ON THE MAGIC OF OVERTONE SINGING

Piero Cosi, Graziano Tisato

*ISTC-SFD - (ex IFD) CNR
Istituto di Scienze e Tecnologie della Cognizione - Sezione di Fonetica e Dialettologia
(ex Istituto di Fonetica e Dialettologia) - Consiglio Nazionale delle Ricerche
e-mail: cosi@csrf.pd.cnr.it tisato@tin.it
www: http://nts.csrf.pd.cnr.it/Ifd

I really like to remember that Franco was the first person I met when I approached the "Centro di Studio per le Ricerche di Fonetica" and I still have a greatly pleasant and happy sensation of that our first warm and unexpectedly informal talk. It is quite obvious and it seems rhetorical to say that I will never forget a man like Franco, but it is true, and that is, a part from his quite relevant scientific work, mostly for his great heart and sincere friendship.

1. ABSTRACT

For "special people" scientific interests sometimes co-occur with personal "hobbies". I remember Franco talking to me about the "magic atmosphere" raised by the voice of Demetrio Stratos, David Hykes or Tuvan *Khomei*[1] singers and I still have clear in my mind Franco's attitude towards these "strange harmonic sounds". It was more than a hobby but it was also more than a scientific interest. I have to admit that Franco inspired my "almost hidden", a part from few very close "desperate" family members, training in Overtone Singing[2]. This overview about this wonderful musical art, without the aim to be a complete scientific work, would like to be a small descriptive contribute to honor and remember Franco's wonderful friendship.

2. THE THROAT-SINGING TRADITION

"*Khomei*" or "Throat-Singing" is the name used in Tuva and Mongolia to describe a large family of singing styles and techniques, in which a single vocalist simultaneously produces two (or more) distinct tones. The lower one is the usual fundamental tone of the voice and sounds as a sustained drone or a Scottish bagpipe sound. The second corresponds to one of the harmonic partials and is like a resonating whistle in a high, or very high, register. For convenience we will call it "diphonic" sound and "diphonia" this kind of phenomenon.

Throat-Singing has almost entirely been an unknown form of art until rumours about Tuva and the peculiar Tuvan musical culture spread in the West, especially in North

---

[1] We transcribe in the simplest way the Tuvan term, for the lack of agreement between the different authors: Khomei, Khöömii, Ho-Mi, Hö-Mi, Chöömej, Chöömij, Xöömij.

[2] This is the term used in the musical contest to indicate the diphonic vocal techniques.

America, thanks to Richard Feynman [1][3], a distinguished American physicist, who was an ardent devotee of Tuvan matters.

This singing tradition is mostly practiced in the Central Asia regions including Bashkortostan or Bashkiria (near Ural mountains), Kazakhstan, Uzbekistan, Altai and Tuva (two autonomous republics of the Russian Federation), Khakassia and Mongolia (Fig. 1), but we can find examples worldwide: in South Africa between Xosa women [3], in the Tibetan Buddhist chants and in Rajastan.

The Tuvan people developed numerous different styles. The most important are: *Kargyraa* (chant with very low fundamentals), *Khomei* (it is the name generally used to indicate the Throat-Singing and also a particular type of singing), *Borbangnadyr* (similar to *Kargyraa*, with higher fundamentals), *Ezengileer* (recognizable by the quick rhythmical shifts between the diphonic harmonics), *Sygyt* (like a whistle, with a weak fundamental) [4]. According to Tuvan tradition, all things have a soul or are inhabited by spiritual entities. The legends narrate that Tuvan learnt to sing *Khomei* to establish a contact and assimilate their power trough the imitation of natural sounds. Tuvan people believe in fact that the sound is the way preferred by the spirits of nature to reveal themselves and to communicate with the other living beings.

Figure 1. Diffusion of the Throat-Singing in Central Asia regions.

In Mongolia most Throat-Singing styles take the name from the part of the body where they suppose to feel the vibratory resonance: *Xamryn Xöömi* (nasal *Xöömi*), *Bagalzuuryn Xöömi* (throat *Xöömi*), *Tseedznii Xöömi* (chest *Xöömi*), *Kevliin Xöömi* (ventral *Xöömi*, see Fig. 13), *Xarkiraa Xöömi* (similar to the Tuvan *Kargyraa*), *Isgerex* (rarely used style: it sounds like a flute). It happens that the singers itself confuse the different styles [5]. Some very famous Mongol artists (Sundui and Ganbold, for example) use a deep vibrato, which is not traditional, may be to imitate the Western singers (Fig. 13).

The Khakash people practice three types of Throat-Singing (*Kargirar*, *Kuveder* or *Kilenge* and *Sigirtip*), equivalent to the Tuvan styles *Kargyraa, Ezengileer* and *Sygyt*. We

[3] Today, partly because of Feynman's influence, there exists a society called “Friends of Tuva” in California, which circulates news about Tuva in the West [2].

find again the same styles in the peoples of the Altai Mountains with the names of *Karkira*, *Kiomioi* and *Sibiski*. The Bashkiria musical tradition uses the Throat-Singing (called *Uzlau*, similar to the Tuvan *Ezengileer*) to accompany the epic chants. In Uzbekistan, Kazakhstan and Karakalpakstan we find forms of oral poetry with diphonic harmonics [6].

The Tibetan Gyuto monks have also a tradition of diphonic chant, related to the religious believes of the vibratory reality of the universe. They chant in a very low register in a way that resembles (see later the difference) the Tuvan *Kargyraa* method. The aim of this tradition is mystical and consists in isolating the 5th or the 10th harmonic partial of the vocal sound. They produce in this way the intervals of 3rd or 5th (in relation to the fundamental) that have a symbolic relation with the fire and water elements (Fig. 14) [4].

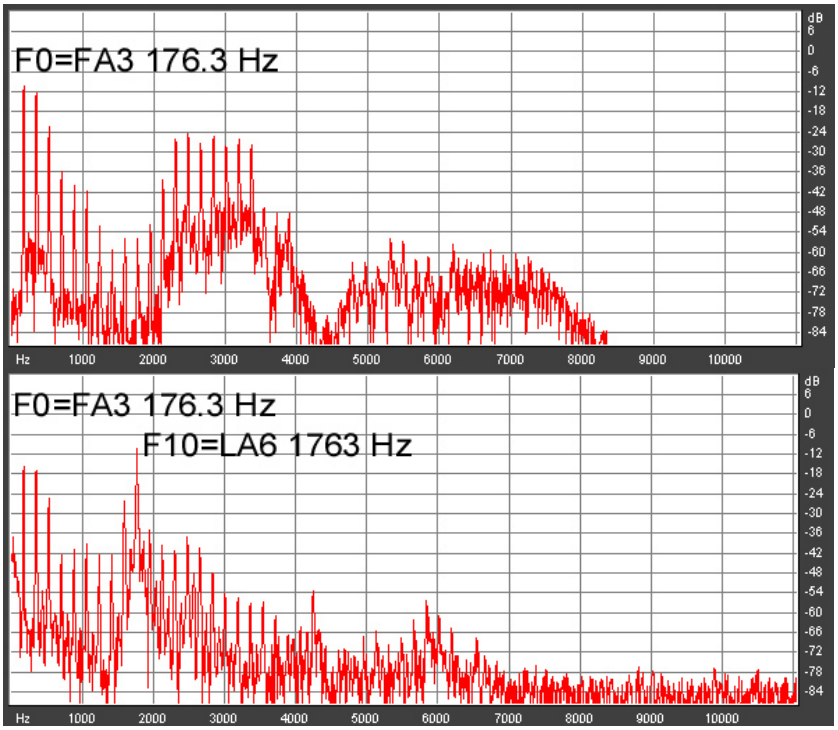

Figure 2. Spectral section of a vocal (up) and a diphonic vocal (down).

3. SEPARATION OF THE AUDITORY IMAGE IN THROAT-SINGING

What is so wonderful in Throat-Singing? It is the appearance of one of the harmonic partials that discloses the secret musical nature of each sound. When in Throat-Singing the voice splits in two different sounds, we experience the unusual sensation of a pure, discarnate, sine wave emerging from the sound. It is the same astonishment we feel when we see a rainbow, emerging from the white light, or a laser beam for the first time.

The natural sounds have a complex structure of harmonic or inharmonic sinusoidal partials, called “overtones” (Fig. 2). These overtones are not heard as distinct sounds, but their relative intensity defines our perception of all the parameters of sound (intensity, pitch, timbre, duration). The pitch corresponds to the common frequency distance between the partials and the timbre takes into account all the partials as a whole. The temporal evolution of these components is what makes the sound of each voice or instrument unique and identifiable.

In the harmonic sounds, as the voice, the components are at the same frequency distance: their frequency is a multiple of the fundamental tone (Fig. 2). If the fundamental frequency is 100 Hz, the 2nd harmonic frequency is 200 Hz; the 3rd harmonic frequency is 300 Hz, and so on. The harmonic partials of a sound form a natural musical scale of unequal temperament, as whose in use during the Renaissance [7]. If we only take into consideration the harmonics that are easy to produce (and to perceive also), i.e. from the 5th to the 13th, and if we assume for convenience a C3 131 Hz as starting pitch, we can get the following musical notes:

Harm. N.	Freq. (Hz)	Note	Interval with C3
5	655	E5	3rd
6	786	G5	5th
7	917	A+	6th +
8	1048	C6	Octave
9	1179	D6	2nd
10	1310	E6	3rd
11	1441	F6+	4th +
12	1572	G6	5th
13	1703	A6-	6th-

The series of 8th, 9th, 10th, 12th, 13th harmonic and the series from 6th to 10th are two possible pentatonic scales to play. Note that the frequency differences between these scales and the tempered scale are on the order of 1/8th of a tone (about 1.5%).

The Throat-Singing allows extracting the notes of a natural melody from the body of the sound itself.

The spectral envelope of the overtones is essential for the language comprehension. The glottal sound is filtered by the action of the vocal tract articulation, shaping the partials in the voice with some characteristic zones of resonance (called formants), where the components are intensified, and zones of anti-resonance, where the partials are attenuated (Fig. 2-3). So, the overtones allow us to tell apart the different vocal sounds. For example the sounds /a/, /e/, /i/, /o/, etc. uttered or sung at the same pitch, nevertheless sound different to our ears for the different energy distribution of the formants (Fig. 2).

The auditory mechanisms “fuse” the partials in one single “image”, which we identify as voice, musical instrument, noise, etc. [8]. In the same way, the processing of visual data tends to group different dots into simple shapes (circle, triangle, square, etc.). The creation of auditory images is functional to single out and to give a meaning to the sonic sources around us.

The hearing mechanisms organize the stream of perceptive data belonging to different components of different sounds, according to psychoacoustics and Gestalt principles. The “grouping by harmonicity”, for example, allows the fusion in the same sound of the frequency partials, which are multiples of a common fundamental. The “common fate” principle tells that we integrate the components of a complex sound, which show the same amplitude and frequency behaviour (i.e. similar modulation and microvariation, similar attack and decay, similar vibrato, etc.) [8]. If one of these partials reveals a particular evolution (i.e. it is mistuned or has not the same frequency and amplitude modulation, etc.), it will be heard as a separate sound. So the Throat-Singing is a marvelous example to understand the illusory nature of perception and the musical structure of the sound.

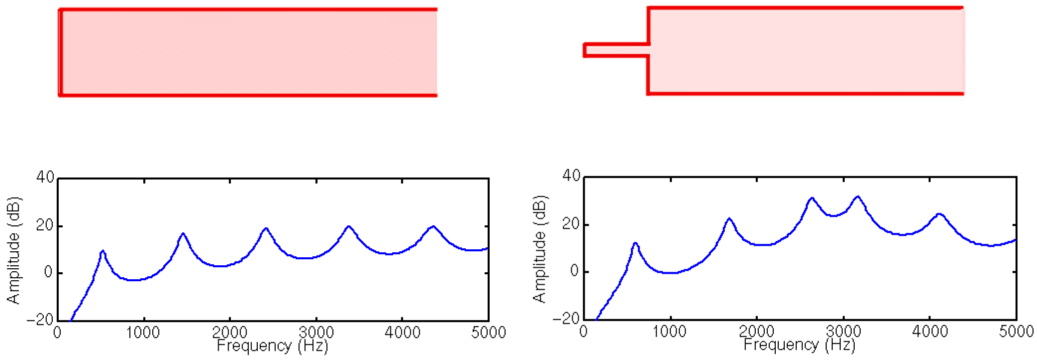

Figure 3. Resonance envelope for an uniform vocal tract (left). A constriction on the pharynx moves the formants so that the intensity of partials in the 2500-3500 Hz region increases (right).

4. FUNDAMENTAL TECHNIQUES IN THROAT-SINGING

In the Throat-Singing the singer learn to articulate the vocal tract so that one of the formants (usually the first or the second) coincide with the desired harmonic, giving it a considerable amplitude increase (even more than 30 dB, see in Fig. 2 the 10th harmonic) and making it perceptible. Unlike the normal speech, the diphonic harmonic can exceed a lot the lower partials intensity (Fig. 2). Soprano singers use similar skill to control the position of the 1st formant, tuning it to the fundamental with the proper articulation (i.e. proper opening of the mouth), when they want to sing a high note [9].

There are many different methods to produce the diphonic sound [5-6], but we can summarize them in two possible categories, called “single cavity method” or “two cavities method”, that are characterized by the use or not of the tongue, according to the proposal of Tran Quang Hai [4].

4.1 SINGLE CAVITY METHOD

In this method, the tongue doesn’t move and remains flat or slightly curved without touching the palate. In this case the vocal tract is like a continuous tube (Fig. 3). The selection of the diphonic harmonic is obtained by the appropriate opening of the mouth and the lips. The result is that the formants frequency raises if the vocal tract lengthens (for example with a /i/) and that the formants frequency lowers, if it extends (for example with a /u/). With this technique the 1st formant movement allows the selection of the partials. As we can see in Fig. 4, we cannot go beyond 1200 Hz. The diphonic harmonic is generally feeble, masked by the fundamental and the lower partials, so the singers nasalize the sound to reduce their intensity [10-11].

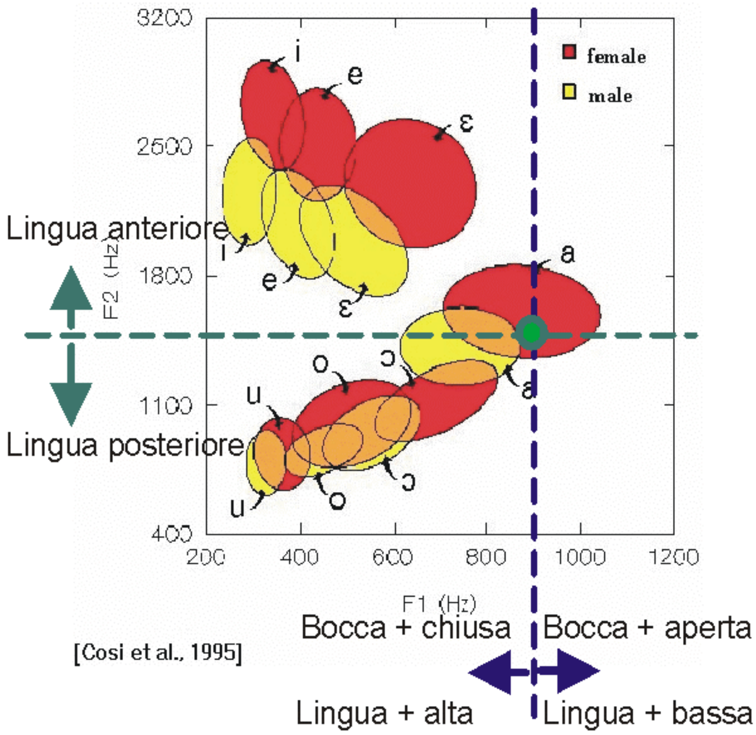

Figure 4. Opening the mouth controls the 1st formant position. The movement of the tongue affects the 2nd formant and allows the harmonic selection in a large frequency range.

4.2 TWO CAVITIES METHODS

In this method, the tongue is raised so to divide the vocal tract in two main resonators, each one tuned on a particular resonance. By an appropriate control, we can obtain to tune two separate harmonics, and thereby to make perceptible, not one but two (or more) pitches at the same time (Fig. 9-12).

There are three possible variants of this technique:

The first corresponds to the *Khomei* style: to select the desired harmonic the tip of the tongue and the tongue body moves forward (higher pitch) and backward (lower pitch) along the palate.

The second is characteristic of the *Sygyt* style: the tip of the tongue remains fixed behind the upper teeth while the tongue body rises to select the harmonics.

In the third variant, the movement of the tongue root selects the diphonic harmonic. Shifting the base of the tongue near the posterior wall of the throat, we obtain the lower harmonics. On the contrary, moving the base of the tongue forward, we pull out the higher harmonics [6].

A different method has been proposed by Tran Quang Hai to produce very high diphonic harmonics (but not to control the selection of the desired component). It consists to keep the tongue pressed by the molars, while singing the vowels /u/ and /i/, and maintaining a strong contraction of the muscles at the abdomen and the throat [4].

The advantage of the two cavities techniques is that we can use the 2nd formant to reinforce the harmonics that are in the zone of best audibility. In this case the diphonic harmonic reaches the 2600 Hz (Fig. 4). Furthermore the movement of the tongue affects the formants displacement in opposite directions. The separation of the 1st and the 2nd formant produces in between a strong anti-resonance (Fig. 2), which helps the perception of the diphonic harmonic.

In all these methods it is useful a slight discrete movement of the lips to adjust the formants position.

5. REINFORCING THE DIPHONIC SOUND

There are three main mechanisms required to reinforce the effect of segregation of the diphonic sound:

- The appropriate movement of the lips, tongue, jaw, soft palate, throat, to produce a fluctuation in the amplitude of the selected harmonic, so that it differentiates from the other partials that remain static. The auditory mechanisms are tuned to capture the more subtle changes in the stream of auditory information, useful to discriminate the different sounds [8].
- The nasalization of the sound. In this way we create an anti-resonance at low frequency (<400 Hz) that attenuates the lower partials responsible for the masking of the higher components [10-11]. The nasalization provokes also the attenuation of the third formant [12], which improves the perception of the diphonic harmonic (Fig. 2).
- The constriction of the pharynx region (false ventricular folds, arytenoids, root of the epiglottis), which increases the amplitude of the overtones in the 2000-4000 Hz region (Fig. 2). This is also what happens in the "singer's formant", the technique used by the singers to reinforce the partials in the zone of best audibility and to avoid the masking of the voice by the orchestra, generally very strong in the low frequency range [9]. For this reason the Throat-Singing technique requires a tuning extremely precise and selective, in order to avoid the amplification of a group of harmonic partials, as in the "singer's formant".

6. VOICE MULTIPHONICS

We disregard in this paper the polyphonic singing that could produces some diphonic effects: for example the phenomenon of the *quintina* in the Sardinia religious singing, where the coincidence of the harmonics of 4 real voices produces the perception of a 5th virtual voice (Fig. 5) [13].

There are in the literature many terms to indicate the presence of different perceptible sounds in a single voice: *Khomei,* Throat-Singing, Overtone Singing, Diphonic Singing, Biphonic Singing, Overtoning, Harmonic Singing, Formantic Singing, Chant, Harmonic Chant, Multiphonic Singing, bitonality, diplophonia, vocal fry, etc.

According to the pioneer work in the domain of the vocal sounds made by The Extended Vocal Techniques Ensemble (EVTE) of San Diego University and bearing in mind that there is little agreement regarding classifications [4], [14-15], the best distinctive criterion for the diphonia seems to be the characterization of the sound sources that produce the perception of the diphonic or multiphonic sound [16].

Following this principle, we can distinguish between **Bitonality** and **Diphonia**:

- **Bitonality**: In this case there are two distinct sound sources that produce two sounds. The pitches of the two sounds could be or not in harmonic relationship. This category includes: diplophonia, bitonality and vocal fry.
- **Diphonia**: The reinforcement of one (or more) harmonic partial(s) produces the splitting of the voice in two (or more) sounds. This category includes: *Khomei,* Throat-Singing, Overtone Singing, Diphonic Singing, Biphonic Singing, Overtoning, Harmonic Singing, Chant, Harmonic Chant.

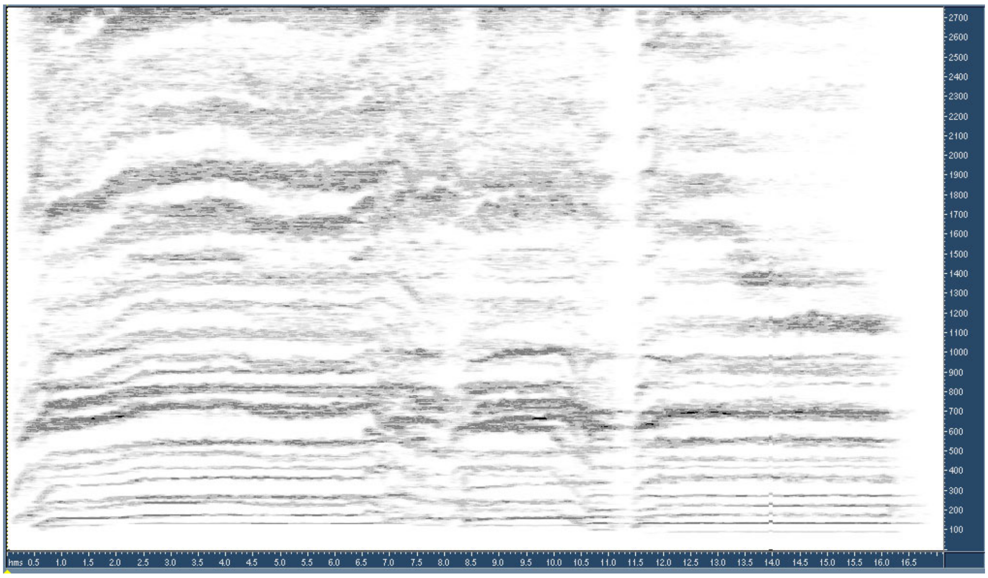

Fig. 5 Sardinia religious folk singing. The pitches of the 4 voices of the choir are F1 88 Hz, C2 131 Hz, F2 176 Hz, A3# 230 Hz. The 8^{th} harmonic of the F1, the 6^{th} of the C2, the 4^{th} of the F2 and the 3^{rd} of the A# coincide at 700 Hz and produce the perception of a 5^{th} voice.

6.1 BITONALITY

Diplophonia: The vibration of the vocal folds is asymmetrical. It happens that after a normal oscillatory period, the vibration amplitude that follows is reduced. There is not the splitting of the voice in two sounds, but the pitch goes down one octave lower and the timbre assumes a typical roughness. For example, assuming as fundamental pitch a C3 130.8 Hz, the resulting pitch will be C2 65.4 Hz. If the amplitude reduction happens after two regular vibrations, the actual periodicity triplicates and then the pitch lowers one octave and a 5^{th}. The diplophonic voice is a frequent pathology of the larynx (as in unilateral vocal cord paralysis), but can be also obtained willingly for artistic effects (Demetrio Stratos was an expert of this technique) [16-18].

Bitonality: The two sound sources are due to the vibration of two different parts of the glottis cleft. This technique requires a strong laryngeal tension [16-17]. In this case there is not necessarily a harmonic relationship between the fundamentals of the two sounds. In the Tuvan *Kargyraa* style, the second sound source is due to the vibration of the supraglottal structures (false folds, arytenoids, aryepiglottic folds that connects the arytenoids and the epiglottis, and the epiglottis root). In this case generally (but not always) there is a 2:1 frequency ratio between the supraglottal closure and vocal folds closure. As in the case of Diplophonia, the pitch goes down one octave lower (or more) [19-21].

Vocal fry: The second sound is due in this case to the periodic repetition of a glottal pulsation of different frequency [14]. It sounds like the opening of a creaky door (another common designation is "creaky voice"). The pulse rate of vocal fry can be controlled to produce a range from very slow single clicks to a stream of clicks so rapid to be perceived as a discrete pitch. Therefore vocal fry is a special case of bitonality: the perception of a second sound depends on a pulses train rate and not on the spectral composition of the single sound.

6.2 DIPHONIA

Diphonic and **Biphonic** refer to any singing that sounds like two (or more) simultaneous pitches, regardless of technique. Use of these terms is largely limited to academic sources. In the scientific literature the preferred term to indicated Throat-Singing is Diphonic Singing.

Multiphonic Singing indicates a complex cluster of non-harmonically related pitches that sounds like the vocal fry or the creaky voice [14]. The cluster may be produced expiring as normal, or also inhaling the airflow.

Throat Singing is any technique that includes the manipulation of the throat to produce a melody with the harmonics. Generally, this involves applying tension to the region surrounding the vocal cords and the manipulation of the various cavities of the throat, including the ventricular folds, the arytenoids, and the pharynx.

Chant generally refers to religious singing in different traditions (Gregorian, Buddhist, Hindu chant, etc.). As regards the diphonia, it is noteworthy to mention the low singing practiced by Tibetan Buddhist monks of the Gyuto sect. As explained before, they reinforce the 5^{th} or the 10^{th} harmonic partial of the vocal sound for mystical and symbolic purposes (Fig. 14). This kind of real diphonia must be distinguished from **resonantial** effects (enhancement of some uncontrolled overtones) that we can hear in Japanese Shomyo Chant [4] and also in Gregorian Chant.

Harmonic Singing is the term introduced by David Hykes to refer to any technique that reinforces a single harmonic or harmonic cluster. The sound may or may not split into two or more notes. It is used as a synonym of Overtone Singing, Overtoning, Harmonic Chant and also Throat-Singing.

Overtone Singing can be considered to be harmonic singing with an intentional emphasis on the harmonic melody of overtones. This is the name used by Western artists that utilizes vowels, mouth shaping, and upper-throat manipulations to produce melodies and textures. It is used as a synonym of Harmonic Singing, Overtoning, Harmonic Chant and also Throat-Singing.

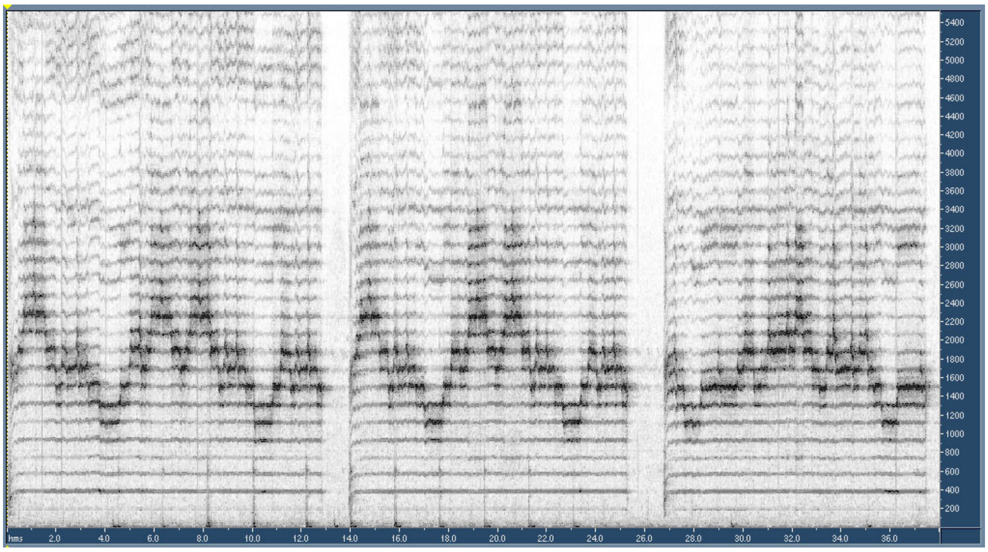

Fig. 6 Tuvan *Khomei Style*. The fundamental is a weak F#3+ 189 Hz. The diphonic harmonics are the 6th (C#6+ 1134 HZ), 7th (E6 1323 Hz), 8th (F#6+ 1512 Hz), 9th (G#6+ 1701 Hz), 10th (A#6+ 1890 Hz) and 12th (C#7+ 2268 Hz).

7. KHOMEI STYLES

Although there is no widespread agreement, *Khomei* comprises three major basic Throat-Singing methods called *Khomei*, *Kargyraa*, and *Sygyt,* two main sub methods called *Borbangnadyr* and *Ezengileer* and various other sub styles.

Khomei means "throat" or "pharynx" and it is not only the generic name given to all throat-singing styles for Central Asia, as underline above, but also a particular style of singing. *Khomei* is the easiest technique to learn and the most practiced in the West. It produces clear and mild harmonics with a fundamental usually within the medium range of the singer's voice (Fig. 6). In *Khomei* style there are two (or more) notes clearly audible. Technically the stomach remains relaxed and there is a low-level tension on larynx and ventricular folds, whereas *Sygyt* style requires a very strong constraint of these organs (Fig. 7). The tongue remains seated flatly between the lower teeth as in the Single Cavity technique, or raises and moves as in the Two Cavities techniques. The selection of the desired harmonic comes mainly from a combination of different lips, tongue and throat movements.

Sygyt means "whistle" and actually sounds like a flute. This style is characterized by a strong, even piercing, harmonic and can be used to perform complex and very distinct melodies (Fig. 10). It has its roots in the *Khomei* method and has the same range for the fundamental. *Sygyt* is sung with a half-open mouth and the tip of tongue placed behind front teeth as if pronouncing the letter "L". The tongue tip is kept in the described position, while the tongue body moves to select the harmonic. This is the same technique described above for the *Khomei* method. The difference is in the timbre quality of the sound lacking of energy in the low frequencies. To produce a crystal-clear, flute-like overtone, characteristic of the *Sygyt* style, it is necessary to learn how to filter out the lower harmonic components, that usually mask the overtone sensation.

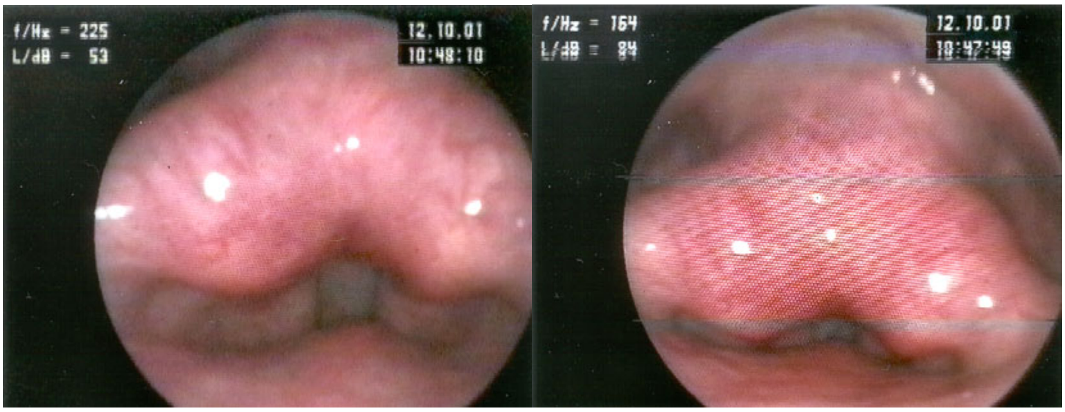

Figure 7. Position of the arytenoids in *Khomei* (left) and *Sygyt* style [21].

Crucial for achieving this goal is a considerable pressure from the belly/diaphragm, acting as a bellows to force the air through the throat. Significant tension is required in the throat as well, to bring the arytenoids near the root of the epiglottis (Fig. 7). In this way, we obtain the displacement of first 3 formants in the high frequency zone (Fig. 3). The result is that the fundamental and the lower harmonics are so attenuated to be little audible (Fig. 10).

It is possible to sing *Sygyt* either directly through the center of the mouth, or, tilting the tongue, to one side or the other. Many of the best *Sygyt* singers "sing to the side": directing the sound along the hard surfaces of the teeth enhances the bright, focused quality of the sound.

Kargyraa style produces an extremely low sound that resembles the roaring of a lion, the howling of a wolf, and the croaking of a frog and all these mixed together (Fig. 9). *Kargyraa* means "hoarse voice". As hawking and clearing the throat before speaking *Kargyraa* is nothing else than a deep and continuous hawking. This hawking must rise from the deepest part of the windpipe; consequently low tones will start resonating in the chest. Overtones are amplified by varying the shape of the mouth cavity and the position of the tongue. *Kargyraa* is closely linked to vowel sounds: the selection of diphonic harmonic corresponds to the articulation of a particular vowel (/u/, /o/, /ɔ/, /a/, etc.), which the singer learnt to associate with the desired note.

This technique is a mixture of Diphonia and Bitonality (see 6.1): in fact the supraglottal structures start to vibrate with the vocal folds, but at a half rate. The arytenoids also can vibrate touching the root of the epiglottis, hiding the vocal folds and forming a second "glottic" source [21]. The perceived pitch will be one octave lower than normal (Fig. 9), but also one octave and a 5^{th} lower [20]. In the case of Tran Quang Hai voice, the fibroendoscopy reveals the vibration and the strong constriction of the arytenoids that hide completely the vocal folds (Fig. 8).

We must distinguish this technique from the Tibetan Buddhist chant, which is produced with the vocal folds relaxed as possible, and without any supraglottal vibration. The Tibetan chant is more like the Tuvan *Borbangnadyr* style with low fundamentals.

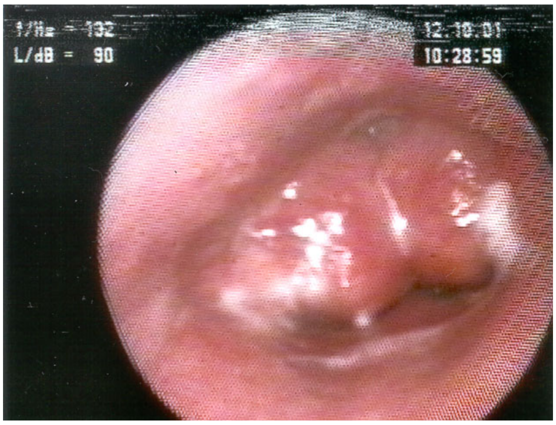

Figure 8. Simulation of the *Kargyraa* style by Tran Quaang Hai: the arytenoids move against the root of the epiglottis and hide the vocal folds [21].

Borbangnadyr is not really a style, as are *Khomei*, *Sygyt* and *Kargyraa*, but rather a combination of effects applied to one of the other styles. The name comes from the Tuvan word for “rolling”, because this style features highly acrobatic trills and warbles, reminiscent of birds, babbling brooks, etc. While the name *Borbangnadyr* is currently most often used to describe a warbling applied to *Sygyt*, it is also applied to some lower-pitched singing styles, especially in older texts. The *Borbangnadyr* style with low fundamentals sounds like the Tibetan Buddhist chant.

Rather the pitch movement of the melody, *Borbangnadyr* generally focuses the attention on three different harmonics, the 8th, 9th, and 10th, which periodically take their turn in prominence (Fig. 11). In this style the singer easily can create a triphonia effect between the fundamental, a second sound corresponding to the 3rd harmonic at an interval of 5th, and the tremolo effect on the higher harmonics.

Ezengileer comes from a word meaning “stirrup” and features rhythmic harmonic oscillations intended to mimic the sound of metal stirrups, clinking to the beat of a galloping horse (Fig. 12). *Ezengileer* is a variant of *Sygyt* style and differs considerably from singer to singer, the common element being the “horse-rhythm” of the harmonics.

8. OVERTONE SINGING IN THE WEST

In the West the Overtone Singing technique has unexpectedly become very popular, starting into musical contests and turning very soon to mystical, spiritual and also therapeutic applications. The first to make use of a diphonic vocal technique in music was Karlheinz Stockhausen in Stimmung [22]. He was followed by numerous artists and amongst them: the EVTE (Extended Vocal Techniques Ensemble) group at the San Diego University in 1972, Laneri and his Prima Materia group in 1973, Tran Quang Hai in 1975, Demetrio Stratos in 1977 [17-18], Meredith Monk in 1980, David Hykes and his Harmonic Choir in 1983 [23], Joan La Barbara in 1985, Michael Vetter in 1985, Christian Bollmann in 1985, Noah Pikes in 1985, Michael Reimann in 1986, Tamia in 1987, Bodjo Pinek in 1987, Josephine Truman in 1987, Quatuor Nomad in 1989, Iegor Reznikoff in 1989, Valentin Clastrier in 1990, Rollin Rachele in 1990 [24], Thomas Clements in 1990, Sarah Hopkins in 1990, Les Voix Diphoniques in 1997.

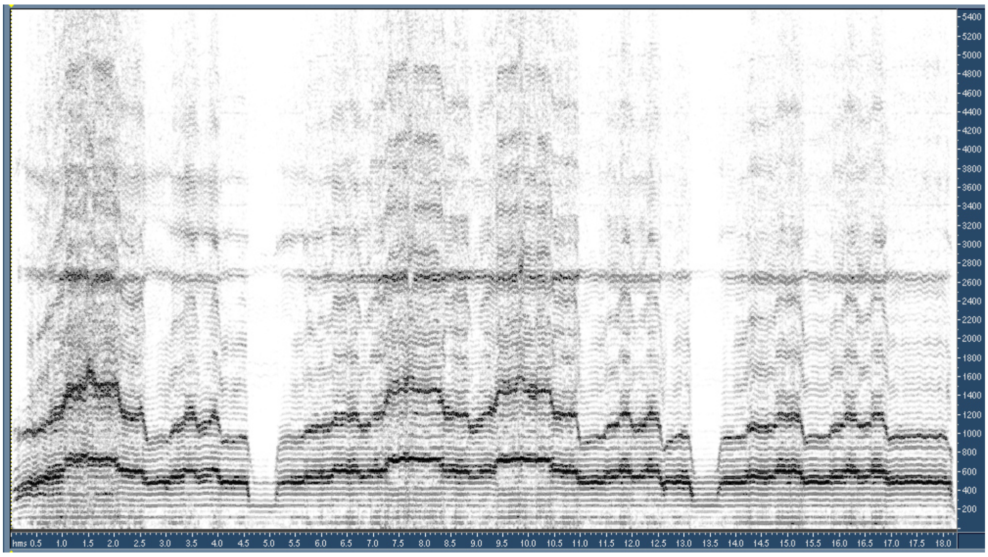

Figure 9. Vasili Chazir sings "*Artii-sayir*" in the *Kargyraa* Tuvan style. The fundamental pitch is B1 61.2 Hz. The diphonic harmonics are the 6th (F#4- 367 HZ), 8th (B4 490 Hz), 9th (C#5 550 Hz), 10th (D#5- 612 Hz) and 12th (F#5- 734 Hz). The diphonic (but not perceptible) harmonics 12th-24th are in octave with the previous one. In the 2600-2700 Hz region, a steady formant amplifies the 43rd and 44th harmonics.

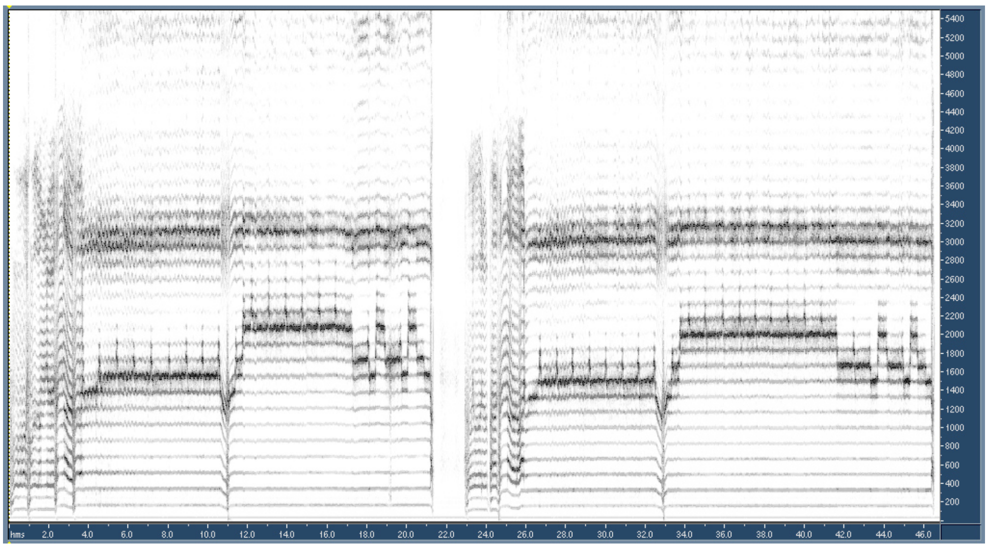

Figure 10. Tuvan *Sygyt* style. The fundamental is a weak E3+ 167 Hz. The melody uses the 8th (E6+ 1336 Hz), 9th (F#6+ 1503 Hz), 10th (G#6+ 1670 Hz) and 12th (B6+ 2004 Hz). There is a rhythmic shift between contiguous harmonics each 900 ms. In the 3000-3200 Hz zone, we can see a second resonance region.

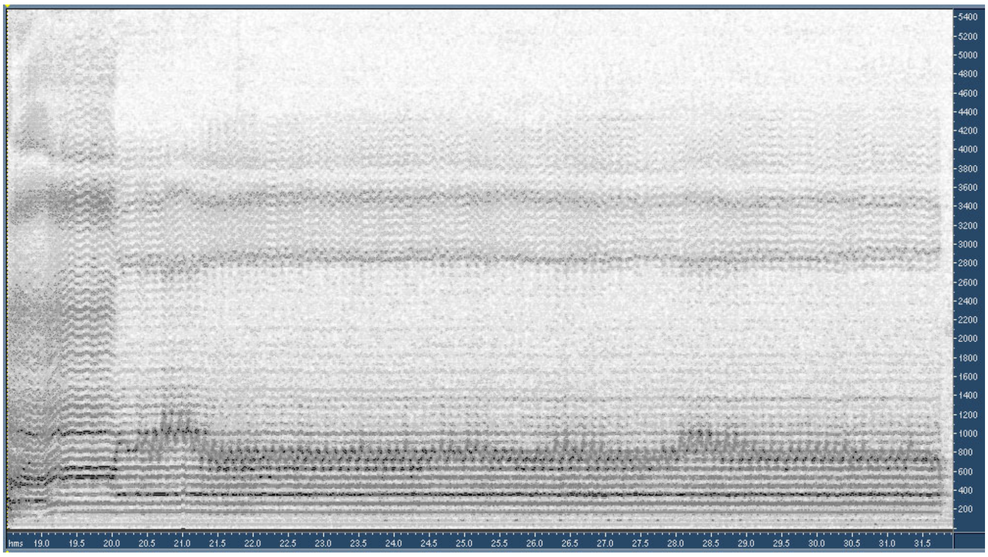

Figure 11. Tuvan *Borbangnadyr* style. The fundamental is a weak F#2 92 Hz. We can see on the harmonics 7-11 the effect of a periodic formantic shift (6 Hz about).

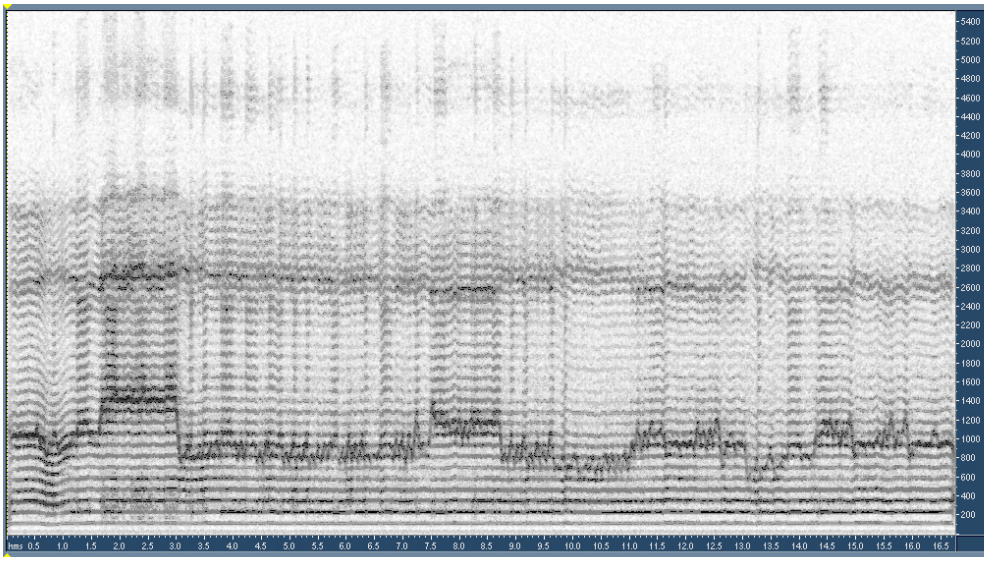

Figure 12. Tuvan *Ezengileer* style. The fundamental is A#2 117 Hz.

The most famous proponent of this type of singing is David Hykes. Hykes experimented with numerous innovations including changing the fundamental (moveable drone) and keeping fixed the diphonic formant, introducing text, glissando effects, etc., in numerous works produced with the Harmonic Choir of New York (Fig. 15) [23].

9. ACOUSTIC ANALYSIS

In the recent past, some work has been done on the analysis of *Khomei*, and more has been done on Overtone Singing generally. The focus on this research has been on the effort to discover exactly how overtone melodies are produced. Hypotheses as to the mechanics of Overtone Singing range from ideas as to the necessary physical stance and posture used by the singer during a performance, to the actual physical formation of the mouth cavity in producing the overtones.

Aksenov was the first to explain the diphonia as the result of the filtering action of the vocal tract [25-27]. Some years later Smith *et al.* engaged in an acoustical analysis of the Tibetan Chant [28]. In 1971, Leipp published an interesting report on *Khomei* [29]. Tran Quang Hai carried out a deep research on all the diphonic techniques [4-5][30]. The mechanism of the diphonia was demonstrated in 1989 by two different methodologies. The first applied direct clinical-instrumental methods to study the vocal tract and vocal cords [31-32]. The optic stroboscope revealed the perfect regularity of the vocal folds vibration. The second method made use of a simple linear prediction model (LPC) to analyse and synthesize the diphonic sound [33-34]. The good quality of the resynthesis demonstrated that the diphonia is due exclusively to the spectral resonance envelope. The only difference between normal and diphonic sound consists in the unusual narrow bandwidth of the prominent formant.

Several researchers seem to agree that the production of the harmonics in Throat-Singing is essentially the same as the production of an ordinary vowel. Bloothooft reports an entire investigation of Overtone Singing, based on the similarity of this kind of phonation to the articulation of vowel [10].

Other authors, on the contrary, argue that the physical act of creating overtones may originate in vowel production, but the end product, the actual overtones themselves, are far from vowel-like [35]. They stated, in fact, that for both acoustic and perceptual reasons, the production of an overtone melody cannot be described as vowel production.

Acoustically, a vowel is distinctive because of its formant structure. In Overtone Singing, the diphonic formant is reduced to one or a few harmonics, often with surrounding harmonics attenuated as much as possible. Perceptually, Overtone Singing usually sounds nothing like an identifiable vowel. This is primarily because, a major part of the overtone-sung tone has switched from contributing to the timbre of the tone to provoking the sensation of melody and such a distorted "vowel" can convey little phonetic information.

10. CONCLUDING REMARKS

All musical sounds contain overtones or tones that resonate in fixed relationships above a fundamental frequency. These overtones create tone color, and help us to differentiate the sounds of different music instruments or one voice and another.

Different cultures have unique manifestations of musical traditions, but, what it is quite interesting, is that some of them share at least one aspect in common: the production of overtones in their respective vocal music styles. Among these, each tradition has also its own meanings and resultants from Overtone Singing, but they are often related to a common sphere of spirituality. Overtones in Tibetan and Gregorian Chant, for example, are linked with spirituality, and even health and well being. Overtones in Tuvan *Khomei* have at least three different meanings: shamanistic, animistic, and aesthetic.

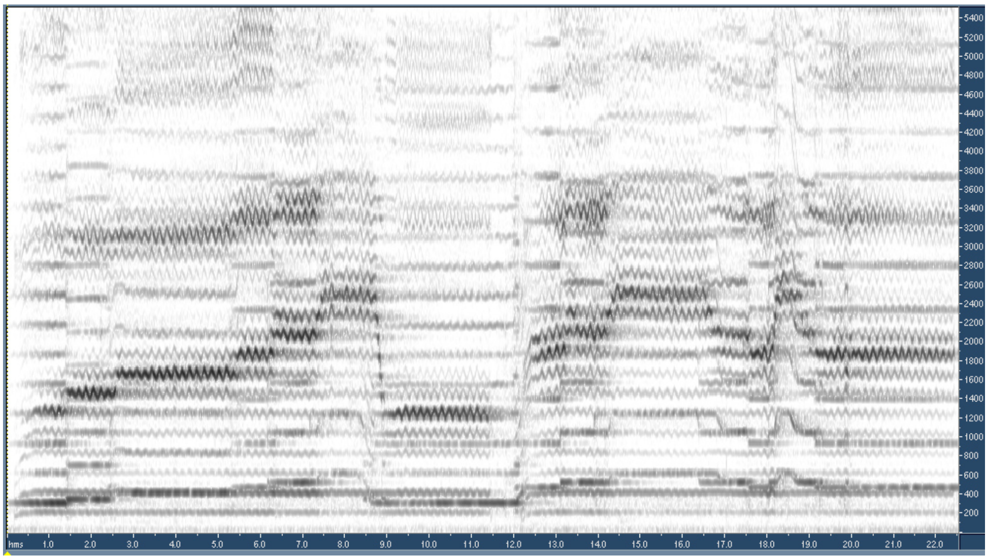

Figure 13. Mongolia: Ganbold sings a *Kevliin Xöömi* (ventral *Xöömi*, similar to Tuvan *Sygyt*.). The pitch is G3# 208 Hz. The diphonic harmonics are 6th (D#6 1248 Hz), 7th (F#6-1456 Hz), 8th (G#6 1664 Hz), 9th (A#6+ 1872 Hz), 10th (C7- 2080 Hz), 12th (D#7 2496 Hz). There is a 6 Hz strong vibrato.

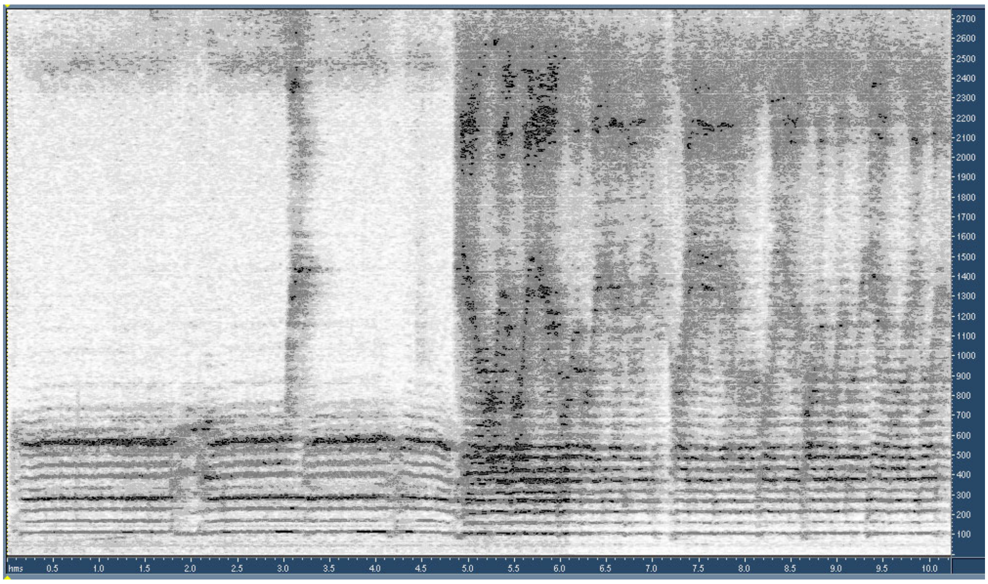

Figure 14. Tibetan Gyuto Chant in the *Yang* style. The pitch is a weak A1 56 Hz. In the beginning, the singer chant a vowel /o/ that reinforces the 5th partial (and the 10th). In the choir part, the articulation of the prayers produces a periodic emerging of all the scale of the harmonics up to the 30th. There is also a fixed resonance at 2200 Hz.

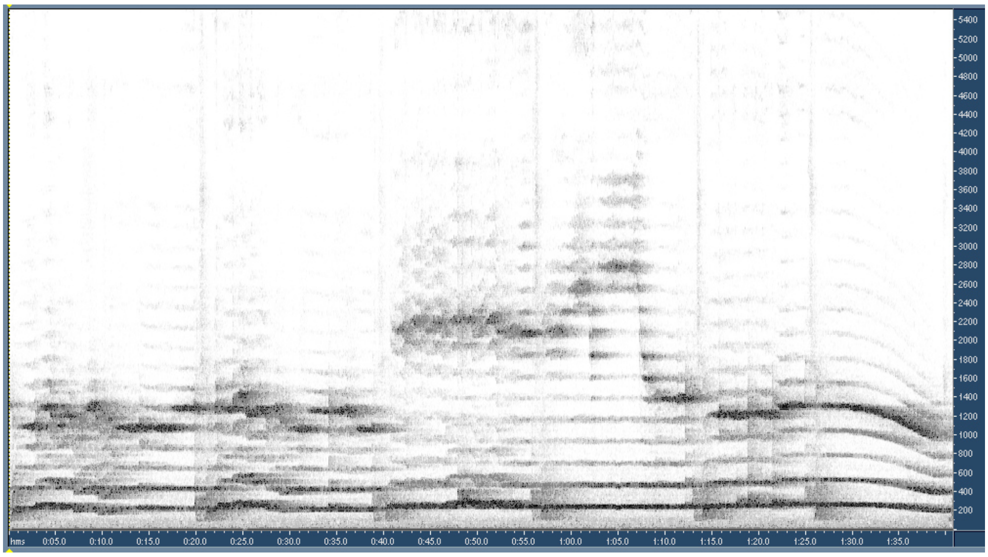

Figure 15. David Hykes and the Harmonic Choir. In this 100 s passage from "Hearing the Solar Winds" [23], the pitch moves slowly from A3, A#3, B3, C4, A3, to the final G3. The diphonic harmonics change in the range 6^{th}-12^{th}.

11. ACKNOWLEDGMENTS

We would like to thank Sami Jansson [36] and Steve Sklar [15] for the useful information they made available to us via their respective web sites.

REFERENCES

[1] Feynman (http://www.feynmanonline.com/), website.
[2] Friends of Tuva (http://www.fotuva.org/), website.
[3] Dargie D., "Some Recent Discoveries and Recordings in Xhosa Music", *5th Symposium on Ethnomusicology*, University of Cape Town, International Library of African Music (ed) , Grahamtown, 1985, pp. 29-35.
[4] Tran Quang Hai, *Musique Touva*, 2000, (http://www.baotram.ovh.org/tuva.html), website.
[5] Tran Quang Hai, Zemp H.,"Recherches expérimentales sur le Chant Diphonique", *Cahiers de Musiques Traditionnelles*, Vol. 4, Genève, 1991, pp. 27-68.
[6] Levin Th., Edgerton M., *The Throat Singers of Tuva,* 1999, (http://www.sciam.com/1999/0999issue/0999levin.html), website
[7] Walcott R., "The Chöömij of Mongolia - A spectral analysis of Overtone Singing", *Selected Reports in Ethnomusicology*, UCLA, Los Angeles, 1974, 2 (1), pp. 55-59.
[8] Bregman A., *Auditory scene analysis: the perceptual organization of sound*, MIT Press, Cambridge, 1990.
[9] Sundberg J., *The science of the singing voice*, Northern Illinois University Press, De Kalb, Illinois, 1987.

[10] Bloothooft G., Bringmann E., van Capellen M., van Luipen J.B., Thomassen K.P., “Acoustic and Perception of Overtone Singing”. In *Journal of the Acoustical Society of America*, JASA Vol. 92, No. 4, Part 1, 1992, pp. 1827-1836.
[11] Stevens K., *Acoustic Phonetics*, MIT Press, Cambridge, 1998.
[12] Fant G., *Acoustic theory of speech production*, Mouton, The Hague, 1960.
[13] Lortat-Jacob B., “En accord. Polyphonies de Sardaigne: 4 voix qui n’en font qu’une”, *Cahiers de Musiques Traditionnelles*, Genève, 1993, Vol. 6, pp. 69-86.
[14] Kavasch D., “An introduction to extended vocal techniques”, *Report of CME*, Univ. of California, San Diego, Vol. 1, n. 2, 1980, pp. 1-20.
[15] Sklar S., Khöömei Overtone Singing, (http://www.atech.org/khoomei), website.
[16] Ferrero F., Ricci Maccarini A., Tisato G., “I suoni multifonici nella voce umana”, *Prooceedings of XIX Convegno AIA*, Napoli, 1991, pp. 415-422.
[17] Ferrero F., Croatto L., Accordi M., “Descrizione elettroacustica di alcuni tipi di vocalizzo di Demetrio Stratos”, *Rivista Italiana di Acustica*, Vol. IV, n. 3, 1980, pp. 229-258.
[18] Stratos D., *Cantare la voce*, Cramps Records CRSCD 119, 1978.
[19] Dmitriev L., Chernov B., Maslow V., “Functioning of the voice mechanism in double voice Touvinian singing”, *Folia Phoniatrica*, Vol. 35, 1983, pp. 193-197.
[20] Fuks L., Hammarberg B., Sundberg J., “A self-sustained vocal-ventricular phonation mode: acoustical, aerodynamic and glottographic evidences”, *KTH TMH-QPSR*, n.3, Stockholm, 1998, pp. 49-59.
[21] Tisato G., Ricci Maccarini A., Tran Quang Hai, “Caratteristiche fisiologiche e acustiche del Canto Difonico”, *Proceedings of II Convegno Internazionale di Foniatria*, Ravenna, 2001, (to be printed).
[22] Stockhausen K., *Stimmung*, Hyperion A66115, 1968.
[23] Hykes D., *David Hykes and the Harmonic Choir*, (http://harmonicworld.com), website.
[24] Rachele R., “Overtone Singing Study Guide”, *Cryptic Voices Productions* (ed), Amsterdam, 1996, pp. 1-127.
[25] Aksenov A.N., *Tuvinskaja narodnaja muzyka*, Mosca, 1964.
[26] Aksenov A.N., “Die stile der Tuvinischen zweistimmigen sologesanges”, *Sowjetische Volkslied und Volksmusikforschung*, Berlin, 1967, pp. 293-308.
[27] Aksenov A.N., “Tuvin folk music”, *Journal of the Society for Asian Music*, Vol. 4, n. 2, New York, 1973, pp. 7-18.
[28] Smith H., Stevens K.N., Tomlinson R.S., “On an unusual mode of singing of certain Tibetan Lamas”, *Journal of Acoustical Society of America*, JASA. 41 (5) , USA, 1967, pp. 1262-4.
[29] Leipp M., “Le problème acoustique du Chant Diphonique”, *Bulletin Groupe d’Acoustique Musicale*, Univ. de Paris VI, n. 58, 1971, pp. 1-10.
[30] Tran Quang Hai, “Réalisation du chant diphonique”, *Le Chant diphonique*, Institut de la Voix, Limoges, dossier n° 1, 1989, pp. 15-16.
[31] Pailler J.P., “Examen video du larynx et de la cavité buccale de Monsieur Trân Quang Hai”, *Le Chant Diphonique*, Institut de la Voix, Limoges, dossier n° 1, 1989, pp. 11-13.
[32] Sauvage J.P., “Observation clinique de Monsieur Trân Quang Hai”, *Le Chant Diphonique*, Institut de la Voix, Limoges, dossier n° 1, 1989, pp. 3-10.
[33] Tisato G., “Analisi e sintesi del Canto Difonico”, *Proceedings VII Colloquio di Informatica Musicale (CIM)*, Cagliari, 1989, pp. 33-51.
[34] Tisato G., Ricci Maccarini A., “Analysis and synthesis of Diphonic Singing”, *Bulletin d'Audiophonologie*, Vol. 7, n. 5-6, Besançon, 1991, pp. 619-648.
[35] Finchum H., Tuvan Overtone Singing: Harmonics Out of Place, (http://www.indiana.edu/~folklore/savail/tuva.html), website.
[36] Jansson S., Khöömei Page (http://www.cc.jyu.fi/~sjansson/khoomei.htm), website.
[37] Leothaud G., “Considérations acoustiques et musicales sur le Chant Diphonique”, *Le Chant Diphonique*, Institut de la Voix, Limoges, dossier n° 1, 1989, pp. 17-43.
[38] Zarlino G., *Istitutioni Harmoniche*, Venice, 1558.