

## More Rhythmic Interactions in Two (or Three) *Aksak* Performances

FERNANDO BENADON[1]

*American University*

**ABSTRACT:** Goldberg (this issue) presents evidence of interactions between percussive timing and song structure in two 3+2+2-meter Balkan performances. Using interonset measurements and checking carefully for statistical significance, he shows that formal and contextual characteristics of each performance coincide with timing changes at the level of the beat and bar, with durations becoming slightly shorter or longer at consistent locations. I review Goldberg's main findings and contribute brief supporting analyses. For one of the songs, I provide additional context regarding verse meter, melodic structure, anacrustic drive, links between timing and breathing, and offsets between taps and syllabic onsets. For the other song, I attempt to tease out the difference between shortened non-isochronous beats (an altered *aksak* ratio) and shortened bars (an accelerated tempo). I end with a brief analysis of a different song by the same performer as the first song, showing that the timing effects Goldberg observed in one are also present in the other.

Submitted 2015 March 2; accepted 2015 September 23.

**KEYWORDS:** *rhythm, Romani folksong, Macedonia, breathing, Muzafer Bizlim*

GOLDBERG'S analysis of percussion patterns in two contrasting *aksak* performances draws out interesting similarities between them. While Goldberg is quick to say that the timing features verge on imperceptibility—the tight rope walked by many a microtiming analysis—and that unbridled exploration of the data would have been best avoided, I find his quantitative rigor persuasive. I also appreciate that Goldberg's statistical offering does not come at the expense of alert listening. Someone casually browsing his paper may notice mostly graphs and tables. The thorough reader is instead rewarded with a thoughtful musical analysis supported by abundant data. In this review I summarize Goldberg's main findings while introducing some observations and analyses of my own.

Both analyzed recordings employ a 3+2+2 metrical framework; see the two rhythms in Goldberg's Figure 3. In one, Muzafer Bizlim finger-taps the accompaniment while singing a Romani folksong in his home in Shutka, Macedonia. In the other, Bulgarian percussionist Mitko Popov accompanies a folk ensemble's rendition of a popular song while audience members line-dance the *lesnoto*. Despite obvious contextual differences between the two performances, parallel timekeeping behaviors emerge.

### MELODIC FEATURES OF *BEŠLI I ČAJ TELE*

Before delving into Goldberg's discussion of timing in "Bešli i čaj tele," it is a good idea to familiarize ourselves with the song's melody. The transcription appears in Figure 1.

The opening notes' dominant-to-tonic gesture distinguishes the top of the form from the remaining measures in two ways: it contains the largest leap (whereas the remaining is mostly stepwise with some third leaps) and it contains the only note well outside the melody's prevailing range (a major pentachord spanning scale degrees 1 to 5). Each line in the transcription corresponds to a line of text and traces a different tonal path: phrase A ascends to scale degree 3 and then descends to 1, while phrase B circles around 3 and then comes to rest on 1. The two middle lines are near-reflections of each other, giving rise to a curious sensation of déjà vu (or écouté) when these lines share the same text in the first three verses.[2]



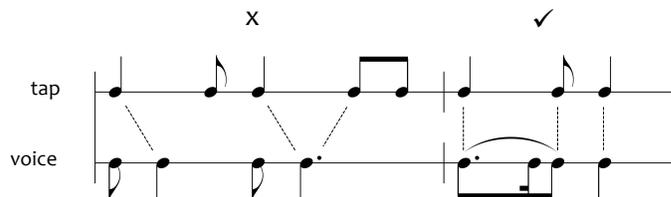
**Fig. 1.** Transcription of “Bešli i čaj tele.” Each 16-bar verse is preceded by a two-bar tapping-only interlude.

Focus now on the melody’s rhythm. The first beat of every odd bar contains a syncopated short-long pattern with an eighth note on the downbeat and a quarter note on the offbeat.[3] This is immediately followed by another syncopation: all odd bars end with a tied eighth note anticipating the third beat. The even bars contain no syncopation. Instead, a dotted melisma leads consonantly from the first beat to the second.

The syncopations in the odd bars are enhanced by syllabic stress. I was fortunate to count on the help of UT Austin’s Ian Hancock in parsing the song’s text and determining the location of syllabic accents. A linguist of Romani origins, Professor Hancock confirmed my impression that the song’s text is in iambic tetrameter. The first two feet are standard iambs, the unstressed syllable is deleted from the third foot (there is no ~ pickup), and in the last foot the ~ rearticulates the previous vowel melismatically instead of introducing a new syllable. Thus:

~    \_        ~    \_        \_        ~    \_  
 beš li    i čhaj    [ ] te    [e] le

How does the iambic nature of the melody fit with the tapped *aksak* ostinato? Not too well, since one is in 6/8 and the other in 7/8. Figure 2 shows that the resulting effect is one of accentual offsets (x) in the odd bars followed by matches (✓) in the even bars.



**Fig. 2.** Rhythmic counterpoint between Bizlim’s singing and tapping. The first measure’s offsets are resolved in the second.

Since the composite rhythm of voice plus tap is identical—except for occasional variants—across all odd bars as well as across all even bars, we have a chain of offset-match pairs forming two-bar groups; each of these groups appears four times in phrase A and four times in phrase B (counting repeats). Borrowing from Cooper & Meyer (1960), we can see this sequence as a fractal-like analogue of the text’s iambic scansion, where  $\sim$  precedes  $-$  in a recurring fashion four times per line. Add to this the tonal trajectories discussed earlier, where a ‘weak’ half-phrase that ends on *mi* precedes a ‘strong’ half-phrase that ends on *do* four times per verse, to get:

4 unstressed-stressed feet per line:	$\sim$	$-$	$\sim$	$-$	$-$	$\sim$	$-$
4 offset-match lines per phrase:	x	✓	x	✓	x	✓	x
4 weak-strong phrases per verse:	<i>mi do</i>						

Operating concurrently at the levels of verse, phrase, and line of text, this three-tiered ‘tetrametric’ framework provides the site on which Bizlim’s timing manipulations take place.

### WHERE IS THE DECELERATION PEAK?

In Bizlim’s and Popov’s performances, a non-sung transition leads to the following sung section. Toward the end of these transitions, the tempo (as determined through bar durations) begins to slacken slightly until reaching its slowest point a bar or two past the sectional downbeat, after which ‘normal’ speed is resumed. The tempo is thus a bit slower at the boundary region between verses.

The decelerations span roughly three bars and appear to cut through the section boundaries, peaking right after (not on) them. Of course, this feature is hardly unique to Balkan timing. As Goldberg mentions, work on tempo rubato shows that structural decelerations, while generally clustered around sectional seams, can follow a variety of trajectories. Crucially, these trajectories need not peak prior to the section boundary, and may spill over to the initial moments of the following phrase.[4]

Goldberg notes that the peaks in Bizlim’s decelerations occur on the second bar of each verse’s first phrase; this is easily seen in Goldberg’s Figure 5. It is tempting to interpret these peaks as deceleration spillovers that extend not just onto the structural downbeat but well beyond it. But in fact the second bar only balloons in size on its last beat, and there is good reason to believe that this beat is lengthened for a different reason (discussed below).

Figure 3a shows that Bizlim’s bars 1 and 2 are equally long once we remove the third beat from the equation.[5] The deceleration endpoint therefore lies somewhere along bar 1.

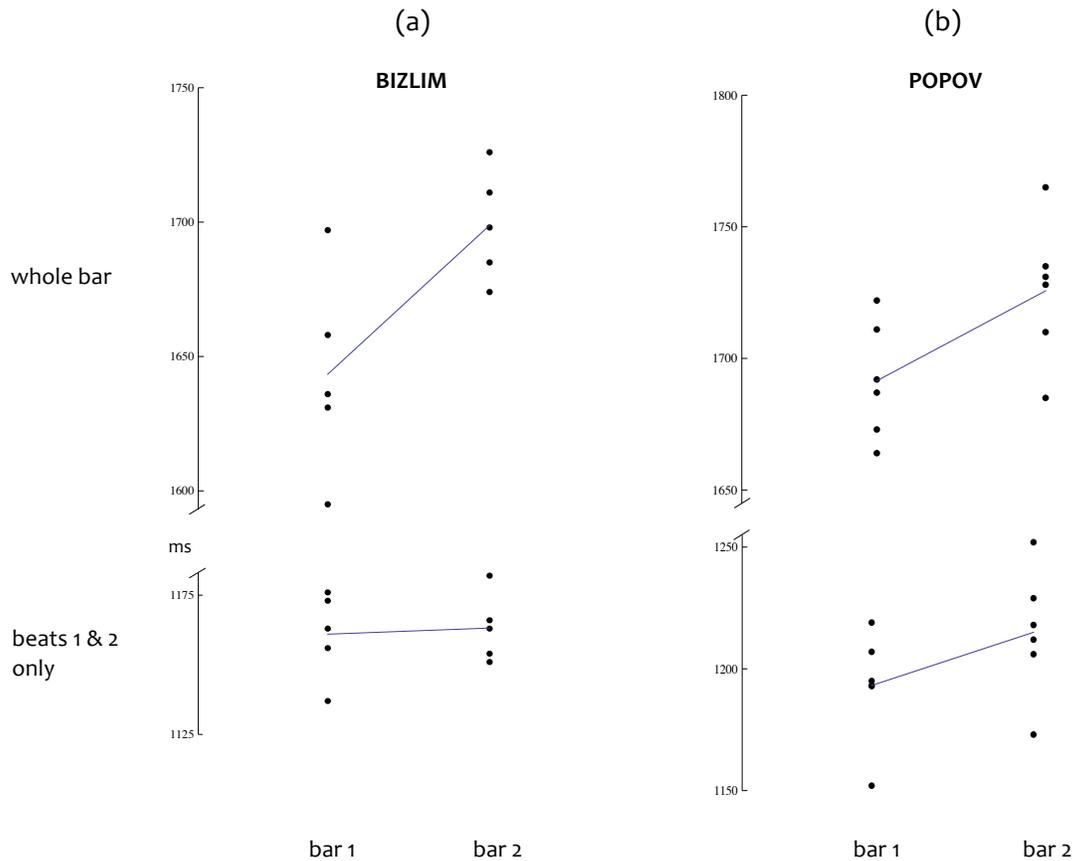
Like Bizlim’s, Popov’s recording features bar duration peaks past the structural downbeat. But while the former case appeared to be an artifact of whole-bar measurement (in the sense that measuring at the beat level revealed an earlier peak), the same may not be true of the latter.

Figure 3b compares Popov’s durations for the phrase’s initial bars with and without their third beats. Unlike the previous figure, a mild deceleration is also visible when only the partial bar is measured. This suggests that the deceleration’s endpoint lies somewhere after bar 1, pointing to a possible spillover past the verse’s structural downbeat.

Both examples remind us that the slowing at the end of a section need not end at the end of the section. When the deceleration continues through the boundary, it may do so to varying extents. Goldberg’s two examples appear to differ from each other in that regard.

### BIZLIM’S AKSAK VARIANTS

As mentioned earlier, bar 2 of every phrase in “Bešli” ends on the mediant while bar 4 ends on the tonic. This feature connects to Goldberg’s finding that the four-bar groups are marked by rhythmic nuances in the finger-tapping. The most consistent of these nuances is the shortening of the first beat on every fourth bar: the tapping is a little faster during the beat that accompanies the cadential *re-mi-re-(do)* figure than during the non-cadential *re-do-re-(mi)*.

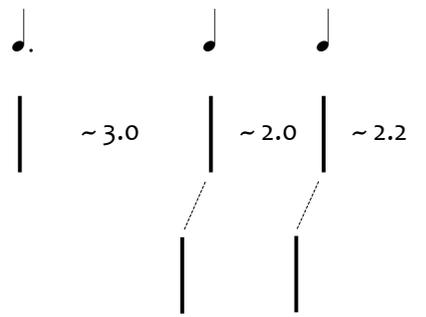


**Fig. 3.** Millisecond durations for the first two bars of phrase A in Bizlim's recording (a) and phrase D in Popov's recording (b). The lines show mean values. The difference observed at the bar level (top graphs) remains when the third beat is removed in Popov's case (bottom right) but not in Bizlim's (bottom left).

Goldberg does not speculate on why Bizlim shortens the first beat when the melody is headed to *do* but not *mi*. Perhaps the shortened first beat on cadential measures heightens their anacrusic drive. Before we see how, consider Goldberg's evidence that the third beat is consistently longer than the second despite their categorical equivalence (both are quarter notes). His Figure 4 shows that this is the case for all bars, not just the cadential ones ("bar 4").

The third beat is long because the start of the following *aksak* downbeat is delayed, not because the third beat arrives early. This can be deduced from the reported inter-beat ratios of 3:2.01:2.19 (instead of 3:2:2).[6] Given this, and after Butterfield (2006), we can qualify the third beat of each *aksak* cycle as continuative (pointing backward in time) rather than anacrusic (pointing forward). In other words, the forward motion begins at the downbeat and comes to rest at the end of the bar. This interpretation is consistent with Goldberg's suggestion that the longer third beat could be an indication of group-final lengthening operating at the level of the bar with each iteration of the *aksak* cycle (p. 312).

Now consider the cadential bars, where the first beat is shortened. Butterfield's (2006) claim that early onsets point forward in time is applicable here. For to say that the first beat is shortened amounts to saying that the second and third beats arrive early, as schematized in Figure 4. This generates forward motion at the tactus level. The tentative conclusion I draw from all this is that the cadential bars have a little extra push leading to the last beat.[7]



**Fig. 4.** Beat placement in regular (top) and cadential (bottom) bars. The shorter first beat in cadential bars results from an early second beat.

The shortened first beat at cadential bars is not the only way Bizlim delineates groups. Goldberg relates the longer third beat at the non-cadential bar (phrase A’s bar 2) to Bizlim taking a breath at the end of the first line of text. I was curious to see if the link between taking a breath and lengthening the third beat happens elsewhere in the performance. Aside from the spot just mentioned and excluding the interlude, Bizlim always takes a breath just before the start of both iterations of phrase B. But this point cannot be used for analysis because it coincides with a major group boundary and results in a confound. This leaves only four instances where Bizlim takes a breath during a within-phrase bar. All occur in the second iteration of phrase B: once at the end of the second bar in the third verse and three times at the end of the third bar in the fourth, fifth, and sixth verses. In three out of four pairwise comparisons, the third beat is longer when it coincides with a breath.[8] While four data points do not give us a lot to go on, when these are considered in conjunction with the consistent effect at phrase A’s second bar, they reinforce the possibility of a link between lengthened beats and taking breaths.

Goldberg goes on to say that while the breathing-related elongation probably lies below conscious perception, its coupling with the “audible boundary ... makes this timing pattern a candidate for group-final lengthening” (p. 318). Another timing-based articulation that is certainly perceivable by the naked ear is Bizlim’s behind-the-beat singing on cadential bars. His vowel onsets on the non-cadential *mi* are generally coincident with his tap, whereas on the cadential *do* they usually occur toward the end of slowly descending pitch glides and are thus delayed considerably. The offset magnitudes for verses 1 through 5 appear in Table 1. (There are no such offsets in the sixth verse.) The time interval between tap onset and vowel onset in non-cadential *mi* bars tends to be perceptually negligible—14ms on average. In contrast, the offset in cadential *do* bars often exceeds 100ms and is detectable without the slightest effort.

verse	A2 ( <i>mi</i> )	A4 ( <i>do</i> )	A2 ( <i>mi</i> )	A4 ( <i>do</i> )	B2 ( <i>mi</i> )	B4 ( <i>do</i> )	B2 ( <i>mi</i> )	B4 ( <i>do</i> )
1	9	92	10	206	18	128	28	319
2	19	86	12	177	3	149	16	244
3	7	58	6	-30	27	88	-10	258
4	50	230	33	139	8	145	21	354
5	-8	98	5	111	17	107	16	276
<i>mean</i>	<i>15</i>	<i>113</i>	<i>13</i>	<i>121</i>	<i>15</i>	<i>123</i>	<i>14</i>	<i>290</i>

**Table 1.** Time intervals (ms) between Bizlim’s tap onset and syllabic onset on the second beat of even bars. Positive values denote tap followed by syllable.

Table 1’s last column shows a more than twofold increase in the tap-to-vowel offset as compared to the other three cadential bars. This demarcates the end of the 16-bar verse, a grouping boundary discussed above in connection to bar durations being longest at the beginning of the verse and shortest at the end.

**POPOV’S AKSAK VARIANTS**

A central difference between the two recordings relates to their *aksak* ratios. Goldberg’s Figure 8 shows side-by-side measurements for the long-short-short beats. The lengthening of the third beat that is so prevalent in Bizlim’s performance is not visible in Popov’s data. But this does not mean that Popov’s beats are deadpan throughout. He, like Bizlim, makes slight changes to beat durations at the end of phrases.

Goldberg’s Figure 10 shows shortened first and third beats on phrase E’s last two bars. (The form is CC+D+EE; C is an instrumental *otsvir*; D and E are vocal phrases; each letter is eight bars except for an extra bar after the second C.) He proposes that these shorter IOIs, in conjunction with accented beater strokes and coincident with the bagpipe entry and the singers’ held final note, could serve to signal the approaching section boundary to the musical participants via acceleratory “cadential anchoring” (Ashley 2002). This is a very reasonable suggestion, although I think a distinction should be made between shortened beats and shortened bars. Clearly, if two bars have equal *aksak* ratios but one bar is shorter than the other, then its beats will necessarily be shorter, too. Is this the case at the end of E, in the sense that the shortened beats are a byproduct of a larger-scale acceleration (and vice-versa)? Or does Popov alter the inherent profile of the *aksak* ratio, as we saw earlier with Bizlim’s cadential bars?

Looking at points 15-16 and 23-24 (corresponding to the durations of E’s two final bars) in each sectional slice of Goldberg’s Figure 6, it is not immediately clear whether these bars are faster than the others. A comparison of phrase E’s bars 3-4 against bars 7-8 for all six verses yields a significant difference, with the latter being shorter.[9] So the final bars are indeed a bit shorter, which explains the shortened beats.

The shifted stacking order of the plot lines in bars 7 and 8 of Goldberg’s Figure 10 indicates that phrase E’s final bars are ratio-modulated in addition to being tempo-accelerated. To confirm this, we can map the *aksak* ratios “chronotopically” (Benadon 2014, Desain & Honing 2003); see Figure 5a. Each point in Figure 5b corresponds to an individual bar, with the bar’s long-short-short ratio determining the point’s location. (That is, the graph plots proportions, not absolute durations.) The orange points correspond to bars 7 and 8, the purple points correspond to bars 3 and 4, and the larger landmark points are deadpan meters. While the two colors show some clustering, the separation is not particularly striking. The clustering comes into sharper focus when we remove bars 3 and 7, as in Figure 5c. The evidence for a transformed *aksak* at the end of phrase E is more compelling for bar 8 than for bar 7.

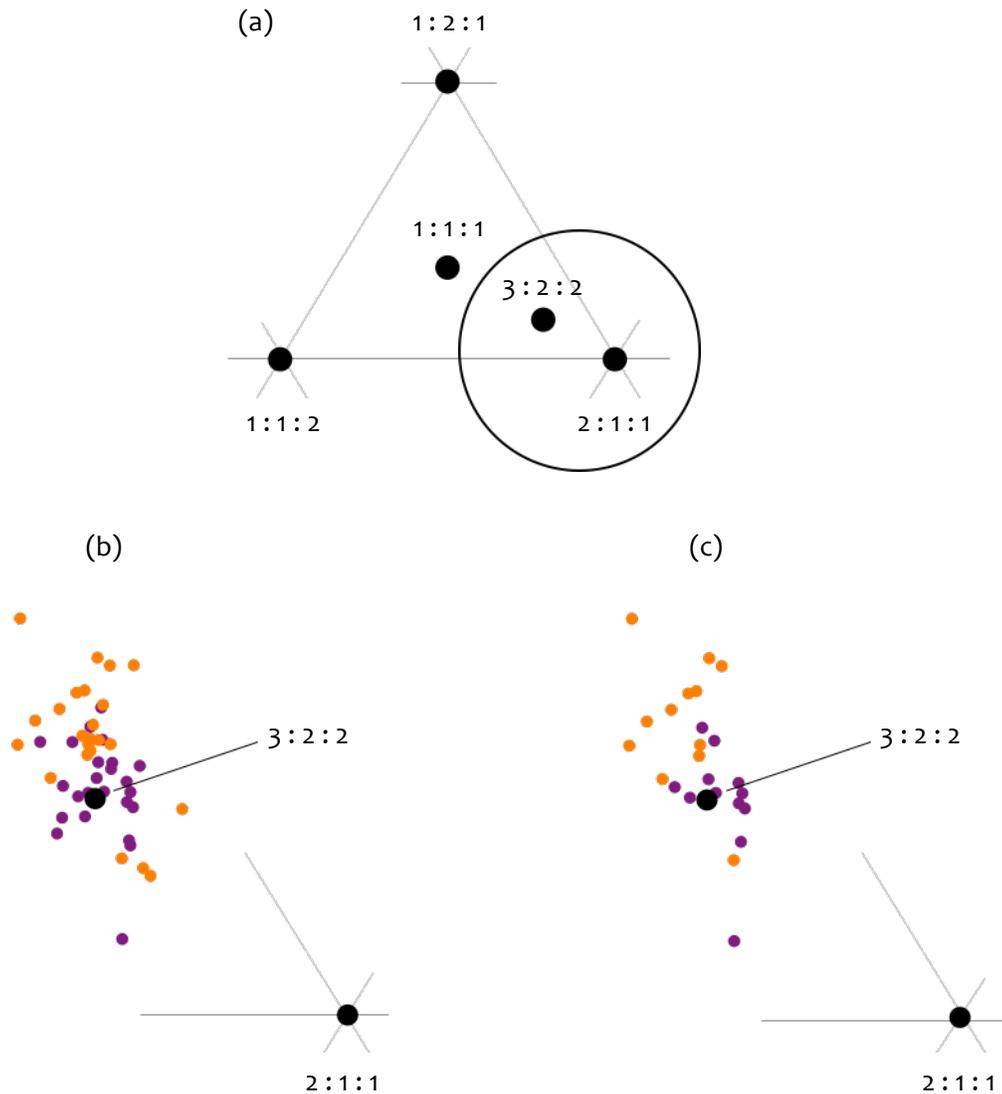
**ENLARGING THE SAMPLE**

Before closing, I will take a modest step toward heeding Goldberg’s call—echoing Huron (2013)—for “tests of much larger samples” (p. 306). As it happens, the CD containing Bizlim’s performance also contains seven others by him, all with finger-tapping accompaniment.[10] “Dzemo Dzemile” uses the same *aksak* pattern as “Bešli,” is also strophic, has almost the same number of bars (104 vs 108), and is only slightly slower in tempo. Structurally, however, the two songs are far from alike. Recall that “Bešli” falls neatly into four-bar multiples, with two-bar transitions. “Dzemo Dzemile”’s phrase lengths of 5, 7, 10, and 11 bars cut a very different profile:

section:	<i>trans.</i>	<i>verse</i>	<i>chorus</i>	$3x$
no. of bars:	2	5 + 7	11 + 10	:

An in-depth comparison being beyond the scope of this review, we can set our sights on whether the main tendencies found by Goldberg in “Bešli” are also present in “Dzemo”—namely, whether:

1. The first beat is shorter at the end of vocal phrases
2. The third beat is consistently longer than the second beat
3. Bars are longer at the beginnings of vocal sections and shorter just before them
4. Taking a breath results in a longer coincident (third) beat



**Fig. 5.** Chronotopic representation of interbeat ratios in Popov’s phrase E. (a) Basic map architecture with circled zoom area. (b) Bars 3 & 4 (purple) and 7 & 8 (orange). (c) Bars 4 (purple) and 8 (orange).

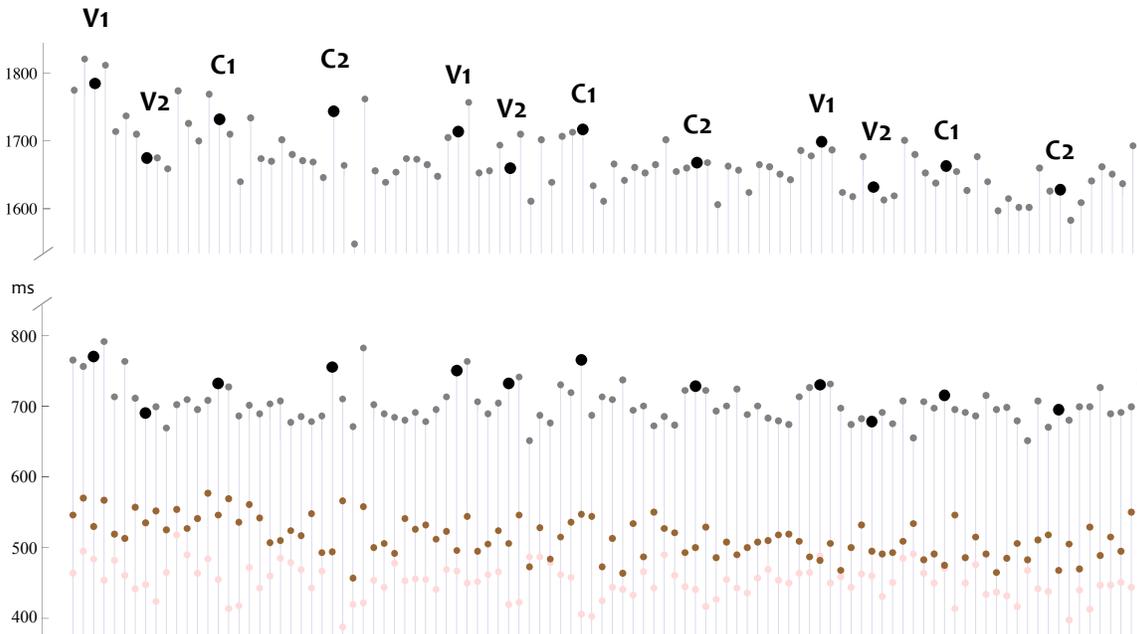
The first three questions can be addressed with one graph. Figure 6 shows bar and beat durations for all bars in “Dzemo Dzemile.”

1. Careful examination of the lower graph and associated values reveals no clear patterns in relation to the duration of first beats. Their shortening in “Bešli”’s cadential bars was probably a feature of that particular performance.

2. The third beat (plotted in brown) is consistently longer than the second (light red), resembling the *aksak* profile in “Bešli.”

3. While some phrase beginnings (larger points) coincide with bar duration peaks in the upper graph, the effect is not nearly as robust as in “Bešli.” We do find a related recurring pattern, however: with the exception of the first chorus, the first iteration of every verse and chorus begins more slowly than the

second. (In the top graph, the second large point is below the preceding large point in five out of six pairs.) This is analogous to the group-initial peaks in “Bešli.”



**Fig. 6.** Bar (top) and beat (bottom) durations in “Dzemo Dzemile.” Larger points mark the start of verse (V) and chorus (C) phrases. In the lower graph, gray = beat 1, light red = beat 2, brown = beat 3.

4. As was the case in “Bešli,” Bizlim takes breaths in consistent locations in “Dzemo,” and always at the end of the bar. The mean duration for the third beat on these bars is 527 ms ( $n = 32$ ), higher than the mean of 509 ms ( $n = 71$ ) for bars containing no breaths ( $p < .01$ ).<sup>[11]</sup> This suggests that the coupling observed in “Bešli” is not unique to that performance and may pop up in other vocal performance contexts. The likely imperceptibility of the effect should not be cause for concern. What matters, Goldberg reminds us, is its “co-occurrence with meaningful musical action” (p. 318).

The comparison between Bizlim’s two performances indicates that most of the effects found in one are also present in the other. We cannot at this point generalize these findings beyond these two songs, let alone beyond this single performer. Nonetheless, the timing tendencies shared by “Bešli i čaj tele” and “Dzemo Dzemile” highlight Goldberg’s overarching argument that timing is inextricably, and often consistently, linked to form and action.

## NOTES

[1] Correspondence can be addressed to fernando@american.edu.

[2] F major is used for convenience. Bizlim’s performance begins in F-sharp and gradually drops two semitones to E by the end.

[3] The grace note at the beginning of the second line mitigates the syncopation, while the ascending leap at the beginning of the verse heightens it.

[4] We are speaking here of melodic phrases that end before the barline. Elisions are of course likely to have the deceleration peak after the barline.

[5] The first verse is omitted because its initial bars are disproportionately slow. The data plotted in this and subsequent graphs are derived from my interonset measurements.

[6] Adjusting the third beat's ratio from 2.19 to a deadpan 2.0 by delaying its onset would increase the ratio of the second beat above 2.01. The only way to obtain the reported ratio is by delaying the start of the next cycle. The mean ratios derived from my measurements are identical to Goldberg's. This does not rule out possible differences between our datasets. One factor that might contribute to such differences is the approach used to mark tap flams. Bizlim's flams are relatively frequent and span anywhere from 10 to 30ms. I marked the second onset; Goldberg's method is not reported.

[7] I am grateful to Matthew Butterfield for his thoughts on this example.

[8] Pairwise comparisons are between breath bars in the second iteration of B and matching non-breath bars in the first. Beat 2-to-beat 3 ratios for the four non-breath vs breath pairs are .95 vs .87, 1.00 vs .88, .93 vs .99, and .98 vs .85. Lower values denote longer third beats. In absolute durations, the four ratio pairs give differences of 42, 53, -29, and 78 ms—e.g., in the first comparison, the third beat is longer than the second by 23 ms when no breath is involved as opposed to 65 ms four bars later when Bizlim takes a breath, a pairwise difference of 42 ms. Note that the smallest absolute magnitude (-29 ms) corresponds to the exception.

[9] I obtained the following mean bar durations for bars 3-4 (first iteration), bars 7-8 (first iteration), bars 3-4 (second iteration), and bars 7-8 (second iteration): 1666 (26), 1635 (23), 1662 (28), and 1626 (20) ms. Bars 3-4 were chosen for comparison because they occupy a similar hypermetric location as bars 7-8 and because they show minimal inter-beat variability in Goldberg's Figure 10.

[10] Four of the songs are in 4/4, one is in 2+2+2+3, and three (including the two discussed here) are in 3+2+2.

[11] The last bar was excluded because it is swept by a substantial song-final deceleration.

## REFERENCES

- Ashley, R. (2002). Do[n't] change a hair for me: The art of jazz rubato. *Music Perception*, 19(3), 311-332.
- Benadon, F. (2014). Metrical perception of trisyllabic speech rhythms. *Psychological Research*, 78(1), 113-123.
- Butterfield, M. (2006). The power of anacrusis: Engendered feeling in groove-based musics. *Music Theory Online*, 12(4). Retrieved from <http://www.mtosmt.org/issues/mto.06.12.4/mto.06.12.4.butterfield.html>.
- Cooper, G. W., & Meyer, L. B. (1960). *The rhythmic structure of music*. Chicago: University of Chicago Press.
- Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception* 32, 341-365.
- Huron, D. (2013). On the virtuous and the vexatious in an age of big data. *Music Perception*, 31(1), 4-9.