# Locus: *rien n'aura eu lieu que le lieu*

Claudy Malherbe

The musical quality of natural languages has long been known. Today, sound analysis tools show us that they contain not only melody but also harmony, and accurately reveal at the same time their inner rhythmic and dynamic qualities. More precisely, this technology, which gives a coherent and controllable representation of a complex sound reality can unite what was formerly distinct. Indeed, with these new possibilities, handling sound material is not irreparably separated from the symbolic representations. On the contrary, the manipulation of musical signs no longer omits concrete approaches to sounds.

*Locus*, for real and virtual voices, was commissioned by IRCAM and written in 1997. In this piece, speech is turned into music by means of its own contents, once this content has been reconfigured in an abstract space. This chapter will detail some aspects of this piece, emphasizing the use of CAC in its composition. The piece was initially created in PatchWork, but many functions used for the composition were ported to OpenMusic.

## The text

The text of *Locus* is an original montage composed of short fragments (in English) taken from various technical works about vocal acoustics, phonetics and phonology. These fragments provide both vocal material and comments that describe the elements and situations concerned (vocal tract, larynx, throat, tongue, lips, breathing, speaking, eating, etc.) Within this context, the choice of English is not unintentional; it takes into account the performers' native language. Perfectly mastered, it enables the precise introduction of the phonetic material, even in situations where it is highly deconstructed. English is also a more accentuated and melodic language than French (another possible alternative), which is of particular interest since it constitutes the essential raw material of the piece.

The text follows an evolution parallel to that of the vocal material and its avatars: speech, song, noise, sinus sounds and reconstructed speech. The combined movements and relations between text and voice form the core contents of the piece. These combinations move from one state to another by mental or material associations: these multiple states are revealed, hidden, embedded or disjoined, by playing around with the physical, semantic, and emotional aspects of text and sound.

At the beginning, for example, the description of the vocal apparatus produces throat and mouth noises generated by the phonemes contained within the description; the meaningless words and phrases constructed with this material provoke laughter, and their jerks produce rhythms. Later on, the laughter is transformed into a wail that lengthens the sound and introduces melody in the piece. Elsewhere, the statement that *eating and speaking employ the same organs* gives rise to a "meal of words", where syllables are cut and grinded, and phonemes chewed up and swallowed.

## Material, form, movement

Five distinct states of the material characterise the development of *Locus*. The first three states can be described as natural: speech, song,[1] and vocal noises (coughs, laughs, tears, etc.); a fourth state presents artificial (more or less distorted) voice reconstructions; while the last state is a synthetic representation of the voice, revealing constitutive parts that are usually hidden. Figure 1 illustrates the constitution of the corresponding structure. (a) The speech (*parole*) is the more common form of vocal flow, and is the starting point of this structure. (b) Song (*chant*) and vocal noises (*bruit*) are then appended, as complementary categories of the human vocal emissions. (c) The fourth category is provided by the computer tools that analyse and reconstruct a spoken flow into a virtual speech (*parole virtuelle*) made of elements from the real speech. (d) Finally, the time/frequency analyses of the vocal material reveal and render some of its constituent and usually hidden elements. The artificial character of this reconstitution, here named *sinus*, places it in another plane, thus introducing a third dimension and creating a tetrahedron which will determine the whole development of the piece.

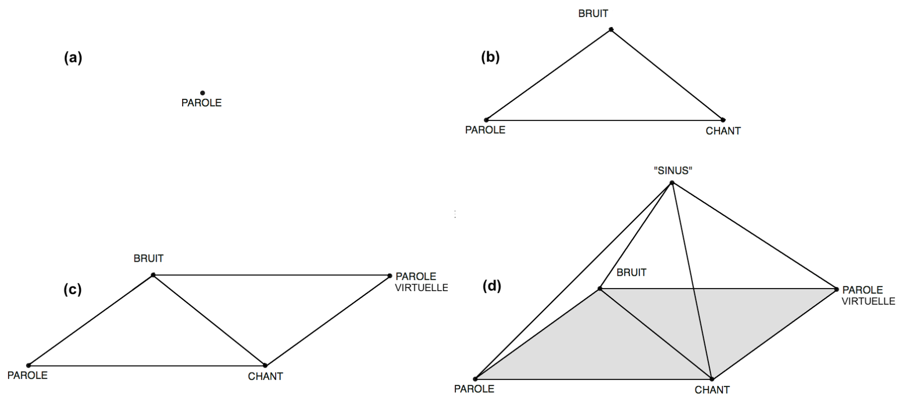


**Figure 1.** Construction of the five-state development of the vocal material.

This geometrical figure defines the different states and processes of the material as the form of the piece: the space between two particular states (the vertices of the tetrahedron) delimits a field (see Figure 2) and an organisation which determines the setting of the syntactic and musical materials (see Figure 3). This organisation is divided into intermediate states (three on each line) which govern the sequencing and transitions (Figure 4).

Each line of the figure is thus composed of five states, two at the vertices and three intermediates (which are of a sufficient density to cover the different vocal possibilities). The padding of this geometrical figure underlines original configurations to be imagined between canonical vocal situations. It also helps avoid any mechanical activity which would reproduce ready-made schemes.

[1] In this context, song becomes a commonplace category, and it is not as dominant as it used to be in western vocal music.

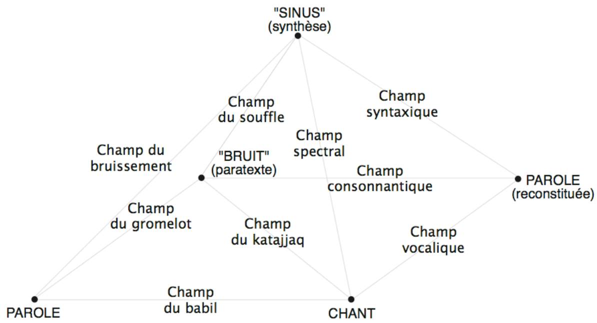


**Figure 2.** . The different fields (*champs*) between the tetrahedron vertices.

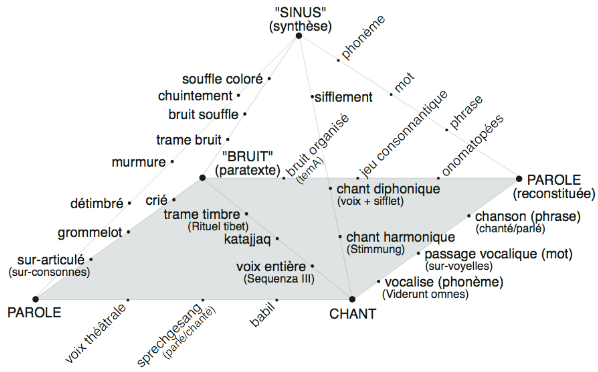


**Figure 3.** The organisations corresponding to each different field.

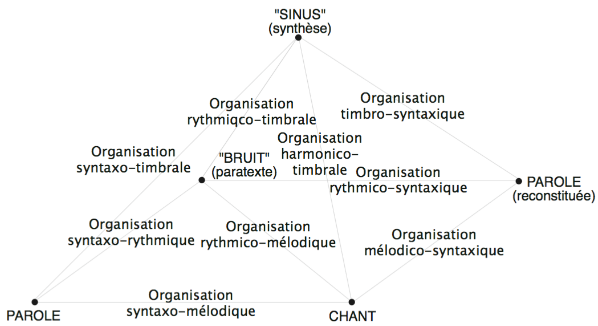


**Figure 4.** Intermediate states.

For example, the *babil* (babble) field located between the speech (*parole*) and the song (*chant*) induces a syntactic-melodic organisation (a mixture of spoken language and sung melody) characterised by three intermediate states that enable a successive transition from speech to song: from theatrical voice (speech with accentuated intonations), *sprechgesang* (sung-spoken), to *babil* (child's speech without precise meaning, close to the song). In a similar way, the *grommelot* field (from grumbling, i.e. a little-articulated speech partially indistinct) is set up between speech (*parole*) and noise (*bruit*). This field is built on a syntactic-rhythmic organisation: whilst the song recalled melody, here the noise recalls rhythm.

Finally, the path which constitutes the form of the piece is derived from the same figure (see Figure 5). It is determined according to the following directive: once an entry point is chosen on a given vertex, it must pass through all the lines of the tetrahedron only once without jumping from one vertex to another discontinuously (one vertex can however be visited several times). With this figure, which contains four vertices connected to an even number of lines, there exist only two possible entry points that respect this rule: real speech and virtual speech.

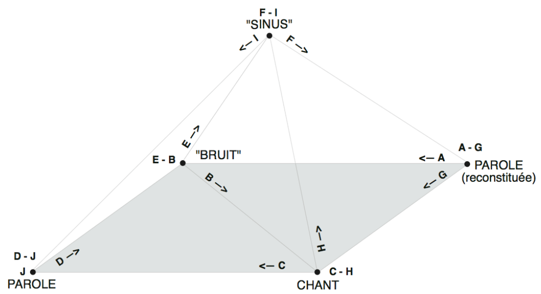


**Figure 5.** The path through the lines of the figure determines the form of the piece. Letters A to J indicate one possibility to go continuously through each of the lines without repeating any of them.

The formal construction is built on a database corresponding precisely to every state of the form. Although it is prior to the composition, it is nevertheless organised as the piece: a precise setting of the parameters of the sequences (durations, tempi, harmonic initializations, etc.) brings it as close as possible to what is going to be written. Here is an example of a set of instructions that defines the content of sequence 15, where speech evolves towards noise:

```
Sequence 15 (database)

A - Speak very quickly the following texts:
    "Intonation expresses emotions and attitudes when we speak."
    "Words do not change their meaning but the tune we use adds the speaker's feeling."
    "Intonation helps to produce the effect of prominence on syllables."
    "Intonation is used to convey our feelings and attitudes."
    1) Normal
    2) With the following characteristics:
          Soprano: ecstatic; dreamy; tender; languorous; joyful.
          Mezzo: relieved; apprehensive; coy; anxious; whimpering.
          Tenor: urgent; tense; nervous; frantic; gasping.
          Bass: impassive; distant; calm; noble; wistful.

B - Alphabets
  Refer to the lists and groups of letters in the alphabet tables in the following tempi:
      - 40, quarter notes
      - 140, quarter notes, sixteenth notes, eightieth notes.
    1) Normal pronunciation, articulated.
    2) Phonemes (noises), articulated.

C - Phonemes
  Same as B, using the phoneme tables.
      - Tempo 40, quarter notes.
      - Tempo 140, quarter notes.
```

Thus, the progression of the piece, composed of a succession of such sequences, goes through various styles and borrows from varied genres, which appear in the background or are used as references. Concerning the text, I must mention the influence of Mallarmé who talked about "*laisser l'initiative aux mots*",[2] or that of Georges Perec's lists and constraint systems. The form of *Locus* is indebted to Raymond Roussel's *machineries* which constructed an entire novel from a simple play on word. The music of the piece sometimes evokes the Inuits' proto-melodies, and is, at other times, inspired by Guillaume de Machaut's evolving harmony. The organised noise in Helmut Lachenmann's style is a point of transition and the liberated voice of Luciano Berio a reference. Finally, and on a more abstract level, the symbolic organisation of the material in Karlheinz Stockhausen's style counterbalances the voice's natural presence.

## From speech to musical material

The use of ready-made material – here, the voice in its more common state (speech) – raises the question of its insertion into a heterogeneous musical network. In the following examples, I describe processes for extracting structures, elements and parameters from the vocal materials. These processes constitute the basis for the musical construction and therefore they also ensure that the material is coherently inserted into this construction.

### From singing voice

The first example starts with an excerpt from Luciano Berio's *Sequenza III*: a sung high-pitched D, which varies according to a sequence of vocalic colours.[3] As is notated on the score, the note emission starts with the mouth shut [+], and then evolves towards the vowel [ø], and then the vowel [e]. Figure 6 shows the sonogram analysis of this excerpt, displayed in AudioSculpt. As can be observed in this figure, the transitions gave rise to intermediate formants which complete the three ones written by the composer. The partial tracking analysis extracted the more significant partials from the previous analysis, thus providing a simplified representation (see Figure 7).

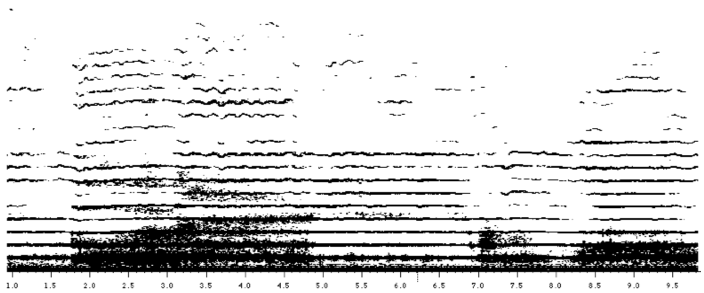


**Figure 6.** Sonogram analysis of a mezzo-spoprano voice singing a high-pitched D with the vocal progression [+ - ø - e] (from Berio's *Sequenza III*).

[2]The title of the present chapter: *rien n'aura eu lieu que le lieu*, is a quotation from Mallarmé's poem *Un coup de dés jamais n'abolira le hasard*.

[3]Measure 4, "Distant and dreamy". Performed by Cathy Berberian.

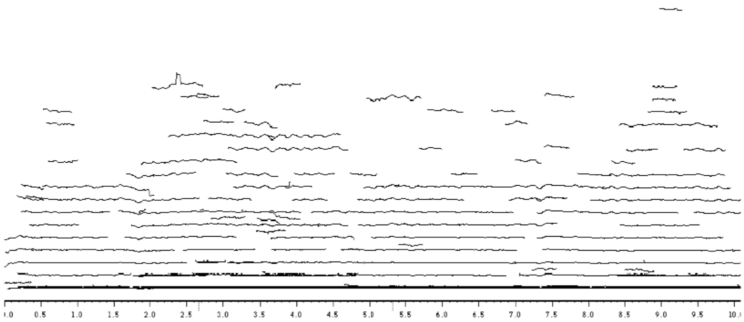

**Figure 7.** Partial tracking: extraction of the more significant partials from Figure 6.

The data from the previous analyses were then transferred to PatchWork or OpenMusic,[4] one more time reduced to their most significant elements (for a musician) and then represented symbolically as a score (see Figure 8). This data could then be used at the various levels of the composition. It could for instance be re-synthesized to produce some new musical material derived from the initial voice sound, or used to produce higher-level musical structures.

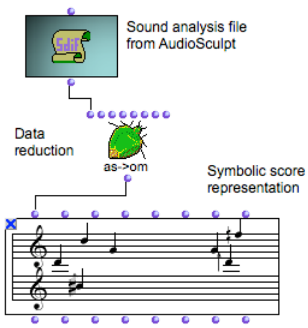


**Figure 8.** Conversion from sound analysis data to symbolic musical data in OpenMusic.

The partials were converted into sequences of chords. Some parameters in this conversion were of a special interest for my purpose. First, a maximum number of simultaneous notes in a chord was set in order to limit the amount of data (the $n$ more important partials are selected). Then, a *delta time* parameter determines an interval of time during which all the partials are collected in a same chord (i.e. if two notes are separated by a time interval lower than this threshold, they will be reduced to a single chord). By progressively increasing the value of this parameter (and with a fixed maximum number of notes), I obtained successive sequences containing fewer and fewer elements. I created four successive sequences following this principle. The *delta time* was increased so as to divide the number of chords by two each time. Figure 9 shows these four successive sequences.

[4]Originally in PatchWork, using text files and the *as->pw* function. The corresponding operations are available in OpenMusic through SDIF files and the *as->om* function.

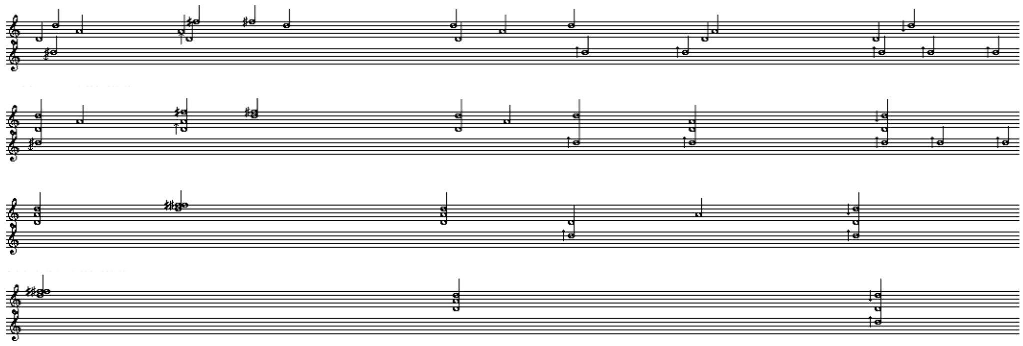

**Figure 9.** Reduction of the sequences to 22, 11, 6, and 3 events (chords).

This simple operation produces a definite musical effect. In the initial sequence, the 22 events are perfectly correlated with the initial sung fragment (no notes have been grouped). Then each reduction presents, each time more clearly, a harmonic autonomy which is nevertheless coherent with the initial object. The reduction of the harmonic density produces "delay" or "pedal" effects, similar to those of tonal harmonic music. The result of the combination of the voice with the last two reductions (to 6 and then to 3 elements) can be compared to an accompanied melody, with the difference that the accompaniment was here not elaborated from an abstract musical construction but directly derived from the spectral components of the melody itself.

## From the spoken flow

In this second example I took as a starting point a spoken phrase: "*When speech sounds are made, the larynx may or may not itself be vibrating to produce an oscillatory flow of air.*" Figure 10 shows the time/frequency analysis of this phrase (sonogram) in AudioSculpt. A syllabic segmentation of this AudioSculpt analysis was carried out using markers (Figure 11). Finally, the mean value of the partial between each marker was computed, in order to make each syllable correspond to a harmonic aggregate (Figure 12).

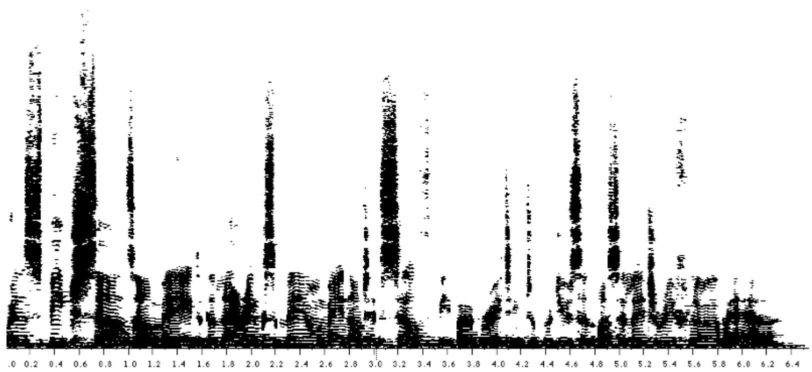


**Figure 10.** Sonogram analysis of a recording of a spoken phrase.

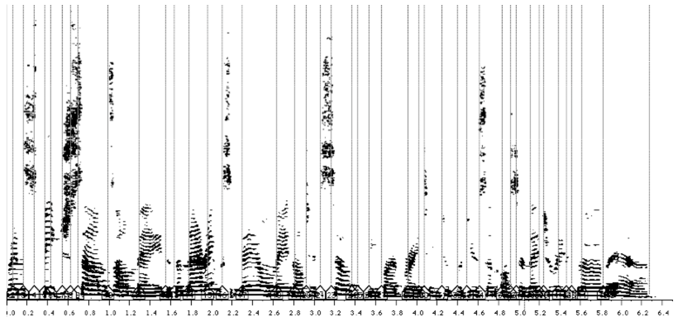

**Figure 11.** The spoken flow is divided into segments corresponding to the successive syllables by the AudioSculpt markers.

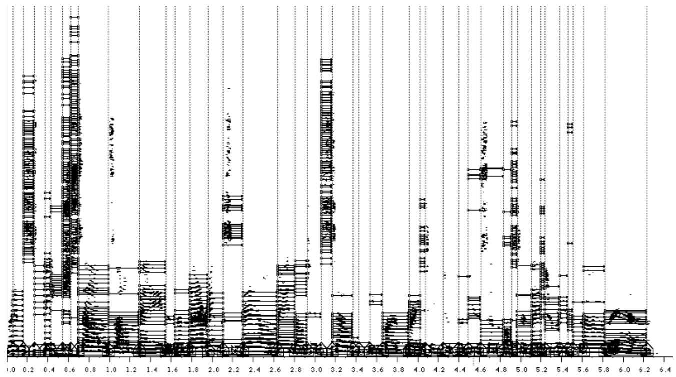

**Figure 12.** Averaging of the partial between the markers (usually called *chord-seq analysis*).

The resulting data was also ported to OpenMusic where it was converted into symbolic musical material. The same process of reduction, detailed in the previous section, was applied. Figure 13 shows the resulting sequences.

**Figure 13.** Reduction of the sequences from 44 to 23, 11, and to 6 elements during the conversion of the analysis data into chord sequences.

The first sequence is made of aggregates (chords) closely related to the initial spoken sequence. (The machine skipped some low-energy syllables and this explains the slightly lower number of elements compared to the number of syllables in the text.) The musical result is even more original in this case, since the spoken voice is harmonised with its own frequency contents.

# From concrete material to symbolic manipulations

## Spectral correlations between durations and pitches

From a scientific point of view, durations and pitches are the constitutive elements of a same vibrating field; durations are related to the domain of the low frequencies while pitches are linked to that of the high frequencies. The ear differentiates between these domains near to 16Hz, when the sensation of periodic durations is perceived as a pitch, or inversely when the continuous perception of a pitch becomes a fast repetition of durations.

Thus, it is interesting to start from this close correlation between these two musically distinct parameters. It restores to the act of score writing equality in the treatment of pitch and duration which is uncommon in traditional Western music. Hence, any spectral form (that is, abstract material referring to the acoustic spectrum model) can be applied either to the pitch domain or to the duration domain (once the initialisations have been determined: a fundamental frequency for the pitches, and a tempo for the durations).

## Concrete – abstract dialectic

The successive processes which build musical objects from raw materials are shown in Figure 14.

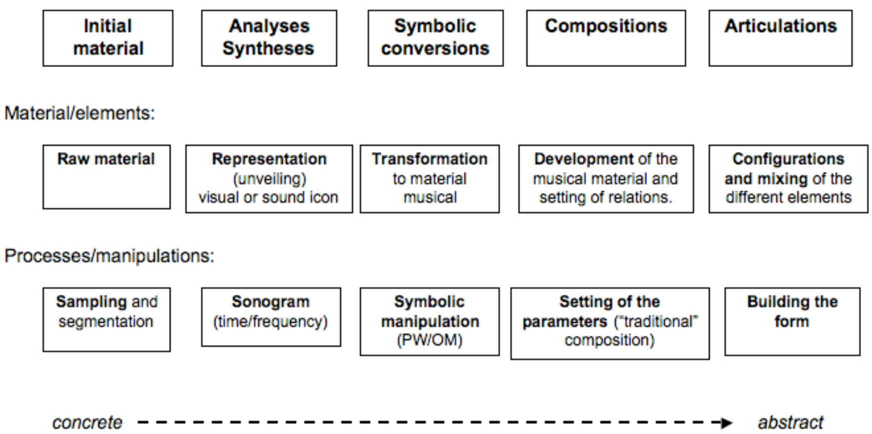


**Figure 14.** Table of the successive steps from the raw material to musical objects.

The main acoustic data (durations/pitches) obtained through sound analysis is structured step by step, eventually producing an abstract and symbolic image of the initial object. During the writing of the piece, these phases may come into play individually or simultaneously: writing brings together the different evolutions of a same material.

The corresponding phases of the transformations from raw material to musical objects can be, respectively for durations and pitches:

1) **Initial material**: raw durations and pitches coming from analyses, expressed in milliseconds and in Hertz, or *natural* durations and pitches;
2) **Analyses/Syntheses**: integration of the durations into a virtual tempo (quantification) and pitch identification within a spectrum (given a virtual fundamental frequency and a temperament), or *fine-tuned* pulses/pitches;
3) **Symbolic conversions**: forced tempo, quantification with a precise given pulsation, and integration of the pitches into a precise spectrum, or *controlled* pulses/ pitches;
4) **Compositions**: durations quantified in a precise tempo and inserted into a *regular meter*, and substitution of the controlled pitches by pitches from a harmonic reservoir, or *defined harmony*;
5) **Articulations**: durations substituted by regular pulses (*duration spectrum*) and harmony substituted by a limited chord set (*basic harmonic language*).

## Work on durations

The events in speech-acts or in vocal materials are in general very short and contain fast articulations, compared to those of the musical material. In order to use vocal materials for musical purposes, some operations are necessary to transform the original durations of these natural objects. In sequence 27 of *Locus*, the soprano voice used as the original material was stretched in order to coincide with the pre-determined duration of the sequence which contains it (these durations were previously defined in the general formal structure of the piece). The initial spoken phrase, of a duration of 5.8 seconds, was stretched to 42 seconds. Figure 15 shows the stretched segments associated with the different phonemes.

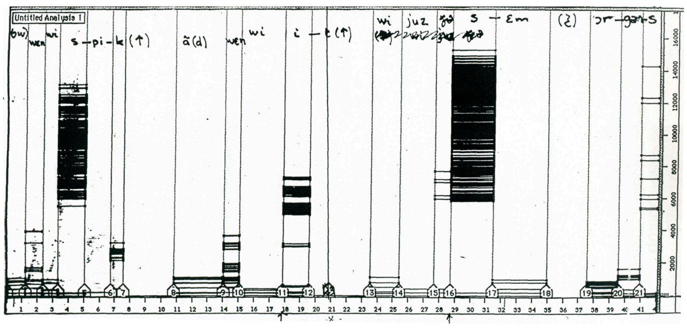

**Figure 15.** The phrase "*When we speak and when we eat we use the same organs*" stretched to 42 seconds.

The inner components (durations, intensity, harmony) of the stretched phrase were then used to structure the other materials (phrases, words, noises) in the same sequence. The initial phrase triggers the same phrase in normal speed, as well as the other synthetic elements, which blend with the initial phrase in a similar way as the upper voices and the *cantus firmus* in medieval music. Hence, these materials, which have their own sound profile, are homogeneously articulated with the help of the stretched speech. Figure 16 shows the score for the beginning of this sequence.

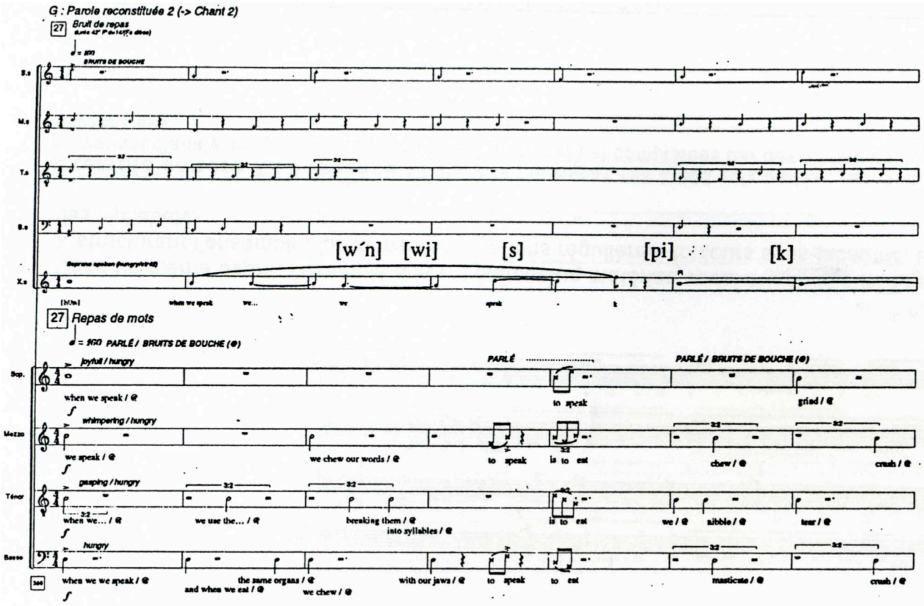


**Figure 16.** Extract from sequence 27, corresponding to the beginning of the phrase from Figure 15.

In many other parts of the piece the vocal material can be transformed and integrated as abstract rhythmic structures. These are derived from the raw durations quantified and organised using the quantification tools in PatchWork and OpenMusic. Figure 17 shows an example of such quantification with the sequences in Figure 9.

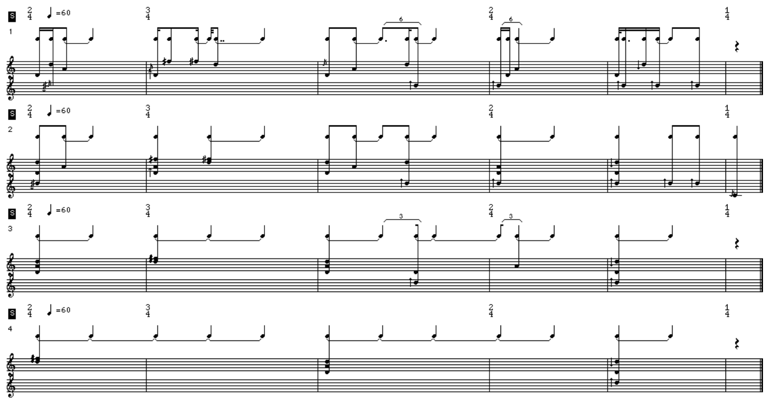

**Figure 17.** Quantification of the sequences in Figure 9: towards a transformation of the raw material into abstract musical structures.

## Work on pitches

Depending on the context, the pitch material can be used in its more basic state (an harmonic spectrum) or in more elaborated configurations (e.g. chordal harmony). Harmonic scales are extracted from the spectra by means of frequency modulations, and set operations on the resulting scales are used to deduce the chords that constitute the harmonic structures.

The composition and articulation operations are then carried out with some functions available in the *RepMus* library.[5] For instance, the *map-chords* function takes a list of chords as a model and then another list of chords as a harmonic reservoir, and selects chords in the reservoir in order to build up a new sequence, as similar as possible to that of the model (see Figure 18).[6] This way, sound analysis material can be converted into organized harmonic material.

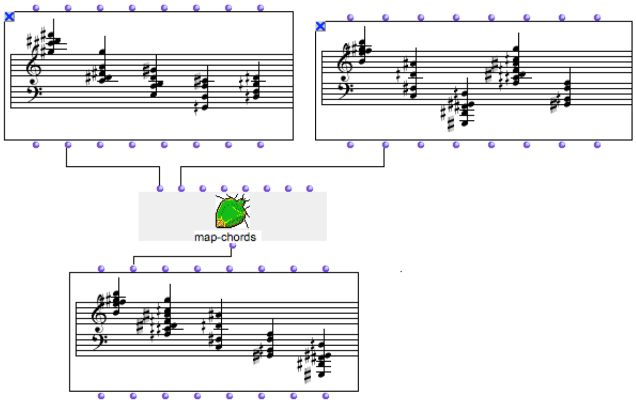


**Figure 18.** The *map-chords* functions build a sequence similar to the model (on the right) using the chords from the harmonic reservoir (on the left).

## Work on durations and pitches

The construction of durations and pitches starting from the initial speech and voice recordings brings forth the dialectic between sound signals and musical forms and structures. These processes are illustrated in sequences 37 and 38 where they are carried out at once. A same phrase is stretched according to two different durations (60 and 80 seconds), and some elements of this phrase (phonemes, syllables, words) are positioned following two different tempi (80 and 40). Meanwhile, the harmony extracted from the sound material is based on two successive fundamental frequencies (corresponding to C and G). The original non-musical elements are thus coherently brought together thanks to the duration and pitch parameters.

[5] Many functions of the *RepMus* library were developed while composing *Locus*.

[6] *Map-chords* uses an Euclidian distance measure between chords in the reservoir and chords in the model, taking into account the number of common notes, the range, the register (the centre of gravity of the chord), and the difference in the number of notes. (The user can give a weighting coefficient for any of these criteria.)

## Conclusion

Every ear, even that of a non-musician, can discern rhythm in a spoken phrase, pitch inflexions, accents and flux that modify intensity, as forms and variations of overall timbre, as elements combined giving rise to meaning. In other words, speech is already composed. The manipulation of these different elements often risks destroying this meaning. A mere deviation can often provoke over-determined connotation, corrupting irreparably the initial material. In *Locus*, the musical work consisted of submitting the voice material to slides and gaps in order to obtain configurations that are improbable and yet believable.

The confrontation of the original material, or its mnemonic manifestations, with the *doubles* resulting from musical manipulation creates gaps that produce meaning (for example a sequence of breathings which becomes unreal once associated to rhythms that are too fast to be uttered by a human voice, or a sequence of vocal noises analysed and reconstructed following a regular meter – in sequence 19). These gaps can also arise from interventions at the semantic level, as in sequences 3 to 6, where virtual and real voices expose a transformation from speech to noise: while the first ones follow this evolution from a material point of view, successively concatenating phrases, words, syllables, and phonemes, the second ones follow the reverse progression, from phonemes to phrases, producing a meaningful noise while reconstituting a kind of imaginary language.

These interactions between concrete sound materials and symbolic musical constructions and concerns could hardly have been carried out without the use of computer-aided sound analysis and composition tools.