The Effect of the Spoken Dialect on the Singing Dialect: The Example of Lithuanian

Rytis Ambrazevičius¹ and Asta Leskauskaitė²

¹ Kaunas University of Technology and the Lithuanian Academy of Music and Theatre
² Institute of the Lithuanian Language

Background in ethnomusicology. It is widely known that dialects of the vocal tradition differ not only in the traits of their musical structures but also in the vocal techniques, articulation, etc. In part, our earlier studies have shown that there are certain dialectal differences between the vocal techniques of northeast and southern Lithuanian (Ambrazevičius, 2001, 2002).

Background in phonetics. Obviously, spoken dialects differ to a greater or lesser extent in their phonetics, lexis and other aspects. For example, numerous studies show that all Lithuanian dialects have particular features. Southern and Eastern Aukštaitian (Southern and Northeast Lithuanian) are similar in different levels of the language. Nevertheless, these dialects differ in the acoustics of the vowels (spectrum, intensity, and pitch) (Leskauskaitė, 2004; Urbanavičienė, 2003, 2004).

Aims. We aim to clarify the possible influences of the spoken dialect on the sung dialect. Hypothetically, the acoustics of speech does not merely affect the phonetics of singing, but also the vocal technique and other traits of singing.

Main contribution. Samples of the speech of eight speakers (three from Northeast Lithuania and five from Southern Lithuania) were analyzed. SPLs, F1 and F2 of vowels were measured, and temporal characteristics of consonants were evaluated. The same procedure was carried out with seven samples of sung performances (three plus four singers, from the corresponding regions). The collation of the results of vowel measurements in speech and singing suggested the following conclusions: 1) a relatively narrow system of speech vowels, in terms of F1-F2, results in covered vowels in singing; however, a relatively wide system of speech vowels remains wide in singing; 2) a wide range of the vowel system corresponds to significant dynamic alterations in the singing performance. Also, it seems that the typical posture of the vocal tract has an influence on the typical (dialectal) ornaments in singing. The collation of the results of consonant measurements led to the following inferences: a distinct enunciation in the spoken dialect correlates with: 1) the accentuation and a marcato quality in the singing; 2) the vocalization of consonants in the singing. Hence, some traits of the singing dialect are neither ‘purely musical’ nor ‘purely vocal’, but rather speech-dependent.

Implications. Ethnomusicology courses should include information on the phonetics of the spoken dialect. Practitioners of traditional singing styles should pay attention to dialectal phonetics, since some relevant features of vocal technique are based primarily on phonetic nuances. A comprehensive study of traditional songs and, particularly of traditional vocal techniques, requires expertise in both the humanities (ethnomusicology, music analysis, linguistics) and the sciences (voice acoustics, statistics).

Keywords: Spoken dialect, singing dialect, vocal technique, marcato, vocalization, ornaments

*Correspondence: Rytis Ambrazevičius, Kaunas University of Technology and Lithuanian Academy of Music and Theatre, Vilnius, Lithuania; e-mail: rytisam@delfi.lt
Introduction

The phonetics of a spoken dialect obviously transfers into dialectal singing. However, from our singing practices\(^1\) we have noticed that speech acoustics possibly does not affect merely the phonetics of singing but also the stylistic features of singing. The present study aims to verify these assumptions. It collates acoustical traits of spoken and sung dialects. Several parameters are chosen: the first two formant frequencies (F1 and F2), and sound pressure levels (SPLs) of vowels are presumed to have some influence on vocal techniques, even the peculiarities of ornament production are presumed to be in tandem with vowel phonetics. Moreover, the distinct enunciation of consonants supposedly transforms into marcato and vocalization (adding an extra vowel to a consonant) in the singing performance.

These phenomena were demonstrated for two Lithuanian spoken/sung dialects. It is quite believable that the phenomena could be generally observed in other vocal traditions and styles as well.

F1, F2, and SPLs in speech performances were measured by Asta Leskauskaitė. The rest of the work was carried out by Rytis Ambrazevičius.

Samples

The samples for the study of speech comprise recordings of eight male Lithuanian speakers (three from Aukštaitija (Northeast Lithuania) plus five from Dzūkija (Southern Lithuania); the duration of each record is from 6 to 30 mins) (see Fig. 1). For the study of singing performance, seven samples (three plus four singers from the corresponding regions; the duration of the songs from two to five mins) of male singing were chosen. The samples were recorded in the second half of the 20th century and nowadays, mostly in the natural environment; all of them are typical of the spoken/sung dialects under investigation.
The Effect of the Spoken Dialect on the Singing Dialect

Figure 1. The ethnographic regions of Lithuania and the location of the samples. x – speech performances; o – singing performances.

Vowels

In this paper, the Lithuanian transcription of vowels is used. Approximate IPA equivalents are shown in Table 1. It is worth mentioning that e, e: are often hardly distinguishable from a, a: (after palatalized consonants), for example, k’e: sounds similar to k’a:, etc.

Table 1. Characters used in this paper and their IPA equivalents.

<table>
<thead>
<tr>
<th>Lithuanian transcription applied in this paper</th>
<th>approximate IPA equivalents</th>
</tr>
</thead>
<tbody>
<tr>
<td>u</td>
<td>ø</td>
</tr>
<tr>
<td>u:</td>
<td>u</td>
</tr>
<tr>
<td>o:</td>
<td>ɔ</td>
</tr>
<tr>
<td>a</td>
<td>ʌ</td>
</tr>
<tr>
<td>a:</td>
<td>ʌ</td>
</tr>
<tr>
<td>e</td>
<td>æ – ɛ, ʌ</td>
</tr>
<tr>
<td>e:</td>
<td>æ, ˈɑ</td>
</tr>
<tr>
<td>è</td>
<td>e</td>
</tr>
<tr>
<td>i</td>
<td>ɪ</td>
</tr>
<tr>
<td>i:</td>
<td>ɪ</td>
</tr>
</tbody>
</table>
F1-F2

Speech
The frequencies of the first two formants of vowels picked out randomly from the recorded speech performances were measured. Only vowels with considerable stationary F1-F2 stages were analyzed. The total set comprised 296 Aukštaitių vowels and 589 Džukai vowels. Then the values of F1 and F2 of phonemic vowel classes (i, i; etc.) were averaged across Aukštaitių/Džukai samples. The results are presented in Fig. 2.

![Figure 2](image-url)

**Figure 2.** F1-F2 charts (averaged values) for speech.

The conclusion is drawn that a slightly more flexible articulation, i.e., slightly greater differences between vowel classes, is characteristic of Džukai vowels.

Singing
The frequencies of the first two formants of all vowel occurrences in the singing performance were measured. Only very short vowels with no stationary stages of F1 and F2 were omitted, as well as vowel diphthongs. Then F1 and F2 of different vowels (in a sense of phonemic quality) in each separate performance were averaged. The difference between ‘short’ (lax) and ‘long’ (tense) vowels (such as i and i:) was not considered, since these vowels are considerably equalized in singing; the corresponding vowels were integrated into one class. The results are presented in Fig. 3.
Figure 3. F1-F2 charts (averaged values) for singing. Top: Aukštaitėsiai samples. Bottom: Dzūkai samples. Vowels of Samples AV and JJ are denoted.

Thus, Dzūkai performances, in comparison with those of Aukštaitėsiai, show a) more different vowel systems, b) wider vowel systems (i.e., they cover larger areas in F1-F2 charts; the vowel classes in a separate sample differ more in phonetic quality), c) shift of vowel systems to higher values of F1 and F2, on average.
Discussion

First, it is natural to expect vowels to be phonetically more similar in singing in comparison with speech. That was generally shown for European academic singing (e.g., Sundberg, 1987: 117) as well as for traditional singing (Ross, 1992). This rule, however, is valid only partly for our case. It is obviously correct if applied to Aukščiai samples (Fig. 4; top): the Aukščiai vowel system covers considerably less space in singing compared to speech. For Džukai samples, however, this is not the case (Fig. 4; bottom). We may make the conclusion that the pronounced difference between spoken vowels remains large in singing (Džukai). Relatively (slightly) ‘covered’ spoken vowels tend to be even more covered in singing (Aukščiai). Thus, singing is likely to be a kind of ‘exaggerated’ speech in terms of covering. In simple words, the Džukai ‘sing as they speak’ (uncovered), and the Aukščiai ‘sing as they exaggeratedly speak’ (covered).

A considerable decrease of F1 and F2, i.e., ‘deep covered voice’ (singing compared to speech; Aukščiai) most probably results from the lowered larynx/protruded lips. It is worth mentioning that F1 and F2 of the vowel u do not actually differ in speech and singing, whereas other vowels tend to shift to u in singing. Thus vowel phonetics are somehow ‘u-based’ or ‘u-attracted’, i.e., ‘low F1-F2-based’.

A considerable increase of F1 and a slight increase of F2 (singing compared to speech; Džukai) can be attributed to a widened jaw opening, while the flexibility of articulation actually remains the same.

In general, ‘u-based’ vowel production corresponds to a lower position of glottis/labialization (Aukščiai singing) in comparison with ‘speechlike’ vowel production (Džukai singing).
The Effect of the Spoken Dialect on the Singing Dialect

Figure 4. F1-F2 charts (averaged values); comparison in speech and singing. Top: Aukštaičiai samples. Bottom: Dzūkai samples.

SPLs

Speech

The SPLs of the same vowels as in F1-F2 measurements discussed above were evaluated. Then the SPL values for each vowel class (u:, u, o:, etc.) in all Aukštaičiai/Dzūkai performances were averaged (Fig. 5).

The conclusion is drawn that the Dzūkai samples are characterized by significant differences in mean SPLs among vowel classes, especially among tense (‘long’)
vowels. Phonetically intense vowels (\(a\); \(o\)) are intense in speech as well, whereas phonetically weak vowels (\(u\); \(i\)) are correspondingly weak in speech performance.

Aukštaičių samples show quite different results. Namely, the differences in mean SPLs are not that pronounced and not systematic (or accidental), i.e., ‘loud’ and ‘soft’ vowels, in a phonetic sense, are not recognized as such in the speech performance.

**Figure 5.** Mean normalized SPLs (to average; in dB) averaged across Aukštaičių/Dzūkai samples (speech). Open diamonds/squares mark lax (‘short’) vowels (\(u, a, e\)).

**Singing**

As in the speech analysis, SPLs of the same vowels as in F1-F2 measurements discussed above were evaluated, and the SPL values for each single vowel class (\(u, o, a, e, ė, i\)) in each performance were averaged. Fig. 6 shows the results.

**Figure 6.** Mean normalized SPLs (to average; in dB) averaged across Aukštaičių/Dzūkai samples (singing).
The Effect of the Spoken Dialect on the Singing Dialect

More or less the same conclusions can be drawn as those for speech performances (see Fig. 5). Phonetically intense and weak vowels can even be recognized from sound signal graphs in the Dzūkai case (Fig. 7; bottom; compare intense a, e, e: with weak u, i,), whereas in the Aukštaičiai case this is hardly possible (Fig. 7; top).

Discussion
Noticeable differences in SPLs among vowel classes transfer from speech to singing performances (Dzūkai; compare Figs. 5 and 6), whereas the leveling of phonetic-dynamic differences in speech are reflected in singing (Aukštaičiai; the same Figs.). It was shown that two different modes of phonation-articulation are at work (Ambrazevičius, 2001: 176-177):

• Dynamic phonation × relatively smeared (equalized) articulation = output dynamics reflects predominantly phonation (Aukštaičiai);

• Dynamically flat phonation × distinct articulation = output dynamics reflects predominantly articulation (Dzūkai).

It should be stressed, however, that these modes represent two poles of dichotomy; individual performances are actually situated on the line of the continuum between these two poles (ibid).
Consonants

Two features connected with consonant articulation were considered: marcato (i.e., stressed attack), and vocalization of end consonants (i.e. addition of an extra vowel to a consonant in syllable-end position; such as i-r(i) ki, ka-d(i) gi).

As we will see, these features are noticeably less expressed and hardly perceived in speech in comparison to singing. In most cases of speech, the notions of marcato and vocalization should probably be put in quotation marks.

Marcato

Speech

The syllables pa- or starting with pa- (such as par-, pal-) were analyzed. The first ten pa(...)- syllables of each sample of speech performance were considered (80 measurements altogether). Two parameters were evaluated: the attack time (AT) and the time from the burst of a consonant to the end of the attack of a vowel (Fig. 8).

Figure 8. Attack time (AT) and time from the burst (beginning of release) of a consonant to the end of the attack of a vowel (BAT).
The Effect of the Spoken Dialect on the Singing Dialect

The conclusion can be drawn that, in comparison with Aukštaičiai performances, those of Džukai are characterized by shorter attack/burst-attack (AT/BAT) time values. Notice that Aukštaičiai/Džukai performances are more differentiated in BATs than in ATs.

**Singing**

The first ten syllables (with syllable-initial prevocalic consonants) of each sample of singing performance were considered (70 measurements altogether). Two parameters were evaluated: the time from the burst of the consonant to the end of attack of a vowel, and the decay in SPL succeeding the attack. In the case of fricatives and sonorants, the beginning of release instead of burst was considered. The results are generalized in Fig. 11.
In comparison with Aukštaitėsai performances, those of Džukai are characterized by shorter attack times and more pronounced attack spikes, formed due to considerable after-attack decays. This creates a more distinct marcato quality.

Vocalization of end consonants

Speech
The parameters of consonant ‘vocalization’ (see example in Fig. 12) in all samples of speech performances were evaluated (Fig. 13). The first ten reliable end consonants in the performances (such as kar-ku, pir-mas) were considered. ‘Reliable’ means that only end consonants that were fluently pronounced, not obscured by background noises, and not succeeded by pauses, were analyzed. The results are presented in Fig. 14.
The Effect of the Spoken Dialect on the Singing Dialect

Figure 13. Inter-onset-interval between the onset of a (potentially) vocalized end consonant and the end of attack of the succeeding syllable (IOI), total duration of ‘vocalization’ of the end consonant (TV), and difference in SPLs between the preceding vowel and the end consonant (ΔSPL).

Figure 14. IOIs, TVs, and ΔSPLs for speech performances.

Thus, there is some difference between the performances of end rs in Aukštaičiai and Dzūkai samples. The time parameters of the consonants show slightly larger values for Dzūkai performances. The decrease in SPL of an end consonant in comparison with a preceding vowel shows considerably lesser values for Dzūkai cases.

All this means that end consonants tend to be more ‘vocalized’ in Dzūkai cases, in the sense that their phonetic quality somehow tends to be shifted slightly towards vowels (larger durations and SPLs). It must be stressed, however, that these differences are slight.

Singing
The parameters of consonant vocalization (see example in Fig. 15, bottom) in all but one sample of singing performances were evaluated. Sample SM was omitted from
the analysis because of weak vocalization due to the specific genre of the song: this is a harvest song performed loudly in the open air, and thus vowels are extended while consonants are weakly articulated.

End consonants of closed syllables in the first melostrophes of the performances were considered, except in Sample VJ: here the consonants in all melostrophes were considered, because of their low frequency of occurrences (five in the entire performance). All in all, 29 occurrences were analyzed. The results are presented in Fig. 16.

**Figure 15.** Examples of non-vocalization (top; JP, Aukštaitija) and distinct vocalization (bottom; VJ, Džūkija) in singing performances. Horizontal: time (span 0.8 s), vertical: SPL (range 30 dB).

Thus the end consonants are vocalized more distinctly in Džūkai performances on average.
Discussion
The quality of syllable-initial attacks as well as that of syllable-closing consonants reflects the distinctness of speech enunciation. Transferred to singing, these qualities result in the distinctness of marcato and vocalization of end consonants respectively. From this point of view, singing also appears as ‘exaggerated speech’: quite faint differences in spoken performances (see Figs. 9 and 14) transform into noticeable differences in sung performances (see Figs. 10 and 16).

Measurements of durations of 288 syllables (in total) in spoken performances show values of 182 ms (Aukštaitija) and 181 ms (Dzūkai), on average. The similar values mean that marcato and vocalization of end consonants are not somehow related to the rhythm of speech.

Ornaments
The samples under investigation are not abundant in ornaments. For instance, only one distinct double-appoggiatura (or mordent) per melostrophe was found in Sample VJ. Nevertheless, the Aukštaitija/Dzūkai differences of the ornaments are worth mentioning. For Dzūkai performances, the mean amplitude of pitch excursions in the ornaments equals 134 cents, whereas for Aukštaitija performances it equals 268 cents and even exceeds 350 cents in some cases (see Fig. 17).

This means that in the Dzūkai cases the pitch excursions can be comprehended as up-and-down sequences of grace notes a second interval apart. In crucial Aukštaitija cases, however, the pitch excursions tend to be perceived as up-and-down glides with indefinite upper grace note. This can be explained by a qualitative trigger of perception mechanism when the interval exceeds three semitones (so-called trill threshold; Miller & Heise, 1950).
Note that the ornamental pitch and SPL tracks are not in phase (Fig. 17). This means that the phonation mode changes somehow in the course of the ornament performance.

At first glance, no direct impact of vowel phonetics on ornament production could be detected. However, based on our own singing experience, we can propose that, with a low and somehow lax glottis, it is easier to produce appoggiaturas of wide interval, whereas, on the contrary, a high and tense position of the glottis results naturally in lesser pitch deviations. Therefore, the types of ornaments can be related to the position of the glottis and, in turn, to the articulation of vowels.

Conclusions

Speech acoustics not merely affects the phonetics of singing but also the vocal technique, musical articulation, etc. This means that some traits of a singing dialect are neither ‘purely musical’ nor ‘purely vocal’, but rather speech-dependent. The fact that, in a sense, singing can be regarded as ‘exaggerated’ speech (or: speech can be regarded as ‘transformed’ singing) is likely to be quite natural from the point of view of ‘musilanguage’ (Brown, 2000).

References


1 Rytis Ambrazevičius has 20 years’ experience as a folk singer.
3 Aukštaitai and Džiūkai are the people of Aukštaitija and Džiūkija.
6 As proposed by Peters, Boves, & van Dielen (1986).
7 Here it was not possible to follow the same procedure as in speech analysis, i.e., to take the first ten *pa*-syllables of each sample. The songs are too short to find ten such syllables in each of them. Although temporal characteristics of attack depend on consonant, it was believed that statistical averaging takes place here.
8 The latter parameter was not discussed in the analysis of speech performances, as it only shows negligible values there. Only BATs (not ATs) are considered, since their values for Aukštaitai and Džiūkai performances are differentiated more clearly.
9 Remember that *e:* and *u:* are very similar.