MODELLING OF LOCAL TEMPO CHANGE WITH APPLICATIONS TO LITHUANIAN TRADITIONAL SINGING

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

ABSTRACT

The present study aims to develop techniques of measurement, mathematical modelling, and evaluation of temporal irregularities (first of all, local tempo changes) in vocal performance, and to test the techniques on examples of Lithuanian traditional singing.

Methods of measurements of note durations (IOIs) in vocal performance are reviewed, their problems including the identification of perceptual attack time and adequate precision based on duration JND are discussed. Three folk song recordings are chosen for modelling of temporal irregularities. The performances are more or less *tempo giusto* so rhythm values are easily identified. Tempo curves of the chosen folk song performances are composed and analyzed: microtiming, in terms of LS/SL divisions of rhythm values, and local tempo changes in longer time spans are evaluated. Three measures of temporal unevenness are introduced; 1) the general unevenness, 2) the note-to-note unevenness, and 3) the unevenness of smoothed local tempo.

The designed model is applied to a set of 40 song recordings (10 songs from each of the 4 Lithuanian main ethnographic regions). The vocal dialects corresponding to the ethnographic regions differ noticeably in timing expressed in terms of microtiming and the three indices. Thus different combinations of the indices are characteristic of different musical dialects. This allows us to conclude that the different parameters of rhythm interpretation in vocal style can serve as more or less reliable markers of a musical dialect.

1. INTRODUCTION

There is a large number of studies on (micro)timing and research techniques, with some ethnomusicological applications (cf. Bengtsson, Gabrielsson, & Thorsén, 1969; Clarke & Cook, 2004; Danielsen, 2010; Ellis, 1991a; Friberg, Bresin, & Sundberg, 2006; Friberg, & Sundström, 2002; Gabrielsson, 1999; Ledang, 1967). Also, there is a considerable number of studies on perception of tempo changes and extraction of tempo changes from (mostly MIDI based) recordings (cf. Ellis, 1991b; Sheldon & Gregory, 1997; Dahl & Granqvist, 2003; Thomas, 2007; Müller et al., 2009; Dannenberg & Mohan, 2011; Yanagida & Yamamoto, 2017). nPVI (normalized Pairwise Variability Index) is an additional technique to eliminate the factor of local tempo changes and to study the small-scale timing phenomena (Grabe & Low, 2002). Yet, to our knowledge, there is a lack of studies on gradual tempo changes in vocal performance and their mathematical modelling.

Concerning the ethnomusicological studies on timing, they usually stop short of discussing patterns of rhythm categories (cf. Čiurlionytė, 1969; for Lithuanian sources). The microtime deviations are usually only presented as lengthening or shortening (microfermatas) of individual notes (cf. Bengtsson, 1974, p. 30; Sevåg & Sæta, 1992, p. 49; Četkauskaitė, 2007; see the examples of microfermatas in Figures 1 and 3) or short note groups (cf. Bengtsson, 1980, p. 303; Czekanowska, 1961; Ledang, 1967; Bartkowski, 1987, p. 69) in transcriptions.

2. METHODS AND PRECISION OF MEASUREMENTS

The analysis of microtiming bases on tempo curves composed from measurements of note durations (Inter-Onset-Intervals; IOIs). IOI measurements are simple and can even be automatized for keyboard performances (and, in general, in performances with short sound attacks), while vocal performances, frequently containing long and smooth attacks, pose some problems. It seems that perceptual attack time (PAT) comes somewhat earlier than the attack peak and there is no clear dependence of the lag on acoustical parameters (cf. Vos & Rasch, 1981; Gordon, 1987).

Therefore the measurements of IOIs in vocal performances can hardly be automatized or objectivized in any way. One should rely on his/her ability to detect PATs while listening to recordings and grasping the PATs from positions of the moving cursor. Fortunately, this technique is, nevertheless, precise enough compared to precision of time perception: in the cases of the best listening conditions, duration JND (just noticeable difference) can even amount 10% (Woodrow, 1951; Michon, 1964; Povel, 1981) while the readings of individual listeners usually differ not more than in 15 ms (Ambrazevičius, 2009). For crotchets and quavers performed in moderate tempo (MM=100 bpm), this correspondingly results in 2.5 and 5%. Thus precision of the IOI measurements can be considered adequate.

3. EXAMPLES OF TEMPO CURVES. LS AND SL CASES

Composition of tempo curves is well described elsewhere (cf. Clarke, 2004). Basically, all IOIs are normalized, i.e., one rhythm value (category) is chosen as a duration unit and other IOIs are recalculated. For instance, if a quaver is the unit, the duration of an individual crotchet in actual performance is divided by two; this gives the corresponding (effective) duration of the quaver. The sequences of the effective durations presented graphically constitute the tempo curves. Therefore, only (at least approximate) *tempo giusto* performances (i.e. which are characterized by identified rhythm values) can be employed for the composition of tempo curves. If listening to a performance results in ambiguous or alternative, different interpretations of rhythm and meter, the composition of tempo curves seems useless. Nevertheless, collation of the alternative tempo curves may facilitate recognition of more adequate interpretations.

Figure 1. Transcription of the first verse of the song *Kad aš dukrelių daug turėčia* (Adelė Kazlauskienė; Gustaičiai, Prienai Dst. Recording: Četkauskaitė, 2002, N76).

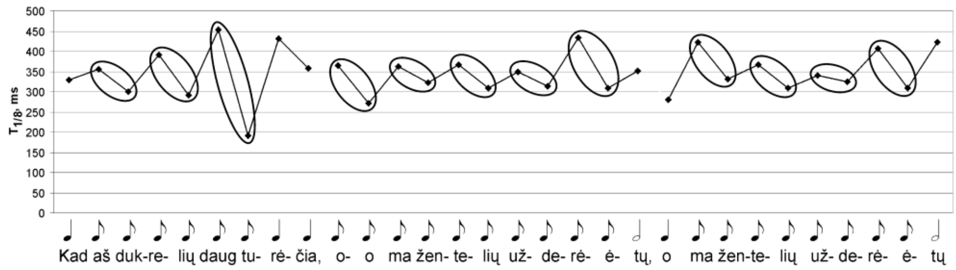

Figure 2. Example of a tempo curve (see the melody in Figure 1). Vertical axis: (effective) durations of eighth notes; LS tendency. Pairs of quavers are circled.

Figure 3. Transcription of the first verse of the song *Jojau pro dvarą* (Vincas Jurčikonis; Babrai, Lazdijai Dst. Recording: Četkauskaitė, 1995, N18).

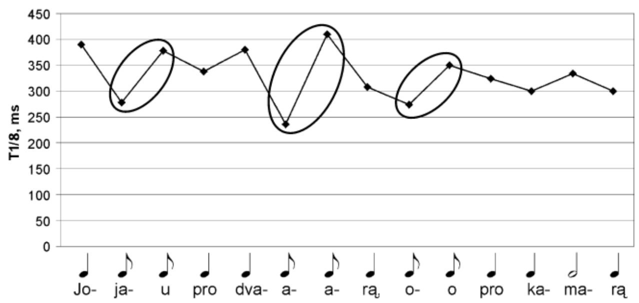

Figure 4. Example of a tempo curve (first four measures of the melody in Figure 3). Vertical axis: (effective) durations of eighth notes; SL tendency. Pairs of quavers are circled.

Figure 5. Transcription of the first verse of the song *Oi, šiaudai šiaudai* (Barbora Buivydaitė; Rūdaičiai, Kretinga Dst. Recording: Baranauskienė *etal*, 2006, N6).

Figures 2 and 4 show typical tempo curves composed from the IOI measurements of two Lithuanian traditional vocal *tempo giusto* performances (Figures 1 and 3). The structural notes were considered; durations of embellishments (appoggiaturas, etc.) were incorporated into the corresponding structural notes. Only three notes in the transcriptions are supplemented with microfermata marks (see the syllables *daug* and *tu-* in Figure 1, and the first *-no* in Figure 3). Consequently, only for three notes is the prolonging or shortening of the rhythm values clearly perceived. Yet a significant fluctuation of the durations is observed in the tempo curves; the performances present two opposite cases of *inégales*. Figure 2 shows a clear LS tendency (long-short division of crotchet into two quavers) whereas Figure 4 shows a reverse SL tendency; the performance is characterized by a somewhat "limping" rhythm. The median of $T1/T2$ ratios (ratios of quaver durations forming one crotchet) for the song *Kad aš dukrelių daug turėčia* equals 1.23 and the interquartile encompass the range 1.17-1.34. For the song *Jojau pro dvarą*, the median is 0.72 and the interquartile is 0.62-0.78.

4. MODELLING OF GRADUAL TEMPO CHANGE

The third song (Figure 5) is chosen for modelling of characteristics of temporal unevenness in longer time spans of a performance. Three measures of temporal unevenness are introduced; 1) the general unevenness ("general rubato index", R_{AAD}), 2) the note-to-note unevenness (the "nPVI rubato index" provided by nPVI technique, R_{nPVI}), and 3) the unevenness of smoothed local tempo ("tempo change index", TV). For the first measure, average absolute deviation of duration is used (Figure 6), instead of previously used standard deviation (Ambrazevičius, 2009; 2018); this facilitates better compatibility with nPVI measures as they also apply absolute deviation. R_{AAD} is evaluated as *AAD* normalized to the mean duration:

$$R_{AAD} = \frac{1}{n\bar{T}}\sum_{k=1}^{n}|T_k - \bar{T}|\ ; \qquad (1)$$

where T_k is the duration of kth note, $\bar{T}$ is the mean duration, and n is the number of structural notes in the melody contour. For the examined performance, $R_{AAD} = 0.122$, i.e. the actual 1/8-durations deviate in 12.2% from the mean, on the average.

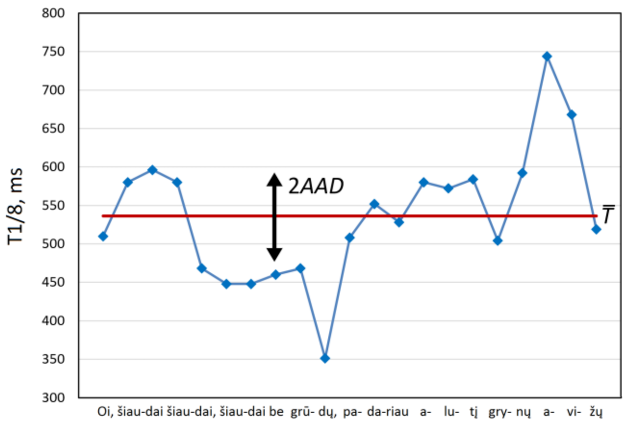

Figure 6. Example of tempo curve (see the melody in Figure 5). Vertical axis: (effective) durations of eighth notes. The horizontal line depicts average duration of eighth note.

However, if one needs to eliminate relatively slow gradual change of tempo and to evaluate note-to-note unevenness (the "jaggedness" of tempo curve), R_{nPVI} is applied instead. This index reflects average deviation from the changing local average of duration (average of two adjacent note durations; Figure 7). Grabe & Low (2002) introduced the "normalized Pairwise Variability Index" (*nPVI*):

$$nPVI = 100 \times \left[\sum_{k=1}^{n-1}\left|\frac{T_k - T_{k+1}}{(T_k + T_{k+1})/2}\right| / (n-1)\right]. \qquad (2)$$

Then the expression for R_{nPVI} follows:

$$R_{nPVI} = \frac{1}{n-1}\sum_{k=1}^{n-1}\frac{|T_k - T_{k+1}|}{T_k + T_{k+1}}. \qquad (3)$$

For the examined performance, $R_{nPVI} = 0.057$, i.e. the actual 1/8-durations deviate in 5.7% from the changing local mean, on the average. Naturally, $R_{nPVI} < R_{AAD}$ (in general; not only for the analyzed particular piece).

First approximation of tempo change in longer time spans can be made by substitution of duration in R_{nPVI} with moving duration average (period = 2; Figure 7):

$$TV_{(d)} = \frac{1}{n-2}\sum_{k=1}^{n-2}\frac{|T_{k+2} - T_k|}{\frac{T_k}{2} + T_{k+1} + \frac{T_{k+2}}{2}}. \qquad (4)$$

However, from the viewpoint of perception, interpretation of continuous tempo change (Figure 8) seems to be more adequate than interpretation of discrete tempo change, in smaller or larger steps. (Of course, this statement is not valid for cases of sudden tempo changes, e.g. when tempo is clearly different in two structural parts of melody.) Following this approach, the "tempo change index" is designed again based on the nPVI logic; only summation of discrete IOI values in "nPVI rubato index" is substituted with integration of IOI equivalents in the continuous smoothed curve of local tempo change:

$$TV_{(c)} = \frac{1}{2(n-1)} \int_{t_i}^{t_f} \frac{|dT(t)|}{T(t)}. \quad (5)$$

After some mathematical procedures we get

$$TV_{(c)} = \frac{1}{n-1}\left[ln\frac{\prod T_{max}}{\prod T_{min}} \pm \frac{lnT_i}{2} \pm \frac{lnT_f}{2}\right]; \quad (6)$$

here T_i and T_f stand for the initial and final values ($T(t_i)$and $T(t_f)$) of function $T(t)$, and T_{max} and T_{min} are its (local) maximum and minimum values. Plus and minus signs in the expression are applied if T_i, T_f present, correspondingly, local maxima or minima. For instance, for the examined performance, minus signs are applied for both $lnT_i/2$ and $lnT_f/2$. For this performance, $TV_{(c)} = 0.032$, i.e. the local tempo changes in 3.2% from note to note, on the average. The performance example chosen here for modelling is characteristic of large local tempo change. Later we will see that usually tempo changes to a lesser extent (see the values for *TV* in Figure 10).

Certainly, the chosen technique of smoothing affects the values in $TV_{(c)}$ expression. Nevertheless, the discrepancies are not significant. For instance, if even an alternative technique of smoothing gives a descending line from the first T_{max} to the second T_{min} in Figure 8 (no first T_{min} and second T_{max}), $TV_{(c)} = 0.0305$ instead of 0.032. Thus the technique is still adequate for generalized evaluations.

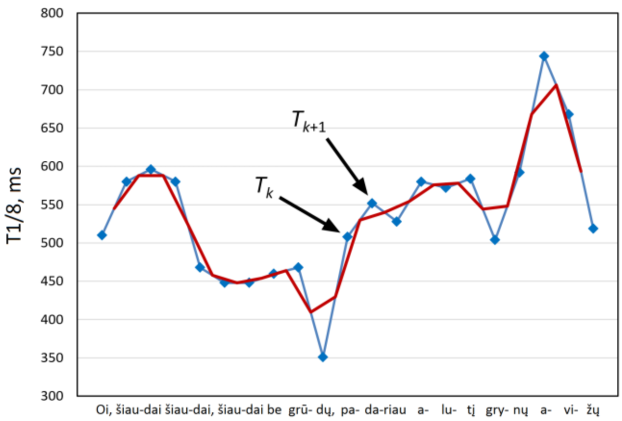

Figure 7. Example of tempo curve, as in Figure 6. The red line depicts changing average duration of an eighth note (average of two adjacent note durations).

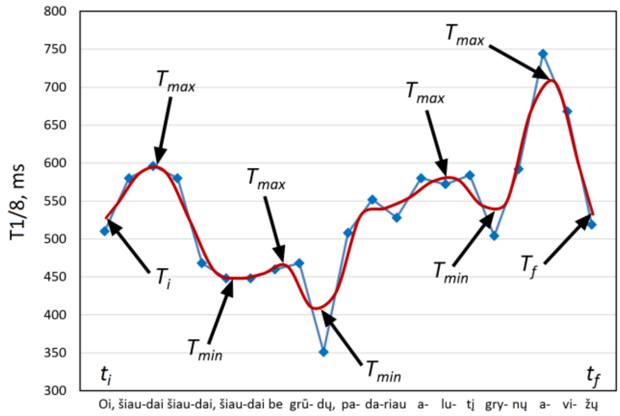

Figure 8. Example of tempo curve as in Figure 6. The smoothed line depicts local change of duration of an eighth note. Characteristic durations for calculation of $TV_{(c)}$ ("tempo change index") are marked.

5. TECHNIQUE TESTING ON SETS OF EXAMPLES OF LITHUANIAN TRADITIONAL SINGING

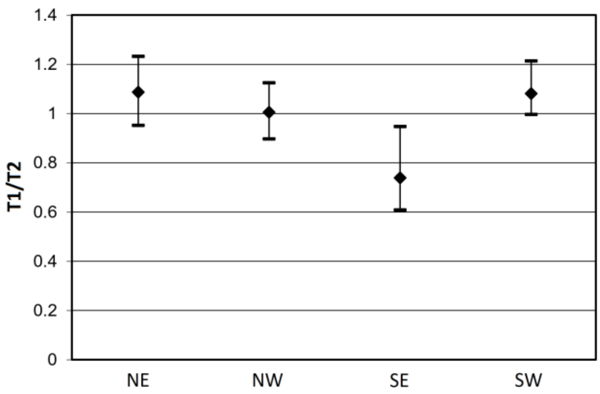

Figure 9. Generalized $T1/T2$ values for the samples representing four Lithuanian musical dialects. Diamonds and vertical lines mark the medians and interquartiles.

The designed model is applied to the set of 40 song recordings (10 songs from each of 4 Lithuanian main ethnographic regions; Aukštaitija, Dzūkija, Suvalkija, and Žemaitija; NE, SE, SW, and NW, correspondingly) used in a previous study (Ambrazevičius, 2018). The vocal dialects corresponding to the ethnographic regions differ noticeably in timing expressed in terms of microtiming and the three indices (Figures 9 and 10). The mean $T1/T2$ ratios range from 1.09 (NE) to .77 (SE) ($p_{SE\text{-}NE}$, $p_{SE\text{-}NW}$, $p_{SE\text{-}SW}$ < .001, $p_{NE\text{-}NW}$ = .035, $p_{SW\text{-}NW}$ = .026), mean R_{AAD} values range from .14 (SE) to .09 (SW), R_{nPVI} values range from .08 (SW) to .12 (SE), and *TV* values range from .005 (SW) to .013 (NW).

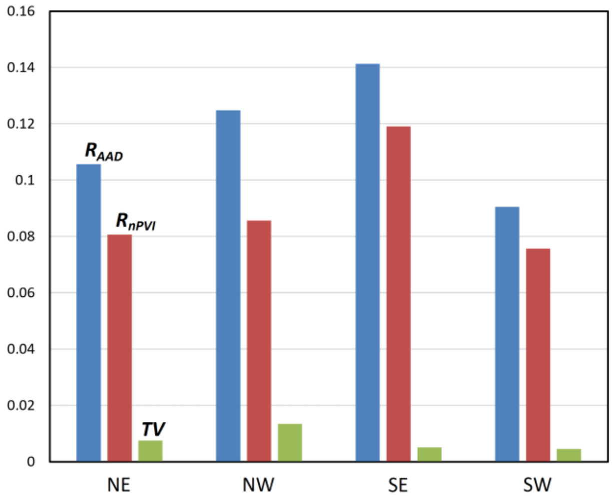

Figure 9. Generalized R_{AAD}, R_{nPVI}, and *TV* values (averages) for the samples representing four Lithuanian musical dialects.

Thus different combinations of the indices are characteristic of different musical dialects. For instance, large note-to-note rhythm unevenness combines with quite negligible changes in overall tempo in SE, whereas noticeably less note-to-note rhythm unevenness combines with large changes in overall tempo in NW ($p_{SE\text{-}NW} < .001$). This allows us to conclude that the different parameters of rhythm interpretation in vocal style can serve as more or less reliable markers of a musical dialect. The positive correlation between R_{nPVI} values in spoken and vocal dialects found in Ambrazevičius, 2018, shows the impact of linguistic rhythm onto musical rhythm.

6. DISCUSSION

One of the most important results of the present study is the proposed and derived index for evaluation of gradual tempo changeability (*TV*). Together with the other two indices (R_{AAD} and R_{nPVI}) and $T1/T2$, it constitutes a system of quantitative parameters for description of general temporal characteristics of a musical performance. It is demonstrated how the proposed system can be applied for revelation of stylistic traits of a performance and, in turn, how different styles can be compared, in terms of temporal performances.

Further the model could be developed in several directions. First, different techniques of tempo curve smoothening could be examined, the most adequate ones could be identified, and techniques for automated extraction of *TV* could be derived. Second, the model could be extended for the analysis of the phenomena characteristic of intermediate time spans (basic rhythm values, measures, etc.). For instance, in certain cases, it would be interesting to know whether the temporal movement is based on the basic rhythm values or rather on their subdivisions (e.g. whether the movement in crotchets is more stable than the movement in quavers). Third, the developed model could be applied for the evaluation of stability of performance repetitions (e.g. similarity of melostrophes). Fourth, the model could be modified for the study of temporally more complicated (non ‘tempo giusto’) performances.

Finally, the model could be applied not only for the study of traditional vocal performance, but also for other vocal styles and instrumental music.

7. REFERENCES

Ambrazevičius, R. (2009). Rhythm interpretation in vocal performance: Lithuanian examples. In *XXV ESEM. ‘Performance’. Abstracts* (p. 10). Milton Keynes, UK: The Open University.

Ambrazevičius, R. (2018). Aspects of timing in Lithuanian traditional singing. In R. Parncutt & S. Sattmann (Eds.), *ICMPC15/ESCOM10: Abstract book (electronic)* (p. 34). Graz, Austria: Centre for Systematic Musicology, University of Graz.

Baranauskienė, V., Zakarienė, V., Juciutė, L., & Ivanauskaitė, V., eds. (2006). *Dainų karaliai ir karalienės. Barbora Buivydaitė*. Vilnius: Lietuvos muzikos ir teatro akademija.

Bartkowski, B. (1987). *Polskie śpiewy religijne w żywej tradicji. Style i formy*. Kraków: Polskie Wydawnictwo Muzyczne.

Bengtsson, I. (1974). On notation of time, signature and rhythm in Swedish polskas. In G. Hilleström (Ed.), *Studium instrumentorum musicae popularis. B. III*. (pp. 22-31). Stockholm: Nordiska Musikförlaget.

Bengtsson, I. (1980). *Folkmusikboken*. Stockholm: Bokförlaget Prisma.

Bengtsson, I., Gabrielsson, A., & Thorsén, S. M. (1969). Empirisk rytmforskning. *Swedish Journal of Musicology, 51*, 49-118.

Četkauskaitė, G., ed. (1995). *Lietuvių liaudies muzika*. Vilnius: 33 Records.

Četkauskaitė, G., ed. (2002). *Lietuvių liaudies muzika. V. 3. Suvalkiečių dainos. Pietvakarių Lietuva*. Vilnius: Lietuvos muzikos akademija.

Četkauskaitė, G., ed. (2007). *Lietuvių liaudies dainų antologija*. Vilnius: Lietuvos muzikos ir teatro akademija.

Čiurlionytė, J. (1969). *Lietuvių liaudies dainų melodikos bruožai*. Vilnius: Vaga.

Clarke, E. F. (2004). Empirical methods in the study of performance. In E. Clarke & N. Cook (Eds.), *Empirical musicology. Aims, methods, prospects* (pp. 77–102). New York: Oxford University Press.

Czekanowska, A. (1961). *Pieśni biłgorajskie*. Wrocław. Polskie Towarzystwo Ludoznawcze.

Dahl, S., & Granqvist, S. (2003). Looking at perception of continuous tempo drift - a new method for estimating internal drift and just noticeable difference. In R. Bresin (Ed.), *Proceedings of the Stockholm Music Acoustics Conference* (pp. 595-598). Stockholm: KTH Speech, Music and Hearing.

Danielsen, A., ed. (2010). *Musical rhythm in the age of digital reproduction*. Surrey, Burlington: Ashgate.

Dannenberg, R. B., & Mohan, S. (2011). Characterizing tempo change in musical performances. In *Proceedings of the International Computer Music Conference* (pp. 650-656). Huddersfield: University of Huddersfield.

Ellis, M. C. (1991). An analysis of “swing” subdivision and asynchronization in three jazz saxophonists. *Perceptual and Motor Skills*, 73, 707-713.

Ellis, M. C. (1991). Research note. Thresholds for detecting tempo change. *Psychology of Music, 19*, 164-169.

Friberg, A., Bresin, R., & Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology, 2*, 145-161.

Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception,* 19, 333–349.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.). *Psychology of music* (2nd ed.; pp. 501-602). San Diego, London: Academic Press.

Gordon, J. W. (1987). The perceptual attack time of musical tones. *Journal of the Acoustical Society of America, 82*, 88-105.

Grabe, E., & Low, E. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology*, 7. Cambridge: Cambridge University Press.

Ledang, O. K. (1967). *Song syngemåte og stemmekarakter*. Oslo: Universitetsforlaget.

Michon, J. A. (1964). Studies on subjective duration: I. Differential sensitivity in the perception of repeated temporal intervals. *Acta Psychologica, 22*, 441-450.

Müller, M., Konz, V., Scharfstein, A., Ewert, S., & Clausen, M. (2009). Towards automated extraction of tempo parameters from expressive music recordings. Retrieved from: https://www.researchgate.net/publication/47863781_Towards_Automated_Extraction_of_Tempo_Parameters_from_Expressive_Music_Recordings

Povel, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 3-18.

Sevåg, R., & Sæta, O. (1992). *Norsk Folkemusikk. Serie II. Slåtter for vanlig fele. B. I. Oppland*. Oslo: Universitetsforlaget.

Sheldon, D. A., & Gregory, D. (1997). Perception of tempo modulation by listeners of different levels of educational experience. *Journal of Research in Music Education, 3*, 367-379.

Thomas, K. (2007). Just noticeable difference and tempo change. *Journal of Scientific Psychology, 2*, 14-20.

Vos, J., & Rasch, R. (1981). The perceptual onset of musical tones. *Perception & Psychophysics, 29*, 323-335.

Woodrow, H. (1951). Time perception. In S. S. Stevens (Ed.). *Handbook of experimental psychology* (pp. 1224-1236). New York: Wiley.

Yanagida, M., & Yamamoto, S. (2017). Effects of the mode of tempo change on perception of tempo change. *Proceedings of Meetings on Acoustics, 28*, (1-10).