The Keller Plan Handbook.* Menlo Park: W. A. Benjamin, Inc., 1984.

Laske, Otto. "Toward an Explicit Cognitive Theory of Musical Listening." *Computer Music Journal* 4/2 (1980): 73-83.

Lee, Lawrence. "The Effectiveness of a Computer-assisted Program Designed to Teach Verbal-descriptive Skills Upon an Aural Sensation of Music." Doctoral diss., East Texas State University, 1975.

Massaro, D. W., Kallman, H. J., and Kelley, J. L. "The Role of Height, Melodic Contour, and Tone Chroma in Melody Recognition." *Journal of Experimental Psychology* 6/1 (1980): 77-90.

Plomp, R. *Aspects of Tone Sensation.* London: Academic Press, 1976.

Plomp, R., and Smoorenburg, G. F., eds. *Frequency Analysis and Periodicity Detection in Hearing.* Leiden: A. W. Suithoff, 1970.

Pollack, I. "Decoupling of Auditory Pitch and Stimulus Frequency. The Shepard Demonstration Revisited," *Journal of Acoustical Society of America* 63 (1978): 202-206.

Rasch, R. A., and Plomp, R. "The Listener and the Acoustical Environment." In *The Psychology of Music*, edited by Diana Deutsch. New York: Academic Press, 1982.

Remez, Robert, Cutting, J. E., and Studdert-Kennedy, M. "Cross-series Adaption Using Song and String." *Perception and Psychophysics* 27/6 (1980): 524-530.

Risset, Jean-Claude, and Wessel, David L. "The Exploration of Timbre by Analysis and Synthesis." In *The Psychology of Music*, edited by Diana Deutsch. New York: Academic Press, 1982.

Ritsma, R. J. "Frequencies Dominant in the Perception of the Pitch of Complex Sounds." *Journal of the Acoustical Society of America* 42 (1967): 191.

Wessel, D. L. "Timbre Space as a Musical Control Structure." *Computer Music Journal* 3/2 (1979): 45-52.