Speech-Based Computer Music: Selected Works by Charles Dodge and Paul Lansky

by Madelyn Byrne
Hunter College and Columbia University
MByrne1204@AOL.COM

Abstract

This paper examines two speech-based computer music compositions: *Word Color* by Paul Lansky, and *In Celebration* by Charles Dodge. The following issues are considered; the composers' treatment of the voice, text interpretation, musical structure, timbre, and use of computer music programs.

1. Introduction

There were two general questions that propelled me toward a study of speech-based computer music, these were a desire to understand the musical and expressive possibilities of speech, and how various computer programs could be used to create music with those possibilities. I choose to analyze pieces that set meaningful texts and that were skillfully crafted by the composer to enhance the speaker's interpretation of those texts. The pieces discussed in this paper are *Word Color* by Paul Lansky and *In Celebration* by Charles Dodge.

Paul Lansky has often stated that he wishes to make explicit the implicit music in speech. In this endeavor, Lansky has often collaborated with Hannah MacKay, an accomplished actress, who has a remarkable speaking voice.

Charles Dodge has cited a long-held interest in acoustics and the voice, and the work of a number of Swedish text-sound composers, as having influenced his work. Dodge has also said that he was interested in the possibilities of computer generated speech, and that he wanted to integrate more of his personality with his music. He found a successful way of doing these things in his speech synthesis pieces. (Dodge has created a number of pieces that use a synthetic voice based upon the qualities of his own voice.)

Among the factors considered in these analyses are text interpretation, treatment of the voice, musical structure, timbre, and the composers' choice and use of computer programs.

2. *Word Color* (1992) by Paul Lansky

Important components of *Word Color* include the musical structure of the composition, the use of comb filters, the careful manner in which the piece alternates between isolated words and fragments of verse 17 of Walt Whitman's *Song of Myself*, and finally, how these factors work together to create a unified whole.

Word Color is a computer music composition based completely on samples of Hannah MacKay's reading of the aforementioned isolated words and Whitman text. Paul Lansky took the isolated words from children's stories; most of them refer to time. This piece was premiered at a festival in Delphi on a cliff overlooking the sight of the oracles. Lansky wanted a text befitting such a setting and chose the Whitman text for its "oracular" quality. He was especially drawn to the Whitman text for its broad-based character and its suggestions that profound truth is both familiar and universal. Universal truth is often thought of as timeless. This may explain the composer's juxtaposition of time oriented words and Whitman text fragments heard throughout the composition.

The composer is also interested in the double meaning of resonance, "*Word Color* is based on the sense that words, as sounds, can ring, and have resonance in our memory. While that memory may be regarded as purely sonic, words themselves inevitably reach more deeply into other areas of our consciousness...."

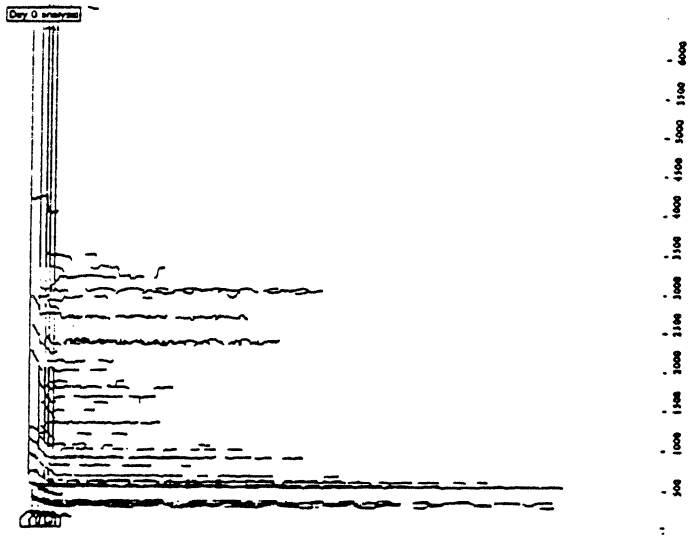
Lansky elaborated on this idea in our interview,

> The piece evolves over a course of time. One of the ideas of the piece was to capture the resonance of sound. For instance, everyone has the experience of saying a word and then saying it again, and again and at a certain point it makes no sense. It becomes a timbral object. I was interested in that. The idea was that these words just hang in space. There are ten second reverbs on words and they're just sort of hanging out there. I was also interested in the arbitrary nature of these words hanging out in space. As the piece evolves, the Whitman text comes to the surface – but more than that – the whole feeling of words as objects comes to the surface. It's a complicated texture. At the end there's a series of repeated ascending chords. I was inspired by Bizet's *L'arlesienne* Suite.... The idea was to take the chords and at the end resolve them into this pattern that goes on but the harmonies change with each four. There is a sense that it gets darker, it moves inward and becomes introspective with the repeated chords.

The composer used comb filters for harmony and bass tones, and combpluck (a CMIX instrument combining comb filters with the "plucked string" algorithm) to create a melodic accompaniment in a high register, somewhat analogous to a descant. These programs are never used to obscure the reader's voice, but rather to accompany her consistently clear and audible reading. With regard to combpluck, Lansky

The First articulation of "Day," 0 -7 seconds.

Figure 1

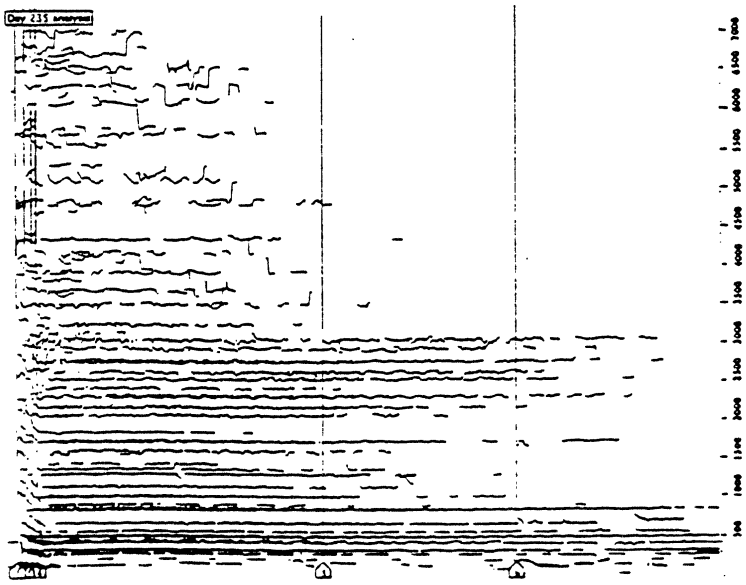


```
nchords 1

pitches = load_array(13,8.00,8.02,8.04)
revtime=9
expenv(0)
count = 0
dur=revtime+1
time=0 track=1 beat=0 beatpoint=-1
beatpoint2=-1
voiceamp=5
setcombs(x=1/cpspch(pch),revtime,1,
y=1/cpspch(pch+.001),revtime,1,
y=1/cpspch(pch+.002),revtime,1)
beat = trunc(count/rhythm)*dur
multicombd(beat,end-skip, skip,voiceamp,
0,.0005,.2,.4)
```

The second articulation of "Day" is at 2:35, after the composer has presented a series of isolated words and the opening fragments of the Whitman text. This setting occurs at the apex of a line spanning a tritone, has a bass motive in Bb underneath (one octave lower than the chord), and is cadential. The time shown below is 2:35 - 2:42.5.

Figure 2



```
nchords 23

pitches=load_array
(13,7.10,8.02,8.05,8.09, 9.00)
revtime=16
expenv(0)
count = 0
dur=revtime+1
time=0 track=1 beat=0 beatpoint=-1
beatpoint2=-1
setcombs(x=1/cpspch(pch),revtime,1,
y=1/cpspch(pch+.001),revtime,1,
y=1/cpspch(pch+.002),revtime,1)
beat = trunc(count/rhythm)*dur
multicombd(beat,end-skip,skip,5,
0,.0005,.2,.4)
```

explains that this program looks for peaks in amplitude and frequency. Combpluck was applied to the recording made by Hannah Mackay and the resultant data was then fed into a comb filter that would transpose the frequencies to a higher range (two and three octaves above middle C). This results in the rhythms and pitches inherent in Ms. MacKay's speech realized at a higher register.

The musical structure of the piece consists of three main sections; section one (0:00-3:18), section two (3:19-10:40), and section three (10:41-12:45). The first section is somewhat analogous to an exposition as it introduces all of the elements of the piece.

The middle section is marked by the prominence of the combpluck accompaniment and bass tones. This section may be further divided into five subsections. Subsections one, three and five focus on the isolated words, while subsections two and four focus on fragments of the Whitman text.

The last section is an augmentation of the introduction. The introduction (0:00–1:12) is composed of isolated words grouped into two units of four. The last section also focuses on isolated words, but in eight units of four. Furthermore, the chord types and voicings of both the introduction and the closing are very similar.

The central melodic motive of the composition was alluded to by Lansky; it is composed of groups of four chords or clusters supporting an ascending melodic motive of whole-steps, or a combination of whole and half-steps.

The word "Day" is the most apparent structural marker in this piece and provides an excellent example of the composer's concept of resonance. "Day" is the first sound of the composition, is heard four times throughout the piece, and often begins musical lines. The only time "Day" closes a line is at 2:35, when it is the last isolated word in that section. Comparing the opening setting of "Day" with the setting at 2:35 demonstrates how Lansky skillfully leads this initially austere composition into a spectrally rich sound world. The excerpts from Lansky's computer files illustrate pitch, detuning, reverberation time, vibrato, and amplitude information. The spectral analyses illustrate the timbral richness in the second occurrence of "Day" as compared with the austerity of the first. Harmonically, the opening chord may be interpreted as an unresolved secondary dominant to the Bb major ninth chord heard at 2:35. (See figures 1 and 2 for file excerpts and spectral analyses.)

3. *In Celebration* (1974) by Charles Dodge

In Celebration by Charles Dodge is an early example of a piece created with speech synthesis-by-analysis. The composer primarily used LPC. Dodge used his own voice, reading the Mark Strand poem *In Celebration*, as a model for the synthetic voice. The resynthesis uses a pulse generator that simulates a glottal wave for voiced speech, and noise that simulates turbulence in the vocal tract for unvoiced speech. The latter technique is sometimes employed by Dodge to create the illusion of a whisper.

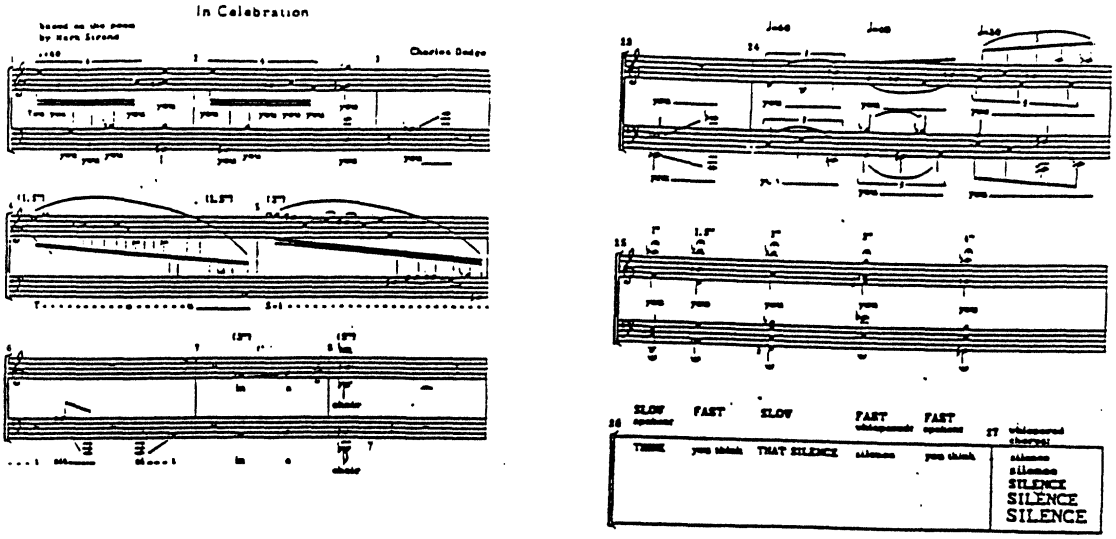
The poem portrays its subject's emotionally dissociated state, and his ultimate decision to celebrate death. The second point may be interpreted as a decision to commit suicide. Dodge's "otherworldly" synthetic vocal timbre, highly chromatic musical language, and thoughtful text setting aid in depicting the macabre and dissociated emotional quality of the poem. Settings of "you" and the role of pentachords illustrate these points.

Pentachords that function on a local level are composed primarily of interval classes 1, 3, and 6. These interval classes saturate the musical landscape of the piece. Cadential pentachords however, are composed of either rolled Bb's over five octaves (as in measure 8) or five articulations of a whispered word (as in measure 27). Measures 26 and 27 introduce the first spoken and whispered text, and the first whispered cadence of the piece. This cadence consists of five whispered articulations of the word "silence"that increase in volume. This treatment enhances the poet's use of that word, as "silence" foreshadows the subject's decision to embrace death. Dodge achieves a stunning dramatic effect in this climax with layering and by adding noise to the synthetic voice. The composer obtains the timbre associated with a whisper coupled with more volume than a natural whisper could carry.

The use of the personal pronoun "you" initially generates some ambiguity as to whether the narrator is speaking about himself, speaking in general terms, or addressing a second person. It quickly becomes clear that the former case is intended as the text becomes more detailed with personal observations. This choice of narration underlines the dissociation that permeates the subject's emotional state. Dodge honors this quality effectively with his settings of "you," and has stated that those settings represent maximum pitch activity with a minimum of text. "You" is the only word set for the first four measures of the piece and it is articulated in a variety of ways: with widely displaced sixteenth-notes in measures one and two, and by pentachords at the end of each of those measures. Measures 15 through 25 also set "you" and are an expansion of the opening two measures. This setting renders "you" a timbral object and diminishes its personal connotations.

4. Conclusion

While both Charles Dodge and Paul Lansky have composed a substantial number of speech–based computer music pieces, their stylistic approaches and goals are quite distinct. The most fundamental difference is in the sound world that the composers



Above are score excerpts of *In Celebration* by Charles Dodge. These excerpts illustrate Dodge's settings of "you" with maximum pitch activity utilizing a minimum of text. Dodge's use of pentachords is also present. Figure 3 is a spectral analysis of measure 8, and shows 6.3 seconds. The spectral analysis in figure 4 is of measure 27 and shows 4.1 seconds. The spectral analyses of the two cadences illustrate opposite extremes of each other. The Dodge analyses also show a heavier concentration in the lower end of the spectrum which is attributable to the synthesis technique employed by the composer. The spectral analyses of *Word Color* display a more tonal orientation and a progression toward a richer spectrum.

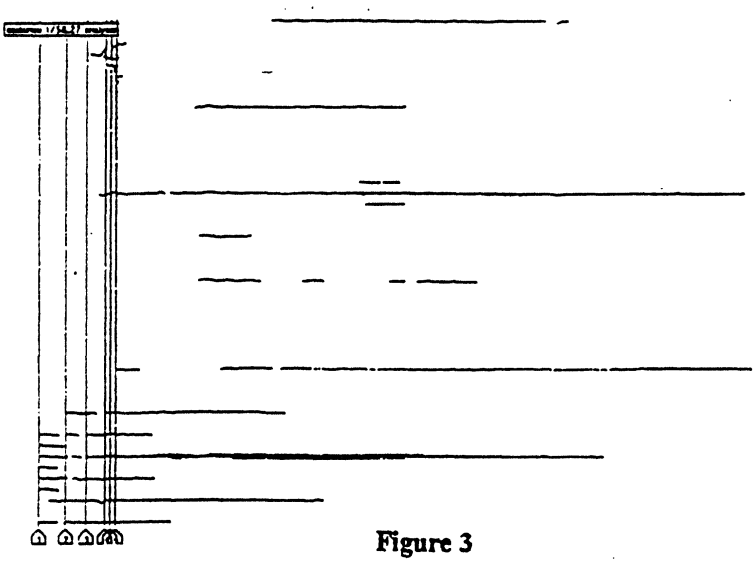


Figure 3

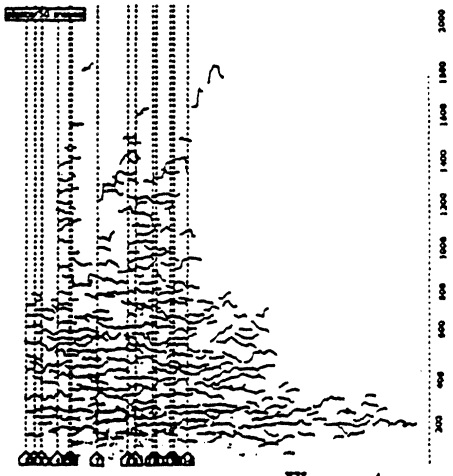


Figure 4

make for their pieces. Paul Lansky endeavors to enhance "real-world" sounds, while Charles Dodge creates an "otherworldly" soundscape in his pieces. The otherworldly quality sometimes attributed to Dodge's speech-synthesis compositions is due, in part, to the fact that the composer will often create a synthetic voice which carries many human characteristics, but can sing and speak in a manner completely impossible for a human voice to perform. This builds a fascinating dichotomy into his sound world. Paul Lansky, however, hears beauty and musical possibilities in everyday sounds. He is interested in writing music that will enhance the aural appeal and inherent musicality that he hears in these sounds in order to facilitate sharing his musical view of the world with an audience.

Acknowledgements and References

The author wishes to thank the staff of Columbia University's Computer Music Center where I have been a guest composer. Special thanks are in order to Brad Garton, Jonathan Lee, and Doug Geers for their invaluable assistance in writing this paper.

Lansky, Paul. Personal Interview conducted by Madelyn Byrne, Princeton, July 1998.

Lansky, Paul. *More Than Idle Chatter*, Bridge Records, Inc., 1994, compact disc liner notes.

Dodge, Charles. "*In Celebration.*" In *Composers and the Computer.* Edited by Curtis Roads. California: William Kaufman, Inc., 1985. 47-74.