Proceedings of the Ninth Triennial Conference of the European Society for the Cognitive Sciences of Music, 17-22 August 2015, Manchester, UK
Ginsborg, J., Lamont, A., Phillips, M., Bramley, S. (Editors)

Dissonance/roughness in Lithuanian traditional *Schwebungsdiaphonie*

Rytis Ambrazevičius

Dept. of Audiovisual Arts, Kaunas University of Technology, Lithuania
Dept. of Ethnomusicology, Lithuanian Academy of Music and Theatre, Lithuania
rytis.ambrazevicius@ktu.lt

ABSTRACT

In contrast to Western art music, the dissonance-like sonorities in *Schwebungsdiaphonie*-cultures are at the core of the tonal structures. These cultures, although not abundant, are found in different locations all over the world (Cazden, Brandl, Messner, etc.). *Sutartinės* are a Lithuanian type of Schwebungsdiaphonie (Račiūnaitė-Vyčinienė, Ambrazevičius & Wiśniewska, etc.). On the one hand, the studies on psychoacoustic roughness and sensory dissonance are really big in number. The notions of roughness and sensory dissonance are usually considered as synonyms. On the other hand, it was proposed that ideal sounding of *Schwebungsdiaphonie* conforms to a maximum dissonance / roughness (Brandl, the diaphony in the Balkans and elsewhere; Ambrazevičius, the Lithuanian *Sutartinės*). In the present study, we analyze the occurrences of the notions of roughness and sensory dissonance in the psychoacoustic studies and define the case of *Sutartinės* in this context. The review of the experimental findings on the intervals corresponding to the maximum values of roughness / sensory dissonance reveals certain discrepancies between the concepts of roughness and sensory dissonance. It seems that, at least for a substantial frequency range, roughness is associated with larger interval sizes (Plomp & Levelt, Kameoka & Kuriyagawa, Terhardt, Zwicker & Fastl, Hutchinson & Knopoff, Vassilakis, etc.). Collation of these results and the findings of acoustical measurements of Sutartinė performances leads to the conclusion that the ideal vocal "clash" in *Sutartinės* most probably corresponds to psychoacoustic roughness, but not to sensory dissonance.

I. INTRODUCTION

Quite a few musical cultures favour dissonances (in terms of physiological acoustics) rather than consonances in their polyphonies. This is described as various types of psychoacoustically based "diaphony of beats" (*Schwebungsdiaphonie*) in some places (although not abundant) throughout the world (Cazden, 1945; Brandl, 1989; Messner, 1989; etc.).

Thus it is important that sonorities in the *Schwebungsdiaphonie*-cultures are governed by the phenomena opposite to those that are characteristic of the Western tonal music: there is a striving for (maximum?) dissonance (or roughness; see below) rather than consonance. In certain cases it could be stated that aesthetic standards and notions are somehow reversed. For instance, strong (in terms of roughness) "clashes" of seconds obtain positive connotations. Thus generally striving for the "native" consonance could be envisaged instead.

II. DISSONANCE AND ROUGHNESS

Sensory dissonance and roughness are two concepts used in psychoacoustic studies almost always as synonyms. The classical study of Plomp & Levelt (1965) could serve as a typical example of the presumed interchangeability of the two concepts. While the authors asked the subjects to judge intervals on the scale "consonant-dissonant" (or, in the case of incomprehension, they substituted the "consonant" with "beautiful" or "euphonious" instead; p. 553), they exploited both notions of dissonance and roughness unambiguously in their discourse. Incidentally, in many other studies, the questions presented to the participants are not revealed and the procedures of the experiments are not (or only faintly) detailed. Therefore the subjective sonic qualities meant and evaluated in the experiments remain obscure.

However, it is also argued that, even though roughness is one of the main constituents of sensory dissonance, it is not the only one. Moreover, several types of roughness are distinguished or in some cases the multidimensionality of roughness is suggested.[1]

Now we will take glance at the results of several studies on sensory dissonance / roughness. For instance, Ernst Terhardt (1968, p. 219) states that "the modulation frequency of maximum roughness increases with increasing carrier frequency initially and reaches a constant value f^*_{mod} = 75 Hz at carrier frequencies above approximately 2 kHz"[2] and presents the corresponding graph (see Figure 1). In his later study (1974), Terhardt claims the approximate identity of dissonance and roughness. However, there is some discrepancy between this claim and the factual results (Figure 2): it is clear that at least in the relevant spectral range the sense of roughness slightly differs from the sense of dissonance. Briefly, roughness is stronger for wider seconds and dissonance is stronger for narrower seconds.

[1] See forthcoming paper Ambrazevičius, 2015, for details.

[2] A number of studies employ AM (amplitude modulated) sine tones, while others use sine tone pairs. However, it is stated that the results do not differ significantly for the two cases (e.g. Terhardt, 1968, p. 219).

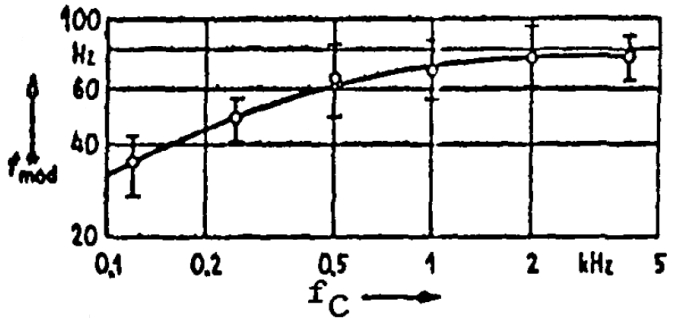


Figure 1. Modulation frequency for maximum roughness f^*_{mod} as function of carrier frequency f_c. Modulation factor is 1, *SPL* = 60 dB (Terhardt, 1968, p. 219).

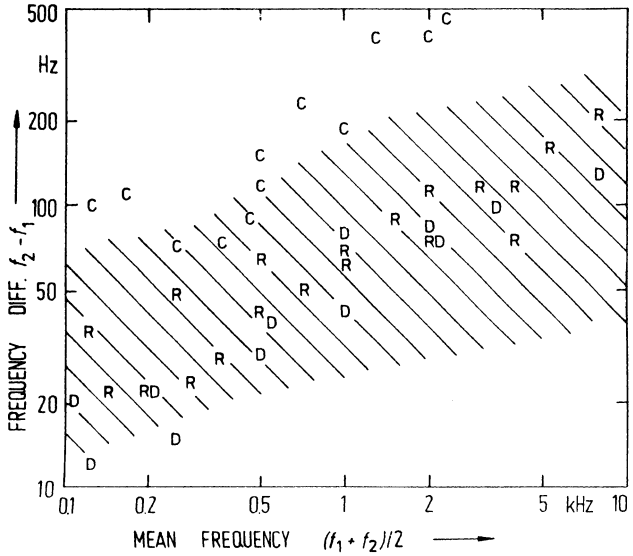


Figure 2. Dissonance, consonance, and roughness; according Terhardt, 1974, p. 1062. Hatching marks the area of pronounced roughness, "R", "D", and "C" mark, correspondingly, the maxima of roughness and dissonance, and the limit of appearance of consonance.

Findings of Andrzej Rakowski (1982) lead to the approximation for the frequency interval for maximum roughness as $2\sqrt{f}$. William Hutchinson and Leon Knopoff (1978) proposed noticeably different evaluation of dissonance. They designed their own approximation for the critical bandwidth as $1.72\ f^{0.65}$ and employed the Plomp's and Levelt's 1/4 CBW-criterion for the maximum dissonance.

The results of Marc Leman's model for roughness (2000) are presented in Figure 3. Pantelis N. Vassilakis (2001, p. 197–198) applied the model proposed earlier by William A. Sethares (1998; see the illustration from the second edition of his book on Figure 4). Finally, consider the evaluations by Fastl & Zwicker (Figure 5).

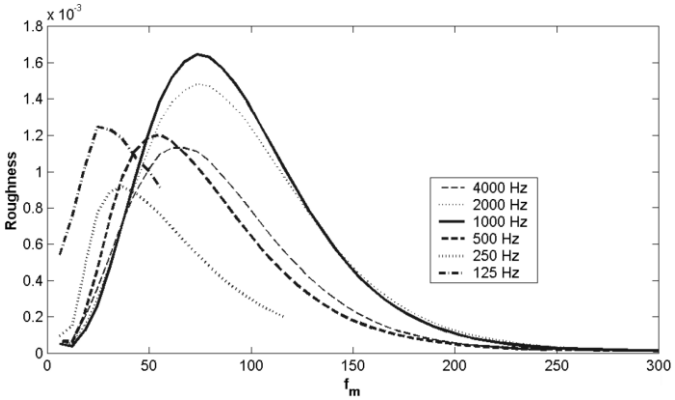


Figure 3. Roughness in function of modulation frequency of different carrier frequencies (modulation factor is 1; Leman, 2000, p. DAFX-5).

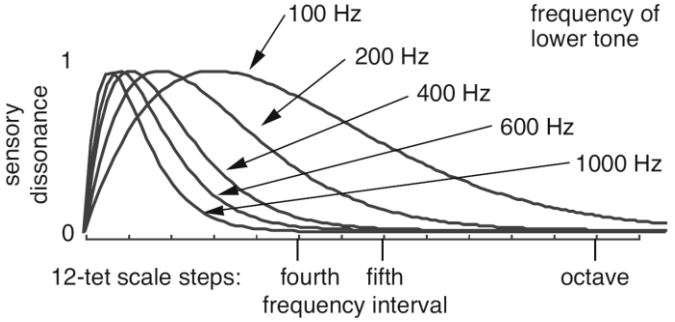


Figure 4. Sensory dissonance in function of frequency interval between two sine tones sounding simultaneously (Sethares, 2005, p. 47). Curves for different frequencies of the lower tone are presented.

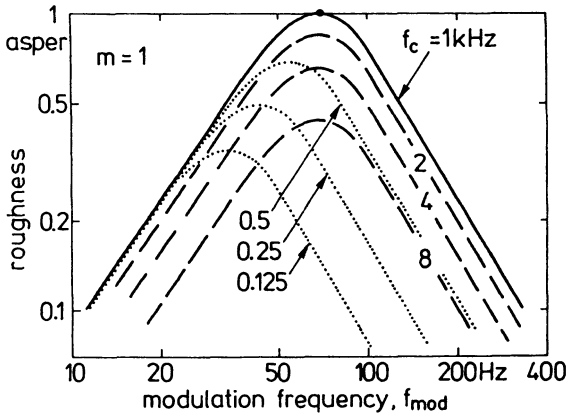


Figure 5. Roughness of 100% amplitude-modulated tones; according Fastl & Zwicker, 2007, p. 259. Curves for different centre frequencies are presented.

Now let's compare the findings of the psychoacoustic studies. The curves in the Figure 6 were composed based on the formulas and interpolations of graphically presented results from the discussed sources. Probably, the confusion between the dissonance, roughness, and its possible types explains why the results of the experiments show significant discrepancies. A

closer examination of Figure 6 reveals that roughness is typically associated with larger interval sizes, and that sensory dissonance is associated with narrower interval sizes. For instance, Terhardt in his experiment asked the subjects specifically to evaluate roughness (1968, p. 216), and the corresponding curve lies higher. On the contrary, as already mentioned, the well-known relating of the maximum dissonance to 1/4 of critical bandwidth (Plomp & Levelt, 1965) refers specifically to dissonance but not to roughness.

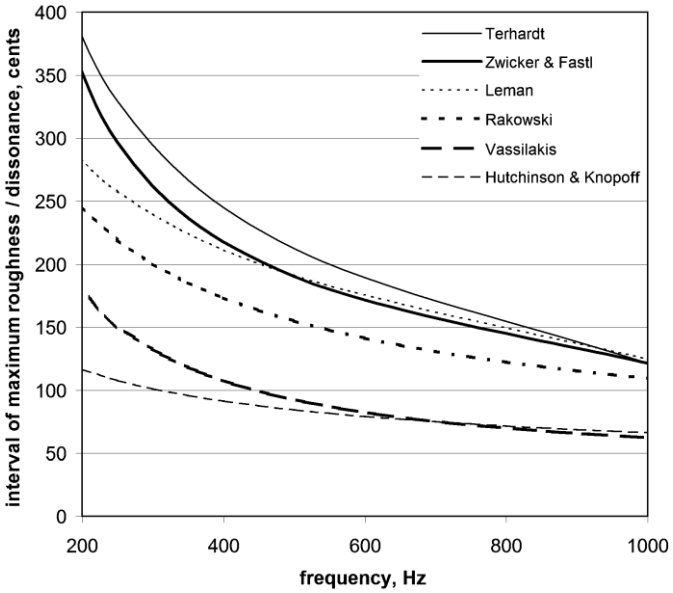


Figure 6. Dependence of maximum roughness / dissonance on the central frequency. See the body text for details.

On the one hand, roughness is typically associated with the perceptual result of rapid fluctuation of envelope of sound pressure amplitude, i.e., with the subjective rate of amplitude change (expressed as product of the subjective modulation depth and f_{beats}; e.g. Fastl & Zwicker, 2007, p. 262), or, briefly, with the "bumpiness of the [subjective] acoustic surface of a sound" (Parncutt, 2006, p. 202). On the other hand, sensory dissonance could be probably connected to the features of critical bandwidth. Certain doubts remain whether the "bumpiness" and critical bandwidth are tightly related. From my purely subjective observations, the (sensory) "dissonance", "unpleasantness", or "annoyance" could be rather equalized to "harshness" and not so much to "roughness". One may therefore speculate that, for instance, a semitone in the middle of a piano keyboard sounds harsher, whereas the whole tone seems to be rougher. Incidentally, the terms such as "harsh" or "turbid" occur episodically when describing non-euphonious, unpleasant, or dissonant sonorities (e.g. Plomp & Levelt, 1965, p. 554; Mashinter, 2006, p. 65, 66).

III. *SUTARTINĖS*: GENERAL REMARKS

Now from the "cosmopolitic" experiments on psychoacoustic roughness and sensory dissonance we move to the Lithuanian ethnic *Sutartinės*. The most distinctive kind of Lithuanian *Sutartinės* present a peculiar type of *Schwebungsdiaphonie*, i.e. diaphony of the *Sutartinės* is based mostly on intervals of the second occurring between the vocal parts which intertwine polyphonically and polyrhythmically.

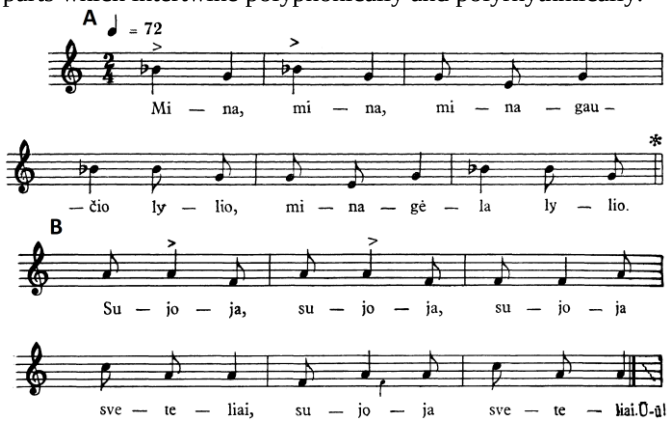
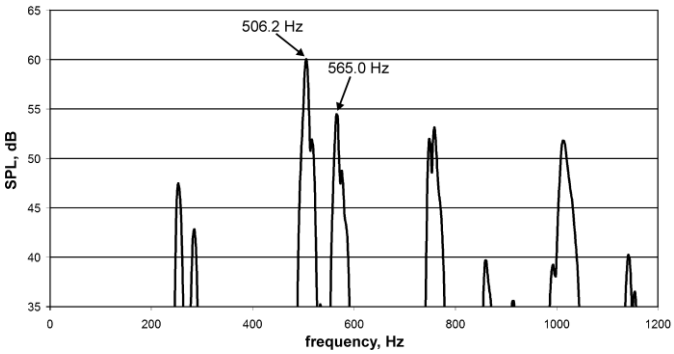


Figure 7. *Sutartinė "Mina, mina, minagaučio lylio"*: original transcription of one part (Slaviūnas, 1958, p. 657 [Nr. 428a]). Entrance of the canonically succeeding voice is asterisked.

Figure 7 shows a typical example of a *Sutartinė*. This *Sutartinė* was performed canonically by three singers in such a way that the two parts A and B (separated by an asterisk in the figure) sound simultaneously, except in the beginning when only one voice (part A) sounds. The lyrics change. Thus mostly intervals of the second occur continuously between the two voices.

When listening to the original recording of this *Sutartinė* (Račiūnaitė-Vyčinienė, 1998)[3], even unsophisticated ear could suggest an idea that the transcription in Figure 7 is actually crude or, at least, quite approximate: the real recording sounds "non-tempered" enough. To reveal the actual scale, the recording was analyzed acoustically: the pitches of dyads were measured and the intervals were calculated (Ambrazevičius, 2005). The pitches were determined from the spectra of the dyads: certain partials were identified as belonging to one or the other of two voices, their frequencies were measured (see the example in Figure 8), and the pitches were calculated. Relatively stable portions of the dyads were considered in terms of spectrum (fortunately, the intrasonic intonation of *Sutartinės* features quite stable segments).



[3] The digitized version of the old recording (from 1930s).

Figure 8. Excerpt from typical spectrum of *Sutartinė "Mina, mina, minagaučio lylio"*.

The statistical distribution of pitches is depicted in Figure 9. It shows approximately 1.8 of tempered semitone between the most frequent intonations. Thus the corresponding most frequent thirds in vocal parts (G3–B3 and A3–C4[4]) equal approximately 3.6 semitones, i.e. they are neutral. Fourth G3–C4 equals 5.4 semitones. D4 occurs very seldom, thus categorical conclusions about its position in the tuning system could not be drawn. Nevertheless C4–D4 equals 1.5–2.2 semitone, i.e. also roughly 1.8 semitones on the average. A3 and H3 are the most stable tones according to the corresponding sharp peaks in Figure 9. This bichord could be treated as certain bitonal nucleus and anchor of the tuning system. G3 and C4 are less stable, whereas F3 and D4 are the least stable. The zones of intonation are quite wide, even for the most stable anchors. Hence, to generalize, the two central steps are intoned relatively steadily in the course of the entire performance thus forming the nucleus of the scale. The marginal steps show greater freedom in intonation.

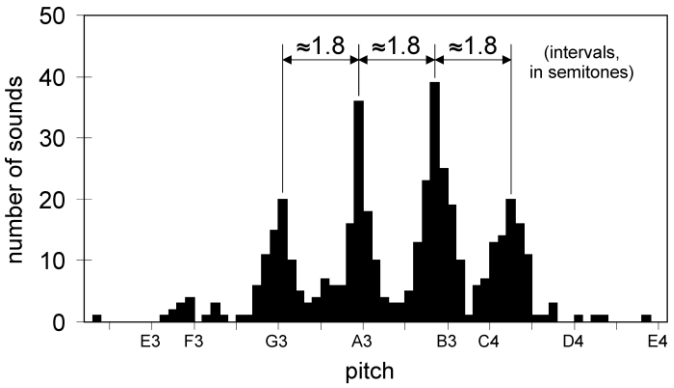


Figure 9. Histogram of pitches in *Sutartinė "Mina, mina, minagaučio lylio"* (Figure 7); all pitches in all parts.

The analysis leads to a conclusion that we have to be very cautious when treating and denominating the scale and tuning system aurally. Western major-minor system and equal temperament work as elements of apperception, which results in "aural ghosts". They lead to misinterpretation that design of the scale is diatonic. Actually the tuning system has nothing in common with diatonics: there is no semitone/whole tone contrast in the sequence of intervals. The scale could be considered as "squeezed anhemitonics", since the intervals between the adjacent pitches are a bit narrower than the tempered whole tone.

One could try to visualize the revealed regularities of the scale in transcription (Figure 10) where a peculiar staff is intentionally applied to avoid associations with the diatonic scale.

[4] Here and hereafter a simplified marking for pitch class is used. For instance, C4 actually could be as high as C#4 or even higher.

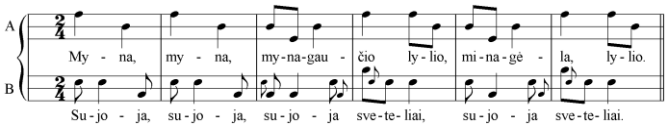


Figure 10. Transcription of characteristic patterns of *Sutartinė "Mina, mina, minagaučio lylio"* on an alternative staff. The petit notes show the most characteristic variants.

In the subsequent study (Ambrazevičius, 2008), a total distribution of dyad-intervals in 25 *Sutartinės* has been also composed. The distribution showed that the majority of the intervals are seconds. The category of the interval is quite wide and does not split into the individual categories of minor and major seconds. As in the case of the separate *Sutartinė "Mina, mina, minagaučio lylio"*, the seconds slightly narrower than the tempered whole tone (around 1.7 semitones) are most preferred.

So, again, we come to a simple conclusion: the intervals of second between the voices in the dyads of *Sutartinės* comprise relatively wide category centred at, approximately, 170–180 cents. What accounts for such a peculiar interval? Let's return to the psychoacoustic studies on sensory dissonance / roughness and collate their results to the findings of the study on intervals in *Sutartinės*.

IV. DISSONANCE OR ROUGHNESS IN *SUTARTINĖS*?

For female voices, frequency of the first formant ranges roughly from 400 to 1000 Hz. So this frequency range is expected to be the most intense range in the spectra of singing voices. This corresponds to the second or third (or sometimes fourth) harmonics. Application of these frequency values to the graphs in Figure 6 leads to an insight that the singers were aiming for maximum roughness: the most intense frequency range corresponds to the wide range of pitch intervals centered at slightly "squeezed" whole tone. Importantly, the aiming for maximum dissonance would lead to significantly narrower intervals, around 70–100 cents, what is not the case of *Sutartinės*. It seems that specifically roughness was meant by the singers of *Sutartinės* when describing the sonorities as "clashing" (clanging, warbling; but not "cutting" which would point to the sensory dissonance and narrower intervals). The "perfect clash" was considered by the singers as an essential quality and marker of a congenial performance. Earlier Brandl already concluded that the psychoacoustic correlate of the ideal ring in *Schwebungsdiaphonie* (found in the Balkans and elsewhere) is of maximum roughness (1989). It is actually dubious whether this statement really works for all traditions in Balkans, as there quite different intervals in the dyads could be registered for different cases (cf. Miljković, 1998; Rihtman, 1969). At any rate, the measurements in our studies support this statement when applied to Lithuanian *Sutartinės*. Therefore it can be credibly stated that the scales of *Sutartinės* are actually determined by psychoacoustic, i.e. by extramusical phenomenon.

Importantly, the noun *Sutartinė* derives from the verb *sutarti* which means "to agree", "to be in concord" ("to live in

concord", "to sing in concord", and so on); in other words, "to sing in consonance". Nowadays the word *Sutartinė* is sometimes even applied to signify a perfect, harmonious performance in general, no matter the kind of the performance. Thus, in the case of the Lithuanian *Schwebungsdiaphonie*, roughness obtains a positive connotation: aesthetically and semantically, the sonorities in seconds are considered as consonances.

However, it should be pointed out that the requirement of maximum roughness is not categorical in the Lithuanian case for the following reasons: the intonational zone of a second is too wide, durations of the sounds are too short to produce exact intervals (initial glides are characteristic), and the partials are, on the average, too different in SPL.[5] All these factors diminish the role of maximum roughness. It could be stated that maximum roughness is a desirable quality, but the zone of the suitable roughness is quite wide; the factor of roughness is possibly reduced by other important factors of articulation.

V. CONCLUSIONS

The close inspection of psychoacoustic studies on roughness / sensory dissonance show significant divergences in their findings. Most probably, this results from different experimental conditions and confusion of notions of roughness and sensory dissonance. Attempts to separate these two notions reveal that, at least for a substantial frequency range, maximum roughness tends to be associated with larger interval sizes, compared to the case of maximum sensory dissonance.

Brandl's insight on aiming for maximum psychoacoustic roughness in performance of *Schwebungsdiaphonie* (exemplified mostly by the examples of Balkan music traditions), most probably, is valid for the case of Lithuanian *Sutartinės* as well. Here specifically roughness and not sensory dissonance is meant; this results from the collation of the findings of psychoacoustic studies on roughness / sensory dissonance and the findings of acoustical measurements of *Sutartinė* performances. The rough quality of the sonorities in *Sutartinės* obtains positive connotations, i.e., in a broad sense, these sonorities are considered as "consonances". The maximum roughness is obtained for the intervals slightly narrower than the tempered whole tone, for the characteristic spectra of the female voices of *Sutartinės* singers. This results in the peculiar scale structures deviating considerably from the twelve-tone equal temperament. As a side product, problem of transcription occurs, making the conventional five-lined staff unsatisfactory for adequate presentation of the scale structures in the roughness-based *Sutartinės*.

REFERENCES

Ambrazevičius, R. (2005). Scale in *Sutartinės*: Psychoacoustic viewpoint. In R. Astrauskas (Ed.), *Traditional music and research in the Baltic area. New approaches in ethnomusicology* (pp. 244-260). Vilnius: Lithuanian Academy of Music and Theatre.

Ambrazevičius, R. (2008). Psychoacoustical and cognitive basis of *Sutartinės*. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima, & M. Tsuzaki (Eds.), *ICMPC10. Proceedings of the 10th International Conference on Music Perception and Cognition. 25-29 August 2008. Sapporo, Japan* (pp. 700-704). Adelaide: Causal Productions.

Ambrazevičius, R. (2015). Dissonance/roughness & tonality perception in Lithuanian traditional *Schwebungsdiaphonie*. *Journal of Interdisciplinary Music Studies*, fortcoming.

Brandl, R. M. (1989). Die Schwebungs-Diaphonie – aus musikethnologischer und systematisch-musikwissenschaftlicher Sicht. In C. Eberhardt & G. Weiss (Eds.), *Südosteuropa-Studien. Bd. 40, Volks- und Kunstmusik in Südosteuropa* (pp. 51-68). Regensburg: Gustav Bosse Verlag.

Cazden, N. (1945). Musical consonance and dissonance: A cultural criterion. *Journal of Aesthetics and Art Criticism, 41*(1), 3-11.

Fastl, H., & Zwicker, E. (2007). *Psychoacoustics. Facts and models*. Berlin, Heidelberg, New York: Springer.

Guirao, M., & Garavilla, J. M. (1976). Perceived roughness of amplitude-modulated tones and noise. *Journal of the Acoustical Society of America, 60*, 1335-1338.

Hutchinson, W., & Knopoff, L. (1978). The acoustic component of western consonance. *Interface, 7*(1), 1-29.

Leman, M. (2000). Visualization and calculation of the roughness of acoustical musical signals using the synchronization index model (SIM). In *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00), Verona, Italy, December 7-9, 2000* (pp. DAFX-1-DAFX-6).

Hutchinson, W., & Knopoff, L. (1978). The acoustic component of western consonance. *Interface, 7*(1), 1-29.

Mashinter, K. (2006). Calculating sensory dissonance: Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff. *Empirical Musicology Review, 1*(2), 65-84.

Messner, G. F. (1989). Jaap Kunst revisited. Multipart singing in three East Florinese villages fifty years later: A preliminary investigation. *The World of Music, 31*(2), 3-51.

Miljković, L. (1998). О тоналној реконструкцији архаичног фолклорно-музичког изржавања [The possibility of reconstructing archaic folkloric music expression]. *Флогистон, часопис за историју науке. Хармонија у природи, науци и уметности кроз историју, 7*, 415-442.

Parncutt, R. (2006). Commentary on Keith Mashinter's "Calculating sensory dissonance: Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff". *Empirical Musicology Review, 1*(4), 201-203.

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38*(4), 548-560.

Račiūnaitė-Vyčinienė, D. (1998). *Sutartinės. Polifoninės dainos. 1935–1937 m. archyviniai įrašai* [Sutartinės. Polyphonic songs. Archival recordings, 1935–1937] (CD). Vilnius: Tautos namų santara, Vilniaus plokštelių studija.

Rakowski, A. (1982). Psychoacoustic dissonance in pure-tone intervals: Disparities and common findings. In C. Dahlhaus & M. Krause (Eds.), *Tiefenstruktur der Musik* (pp. 51-67). Berlin: Technische Universität Berlin.

Rihtman, C. (1969). Le microton dans les aspects les plus anciens de la musique traditionnelle en Bosnie-Herzégovine. In P. Stajnov (*et al*) (Eds.), *Bulletin de l'Institut de musique. T. XIII* (pp. 293-301). Sofia: Izd-vo na Bŭlgarskata Akademiia na Naukite.

[5] Roughness shows substantial dependence on the ratio of amplitudes of the "clashing" harmonics. The strongest roughness occurs for equal amplitudes. The more different in SPL are the harmonics, the weaker is the sense of roughness (Terhardt, 1968; Vogel, 1975; Guirao & Garavilla, 1976).

Sethares, W. A. (1998). *Tuning, timbre, spectrum, scale*. London: Springer.

Sethares, W. A. (2005). *Tuning, timbre, spectrum, scale* (2nd ed.). London: Springer.

Slaviūnas, Z. (1958). *Sutartinės: Daugiabalsės lietuvių liaudies dainos* [*Sutartinės*: Lithuanian Polyphonic Folk Songs]. Vol. 1. Vilnius: Valstybinė grožinės literatūros leidykla.

Terhardt, E. (1968). Über akustische Rauhigkeit und Schwankungsstärke. *Acustica, 20*, 215-224.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America, 55*, 1061-1069.

Vassilakis, P. N. (2001). *Perceptual and physical properties of amplitude fluctuation and their musical significance* [PhD dissertation]. Los Angeles: University of California.

Vogel, A. (1975). Über den Zusammenhang zwischen Rauhigkeit und Modulationsgrad. *Acustica, 32*, 300-306.