Scales in traditional solo singing: models and results

Rytis Ambrazevičius
Department of Ethnomusicology, Lithuanian Academy of Music, Vilnius, Lithuania
rytisam@delfi.lt

In: R. Parncutt, A. Kessler & F. Zimmer (Eds.)
Proceedings of the Conference on Interdisciplinary Musicology (CIM04)
Graz/Austria, 15-18 April, 2004 http://gewi.uni-graz.at/~cim04/

Background in psychoacoustics. Description of pitch scaling in solo singing without accompaniment is problematic because of slightly floating tonality and zonal intonation. It is well known that zonal intonation results from categorical perception. Usage of the zone depends on aspects of performance (e.g. Sundberg 1999) as well as on melodic context (Ambrazevičius 2001). Slightly floating tonality and zonal intonation are especially inherent for archaic traditional solo singing (e.g. Alexeyev 1986); by 'archaic' we mean 'unbound to the Western musical thinking'.

Background in ethnomusicology. It is widely known that archaic traditional solo singing features pre-diatonic scales different in interval structure from the contemporary pitch scales. These scales are mostly based on contrast of voice registers and timbre or made of approximately equidistant steps (α- or γ-types, by Alexeyev 1986). Perception of pitch scale of a distant musical culture results in distortion of the original structure (e.g. Chenoweth 1972). It is a manifestation of also widely known collision of two phonemic systems (e.g. Nettl 1964) that bears inadequacy in transcription (Reid 1977).

Aims

- To develop mathematical model aiming for recognition of an insider's phonemic scaling, i.e. for its reconstruction from acoustical pitch measurements (elimination of masking factors of slightly floating tonality and zonal intonation).
- Also, for comparison, to develop the model for an outsider's (here, researcher's with Western background) perception of a traditional scale. This model is based on the perceptual seek for the best fit of sounding material to the template of equal temperament.
- To study, by means of the models, the regularities of scaling in one idiolect and its perception.

Method. The mathematical models described above were developed. Pitches of songs belonging to one idiolect of Lithuanian traditional male solo singing were measured (computer software of acoustical analysis 'Speech Analyser' was used). The results of acoustical measurements were reconsidered by means of the modeling.

Results and conclusions. Reconstructed scales show different interval structures characteristic of the ancient pre-diatonic and modern diatonic (or even 'equally tempered') musical thinkings. It means that there are different historical layers of musical thinkings manifested in the idiolect.
It is also demonstrated how 'aural ghosts' conditioned by diatonic templates work in perception of traditional scaling by a contemporary outsider and, consequently, how they mislead to false systematization of traditional scales.
Resemblances to interval structures in Lithuanian vocal *Schwebungsdiaphonie* are discussed.

Samples and measurements

Repertoire of one male singer was chosen for examination of scale phenomena. The singer Jonas Jakubauskas (1908-2000) was a prominent bearer of the *Dzūkai* (Southern Lithuania) vocal tradition.

Software Speech Analyser was applied for acoustic measurements. Pitches of structural sounds of 12 recorded songs were measured. Here one of the songs (Fig. 1) was chosen for illustration of the modeling and recognition of scale features.

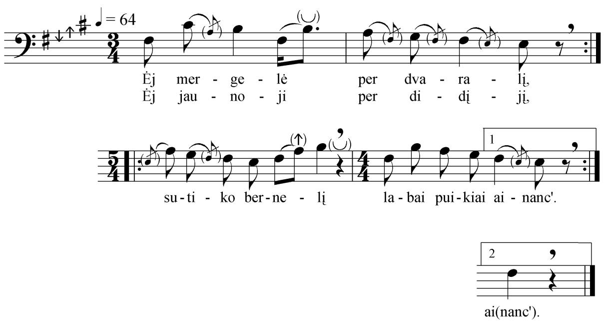

Figure 1. Conventional transcription of the song *Ėj mergelė per dvaralį*. The scale given by key signature is specified according to the mean values of the acoustic measurements. The arrows mark slight deviations of pitch.

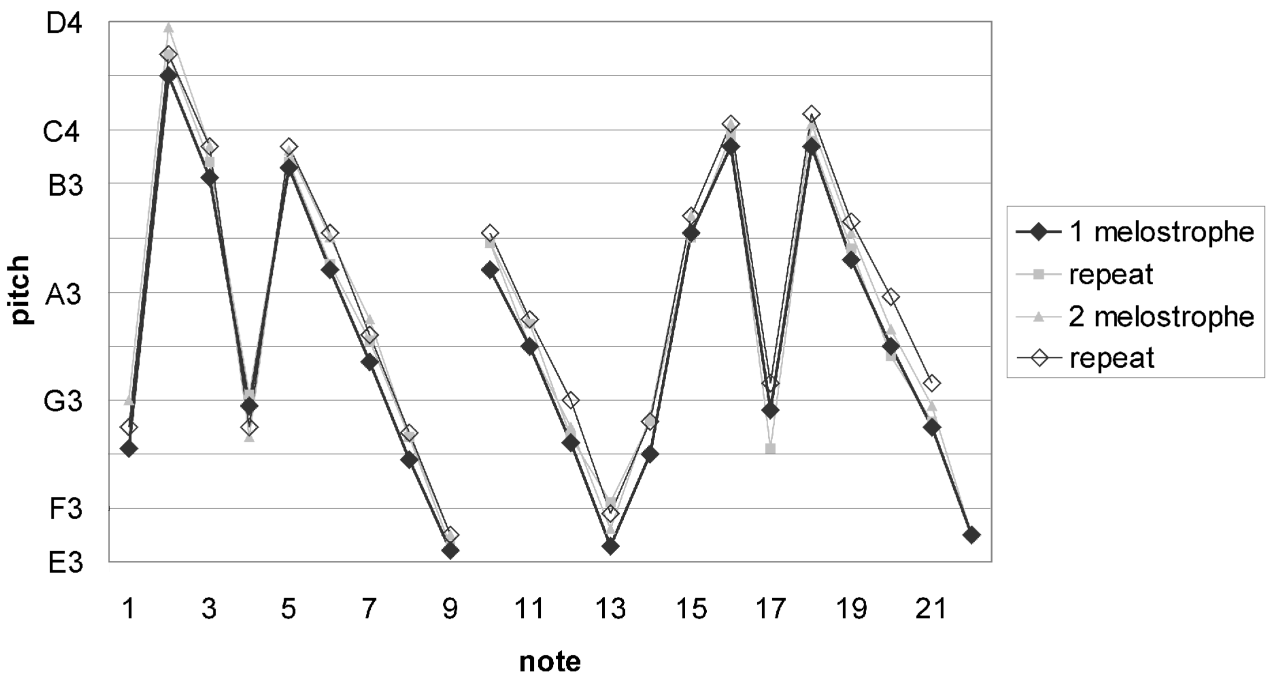

Figure 2. Measured pitch tracks of the structural notes (see Fig. 1). Two melostrophes with repeats are presented. The first and the last tracks are picked out.

Insider's scale

Gradual transposition. Fig. 2 shows that the last realization of the melody contour is slightly higher in pitch than the first one. This fact depicts the phenomenon of gradual transposition (usually gradual rise of pitch) common in traditional singing without accompaniment. More thorough examination reveals that the fluctuating gradual transposition manifests not only when proceeding from one melostrophe to the next one, but inside a separate melostrophe as well.

Since a performer does not perceive the transposition of scale as change of scale, the results of pitch measurements cannot be applied directly for the reconstruction of the performer's scale. They need additional mathematical modeling providing elimination of transposition.

Modeling. The model is based on the following consideration. A performer presents an intended invariant of pitch track, with slight and tolerable deviations, plus slightly fluctuating transposition. The assumption is made that the realizations of the invariant are "as precise as possible". In mathematical terms, it means the problem of minimizing.

Say, a_{ij} marks absolute pitch of the *j*th note in *i*th realization (melostrophe or its repeat). For instance, if refer to Fig. 2, a_{25} marks pitch of the fifth note in the repeat of the first melostrophe. Say, t_{ij} stands for the pitch of a reference tone (e.g. tonal center) in *ij*th position. Change of t_{ij} from note to note means the fluctuating transposition. The relative pitch (i.e. if transposition is eliminated) $b_{ij} = a_{ij} - t_{ij}$.

Say then, $<b_j>$ is the invariant pitch of the *j*th note in melody contour. Then the task of minimizing is presented by the following expression:

$$\sum_{i=1}^{M}\sum_{j=1}^{N}(a_{ij}-t_{ij}-<b_j>)^2+\sum_{i=1}^{M}\sum_{j=1}^{N-1}(t_{ij+1}-t_{ij})^2 = \text{min};$$

here *M* stands for the number of realizations (e.g. *M*=4 in Fig. 2) and *N* stands for the number of structural notes in melody contour (e.g. *N*=22 in Fig. 2). The first member on the left side represents the imprecision of intonation, and the second member represents the gradual transposition[1].

Results. The model based on the algorithm of minimizing allows reconstruction of the invariant pitch track. Consequently, the inherent scale can be evaluated. Fig. 3 shows that there is no diatonic contrast of whole tone / semitone. Mean intervals between the adjacent scale steps range from 162 to 200 cents, approximately 180 cents in average. Thus the scale can be treated as "squeezed anhemitonic".

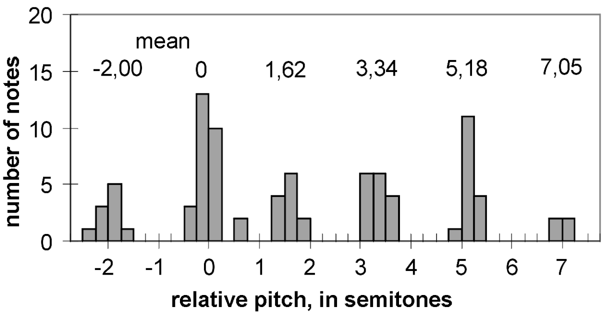

Figure 3. Pitch distribution based on data of Fig. 2 and recalculated by means of the mathematical modeling of insider's scale. Relative pitch in tempered semitones to the mean intonation of the tonal center is shown.

Outsider's scale

Modeling. An outsider, i.e. listener perceiving and translating the scale of traditional music, is usually musicologist with Western musical background. Thus he or she is bound to the equal temperament, and the perceived scale tends to be forced to the template of the equal temperament. It can be reasonably assumed that perception seeks for the best fit of a sounding material to the equal temperament. Tones that make intervals close to tempered values work as sort of perceptual supports. The rest tones are perceived as being "out of tune"[2].

Evaluation of a certain pitch bases on comparison with neighboring tones (relative pitch) or/and with a reference tone (absolute pitch). If the relative recognition prevails, the nearest neighbors have naturally the weightiest impact on the evaluation.

The present model considers the relative recognition and it is of an illustrative character only.

Initially weight $w_0 = 1$ is fixed for each note. Intervals are said to be close to tempered values if the deviations from the tempered equivalents do not exceed 20 cents[3]. The corresponding notes work as pitch supports; the weights of both notes are increased by 1 if these notes are adjacent in melody contour, by 1/2 if there is one note in between, by 1/4 if there are two notes in between, and so on (1/8, 1/16,..). If a certain note makes several (almost) tempered intervals with different notes, the corresponding weights are summed up.

Evaluation of subjective pitches b_{ij} is then carried out:

$$b_{ij}=\frac{\sum_{\substack{k=1\\k\neq j}}^{N}2^{-|j-k|}w_{ik}(s_{ik}+a_{ij}-a_{ik})}{\sum_{\substack{k=1\\k\neq j}}^{N}2^{-|j-k|}w_{ik}};$$

here a_{ij} and a_{ik} stand for absolute pitches; w_{ik} means note weight; the factor $2^{-|j-k|}$ expresses the additional weighting decreasing with the increasing time span between notes. So the contribution of the closest neighbors is boosted. Value s_{ik} means perceived semitonal pitch class, e.g. $s_{ik} = 0$, 3, and 4 denote, correspondingly, the tonal center, minor and major thirds.

If the calculated value of b_{ij} is lower or higher than the tempered value, but not more than 20 cents, it is perceived as tempered pitch. If b_{ij} belongs to the range of 20 cents around the midpoint between the tempered values, the pitch is uncertain[4]. For instance, $b_{ij} = 2.8–3.2$ gives minor third, =3.2–3.4 gives "high" minor third, $b_{ij} = 3.4–3.6$ gives uncertain or neutral third, $b_{ij} = 3.6–3.8$ gives "low" major third, and $b_{ij} = 3.8–4.2$ gives major third.

Results. Distribution of outsider's subjective pitches (Fig. 4) shows that tonal center, fourth, fifth and subsecond are mostly perceived as tempered equivalents. The third scale step is perceived as tempered or raised minor third. Perception of the second scale step is probably the most complex. The second splits into two versions: major (tempered or lowered) and minor (raised). It means that alternation of Aeolian and Phrygian modes is perceived, i.e. chromatic change occurs.

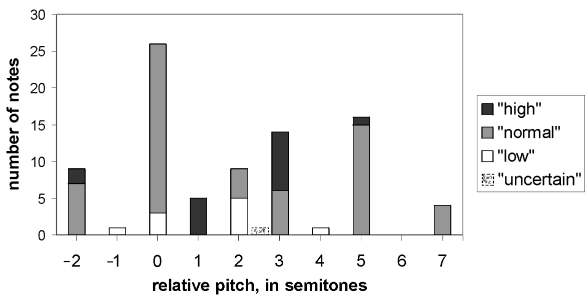

Figure 4. Pitch distribution based on data of Fig. 2 and recalculated by means of the mathematical modeling of outsider's scale. Deviations from the equal temperament are generalized as lower, higher or uncertain intonations. Relative pitch in tempered semitones to the mean intonation of the tonal center is shown.

Discussion

The inherent scale structure of one traditional vocal solo performance is studied. It shows approximately equidistant intervals between scale steps, 180 cents in average. The zone of intonation is rather wide. The detailed study of most samples belonging to the repertoire of the same singer results in similar conclusions. Some of the samples, however, show strong contribution of modern diatonic musical thinking (Fig. 5). Here the diatonic contrast of whole tone / semitone is clearly observed.

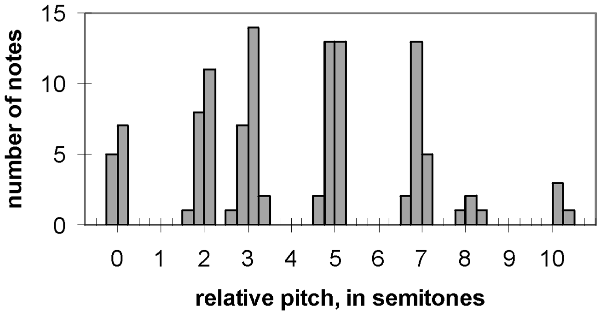

Figure 5. Pitch distribution of song *Arškėtinė šamet rožė* from Jakubauskas repertoire. The histogram is based on data of acoustical measurements and recalculated by means of the mathematical modeling of insider's scale. Relative pitch in tempered semitones to the mean intonation of the tonal center is shown.

In several samples even sort of mixture of scale templates is observed. For instance, the lower part of the scale shows ancient pre-diatonic "squeezed anhemitonics" whereas the upper part resembles to modern diatonics[5]. Thus different historical layers of musical thinking manifest in the repertoire of the singer.

The revealed features of scale show straight correspondence to Alexeyev's (1986) theory of development of modality. The results of the present investigation serve as objective verification (based on acoustical measurements) of the theory (based on aural experience).

My recent study shows similar results for vocal polyphonic *sutartinės* (Lithuanian type of *Schwebungsdiaphonie*) as well as for tunings in Lithuanian archaic instrumental music. For instance, big part of *sutartinės* is characteristic of so-called parallelism of seconds, i.e. mostly seconds appear between two voices sounding simultaneously. The second is neither minor nor major; it equals 180 cents in average. So this is the value found in the present investigation as well. On the other hand, psychoacoustic origin can be envisaged as well. It seems that maximum roughness occurs for these interval values, and this corresponds to the ideal of sounding in *Schwebungsdiaphonie* described as perfect "clash" of voices (Brandl, 1989).

The comparison of the inherent (insider's) scale and its aural interpretation by a musicologist (outsider) reveals the process of making "aural ghosts". Different fictitious diatonic scales are construed, and chromatic change is found due to diatonic apperception. It can result in false transcriptions and, consequently, in false theoretical conclusions on the scale system in traditional music.

References

Alexeyev, E. (1986). *Rannefolklornoje intonirovanije. Zvukovysotnyj aspekt .* [Early folkloric intonation. Pitch aspect; in Russian] Moscow: Sovetskij kompozitor.

Ambrazevičius, R. (2001). On Non-Tempered Scaling in Lithuanian Traditional Singing. *Tiltai*, **8**, 14–24.

Brandl, R. (1989). Die Schwebungs-Diaphonie – aus musikethnologischer und systematisch-musikwissenschaftlicher Sicht. In: C. Eberhardt and G. Wei, eds. *Südosteuropa-Studien. B. 40, Schriftenreihe der Hohschule für Musik in München. Bd. 9. Volks- und Kunstmusik in Südosteuropa*. Regensburg. 51–67.

Chenoweth, V. (1972). *Melodic Perception and Analysis. A Manual on Ethnic Melody*. Papua New Guinea: Summer Institute of Linguistics Ukarumpa, E. H. D.

Locke, S., and L. Kellar (1973). Categorical Perception in a Non-linguistic Mode. *Cortex,* **9**, 355–369.

Nettl, B. (1964). *Theory and Method in Ethnomusicology*. London: Free Press of Glencoe.

Reid, J. (1977). Transcription in a New Mode. *Ethnomusiclogy*, **21/3**, 415–433.

Sundberg, J. (1999). The Perception of Singing. In: Diana Deutsch, ed. *Psychology of Music*, 2nd edition. San Diego, London: Academic Press. 171–214.

[1] Here the simplest case of invariant is discussed. Generally, a melody contour has several distinct versions, thus b_{ij} splits into several distinct values.

[2] They nevertheless fall into the semitonal pitch classes – even they are "very much out of tune". An uncertainty of pitch class occurs in narrow pitch range close to midpoint between centers of the adjacent classes. Locke and Kellar (1973) evaluated that the zone of uncertainty between minor and major thirds is some 20 cents only.

[3] As it was already mentioned, here all the numerical ratings are of an illustrative character only.

[4] In fact it hardly happens unless no choice of s_{ij} corresponds to b_{ij}.

[5] Possibly it results from associations with different melody patterns characteristic of different historical layers of the repertoire.