# EXPLORING LANGUAGE

ALCIDES LANZA
Electronic Music Studio
McGill University
Montreal,PQ H3A 1E3
Canada

ABSTRACT

This paper refers to the use of digital synthesis in the analysis of voiced words and its applications in music composition as generative elements for phrasing and form. Specifically, refers to the creation of the tape part and the compositional aspects of sensors IV (1983-V), for choir and computer tape. The methodology of work included sampling, spectral analysis, digital synthesis and re-synthesis. The Electronic Music Studio at McGill University is equipped with a 32 voices Synclavier II Digital Synthesizer with an Able 60 computer having a core memory of 64Kw. The hardware for the Signal File Manager (SFM) used with the Sample-to-Disk System consists of 16 bit ADC,DAC.

## 1. MODUS OPERANDI

The digital aspects of sensors IV were realized on a Synclavier II Digital Synthesizer, with particular emphasis in the utilization of the Sample-to-Disk and Script systems. The basic modus operandi consisted in the utilization of the Sample-to-Disk system for the conversion of an analog vocal signal into a series of digital samples by 16-bit data converters. Once stored on the Winchester disk the samples were available for display, manipulation, analysis or immediate conversion back into analog form. This was followed by fragmentation of the digitally recorded vocal sounds as explained below. This permitted the creation of the "clipped consonants" section (see rehearsal (E) in the score) both for the tape sounds and for the choir.

The other principal operational mode, more directly related to the analysis of the digital samples, was the creation of totally digital sounds that would have a "vocal quality". Once the general characteristics of the vocal sounds were determined by Fourier analysis of the steady state of the original signal (McNabb, 1981) they were re-synthetized using additive synthesis.

Some of the aspects studied were:

1- the energy content (amplitude variations) of the complete sampled material. In the case of the original sampled word "me-mo-ry" the general contour observed in the print out of these amplitude variations was related to the final form... the composition. (See fig. 1 and 5)

2- the energy content of fragmented material, subjected to arbitrary or organized divisions (syllabic, phonemic and microsegmentation, including the creation of new envelopes).

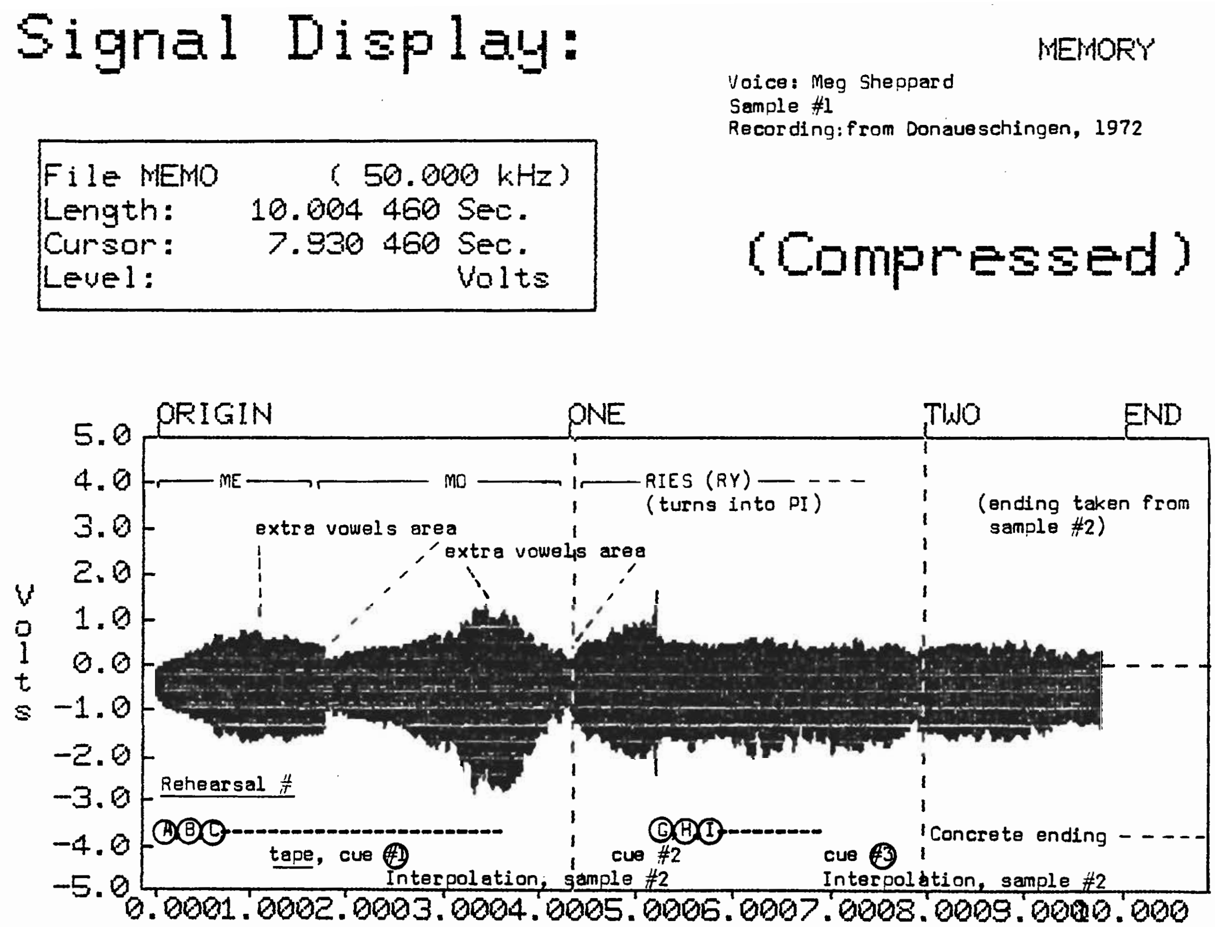
3- re-organization of fragmented elements, via re-synthesis. In this case it was noted that, when using SFM (Signal File Manager, the software used to operate the Sample-to-Disk System) some limitations were encountered. At the time, our system had only monophonic possibilities for SFM sound files. Consequently the Overlay Output mode was used in conjunction with a multitrack tape recorder, extracting a monophonic line of notes on each pass, but doing this with substituted timbres taken from the Timbre Bank. The sequence stored on memory would have the notes just recorded but not the timbre. Correct timbres were obtained by using the SMT (Select Memory Timbre) and proper Track buttons.

### 1.1 Analysis

The spectral analysis of some of the materials described above was based on the Fourier theory. This approach allows for the analysis of a complex waveform in terms of a sum of sinusoidal waveforms, having different frequencies and amplitude levels varying in time. The SFM spectral analysis package computes the frequency components of a sample of a waveform in terms of spectral density (Synclavier II SFM manual, 1982). It was found that the Spectral Display could provide relevant information concerning the frequencies and spectral density for the fundamental and each harmonic present.

By moving the cursor, and at times the vertical and frequency scales, it was possible to look at different aspects and areas of the spectral displays.

The spectral information was derived from a Fourier Transform of a fragment of 200 msec. It was found (also McNabb, 1981) that it was possible to obtain a synthesized "vocal" sound that would blend well with the live sounds (in this case, the choir) by using this partial data obtained from two-dimensional analysis, a considerable reduction from a complete three-dimensional analysis -(time, frequency, amplitude)- such as was done with the "phase" vocoder (Portnoff, 1976). The very important voice qualities that are associated with vowel sounds are determined by the formant frequencies (for example the vowel ë, as in "eat", has formants in the areas of 300 and 2300 Hz.) (Backus, 1977). The simultaneous running of several spectra, giving a three-dimensional representation of the changes of spectral density in time for some frequency components, was attempted but it was not used in connection with sensors IV. It was time consuming and it did not provide any extra relevant information.




Sampled word "memories" (sample #1) showing correlation with rehearsal letters, tape cues, and indication for interpolated areas. Diagram shows where the extra vowels area and consonant transfer are located.

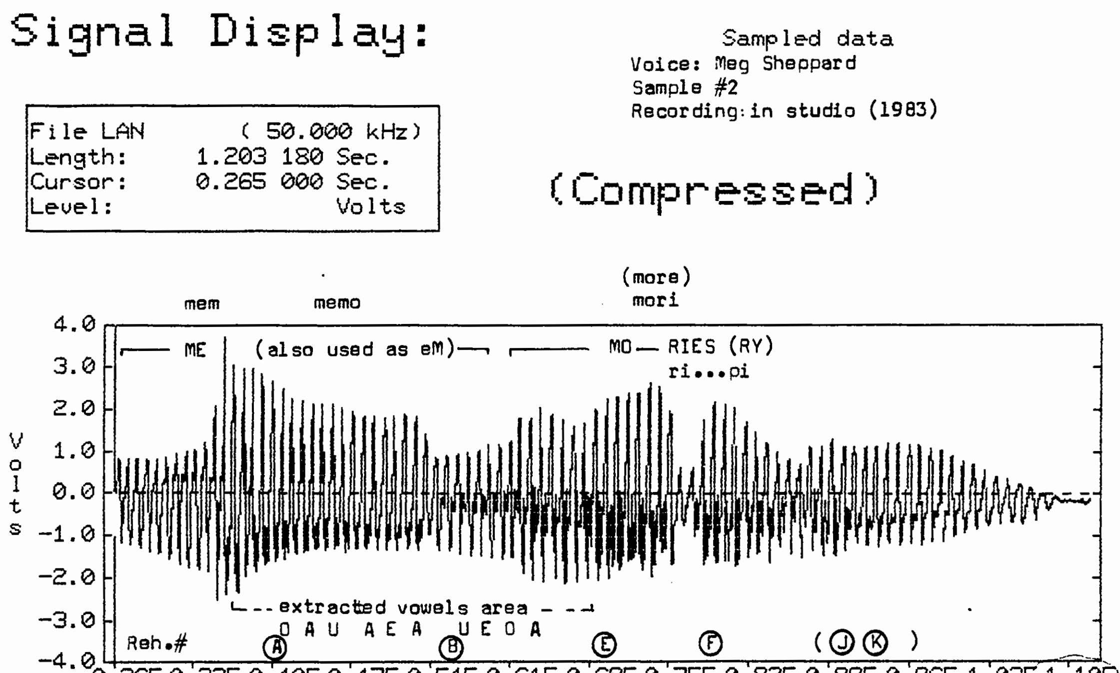
Fig. 1

## 2. COMPOSITIONAL APPLICATIONS

The compositional applications were many, but due to the scope of this paper, only a few will be considered. The question of having or not having a text, (if having one, the choice of text in one or more languages or even in an invented one), has been a rather important one for this composer for a number of years. I can say that some of the limitations inherent to a sampling methodology as it existed in 1983(50 kHz sampling rate, use of anti-aliasing filters, maximum length of sampled material that could be stored on hard disk, or the monophonic restrictions discussed above) have directly influenced the compositional results. In this case the "working text" was finally reduced to one word, "memory". Phonemic units derived from this word constituted the "text" for the choir. Characteristics found on the sampled and fragmented sounds were "applied" to the writing for the voices. The form of the piece was also extracted from the signal display of the complete envelope, with interpolated fragments of extracted samples data.

### 2.1 Creation of the Choir part

A good deal of the music written for the choir follows the idea of sharing or transferring information from singer to singer. For more examples of works by the same composer dealing with similar approach, see sensors I (1976-I), for 4 percussion players; sensors II (1980-I), for multiple trombones, sensors III (1982-II), for organ and two percussion players (Lanza, 1982, 1983). A good example of this "transference" is the cluster treatment; how to have control over microtonal differences in intonation and how to be able to move cluster-chords in massive glissandi. It has been suggested that the area of non-harmonic frequency relationships may nevertheless contain a system of hierarchical structural functions (Dashow, 1980). In sensors IV some non-harmonic chordal formations were created digitally first, to be included in the tape part. This was accomplished by controlling the FM index when working in the Script system, and at times it was done by using speed variation on analog tape. A parallel was established during the writing of the vocal part as follows: a hierarchical function was assigned to each principal voice in the vocal family group, and these singers were given an axial note, in the center of the intended cluster group. The other singers, according to the score, first were to tune in with the same pitch and then, gradually, depart from it in an ascending or descending motion. The intervallic departure from the axial note was controlled with a system of numbers: for example, singers 2,4,6, will move down, the size of the micro-interval depending on the number given to each singer; similarly, singers 3,5,7, will move up.




Sample #2, showing in more detail the extracted vowels area, reversed envelopes (ME (used as eM) and another transfer of consonants from Ri to Pi. Formal ending for the piece.
Fig. 2

A number of treatments were derived from electronic music practices; filtering situations (rehearsal (C) ), looping techniques (rehearsal (F) ), and moments of random frequency and amplitude modulation (the "mumbling" technique required and the "modulated glissando", rehearsal area (A) to (B) ) were integrated in the vocal writing.

The choral music between rehearsals (E) and (F) introduces consonants for the first time: the field has moved from only vowels with almost no consonant content (except for the enunciations UAm and UEm!, examples of envelope transfer and reversal) to "exaggerated consonants". These appear first on the tape part and were produced using the sampling method, as already explained under 1. and 1.1. Two recordings of the word memory were used, #1 was extracted from penetrations VII (Lanza, 1973) as sung by Meg Sheppard for whom the piece was written. On the other hand, #2 was a simple recording of the pronounced word. The first, half sung according to instructions, lasted about 10", the other, spoken, slightly over one second. The segmentation procedure was done on both sampled words, normally by "labelling" smaller fragments, with little or no help at all from the addition of tapered volume segments ( in milliseconds) corresponding to the attack and decay portions. It was observed that in all cases phonemic and syllabic correspondence was maintained but the semantic level was altered. For example, the words "memo", "mory" (sounding "mori", to die, in spanish), "mor" (sounds as "more", in english), were found to exist within the sampled one. When the phonemic level was exceeded by microsegmentation, fragments of vowels and consonants were isolated, analyzed and played upon. The division of the fragment "mem" into "m(e)"-"m" with little vowel content -does not have a direct correspondence with "(e)m". Even more than that, if we compare "m(e)" with "e(m)" played backwards we will see that it does not sound the same; if we join "m(e)" with itself played backwards it will sound completely different if compared with the original fragment "mem". This multiplication of the Extract Procedure by a trapezoidal window was used extensively, generating a number of "clipped consonants" which were used in section (E). One interesting observation was the "transformed consonant" phenomena: when "mory" was reduced to "ry", any subsequent reduction - by the trapezoidal window method - of the tapered volume segments, mainly the attack portion, takes the "r" sound to approximate a "p" sound. In short, a liquid consonant -"r"- has changed into a plosive "p" (Denes, Pinson, 1973). When the final result is heard the analog that comes to mind is string pizzicati, and is actually rather difficult to think of vocal sounds having been the origin. The notation for the choir required special performing techniques by which those "clipped consonants" could be produced live.

Good approximations to different vowel sounds were found to exist in the area between (m)e(m), that is, the area from 350 ms to 615 ms. This corresponds to the moment when "e" is fully established and the apparition of "m" leads to "mo". It was observed that the closing of the mouth and changes in tongue positions from "e" (half open) to the closed one needed for "m", produced a variety of different vowel colorations, quite evident when the samples were transposed down. The beginning of sensors IV was directly derived from that phenomena.

## 2.2 Creation of digital sound complexes

This corresponds to the tape part from the start to (E); then from cue (2) (just before (G) ), all the multiple glissandi area, and ending at (I). The procedures depicted at 1. and 1.1 were used, and the sounds were then created by additive synthesis. Timbres were created following the Define procedures, and the actual music performed on the Synclavier II keyboard and stored as sequences. Reverse compilation possibilitated the transfer from one format to the other. The initial cluster build up was realized with the idea that it should gradually blend with the cluster being sung by the choir, the tape part to be heard clearly only during the windows (𝄐), according to the score. Similar timbrical situations were used for the multiple glissandi area, the glissandi realized in most cases via gradual changes in the FM settings, bringing the multiple differentiated lines to slowly converge on the final D3, after extra pitches have ended, and after other Ds (D5,D4,D2) have also ended.

```
3260 Define Partial Timbre6.Partial3
3270    Volenv            0     159    4461      41
3280    FMenv             0      17      18     126
3290    Harmonics     100.0    50.3    59.5     5.5
3300                   10.4     3.0     1.9     0.8
3310                    0.3
3320    Vibrato        Tri2    5.56    0.11     281
3330    FMRatio       4.000
3340    Stereo          -50       0     0.0       0
3350    Decay         0.410
3360    Kbdenv            0       1      61       0
3370 End Timbre6.Partial3
3380
3390 Define Timbre Timbre6
3400    Partial Timbre6.Partial1
3410    Partial Timbre6.Partial2
3420    Partial Timbre6.Partial3
3430    Numvoices  32
3440    Rte  P1 P2 P3 P4  Prate
3450 End Timbre6
3460
3470 Notelist  Timbre1
3480    6.525  B5  18.160  (99.6)
3490    6.545  G4  23.040  (99.6)
3500    6.560  D#5 22.880  (99.1)
3510    9.665  D3  32.000  (99.6)
3520   21.625  B1  20.320  (99.6)
3530   33.725  D1  17.760  (99.6)
3540 End
3550
3560 Notelist  Timbre2
3570   50.940  D3  40.800  (99.6)
3580   50.970  D4  24.320  (99.6)
3590 End
3600
3610 Notelist  Timbre3
3620    1.310  D3  23.360  (99.6)
3630 End
3640
3650 Notelist  Timbre4
3660   58.845  D5  23.040  (99.6)
3670   66.330  D2  22.720  (99.6)
3680   69.575  D3  33.280  (99.1)
3690 End
3700
3710 Notelist  Timbre5
3720   55.010  D5  18.800  (99.6)
3730   55.010  D3  40.800  (99.6)
3740   55.075  D4  36.000  (99.6)
3750   89.970  C3   6.560  (99.6)
3760 End
3770
3780 Notelist  Timbre6
3790   80.910  D3  40.800  (99.1)
3800 End
```

**Notelist indicating the convergence on the note D3. This moment is three bars before Ⓘ, leading to the improvisation on the unison.** Fig.3

2.3 Final aspects of the tape part

The digital sections on the tape having been described above, it has to be added that there was a concrete element reserved for the final tape part. Starting at tape cue ③ (just before Ⓚ) the tape has a direct quotation from penetrations VII (Lanza, 1973) (from which one of the sampled words was taken). This was a slowed down version of the word memories, sung in real time. sensors IV requires that each singer opens a "window" in his or her memories, and to randomly move from whispering rapidly to practically shouting. The tape part then, to be consistent with the motivation behind this composition, was again done with the voice of Meg Sheppard. The singer had to make four one minute long recordings of whatever came to her memory. The audition of the tape can better exemplify the materials so encountered. Some electronic treatments were applied afterwords, mainly frequency shifting, ring modulation, phasing and mixing, adding the digital "vocal" sounds again near the end.

See Fig. 4 on next page

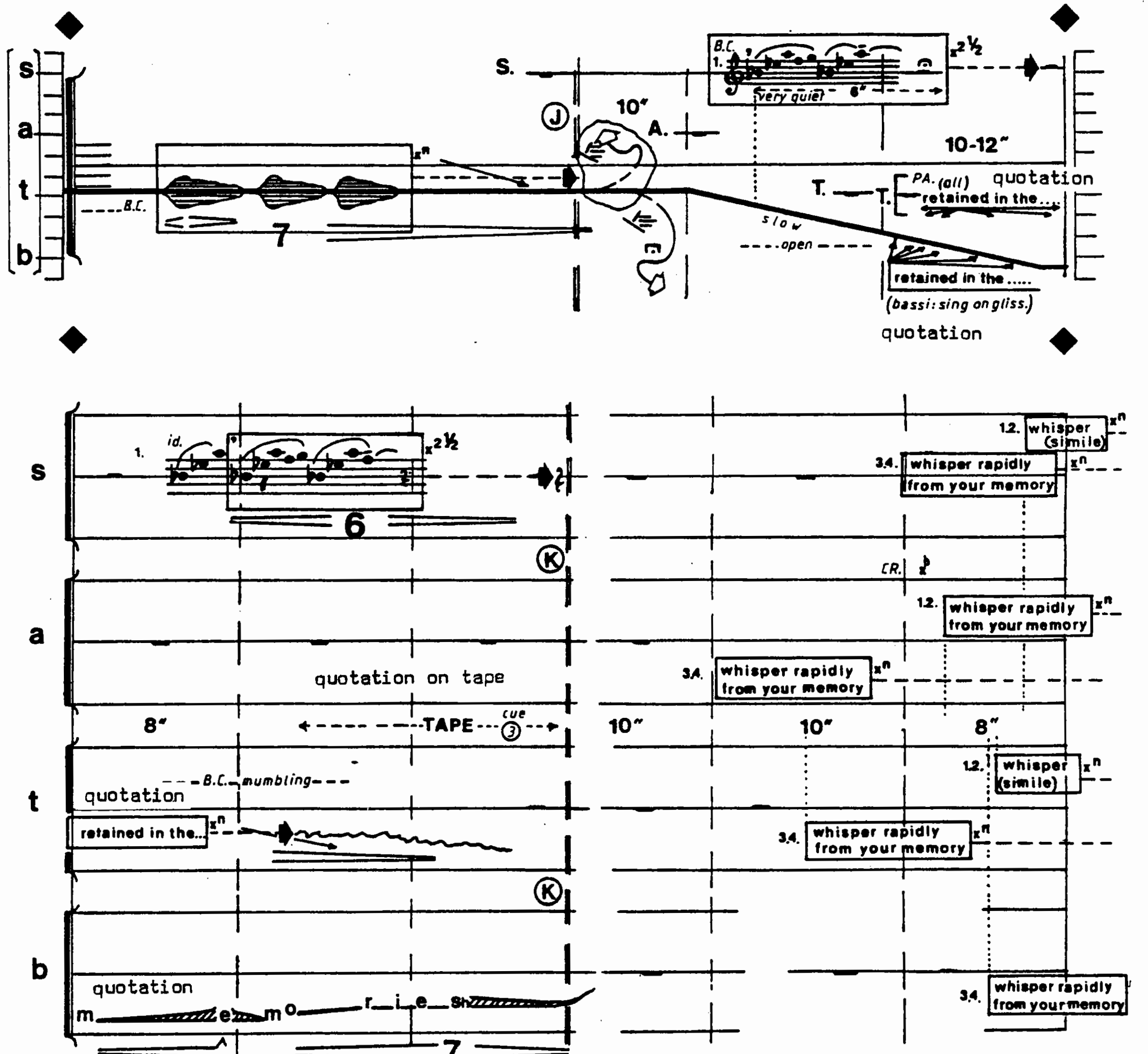
2.4 Form

The form of the piece was related to the Signal Display diagrams mentioned, which were freely interpreted for this task. The general motion intended was to go from continuous to discontinuous materials, shown in Fig. 5. At the same time the text had to go from vowels, with almost no consonant content, to only consonants (with almost no vowel or other consonant content). Another parallel process observed was the changes from a first syllabic level (3 after Ⓔ), leading into words with two syllables, three, encountering "memory"s in different languages and several other words with different semantic content. Ⓖ sees a return to "Em" (somewhat similar to the beginning of the piece) and the multiple glissandi in "mo-ria" (it was dying, in spanish). At Ⓙ the text says "retained in the..." ( from penetrations VII (Lanza, 1973), and after Ⓚ the text is improvised by each singer -as described in 2.3, ending the composition with "UAm!", as it had started.

Phonemes are the basic linguistic units from which words and sentences are put together. On their own, phonemes do not carry any conceptual meaning (Denes, Pinson, 1973).

Semantics, on the other hand, is the study of word meanings, so the linear motion from phonemic level (no semantic content) to word and sentence level (high semantic content) is most relevant to sensors IV.

See Fig. 5 on last page




by permission SHELAN Publications, Montreal, Canada

sensors IV, page 9: all quotations indicated were taken from penetrations VII (SHELAN). Score shows beginning of last climax (from whispering to shouting), and special performing techniques, clusters (1 before Ⓙ), from B.C. (boca chiusa, closed mouth) into mumbling (tenors), slowed down "memories" (basses).
Fig. 4

## 3. RELATIONSHIP WITH EARLIER WORKS

There are constants in my vocal production which should be mentioned here. Actually, a listing of earlier vocal works with brief additional comments, will be very important for the understanding of the motivation behind the writing of sensors IV, the use of the Synclavier II system, and this paper.
Text authorship: in almost all my vocal production, certainly for the compositions created from 1963, I have written my own texts;
Memory: the concept of, or just the word memory, is a part of each of the compositions listed. In the case of ekphonesis V (1979-I) and sensors IV (1983-V) "memory" is the nucleus behind each piece.

three songs (1963-IV), for voice and chamber ensemble (english text)

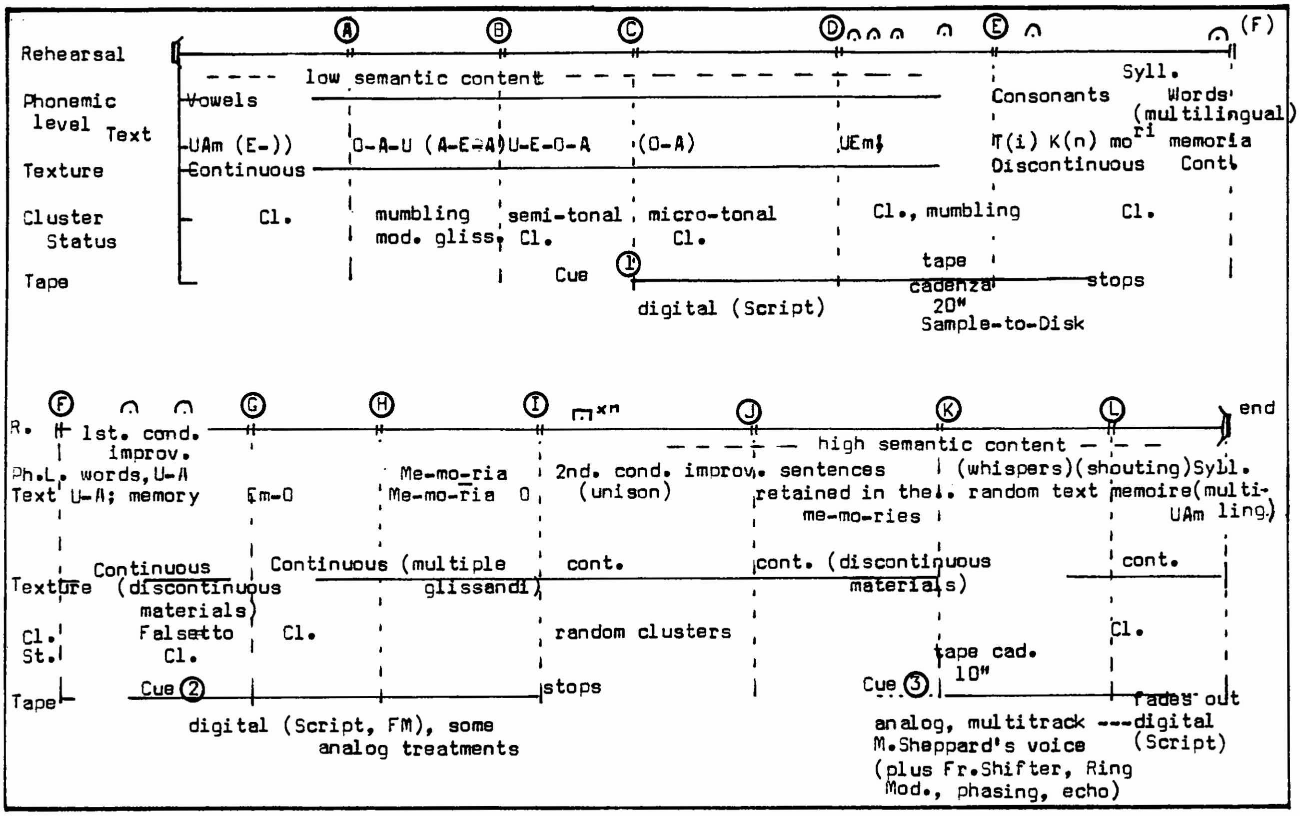
ekphonesis II (1968-III), voice and pianist (english, italian, spanish, french, german, invented utterances) with analog tape

penetrations VII (1972-III), voice, lights, electronic music and electronic extensions (english, invented language)

kron'ikelz, 75 (1975-I), for chamber ensemble, with voices, electronic sounds and electronic extensions (invented language)

ekphonesis V (1979-I), for voice, lights, electronic sounds and electronic extensions (mostly spanish and english) (the piece explores the "library" of memories supposedly encountered if the soloist could enter the brain of the composer)

módulos II (1982-I), for guitar, with optional vocal part and electronic sounds (no text, just vowels and consonants; a modulated vocalise)




sensors IV: linear structural analysis showing the correlation among tape and choral music, text and texture, and microtonal or semitonal clusters.

Fig. 5

It must be mentioned how important for the composer was the analysis of works for voice: Schumann's Dichterliebe, Schoenberg's Pierrot Lunaire, Berio's Omaggio a Joyce, Stockhausen's Gesang der Jünglinge. Also relevant was the study of Concret Poetry and Text-Sound compositions, from the Audiopoems by Henri Chopin, to the recent works by Lars-Gunnar Bodin.

In closing, some additional comments on memory. The fact that memory became such an important building block for sensors IV is logical; after all, it is just the realization that music and languages could not exist unless memory is exercised and utilized. Could we conceive of the Sonata Form, or even a Lied Form, unless memory is exercised? It is also the realization that there are many types of memories, and that information storaging devices, such as long playing records, tape recorders, books, paintings, digital recorders, floppy disks, hard disks, etc., are all again....memories.

sensors IV was commissioned by Christopher Reynolds for the McGill Concert Choir, with a grant from the Canada Council for the Arts.

References

1. McNabb, Michael (1981) Computer Music Journal. 5(4):36-53
2. Synclavier II, SFM Manual (1982). A Musician's Guide to the Sample-to-Disk. New England Digital. Vermont.
3. Portnoff. Implementation of the digital phase Vocoder using the FFT. IEEE proceedings on Acoustics, Speech and Signal Processing. 24:243-248.
4. Backus, John (1977) The Acoustical Foundations of Music. 254-255. New York: Norton.
5. Lanza, Alcides (1982,1983). Editions Shelan Publications. Montreal, Shelan.
6. Dashow, James (1980) , Computer Music Journal. 4(1):43-52.
7. Denes, Peter and Pinson, Elliot (1973). The Speech Chain. 74. New York: Anchor Books.
8. Lanza, Alcides (1982) Penetrations VII (1972-III) 3. Montreal: Shelan.