



香港城市大學
City University of Hong Kong

專業 創新 胸懷全球
Professional · Creative
For The World

CityU Scholars

Preparing for TreeTorika Computer-Assisted Analysis of Mao's oratory Lindborg, PerMagnus

Published in:
SMC'07 4th Sound and Music Computing Conference

Published: 01/07/2007

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

License:
CC BY

Publication record in CityU Scholars:
[Go to record](#)

Published version (DOI):
[10.5281/zenodo.849457](https://doi.org/10.5281/zenodo.849457)

Publication details:
Lindborg, P. (2007). Preparing for *TreeTorika*: Computer-Assisted Analysis of Mao's oratory. In C. Spyridis, A. Georgaki, G. Kouroupetroglou, & C. Anagnostopoulou (Eds.), *SMC'07 4th Sound and Music Computing Conference: PROCEEDINGS* (pp. 267-274). (Proceedings of the Sound and Music Computing Conference, SMC). National and Kapodistrian University of Athens. <https://doi.org/10.5281/zenodo.849457>

Citing this paper

Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights

Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission

Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy

Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.

Preparing for *TreeTorika*: Computer-Assisted Analysis of Mao's oratory.

PerMagnus Lindborg*

School of Contemporary Music, LaSalle College of the Arts, Singapore.

<mailto:permagnus.lindborg@lasalle.edu.sg>, <http://www.pmpm.tk>

Abstract — This paper examines computer-assisted analysis techniques in extracting musical features from a recording of a speech by Mao Zedong. The data was used to prepare compositional material such as global form, melody, harmony and rhythm for *TreeTorika* for chamber orchestra. The text focuses on large-scale segmentation, melody transcription, quantification and quantization¹. It touches upon orchestration techniques but does not go into other aspects of the work such as constrained-based rhythm development. Automatic transcription of the voice was discarded in favour of an aurally based method supported by tools in Amadeus and Max/MSP. The data were processed in OpenMusic to optimise the notation with regards to accuracy and readability for the musicians. The harmonic context was derived from AudioSculpt partial tracking and chord-sequence analyses. Finally, attention will be given to artistic and political considerations when using recordings of such an intensively disputed public figure as Mao.

I. INTRODUCTION

A. *In TreeTorika*[10], I deal with rhetorics through recordings of Mao Zedong speeches, pursuing work from recent pieces, in particular *ReTreTorika* for quartet and computer, *ConstipOrat* for loudspeakers and Mao-variations for trio. In these pieces, transcriptions of the voice are taken as 'found'² musical material. In the first of them, bits of un-edited recording fuse with the saxophone and the ensemble, making Mao's voice very much part of the sonic image. In succeeding pieces, the original material is segmented, transformed and re-composed, thus distancing it from the source. The challenge I set for myself with *TreeTorika* was to make it a faithful transcription of the recording in material terms and at the same time create a composition of maximal integrity in poetic terms. When the source material stems from such a complex personality as Mao Zedong, I would have succeeded when it is possible to listen to the music without being reminded about what he was.

B. Rhetorics as a metaphor for composition

As a musician, I am interested in prosody: the way someone forms and phrases a vocal delivery. Rhetorics is

not about *what* is being said; it is about understanding *how* something is said.³ Oratory is not about clarity in public address; it is about manipulating an audience. Rhetorics is the study of oratory and oratory is the subject of rhetorics. There is no such thing as written oratory. There can be preparatory notes for a speech, and an article can be read aloud in plenum – but even good writing can make bad oratory. Further, rhetorics is associated with a Greek tradition, codified by Aristotle, and oratory with a Roman tradition, codified by Herennius. I am interested in the way speakers use the rhetorical situation, *kairos*: the particular moment and necessity calling for someone to speak in public. Kairos demands that the orator (lawyer, teacher, politician...) gauge the situation and respond to it by adapting to an adequate mode of delivery. The art of speech-making lies in understanding dynamical relations in the speaker+topic+listener system and using this knowledge to affect the listeners' mindset. Great orators craft the situation to their advantage: it is all about *convincing* the listeners. Rhetorics can be a metaphor for composition. But does it mean that a work of music is *oratory*? The answer is no. As much as I like to think that *TreeTorika* is capable of moving the listener, it does not attempt at convincing people to adopt any particular point of view. It is an abstraction: a musical drama involving aspects of rhetorics.[1],[7]

C. The orator

Mao spoke Hunanese, which many Chinese quite clearly find hard to follow. Several writers have noted the particularity of Mao continuing to use a strong rural dialect even after coming to power in Beijing in 1949. His pronunciation must have sounded flavoured to urban ears⁴, but he never gave it up and was obviously concerned with showing "peasant pathos" in order to maintain a strong association with his original rural powerbase, at least up to the end of the 1950s. It can be argued that Mao Zedong, a singularly powerful individual with relatively few public appearances, did not see oratory as a central means to his exercise of power. An analysis of China's leadership's passion for calligraphy can be found in [8]. All his life, Mao seems to have preferred writing as the main channel for his exercise of power.⁵ After leaving

¹ The former refers to rhythm, the latter to harmony.

² As in "found object", *objet trouvé*.

³ Those negative to argumentative techniques would say that they teach how *anything* can be said convincingly.

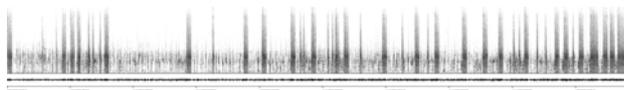
⁴ Taken out of its context, his voice seems to be slightly amusing to contemporary ears, at least with the younger generation.

⁵ In this he differs from politicians such as Martin Luther King, Olof Palme or Benito Mussolini, people for whom the scene, the microphone and the camera consisted the platform of a public mandate.

Shanghai in 1927, Mao spent most of the next two decades in poor rural regions, first with guerrilla warfare and then in Yan'an. At the time of the Guomintang retreat and the Communist seizure of power in 1949, Mao was the undisputed leader and from this point on, his every word was recorded. [4], [15], [16] The recordings in [11] that I have studied are of seven major speeches given between 1949 and 1956. I have not been able to find earlier audio material of Mao. Although it is intensively disputed, I think Mao was a skilled orator.⁶ His delivery is not flamboyant or aloof, but rather, from a musical point of view, based on rhythmic stability and lively melodious prosody with no extreme fluctuations of dynamics or register. The pauses between phrases are carefully measured for maximal effect, with energy accumulating over time. Coughing, grunting and clearing of the throat are not muted, but become an integral part of vocal style, somehow boosting the alpha-male image of “the Great Helmsman”.⁷

D. Preliminary observations of the recording

The recording at the base of *TreeTorika* is the speech Mao gave at Tiananmen in Beijing on October 1st, 1949, declaring the communist People's Republic of China. The duration of the recording is just under 20 minutes⁸; it is narrow-band and noisy⁹. Before going further, I stress that my work is not about the text¹⁰, but focuses entirely on a musical understanding of Mao's oratory. A discussion of the text and its political context is beyond the scope of this article. In any rhetorical situation, the large-scale form must be understood as a dialogue between speaker and listeners. I will start by making some general observations of the present recording. There are two dominant sounds: Mao's voice and the audience clapping their hands. The pattern of their interaction may be observed in a sonogram, shown in Fig. 01.



⁶ Chan & Halliday clearly think he was not.

⁷ At the beginning of the work, I included Mao's guttural noises in the transcription. They appear in all the speeches I have studied. It is likely that the grunting had become an integral part of his style of delivery, somehow adding power and virility (rather than the opposite.) Listening to persistent coughing simultaneous with a round of applause, it struck me that there had been edits. In fact, several applauses in different parts of the speech contained the same coughing. Apparently, some equally dedicated and unscrupulous Chinese Communist Party sound engineer had engaged in copy-pasting. Clearly, if the present recording is different from the actual event in local timing, it may be a construct in overall duration and even large-scale form. Only a study of the raw recording will tell the influence of such edits on rhetorical style, and is beyond the scope of this article. The discovery made me hesitate, but eventually I decided that the potential difference between recording and “real” speech was irrelevant, once I had accepted the recording as ‘found material’. Moreover, who knows if the engineer's version isn't *better* than Mao's?

⁸ It may not be the duration of the speech Mao gave.

⁹ The modern listener is estranged from the reality of the event by the low-fi qualities of the recording. However, this opens up for using the noisiness expressively, as the pieces mentioned in the Introduction show; samples of Mao Zedong's voice appear ‘clean’ (i.e. de-noised), ‘raw’, ‘de-voiced’, deformed by granulation, time-dilated or filtered in various ways. The more transformed the sample is, the less strongly it connotes the source.

¹⁰ Joyce Beetuan Koh has kindly assisted me in studying a transcript.

Fig. 01. Sonogram of Mao's October 1st, 1949 speech. [sound example]

The register of Mao's voice is that of a tenor, mainly covering a range from low *e* (MIDI 53) to high *g* (MIDI 67), at the end reaching a high *c*. Particular emphasis to a word is often given by a long glissando, covering more than an octave, from a lower register around *Bb* (MIDI 46). The way the last few words of a section are drowned in intensive handclapping indicates that the audience was familiar with his timing.¹¹ In the sonogram in Fig. 01, the applauses appear as clearly demarcated bands; there are 44 of them. Most of the applauses are approximately 4 seconds in duration and intersperse the orator's discourse, while larger applauses, around 8 seconds long, seem to mark the end of major parts of the speech. Interestingly, applauses, either ‘big’ or ‘small’, stay the same duration with remarkably consistency, except for the final applause, which is the longest. The interaction pattern gives clues to a large-scale form: the orator dominates the beginning; there is a densification of applauses towards the end; throughout the speech, the phrases of the orator shorten in duration. All in all, these lines of force create a forward-leaning drive. After some of the large applauses, the voice picks up at a significantly slower rate and lower pitch. It then works its way with increased intensity to a local climax, drowning in another large applause. This scheme, repeated six times, each lasting approximately three minutes, suggested a division of the speech in six segments followed by a short coda. I then focused on small differences in vocal sound qualities between the segments and decided to expand those in the orchestration. The first segment is *iskhnos*, ‘even’ in character; the second is imposing; the third is energetic but contained; numbers four and five are grumpily dramatic; the sixth is triumphant, *deinos*. A preliminary sketch is shown in Fig. 02, where segments are grouped so as to create a piece in three movements (as it were, an enlarged version of *Mao-variations*, in which the central movement was composed with material from the present speech). This initial tripartition was later abolished for a grouping in four parts (1+2, 3, 4+5, 6+coda) but, nonetheless, the general directionality, speed and timbre of segments remain essentially as drafted. The segments were composed in chronological order, except for the third, which was finished last. The analysis methods developed according to the requisites of the segment at hand.

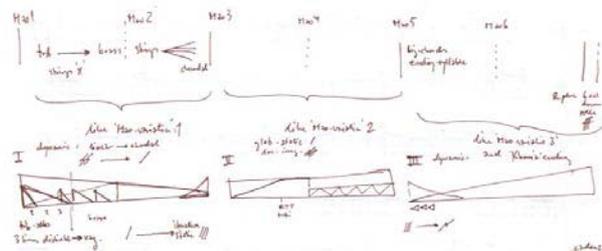


Figure 02. An early sketch of the large form, with preliminary ideas for orchestration that influenced the analysis methods.

II. ANALYSES

After the large-scale segmentation, the ensuing analyses concern the voice: first, the phrases, in particular the rhythm between speaking and silence; second, the

¹¹ It is not difficult to imagine that officials on the podium *cued* ovations.

transcription of the vocal line into a melody; third, the phrases and the melody are used to determine tempo and quantification; fourth, extraction of harmonic material from the vowels. Fig. 03 outlines the processes and dataflow as well as the supporting applications.

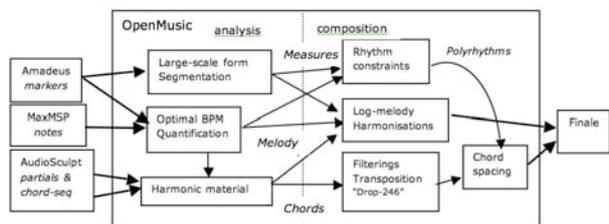
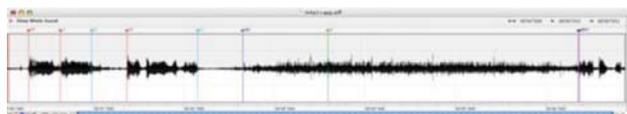


Fig. 03. Overview of analysis and composition phases.

A. Marking the phrases

The low quality of the recordings, Mao's vowel-rich pronunciation, sliding vowel changes and not much articulated consonants pose problems for the transcriber. I initially worked with AudioSculpt's [3] spectral differencing methods to generate automatic segmentation, but did not arrive at satisfactory results.¹² Using the ears turned out to be more precise. More importantly, a "manual-aural" method, obliging me to listen repeatedly and in great detail to the sounds, allowed me to integrate the musicality of the voice.¹³ The software Amadeus [5] provided comfortable navigation in long soundfiles and editing of markers with different name and colour.¹⁴ The entire file, a short extract of which is shown in Fig. 04, contained eventually some 800 markers of different kinds, e.g.:

- **m** = 'Mao', the start of a vocal phrase
- **si** = 'silence', i.e. the end of a phrase
- **app, APP** = 'smaller applause' and 'bigger applause'
- **appsi, APPsi** = silence after applause
- **x** = Mao coughing or clearing his throat
- **>** = 'accent', syllable given particular weight



¹² Apart from the difficulty in finding parameters to produce a correct automatic segmentation, errors appeared when exporting files containing both generated and hand-added markers.

¹³ This side effect of an (admittedly tedious) analysis method – though with the merit of gratifying the musician in the musicologist – later served to facilitated the work of the orchestrator in the composer. Fabien Lévy writes in [9] that he preferred doing the "transcription/quantification by hand, despite the slowness of the work, printing the chord sequence onto graph paper and then copying it again slowly. In fact, it is this manual operation that permits me to understand, to control – in short, to listen to this primary material with my inner ear, in order to render it musical subsequently." I sympathise entirely with the approach, willingly using 'slow' techniques in the transcription, while on the other hand preferring automatic rhythm quantification as far as possible. Nonetheless, like Lévy, I did a fair amount of manual adjustment to rhythms and chord spellings in order to accommodate impractical notation, guided by orchestration techniques, intuition and experience.

¹⁴ There is a built-in function for automatic generation of markers for "silence", but it is not reliable.

Figure 04. Excerpt from segment 4 with markers showing start and end of two short phrases, followed by clapping (during which Mao coughs).

Exporting the markers as a text file and using the data in OpenMusic [2] is straightforward (see the upper left part of the patcher in Fig. 12). We may now visualize the markers in various ways. For example, let us take a closer look at the output from segment 2, the longest stretch uninterrupted by applauses, lasting approximately two and a half minutes. The middle (thicker drawn) curve in Fig. 05 plots the duration of the spoken part of phrases against their absolute onset time (as given by 'm' markers). Below is another curve showing the duration of *quasi*-silences (as given by 'si' markers). The uppermost curve shows the duration of whole phrases (i.e. "enunciation followed by silence"). These curves are partial results from the patch in Fig. 12 (left side).

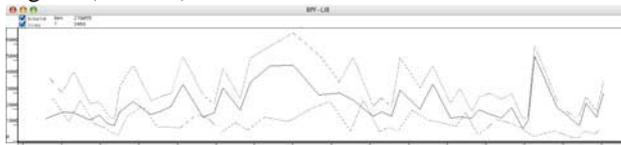


Figure 05. Durations of enunciations and silences in the second segment.

From the diagram, we see that the longest phrases (around four seconds) appear before the middle point. At the outset, silences are long – even longer than Mao's enunciations – but towards the end, the silences are very short, for a gulp of air. As Fig. 06 shows, the proportion between length of phrase and length of silence over the section is quite consistent (around 2:1, except for the end).

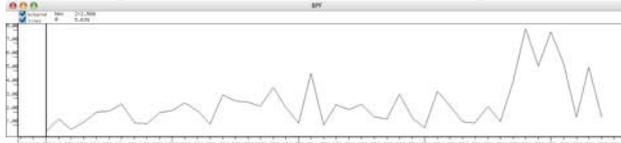


Figure 06. Proportions between speaking and pausing ('silence') parts of the phrases in segment 2.

Interestingly, there is a tendency, in particular the third and sixth segments, for the orator to alternate between long and short phrases, creating a rhythm on the level of phrases. An example is shown in Fig. 07.

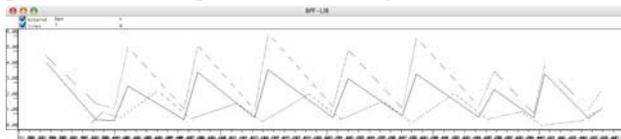


Fig 07. The third segment (here, the second bit out of eight) shows a regular pattern of alternating long and short phrases. The thick line indicates the duration of 'm' (Mao) bits, the dotted one below indicates the 'si' (silence) bits.

B. Transcribing the melody

The next step of the work was the transcription of the vocal line as a *melody*, treating each syllable¹⁵ as a note. The earlier remark about the advantages of a 'manual-aural' method over an automatic segmentation

¹⁵ While ancient Chinese was essentially monosyllabic, *putonghua* is semantically polysyllabic to a high degree. However, the pronunciation, giving each syllable almost identical weight, makes it sound monosyllabic. As can be expected, this happens more in formal speech than in casual.

method applies here as well. The Max/MSP [17] patcher, shown in Fig. 08, supports the transcription process. A syllable is visually and aurally determined by selecting a portion in the *waveform~* window. While it loops, the user decides the pitch either by relying on the *fiddle~* [14] estimation or by checking the sound against a note on the keyboard (if necessary adjusting it by a quartertone).

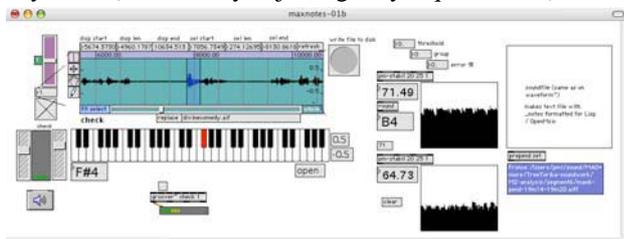


Figure 08. Aural transcription assisted by a simple Max/MSP tool. At this point, *fiddle~* estimates the pitch as B4, but the ear says it is F#4. The amplitude for each note was determined automatically, by *fiddle~*'s *velocity* detection (here 64.7).

As with the phrase markings discussed above, a low-level familiarity with the 'found' material allowed me to integrate the musicality of the source material.¹⁶ After having worked through all the syllables in a section, the program saves onset, offset, pitch and velocity for all the notes in a text file, which is subsequently transferred to Open Music (as shown in the upper right half of the patch in Fig. 12). *Maquettes* such as the one in Fig. 09 were used at different stages to validate the transcription.

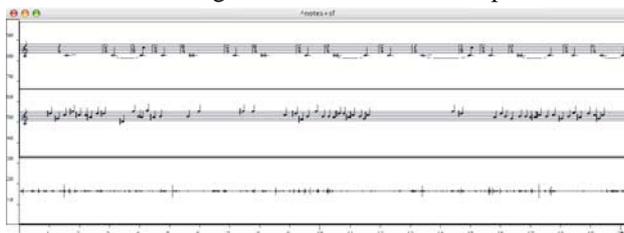


Figure 09. A maquette of segment 3, phrase 2 used to check the 'Max/MSP-notes'. It contains the 'm' markers from Amadeus transcribed as downbeats, the voice transcribed as melody and a soundfile excerpt. [sound example]

Interesting conclusions about Mao's vocal delivery could be drawn from a thorough study of the data. Fig. 10 shows normalized pitch and amplitude curves for a section. They appear linked; indeed for 85% of the notes the correlation is 0.9 or greater¹⁷.

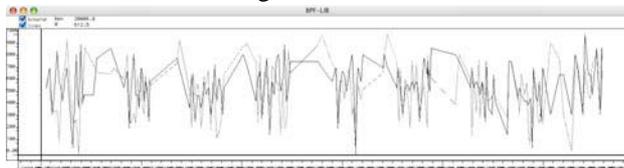


Figure 10. Highly correlated pitch and amplitude in vocal line in segment 3, section 2.

The succeeding section, much shorter, has an even higher degree of correlation. One may notice a pattern

¹⁶ Mao's dialect, Hunanese, is a language of tones, and in particular the many glissandos pose problems for the transcriber. In *TreeTorika* (with the exception of segment 3), I interpreted sliding syllable as two or three fixed notes. I am aware of that this is a limitation and intend to improve the method in future work.

¹⁷ The square of the difference between normalized curves is used as correlation measure.

wherein a phrase typically consists of several short notes leading to a longer, louder note.

When studying the phrases, we noticed that segment 3 is the most regular of all. It is also the segment where the sections have the highest values for correlation between pitch and amplitude. By contrast, segment 5 does not show a consistent degree of correlation. Towards the end, the correlation breaks down, as hinted in Fig. 11. At this point, less than half of the notes show a correlation of 0.9 or more. This means that in segment 5, the make-up of notes is more varied than in segment 3; i.e. there are more of the 'high-soft' and 'low-loud' syllables. In other words, one might say that the orator is "wilder" or that he uses a larger expressive palette. Listening to the recording supports this observation.

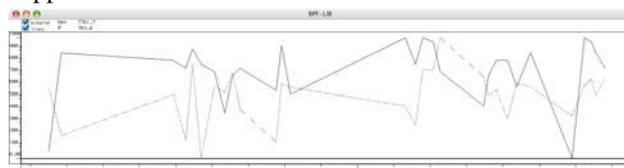


Figure 11. Less correlated pitch and amplitude in segment 5, section 5.

C. Beat speed

After having extracted phrases and notes, the next step was to 'tame' the data in order to make an easily readable score, while maintaining a high degree of accuracy. The crux of the problem is finding a good pulsation speed, in relation to which quantification can be made. The demand for high accuracy has to be negotiated with the musicians' demand for efficient learning and use of resources.¹⁸ Every performer employs some form of "inner clock", with subdivisions ticking away. Since a given duration sequence may be notated in any pulsation speed (with or without changing meters), the composer can act on the efficiency of the inner clock(s) set in motion. The perception of a sequence of durations as rhythm is a psycho-motoric phenomenon¹⁹. A different notation signals a different rhythmic context, even if such context is not explicitly stated. In a subtle way, the musician's inner rhythm influences the listener's perception of the music.²⁰ The beat speed analysis is made in the main analysis patch, shown in Fig. 12. In a first step, the left part of the patch is used to choose an optimal beat speed and calculate a sequence of measures. In a second step, the right-hand side functions are used to quantify the 'Max-notes' melody according to the BPM and measure markings determined in the first step. In addition, the BPF marked "melody's pitch distribution" informs about the orator's use of ambitus.

¹⁸ For composers, logistics such as limited rehearsal time can be a frustrating real-world parameter to deal with. When it comes to complex rhythms, my experience is that the result is highly dependent on notation style.

¹⁹ "We connect the perception of musical motion at the ecological level to human motion. This suggests that musical perception involves an understanding of bodily motion -- that is, a kind of empathetic embodied cognition." [6]

²⁰ Working with relations between composer's notation and musician's "inner clock" is the basic idea of my percussion piece *Danses Condensées* (1997/2005). The durations of individual notes in a ten-second phrase are kept unaltered over 29 repetitions, but are notated within a superimposed process of a large-scale accelerando going from "non-beat" proportional notation, via slow and irregular beats to increasingly high beat speeds.

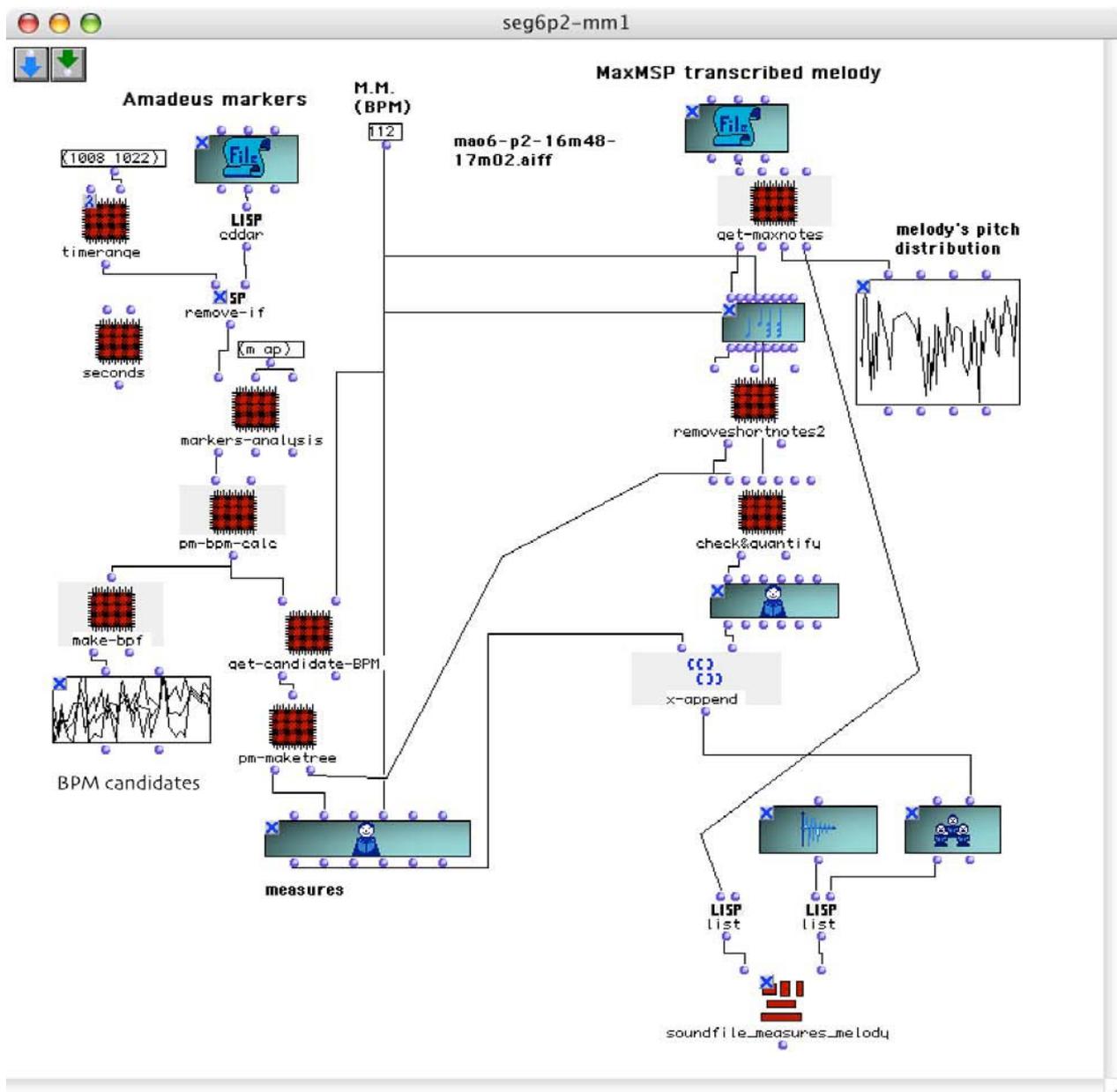


Figure 12. Patch used to choose optimal beat speed, calculate sequence of measures and quantify the transcribed melody. The BPF at the right, marked "pitch distribution", gives additional information, indicating how the orator uses the ambitus in a section.

The phrase markers are used to assist the choice of beat speed for each section. I developed an algorithm that determines the fitness of BPMs²¹ (between 40 and 180) according to *accuracy* and to *readability*.²² Each 'candidate BPM' produces an approximated sequence of measures. The fitness value according to *accuracy* uses the sum-square distance between the 'raw' and approximated duration sequences. The algorithm for calculating *readability* favours simple measure markings such as 4//4 and disfavours short and oblique ones such as

7//16.²³ The OM patch to calculate the fitness is shown in Fig. 13.

The resulting values are displayed as curves in a *BPF-lib*, such as the one shown in Fig. 14. In general, the curve for accuracy climbs steadily over the range, since the time resolution increases with BPM speed. The readability curve is more salient, with peaks and valleys. The algorithm also gives the *average* of the first two curves normalized; the peaks of this curve indicate optimal beat speeds.

²¹ Beats Per Minute, equivalent to M.M. (*Mälzels Metronom*)

²² Experimenting with Benoît Meudic's *OMKant* [13],[12] did unfortunately not lead me to a convenient way of working with changing meters. My algorithm is simple, and may not be extendable beyond the present task, though it could eventually be integrated with *OMKant* as a user-defined function.

²³ An improvement to the algorithm would be the inclusion of a fitness value for *metric stability*.

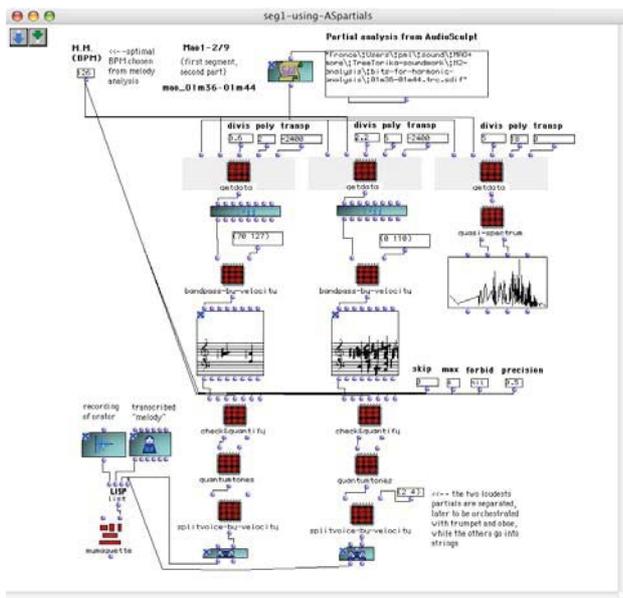


Figure 17. Patch used to filter and quantify a partial tracking from AudioSculpt. Note that quantification is entered in relation to the BPM chosen for the section.

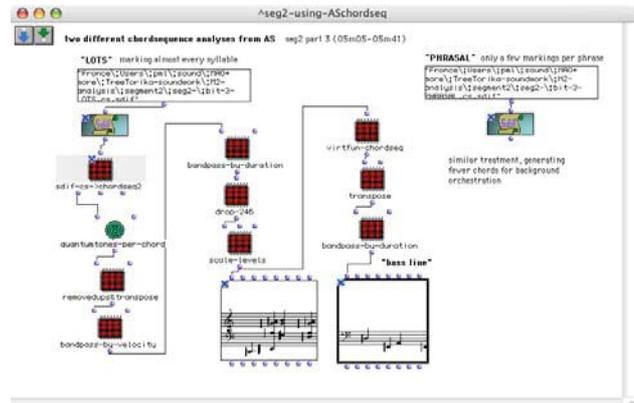


Figure 18. A patch applying harmony quantization and filtering to chord-seq SDF data. “Drop 246” is a chord spacing technique used in big band orchestration to distribute diatonic chords (good-sounding on a piano) to a wind instrument section. A ‘bass’ line (which could be scored for Piccolo or Gongs or some other instrument) is produced with the virt-fun or the best-freq functions.



Figure 19. Excerpt from the end of segment 2, woodwinds and brass (measure 118 ff). The chords for each instrument group were generated from two concurrent chord-seq analyses. The saxophone plays the ‘Max-notes’ transcribed melody. [sound example]

F. Quartertone quantization

The AudioSculpt analysis produces an agglomerate of pitches, a ‘raw chord’, containing potentially a large number of pitches. I typically wanted chords to consist of 6 to 15 notes. To reduce the size of the raw material, some notes were filtered out if they could be considered too faint or short, depending on context. The patches above employ the function *quantumtones*. It works by allowing a certain number (between 15 and 20 percent per agglomerate) of those pitches furthest away from on-pitch temperate intonation to be approximated to tempered quartertones, while the rest become semitones. The

psychological concerns discussed above in the passage about rhythm notation are valid in relation to quartertone harmony²⁷, in that the spelling of chords influences rehearsal efficiency. To my ears, chord sequences are the most convincing and understandable when quartertones are in minority. The way the quantization is made in *TreeTorika* negotiates truthfulness to the *sound* of the recording with demands for unproblematic ensemble intonation.

²⁷ While *TreeTorika*’s harmonic fields are based on spectral techniques, the orchestration does not emphasize instrumental anonymity, as in purist spectral composition-orchestration.

III. CONCLUSION

With *TreeTorika*, I attempted to stay analytically faithful to a given structure (in this case, a sound recording) while retaining compositional freedom. Upon hearing the piece, listeners do not immediately associate it with rhetorics and are surprised to learn how much the composition draws on one orator, from one unique speech. I would like to think it suggests that the project has been successful; at least in regards to the challenge I had given myself. After seven pieces based on Mao's voice, people have asked me whom the "next dictator project" will be centred on... Stalin, perhaps? The darker side of human behaviour has an attraction on many people, and I am no exception in this respect. More importantly, if my work can illuminate and assist in understanding the workings of *musical* power within the field of politics, my research would have come a long way. In future work (involving a less than tyrannical leader) I will concentrate on improving analytical methods in OpenMusic, with the inclusion of a fitness value for metric stability in the beat speed optimization algorithm and integrate it with the AudioSculpt-OMKant tools. I am also pursuing Max/MSP tools to assist transcription of melody, rhythm and chords, employing editable automatic segmentation, to profit from the advantages of real-time interactive analysis.

IV. ACKNOWLEDGMENTS

TreeTorika for chamber orchestra [1111-sax-1111-acc-2prc-pf-2111] (22', 2006) was commissioned by Ensemble Ernst with support from Komponistrådet (Composers' Council), Norway and is dedicated to conductor Thomas Rimul and Ensemble Ernst²⁸. Lars Lien performed the saxophone part at the first performance on October 14, 2006, at Ultima Festival in Oslo. I thank Dr Chan Hing-yan and Dr Joyce Beetuan Koh for stimulating discussions about music and rhetorics.

REFERENCES

- [1] Andersen, Øyvind: *I retorikkens hage*. Universitetsforlaget, Oslo, Norway, 1995.
- [2] Assayag, G. and Agon, C.: *OpenMusic*. Ircam 1998-2007. <http://freesoftware.ircam.fr> (June 2007).
- [3] Bogaards, N., Eckel, G. and Roebel, A.: *AudioSculpt*. Ircam 1994-2007 <http://forumnet.ircam.fr/349.html> (June 2007).
- [4] Chan, Jung and Halliday, Jon: *Mao, the Unknown Story*. Knopf (Random House), 2005.
- [5] Hairer, Martin: *Amadeus*. <http://www.hairersoft.com> (June 2007).
- [6] Iyer, Vijay S.: *Microstructures of Feel, Macrostructures of Sound: Embodied Cognition in West African and African-American Musics*. (PhD Thesis), University of California Berkeley, 1998, at <http://www.cnmat.berkeley.edu/People/Vijay/%20THESIS.html> (April 2007)
- [7] Kennedy, George: *A New History of Classical Rhetoric*. Princeton Press, 1994.
- [8] Kraus, Richard K.: *Brushes of Power. Modern Politics and the Chinese Art of Calligraphy*. University of California Press, 1991.
- [9] Lévy, Fabien: "When the Computer Enables Freedom from the Machine". In Agon, Assayag & Bresson (Ed.): *OM Composers' Book 1*, Delatour/Ircam-Centre Pompidou, Paris 2006.
- [10] Lindborg, PerMagnus: *TreeTorika for chamber orchestra*. Music Information Centre, Oslo, 2006. <http://www.mic.no> (April 2007)
- [11] Mao, Zedong: *[1st October 1949 People's Republic of China Declaration Speech]*. Juren zhi sheng. Shenzhen Shi xian ke yule chuanbo youxian gongsi, Zhongyang danganguan, Zhongyang wenxian yanjiushi (CD). Zhonghua Book Co., Hong Kong 1994.
- [12] Meudic, Benoît: "Librairie OMKant 3.0". Ircam, Paris, 2003.
- [13] Meudic, Benoît: "Modélisation de structures rythmiques." DEA ATIAM, Ircam - Université d'Aix-Marseille II, 2000.
- [14] Puckette, M., Apel, T. and Zicarelli, D.: "Real-time audio analysis tools for Pd and MSP". ICMC 1998.
- [15] Short, Philip: *Mao, a Life*. Hodder and Stoughton, 1999.
- [16] Spence, Jonathan D.: *The Search for Modern China*. W.W. Norton & Company, 2001.
- [17] Zicarelli, D.: "An extensible real-time signal processing environment for Max" ICMC 1998. Max/MSP is developed by Cycling74, <http://www.cycling74.com> (May 2007).
- [18] Wright, M., Chaudhary, A., Freed, A., Khoury, S., and Wessel, D.: "Audio Applications of the Sound Description Interchange Format Standard". AES 1999.

Sound examples are available on the author's webpage, <http://www.pmpm.tk> (June 2007), under [publications].

²⁸ The title has little to do with dendrochronology (although a working-title was the old carpenter's adage "measure twice, cut once") but is a contraction of the initials of those to whom it is dedicated, Thomas-Rimul-Ensemble-Ernst, and the word 'rhetorics'.