ORCHESTRAL INSTRUMENTS: ANALYSIS OF PERFORMED TRANSITIONS

John Strawn
The Droid Works*
P. O. Box CS-8180
San Rafael CA 94912, USA
Tel. (415) 485 5000

ABSTRACT

This study concerns the transition between notes as performed by professional (classical) musicians. Nine instruments (flute, piccolo, bass flute, clarinet, bassoon, oboe, trumpet, violin, cello) were recorded. On each instrument, the performer used two playing styles (such as with or without bow change) to connect notes spanning a variety of intervals from major second to minor seventh, ascending and descending. The time-varying spectrum and time-varying power of each transition was analyzed. The resulting data were analyzed to provide insight into performed transitions.

INTRODUCTION

At least since Helmholtz, musical notes have been broken down into convenient parts descriptively labelled "attack," "steady-state," and "decay." Such regions are indeed often fairly easy to identify in recordings of isolated tones.

When a musician is playing a melodic sequence, the resulting signal can be crudely approximated as a concatenation of individual tones. As the musician is playing the decay of one note, the amplitude of the note falls off, as does the spectrum—then a new pitch sounds as the amplitude rises and the spectrum of the new note quickly becomes richer in the attack. This paper concentrates on a time span of a hundred milliseconds or so in which these changes take place.

Learning to control the juncture between notes is part of the training of a professional musician, classical or popular. It is not adequate simply to play notes one after another, as the model of the previous paragraph would suggest. Successive notes must be joined in one manner or another. Inertia in the instrument plays a rôle too. In other words, the new note does not begin by itself: It must be helped along.

Wind and brass players are taught the technique of "tonguing", in which a syllable such as "ta" or "da" is "spoken" inside the mouth right as the new note begins. Use of this technique is optional: the wind or brass player can and may start a note with no tonguing at all. There are of course other options for starting a note as well. Likewise, the string player has the choice of continuing to move the bow in the same direction when starting a new note, or changing bow direction. Here, too, there are other options, such as varying the velocity of the bow while keeping it moving in the same direction. When to use these techniques, and how much separation to allow between notes, is a matter of applying long training seasoned with good taste to the particular musical passage at hand.

In this work, I chose to study the contrast between, say, tongued and untongued transitions. The purpose is not to learn more about the physical correlates of these playing techniques *per se*; rather, it was convenient to use two kinds of playing techniques which are well known to both performer and listener and which one may readily assume may be distinguished from each other. Subsequent research (Strawn 1985b), which will not be discussed here, showed this to be the case.

Most of the research on instrumental sounds to date has concentrated on analysis of just one note at a time. (Some studies have been done on musical structures longer than one note, but none of these has concentrated on analyzing the transition between notes). Much of an instrument's tell-tale "sound" lies in how the notes are connected, and in the instrument's "signature" as long musical phrases are created. It is thus important to examine more than one note at a time if the nature of musical sound is to be fully understood. The work reported here is the first comprehensive study of these questions ever published. Research along these lines also has implications for auditory theory in general (e.g., categorical perception), but these will not be discussed here (see Strawn 1985b).

*The Droid Works is an affiliate of Lucasfilm and Convergence. The work described here was performed at the Center for Computer Research in Music and Acoustics (CCRMA), Stanford University. "This research was supported by the National Science Foundation under grant NSF MCS 82-14350 and the System Development Foundation under grant SDF #345. The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of Stanford University, any agency of the U.S. Government or of sponsoring foundations."

Table 1. Instruments Recorded.

Family	Instrument	Base Pitch
Air Reed	Flute	A220
	Piccolo	A1760
	Bass flute	A220
Single Reed	Clarinet	A220
Double Reed	Oboe	A440
	Bassoon	A220
String	Violin	A220
	Cello	A220
Brass	Trumpet	A220

RECORDINGS

From 1979 through 1984 I made digital recordings (either directly onto computer disk or onto Sony F1 tape) of several instruments, as shown in Table 1. Each instrument played the following intervals, ascending and descending: major second, major third, perfect fifth, and minor seventh. For all of these intervals, the lower tone was A, and in fact A220 wherever possible; exceptions are given in Table 1. (Only the seconds and thirds were recorded on the strings). The two notes were played at about MM=60 (i.e., one note per second). Each interval was recorded with two playing styles for the beginning of the second note: tongued (with bow change), and untongued (without bow change). For some intervals on some instruments, as many as five separate recordings were made. Sony F1 recordings were transferred digitally to computer disk. All recordings were sampled down to 25.6 kHz and high-pass filtered to remove some recording artifacts. In all, several hundred transitions were recorded and analyzed.

Figures 1–4 show plots of some typical transitions; each horizontal plot shows 300 msec from a transition.

ANALYSIS

After preliminary analysis, I chose two methods for analyzing the data: time-varying power, and time-varying spectrum. For the first, I developed a technique which I call period-synchronous power analysis. Using a technique due to Julius O. Smith (1984), the peaks of the waveform were identified, one for each period. Minor errors in this output were corrected by hand using a graphics-based editor designed for this purpose. For

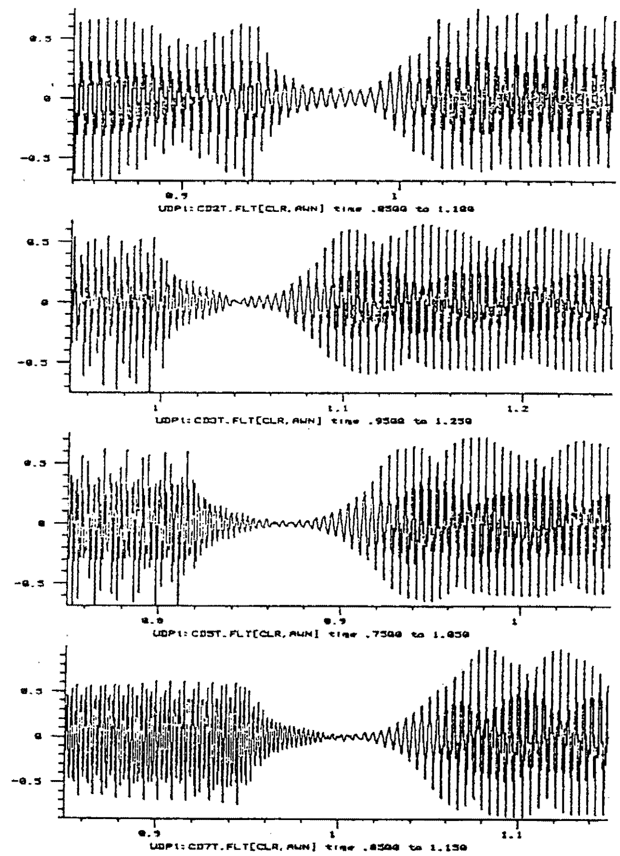

Figure 1. Recorded transitions between two clarinet notes; each plot shows 300 msec. The second note is untongued in each case, and is played on A220; all intervals are descending to that pitch. From the top: major second, major third, perfect fifth, minor seventh.

each period, power was calculated according to the formula

$$P(n) = \frac{1}{T(n)} \sum_{t=n}^{n+T(n)} y^2(t)$$

where $T(n)$ is the length of a period (peak-to-peak) beginning at sample number n, and $y(t)$ are the samples in the recording. $P(n)$ is thus measured once per period. Figures 5–7 show some typical period-synchronous power analyses of the trumpet. In general, note that the most significant differences occur between the untongued cases of Figures 5 and 6 on the one hand, and the tongued case in Figure 7 on the other. The differences for different interval sizes or directions are much less significant. With one major exception (to be discussed below), this was the case for all of the instruments analyzed.

These recordings were also analyzed using the short-time Fourier transform, in the form of the phase vocoder

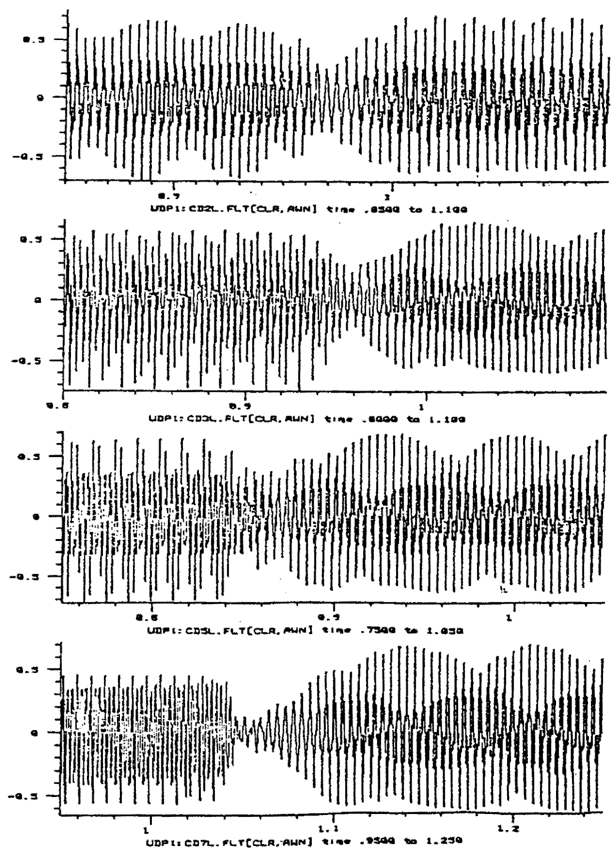

Figure 2. The same intervals as in Figure 1, this time with the second note untongued.

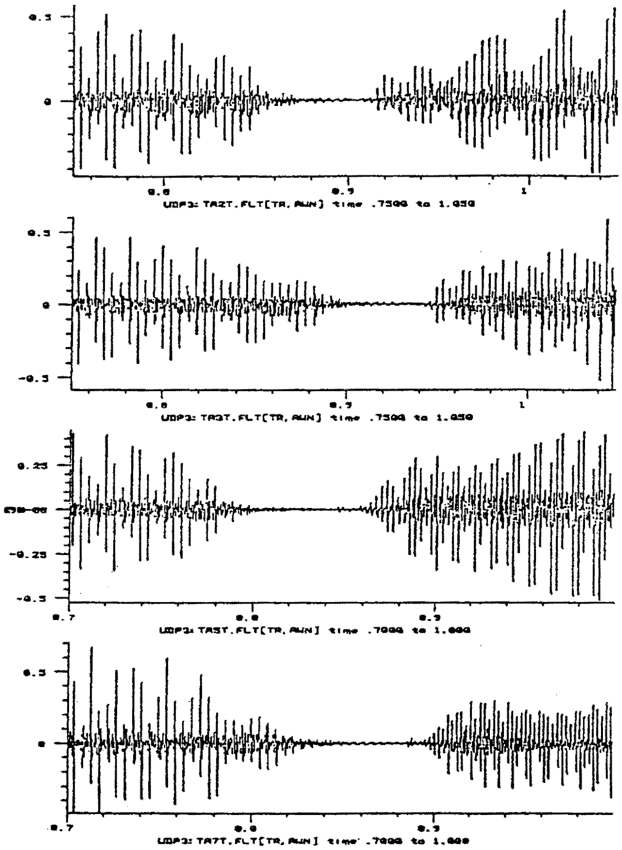

Figure 3. Recorded transitions between two trumpet notes; each plot shows 300 msec. The first note is A220 in each case; the second note is tongued, and played on a higher pitch. Again, from the top, the intervals are: major second, major third, perfect fifth, minor seventh.

(Portnoff 1978; Moorer 1978; Holtzman 1980; Gordon and Strawn 1985). In effect, this places a bandpass filter around each harmonic; the outputs of the filters give a time-varying spectral representation of the signal. The problem with using the phase vocoder for analyzing transitions is that the center frequencies of the filters remain fixed once set, so that the harmonics of the new note no longer fall onto the analysis channels in a useful way. Also, since the amplitude of the signal drops several tens of decibels during many transitions, the frequency traces are difficult to interpret, because the frequency trace becomes unstable at very low amplitudes. Furthermore, if the frequency traces are moving too quickly, it can be shown that the phase vocoder does not track them accurately, which might lead to distortion. These problems are discussed further in (Strawn 1985b).

I solved the problem of the displaced analysis channels by running the analysis twice, once with the filters set for the first note, and again with the filters set for the second note. It was necessary to expand my spectral editor (Strawn 1985a) to handle these two analyses properly. In particular, it was possible to create three-dimensional spectral plots by "splicing" between the two analyses at an appropriate point in the transition.

Some typical results are shown in Figures 8–9. Note the difference in the gap between the untongued transition of Figure 8, and the tongued transition in Figure 9. The "valley" between the notes extends "further down in" to the lower harmonics in the tongued case, and the gap between the notes in the higher harmonics is much wider. Again, similar behavior was observed for all of the instruments analyzed.

CONCLUSIONS

I was not able to develop empirical methods for analyzing the plots resulting from these two kinds of analysis.

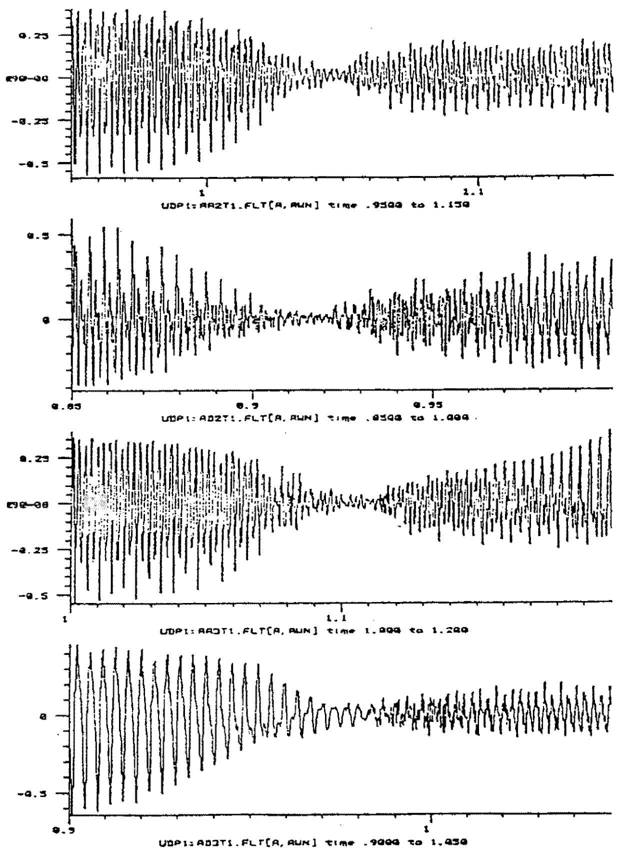

Figure 4. Recorded transitions between two violin notes played on the same string (G); each plot shows 300 msec. The lower note is A220 in each case, and the second note is played with a change of bow. From the top: major second ascending, major second descending, major third ascending, major third descending.

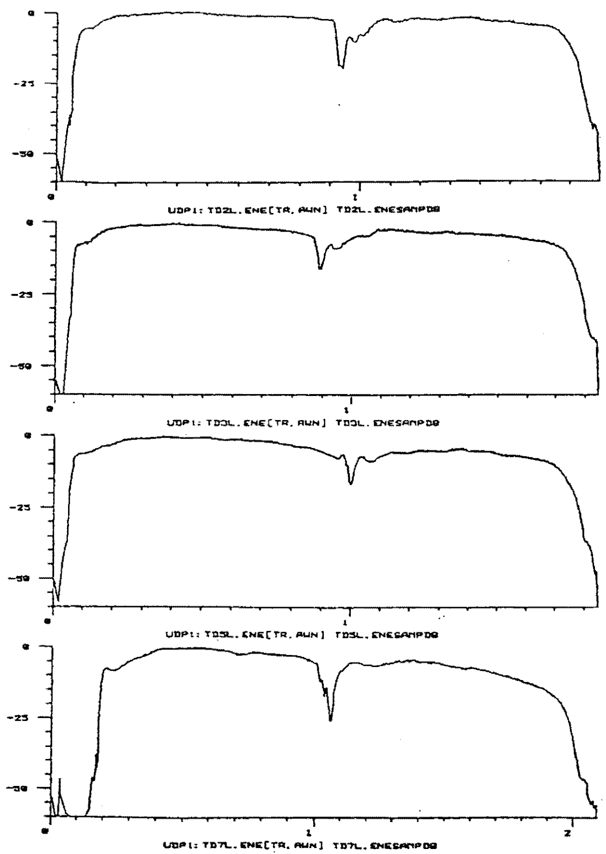

Figure 5. Period-synchronous power analyses of four recordings of two-note pairs played by the trumpet. The second note is played at A220 in each case, and was not tongued; the first note is at a higher pitch. From the top: major second, major third, perfect fifth, minor seventh.

However, examination of these plots suggests the following conclusions, some of which were borne out in later research (Strawn 1985b). (In the following, "tongued" as a generic term includes the "with bow change" case for the strings.)

1. Except for the cello (see below), there is a characteristic drop in overall amplitude between two notes, as one would expect. This should be clear from Figures 1–7.

Likewise, at the end of the first note there is a gradual spectral rolloff (see Figures 8–9). The transition region consists in general of a low-passed version of the tail of the first note. During the attack of the second note, the higher-frequency components re-enter.

2. The transition from one pitch to the next occurs very quickly. A good player can make the transition between notes in the time required for just a few periods of the waveform (see, for example, the nontongued transitions in Figure 2).

3. The skilled player can easily replicate a given articulation. Power and spectral analyses of as many as five repetitions of a given size/direction/playing style combination showed no significant differences. To give one example, Figure 10 shows period-synchronous power analyses of four different recordings by the same bassoonist. The durations of the notes surrounding the transition are slightly different in each case. But the shape of the transition (e.g., the slope of the decay of the first note, the slope of the attack of the second note, the very short trough between the two notes) is surprisingly consistent.

4. Looking more closely at the general behavior given above under #1, I have concluded that the tongued

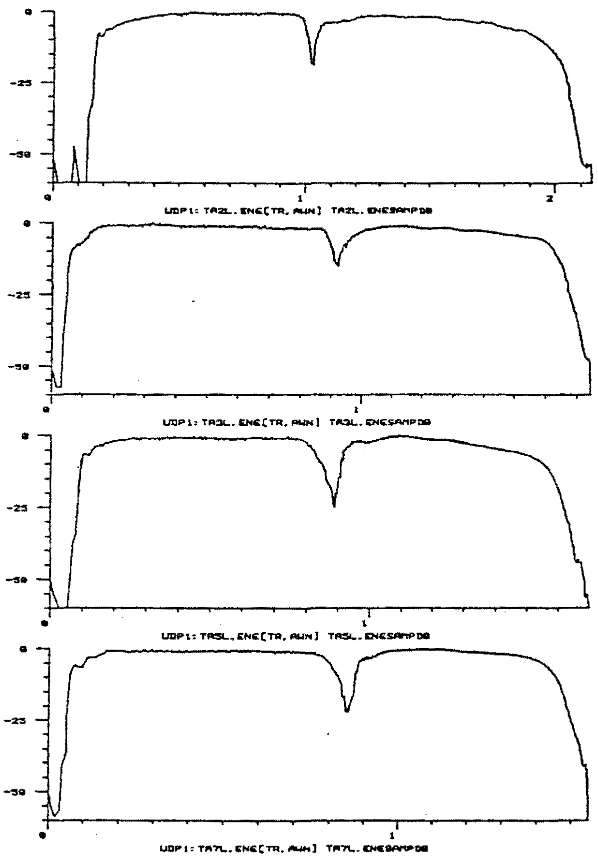

Figure 6. Period-synchronous power analyses as in Figure 5. Here we have ascending intervals instead of the descending intervals of Figure 5.

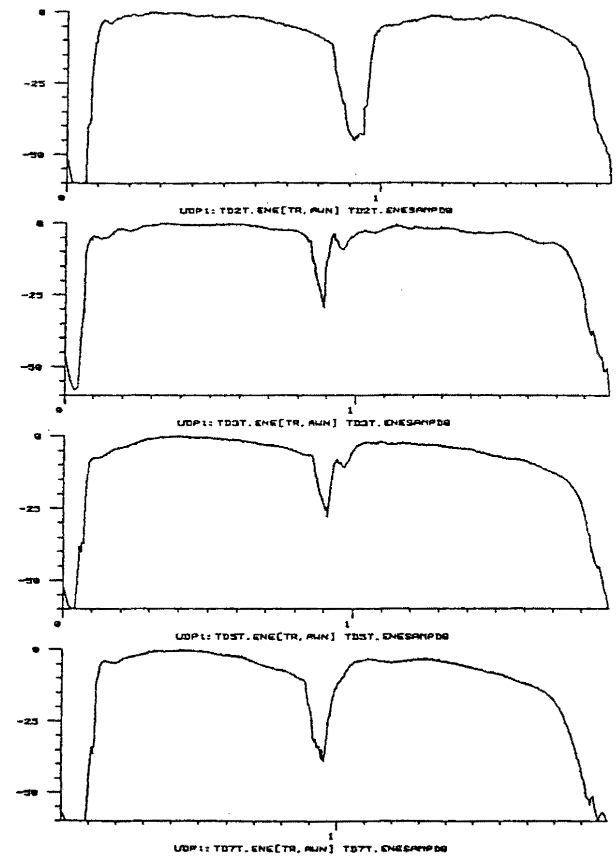

Figure 7. Period-synchronous power analyses of four recordings of two-note pairs played by the trumpet. The second note is played at A220 in each case, and *was* tongued; the first note is at a higher pitch. From the top: major second, major third, perfect fifth, minor seventh.

case often exhibits a wider gap between the two notes, and a greater dip in amplitude, than for the nontongued case. In no case did the amplitude between notes for the tongued case reach or stay at 0.0, as one might expect. This may be due in part to room resonance. (One would not expect a drop in amplitude to 0.0 for the nontongued case, nor did it occur there).

Likewise, the spectral rolloff is in general deeper for the tongued case than for the untongued.

5. There is no systematic difference between ascending and descending intervals for a given instrument, nor for intervals of varying sizes. This finding runs counter to what one might expect. Of course, some intervals on some instruments are harder to play than others, but no systematic differences were found.

6. The tongued transition in the woodwinds and brass has, as one might expect, a small amount of noise right at the attack of the second note. In the strings, one might expect "bow change" to produce a more abrupt attack on the second note. However, the "no bow change" produces its own abrupt attack on ascending notes, because the finger "thwacks" the string to make it shorter, producing a characteristic sound which is probably not noticed by any listener in a real listening environment.

7. In spite of instructions to these players to play the notes at the same loudness, the amplitudes of the two notes are often quite different (see Figure 10). As one might expect, the decay time of the first note is often different from the attack time of the second note. The attacks and decays often follow a two- or even three-tiered pattern, which was unexpected. Some instruments showed characteristic swellings on some notes.

8. Only in the string instruments does the size of the instrument seem to have any effect on the overall behav-

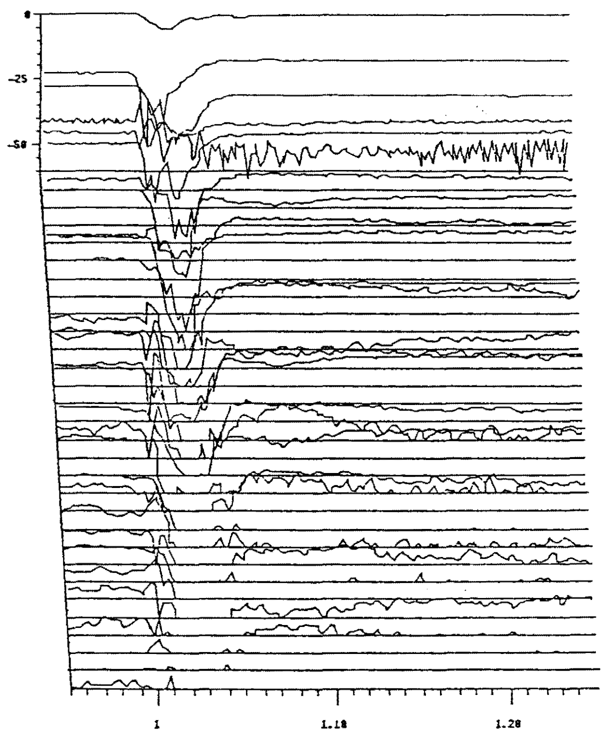

Figure 8. Three-dimensional representation of the amplitudes of the harmonics during a transition between two performed clarinet tones. The first note was A220; the second note was at C-sharp above middle C, and was played without being tongued. The entire plot covers 300 msec (x axis). The y axis shows amplitude on a scale of 0 to −50 dB. The fundamental is at the rear of the plot; 30 harmonics are shown.

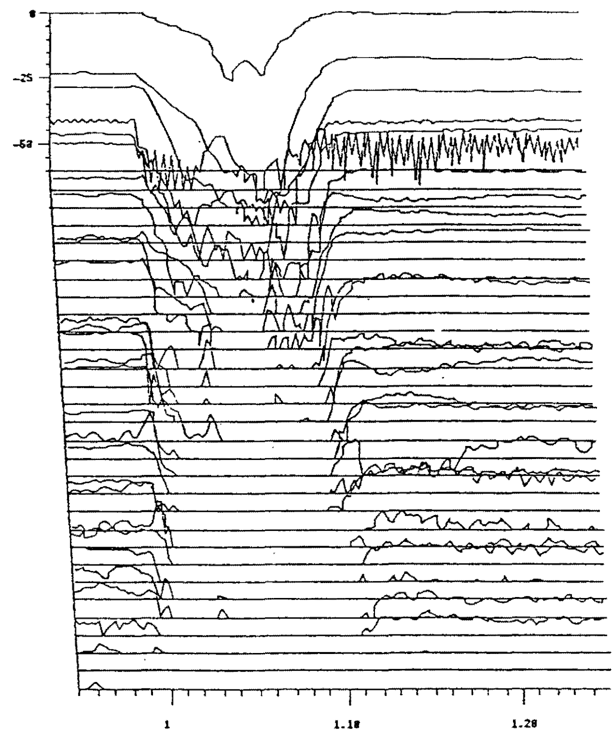

Figure 9. This plot shows the same transition as in Figure 8, except that the second note was tongued.

ior of the transition, and then only in the power analyses. In the cello, it is difficult to distinguish the amplitude dips of the bow-change and no-bow-change cases. I suspect that greater resonance of the cello body, increased mass of the cello string, and the like contribute to inertia which causes the first note to carry over a little into the second note, blurring the transition.

REFERENCES

Gordon, John W., and John Strawn. 1985. "An Introduction to the Phase Vocoder." In John Strawn, Ed. *Digital Audio Signal Processing: An Anthology.* Los Alto, California: William Kaufman, pp. 221-270.

Holtzman, Samuel. 1980. "Non-Uniform Time-Scale Modification of Speech." M. Sc. and E. E. Dissertation, Department of Electrical Engineering and Computer Science, MIT.

Moorer, James A. 1978. "The Use of the Phase Vocoder in Computer Music Applications." *Journal of the Audio Engineering Society* 26(1/2):42–45.

Portnoff, Michael R. 1978. "Time-Scale Modification of Speech based on Short-Time Fourier Analysis." Doctoral Dissertation, Department of Electrical Engineering and Computer Science, MIT.

Smith, Julius O. 1984. Personal communication, 4 March.

Strawn, John. 1985a. "Editing Time-varying Spectra". Paper presented at the 78th AES Convention, Anaheim. Preprint 2228 (A-16).

Strawn, John. 1985b. *Modeling Musical Transitions.* Ph.D. Dissertation, CCRMA/Department of Music, Stanford.

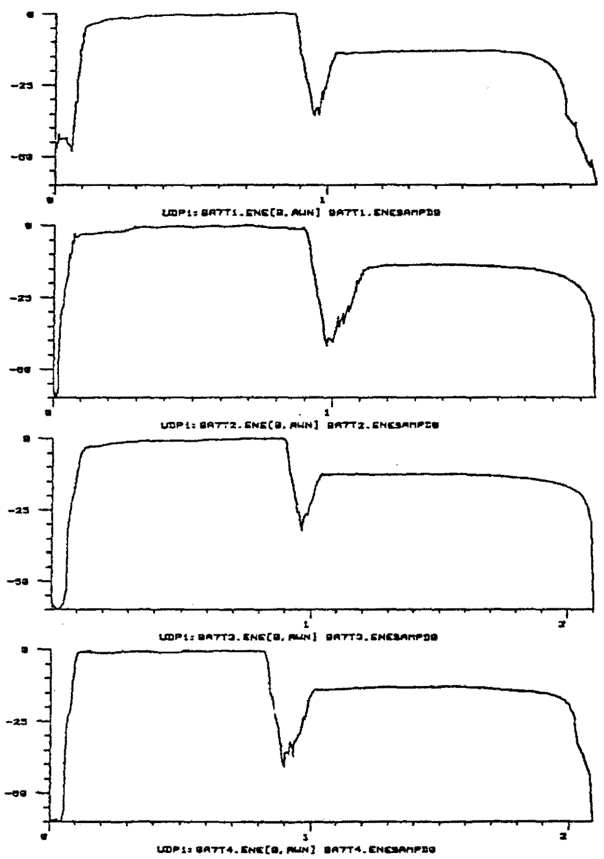

Figure 10. Period-synchronous power analyses of four different recordings of the same two-note pair played on the bassoon. The first note is A220, and the second is the G above middle C, played tongued.