

# Liminal Utterances and Shapes of Sadness: Local and Acoustic Perspectives on Vocal Production among the Yezidis of Armenia

ESTELLE AMY DE LA BRETÈQUE, BORIS DOVAL, LIONEL FEUGÈRE,  
LOUIS MOREAU-GAUDRY<sup>1</sup>

The Yezidis are a Kurmanji-speaking (northern Kurdish<sup>2</sup>) religious minority living scattered across northern Iraq, Syria, the Caucasus (Armenia and Georgia), and western Europe.<sup>3</sup> Among the Yezidis of Armenia, “speech” (*axavtin*), “song” (*stran*; pl. *stranen*), and “words about” (*kilamê ser*; pl. *kilamen ser*) constitute the three main categories of vocal production. This article is a detailed exploration of the acoustic characteristics of these categories of vocal production with a special focus on “words about.” At first sight, one may question whether “words about” (*kilamê ser*) actually amounts to a separate vocal category. These words lack any name that would set them clearly apart from *axavtin* (speech), which is clearly also made up of “words” (*kilamen*). Furthermore, the Yezidis never refer to an abstract category of “words about” but only to “words about” *something*: “words about the dead” (*kilamen ser mirya*), “words about exile” (*kilamen ser xeribîê*), “words about the hero” (*kilamen ser mêranîê*), and so forth. What these topics share in common is their association with tragic events and/or feelings of sadness and nostalgia; “words about” are thus lamenting utterances.<sup>4</sup> When Yezidi people (especially women) say “words about” such things, they often resort to a special tone of voice. This tone may be described as a kind of chanting that combines characteristics of *axavtin* and *stran*. As with *axavtin*, there is neither metre, rhyme, nor steady beat. And as with *stran*, there are sustained pitches and an

---

1. The authors' names are in alphabetical order. We are grateful to Victor A. Stoichiță and to the anonymous reviewers for their inspiring comments.

2. Kurdish is an eastern Iranian language. The major dialects of Kurdish are Sorani, spoken in most of the Iraqi and Iranian Kurdish region, and Kurmanji, spoken in Turkey, the northernmost parts of Iraqi and Iranian Kurdistan, Syria, the Caucasus, and Khorasan.

3. On the origins and intricacies of the Yezidi religion see, for example, Kreyenbroek (1995) and Kreyenbroek and Rashow (2005).

4. *Kilamen ser* include what are commonly understood as laments, and are inclusive of other kinds of utterances that register sorrow, pain, and loss.

interactional salience. These differences are all addressed in more detail below.

The present study reflects the joint work of one ethnomusicologist and three acousticians. Yezidi vocal practices previously have been addressed by Amy de la Bretèque from anthropological and ethnomusicological perspectives (2010, 2012, 2013). We have extended this work by inquiring whether analysis of precise acoustic features might shed light on the way Yezidi people make use of and conceive their voices. Ethnomusicologists have long borrowed concepts and tools from acousticians. In relation to vocal production, these tools are generally deployed to describe certain aspects of sound that might theoretically be heard, albeit not by untrained ears (e.g., Charron 1978; Vaughn 1990; Fales 2002). Spectrograms, for instance, are favourite ways of representing characteristics such as overtones, glottal stops, or minute ornaments (e.g., Zemp 1996). Here, however, our approach differs from what is typically found in ethnomusicological studies. Extracting “low-level” features is common in the literature on acoustics, and many relevant calculations may be carried out with user-friendly, free software.<sup>5</sup> However, such computations are not usually part of the ethnomusicologist’s toolkit. Thus, in addition to the initial aim of better understanding Yezidi vocal practices, this paper also serves as a trial for analytical methods seldom applied in ethnomusicology.

Our study takes as its focus a single five-minute field recording of a casual conversation that took place in September 2006 in the village of Alagyaz near Mount Aparan in Armenia. In this recording, one can hear *axavtin*, two instances of *kilamen ser*, and one short *stran*. The first section of this article describes the ethnographic context in which the example was recorded and the local typology of vocal productions. We then proceed with an analysis of the recording from both musical and acoustic perspectives. We measure the distribution of many “low-level” acoustic parameters across the three vocal categories. From all the features measured, we selected a sample that ensures a detailed and focused description of all the utterances contained in the recorded example. This restricted set of acoustic parameters enables us to compare the three vocal categories distinguished by Yezidis and put forward a new hypothesis regarding their interactional relevance, linked with ethnographic observations. Finally, the last section summarizes some of our findings and discusses their relevance for the anthropology of laments.

## A SHORT CONVERSATION WITH HBO

In 2006, Hbo was more than sixty years old. She spent her entire life in Alagyaz, a village located in Armenia on the road running from Yerevan to Tbilisi. She raised two daughters and had no sons. Her husband died a few years after their

---

5. We use the open-source software Praat (see <http://www.fon.hum.uva.nl/praat/> for more details).



Figure 1. Hbo (left) with her daughter (middle/back) and granddaughter (right), in Alagyaz, Armenia (photo: Estelle Amy de la Bretèque, September 2006).

wedding. Without any male descendants, Hbo says she feels a permanent pain in her heart. As with many Yezidi women of her generation, Hbo defines herself as a “burning heart” (*dilşewat*) and says that she has no hope of treating her “illness”—the deep sorrow that burns her heart. *Dilşewat* is not (just) an emotional condition, but extends to a complex social status that reaches well beyond the embodiment and enactment of certain feelings. Hbo is known in the village for her active participation in funerals, where she always utters “words about the dead” (*kilamen ser mirya*), even when the deceased is not of her own family (see figure 1). She says that, as a *dilşewat*, she feels sincere pain in her own heart. This contrasts with the professional male mourners who are sometimes hired for such events. However, beyond her own feelings at that moment, she is also entitled to a kind of public voice, which other Yezidi women are not granted. When it comes to sadness and tragic events, women with “burning hearts” talk and are listened to in a distinct way that relates to their ongoing, internal martyrdom (on the status of *dilşewat*, see Amy de la Bretèque 2013, 2016).

The recording analysed in this article was made at the house of Hbo’s younger daughter.<sup>6</sup> Her nine-year-old granddaughter was in the room

6. The entire recording is available online at <http://www.ictmusic.org/ytm2017> and <http://www.ebreteque.net/liminalutterancesandshapesofsadness>. We have included a file of the entire five-

throughout the recording. Her daughter and her six-month-old grandson were also present in the beginning (up to 1:33 in *sound\_entireFile.wav*, the entire sound file), then they left the room before returning a few minutes later (at 3:30). Two more people were in the audience: Cemile, the village nurse, and one of the authors, Estelle Amy de la Bretèque (who accompanied Cemile on her weekly visit to monitor Hbo's blood pressure).<sup>7</sup>

In the first part of the recording (*sound\_part1\_speech.wav*), Hbo and Cemile converse about a young man who killed himself some time ago. They ask, "Why did he do that to himself? He had everything. What did he lack?" Then Hbo turns towards her baby grandson, checks whether his temperature is alright, and sings a short *stran* for him (*sound\_part2\_song.wav*; see figure 2). The two women go on talking about whether the baby will start walking this year and about his present unwillingness to fall asleep. Then Hbo's daughter, the baby's mother, takes him out of the room to prepare some coffee. Hbo turns toward her granddaughter, gives her a few warm, cuddly words, and calls for her to draw nearer in order to hear a few *kilamen* (*sound\_part3\_speech.wav*). Hbo continues by telling "words about" a man who dies leaving his kin in sorrow (*sound\_part4\_lament.wav*; see figure 3). The *kilamen ser* ends as her daughter returns to the room with some coffee and sweets. The baby also comes back with her. Cemile talks about the difficult times when there was no gynaecologist in the village. She remembers how Hbo's mother-in-law had acted as a midwife in times past. Hbo then recalls that once she came home and found her mother-in-law attending to Cemile, who was lying ill on the bed (*sound\_part5\_speech.wav*). This prompts Hbo to utter other *kilamen ser*, this time evoking a wounded mother and the medicine necessary to cure her (*sound\_part6\_nonMelodizedLament.wav*; see figure 4). As the baby is in the room, Hbo does not melodize her speech because it is deemed dangerous for infants. The *kilamen ser* ends with Hbo recalling not only that her mother-in-law was very good at performing *kilamen ser*, but that she was also big, strong, liked to eat a lot, and that she sometimes encountered problems getting through doorways.

The recording of Hbo conveys only a partial illustration of what song, speech, and "words about" can sound like among Yezidis in Armenia. A larger collection of audio and video documents is available online at <http://ethnomusicologie.fr/parolesmelodisees>.<sup>8</sup> Nevertheless, this short excerpt

---

minute field recording (*sound\_entireFile.wav*) as well as six shorter segments from this master recording that correspond to the voicing types used by Hbo (see List of Audio Files). Unless specifically indicated, references in this article are to the entire sound file and are indicated, in parentheses, with a time stamp.

7. Another of Hbo's stories is analysed in Amy de la Bretèque (2012).

8. The collection is mirrored at <http://ebreteque.net/parolesmelodisees>. Docs 19 and 20, in particular, illustrate song; docs 7, 50, and 51 are examples of "words about."

remains interesting in that it offers the opportunity to study types of vocal production as deployed by the same person in the same daily setting, a circumstance most favourable to comparing their acoustic features.

## A LOCAL TYPOLOGY OF VOCAL PRODUCTION

*Axavtin* has roughly the same meaning as in English: speech that is used in daily life, often in dialogic utterances. It also has no stable pitch, no rhyme, and no metre. Here is a short transcription of *axavtin* (from 4:27 to 4:47 in *sound\_entireFile.wav*):

**Hbo:** *Davete ser roste zavitwana. Sere bavé xwe. Davete ser xwa. Kevanitiya xwe.*

**Hbo:** She<sup>9</sup> used to improvise words about different people, about her life, about her talents in housekeeping.

*Xasi ya mi yeke sag girederani bu. Wan deriyan ra ranedat!*

My mother-in-law was very strong and in good health. She was so tall that she could not pass through the doorway!

**Cemile:** *Aliye xwarian ra da bas sindirli bu!*

**Cemile:** She also liked her food very much; she was eating from the pan while cooking!

*Stran* is delivered with an isochronic beat, metred verses, and with rhyming text. It is typically heard at gatherings and joyful events. Melodic lines are repeated many times, and the sung words are generally quite simple. Here are the lyrics of the short *stran* performed by Hbo (from 0:38 to 0:46):

*Dexmano bê xwedine.*

I swear he is so pretty.

*Efsune bê te dine.*

Everybody would become crazy without you.

*Emê qurbane qalkê.*

We will be sacrificed for you.

*Kilamê ser* are uttered with “free” rhythm and stable pitches. The words are neither metred nor rhymed. They often tell a story linked to sorrow, death, pain, or exile. They are heard during funerals and mourning ceremonies, where they are performed by bereaved women or professional male singers. They are also uttered by men in casual gatherings to narrate epic or heroic stories. *Kilamen ser* are often inserted seamlessly into daily speech, as in the conversation excerpt analysed here. This is a transcript of Hbo’s first *kilamê ser* (from 1:35 to 2:14):

---

9. “She” refers to Hbo’s mother-in-law, who was also the village’s midwife.

<i>Ezê bejim, minê got: “Ax digiri dilê min.”</i>	I would say, I said: “Ah my heart is crying.”
<i>Feyda tine, bire min.</i>	There is no healing, my brother.
<i>Nizanim çawa hale min.</i>	I don’t know how my situation is.
<i>Gelo kê bisekine be serê min?</i>	Oh people! When will my headache end?
<i>Xali bavên ser çavên min.</i>	Cover my eyes with the carpet.
<i>Cercewê tera çi bejim? Şêrê bavo! Bavêmin şerîn çavo!</i>	What should I say to your orphans? Lion of the fathers! My sweet father with sweet eyes!
<i>Nikarim lo law bi derd xwe bi sekinim</i>	I can’t stay because of my pain.
<i>Ax wexte birê xwuske dimre, bila xwuskê pêra bîmre.</i>	Ah! When the brother dies, the sister should die together with him.

Yezidis say that *kilamen ser* and *stranen* are two different modes of enunciation that belong to opposing sound registers (Amy de la Bretèque 2012, 2013). They would rather link “words about” to speech than to song. Yezidis are well aware that there are also important differences between *kilamê ser* and *axavtin*: first, “words about” are uttered using a recognizable melodic mould rather than normal speech intonation; second, they make extensive use of metaphors, direct reported speech (I say: “...”), and time shifts;<sup>10</sup> and, finally, “words about” break the logic of conversational interactions. When someone utters *kilamen* in this way, the audience is supposed to listen carefully without interrupting the speaker. There are no contradictions, questions, or comments until the end of the utterance (sighs of approval are common, though). Even in daily conversations, such as those analysed here, when a speaker starts a *kilamê ser*, she typically speaks alone for a few minutes. She will stop on her own or because of something that happens in the room, but not because someone else has something to say.

In the five-minute recording analysed here, Hbo performed two *kilamen ser*. The first is the normal way to perform such utterances (from 1:35 to 2:14). The second was uttered with no stable pitch (from 3:48 to 4:26). Hbo said that she performed her second *kilamê ser* in this way to protect her grandson, who was present in the room, from an utterance otherwise potentially harmful to him. Indeed, the Yezidis consider laments to be dangerous for both the young and the weak (for comparative examples from around the world, see Amy de la Bretèque 2005; Andreesco and Bacou 1986; Delaporte 2010; Holst-Warhaft 1992; Tolbert 1990; and Wilce 1999). Unmelodized *kilamen ser* occur quite rarely, but nevertheless reveal that the melodization of an utterance is considered more

---

10. “Time shift” means changes between past, present, and future.

dangerous than the semantic content (which can still be said). Listening to this recording with other Yezidis, all of them agreed that this was not *axavtin*. Even without the specific treatment of pitch, this sort of utterance clearly remained identifiable as *kilamê ser*. The following is an abstract of Hbo's non-melodized *kilamê ser* (from 3:48 to 4:27):

*Got: "Hekimo ti bederi."*

I said: "Doctor you are the healer."

*Kutiye di dermanan helde, qaseki, were wi beri!*

Come here, and bring two boxes of medicine with you!

*Brina dayika effendi, ewê qemet bu! Dayika piro ana qimet bu.*

The wound of Mother Effendi<sup>11</sup> is getting worse! Mother Piro was so respected at that time.

*Hele dare, digotin dayika effendi!*

That's why people used to call her Mother Effendi!

*Brina dayika effendi, birinekê kurê. Kafirê hekim nake sere deri!*

The wound of Mother Effendi is a hard wound. The godless medicine cannot cure!

*Dane qutiya du dermanan, ka ser brina dayika effendi rew!*

He put two boxes of medicines over the wound of Mother Effendi, and left!

*Di got: "Ezi bêjim, mezeli wê mezeli!"*

She said: "I would say 'oh grave!' in the graveyard!"

*Ezê herim bawla ko baxêe hafa mina beri!*

I'll go with my suitcase to my old ancient garden!

*Ezê derê şaristana gran wekim! Binerim, sere mala bavêmin tede bune axa xali! Min nas nekir!*

I would open the gate of that biggest country! I would check my old family house, which is turned to ruin and dust! And I would not recognize!

The words uttered in Hbo's non-melodized *kilamê ser* are similar to the words uttered in melodized "words about." The general topic of the utterance is pain and loss, making abundant usage of direct reported speech, metaphors, and time shifts.<sup>12</sup>

Distinguishing between *axavtin*, *stran*, and *kilamen ser* does not really pose a challenge to Yezidis. They never argue or show the slightest doubt about the

11. *Effendi* literally means lord or master.

12. Time shifts can be observed in the changes of tense used to reference the wounded mother. For example, "The wound of Mother Effendi *is getting* worse! / Mother Piro *was* so respected at that time. / That's why people *used to* call her Mother Effendi!"

type of utterance expressed. In all likelihood, the characteristics listed above would also enable an unfamiliar observer to correctly recognize the three modes of enunciation. The question, rather, is why Yezidis deploy a distinct vocal category—*kilamen ser*—to speak about sad matters. Around the world, people express all kinds of feelings through language. Patterns of intonation convey emotions in conjunction with words (see Banse and Scherer 1996; Goudbeek and Scherer 2010). This also holds true in the Kurmanji Kurdish spoken by Yezidis. They *can* say sad things without voicing them as *kilamê ser*. So why do they regard the voicing of “words about” to be more effective than customary speech in conveying this kind of emotion? One may also ask why Yezidis think that *stranen* are unsuitable for this purpose. In many other cultures, singing serves to express sad and happy feelings. However, among Yezidis, singing is exclusively a joyous performance. Does this arise because Yezidis define singing arbitrarily and in a limited way? Or does their definition reflect something deeper about how they understand vocal communication? To shed light on such questions, we submitted the recording of Hbo’s voice to detailed aural and acoustic analysis.

## ACOUSTIC CHARACTERISTICS OF AXAVTIN, STRAN, AND KILAMEN SER

### An aural analysis

Prior to an acoustic analysis, the entire recording of Hbo’s conversation was manually annotated for every breath group she uttered; a “breath group” is the vocalization produced between two inhalations. The recording contains sixty-eight breath groups: forty-nine for *axavtin*, two for *stran*, seven for melodized *kilamê ser*, and eight for non-melodized *kilamê ser*.<sup>13</sup> From this segmentation, we conducted a general aural analysis. Hbo’s song repeats the same small melodic sequence three times in her mid-register (transcribed in figure 2). The rhythm is regular, but the three repetitions of the sequence are not aligned by any common pulse (there are small gaps between them). This is the main difference with longer

Dex-ma - no bê xwe - di - ne  
Ef - su - ne bê te di - ne  
E - mê qur - ba - ne qal - kê

Figure 2. Musical transcription of Hbo’s *stran* (song).

performances of Yezidi *stran*: longer songs normally have a steady pulse running throughout and one could dance to them. Otherwise, Hbo’s short performance is representative of the lyrics, rhythm, and small repetitive patterns of the genre.

Hbo’s *kilamê ser* also reveals a melodic pattern, although it is

13. Two of the notated breath groups were not taken into account because the same breath groups feature a transition between *axavtin* and *stran* or *axavtin* and *kilamê ser*.

Feyda ti-ne bi-re min

Ni-za-nim ça - wa ha-le min

Ge-lo kê bi - se-ki - ne be se-rê min

[unclear] Xa-li ba-vên ser ça-vên min

Cer-ce-wê te-ra çi be-jim şêrê ba-vo ba-vêm şe-rin ça-vo

Ni-ka-rim lo law bi derd xwe bi se - ki-nim

Ax wex-te bi-rê xwuş-ke dimre bi - la xwuş-kê pê - ra bim-re

Figure 3. Musical transcription of Hbo's *kilamê ser* ("words about").

less precise in terms of fundamental frequencies and rhythms than her song. Figure 3 represents an attempt to transcribe it in conventional notation, but the durations and pitches are only approximate. One remarkable phenomenon is the last-syllable lengthening of each breath group and the melisma on this syllable, which is sung on two or three notes; for the first five breath groups of a total eight, this comes on the “*min*” syllable. The melodic curve of each breath group is descending, and the overall melodic curve of the breath groups traces a descending pattern divided into two blocks: the first five breath groups, then the three last (cf. figure 3). By contrast, Hbo's non-melodized *kilamê ser* does not reveal any apparent melodic structure. There are, however, two or three tone areas for most breath groups; a tone area refers to a pitch with a high degree of variability. In addition, her non-melodized *kilamê ser* reveals some rhythmic

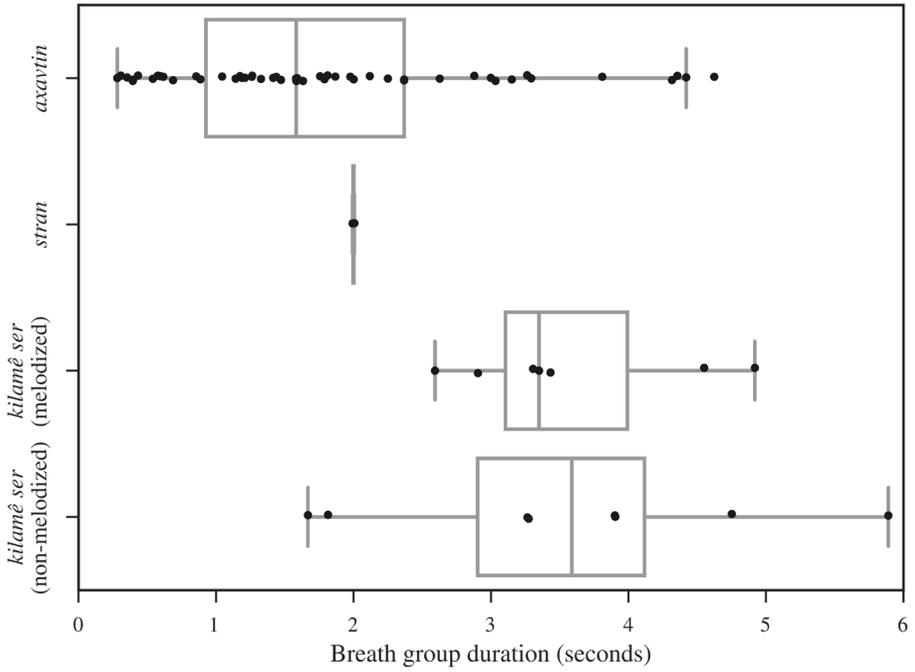


Figure 4. Distribution of breath group durations for four types of vocal production.

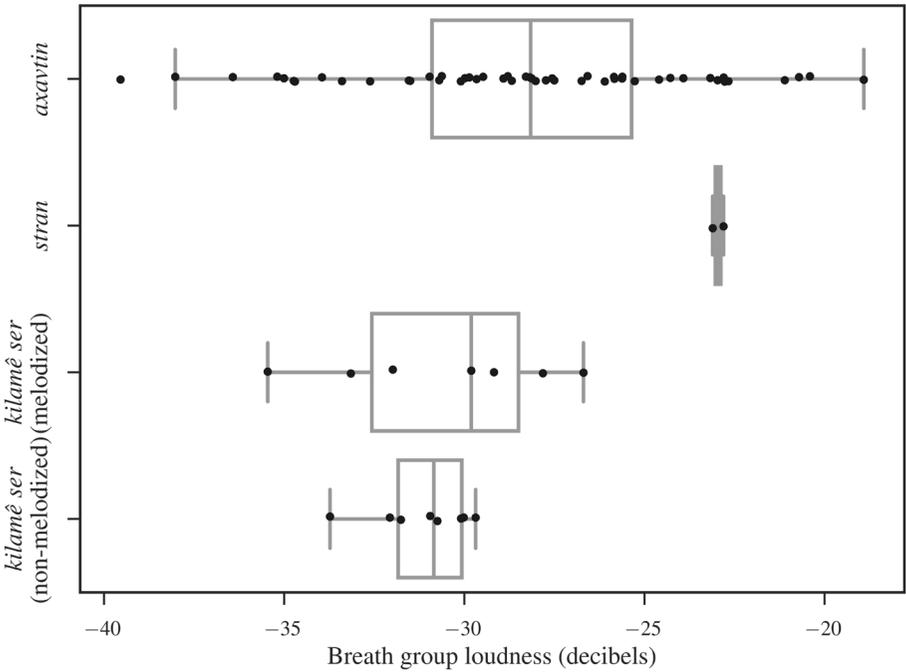


Figure 5. Distribution of breath group loudness for four types of vocal production.

structure and some regularity of syllable rhythm and breath group durations. In contrast to both *stran* and *kilamen ser*, Hbo's spoken performance shows great variability in terms of breath group durations, variability of pitch curves, and the accentuation of some syllables inside the breath groups.

### An acoustic analysis

Following aural analysis, we computed a set of acoustic features from the same segmentation of breath groups. These features include the duration, loudness, pitch range, and median pitch of breath groups for *axavtin*, *stran*, and *kilamê ser*.<sup>14</sup> This approach allowed us to compare and contrast the features of these different categories of utterance.

The first characteristic we considered is the duration of each breath group. The distribution of these durations is shown in figure 4. In layman's terms, the graph describes how the breath groups vary in duration for each type of utterance. Each point represents a breath group and its associated duration (the raw data), while the box and the lines are a representation of the point distribution. The boxes represent half the data and the lines extending from the boxes represent the remaining data (except for the outliers). The vertical line inside the box corresponds to the median value. A long horizontal "whisker" line extending from the side of the box indicates variation within the breath groups; a short "whisker" line indicates similarity within the breath groups. We can see from figure 4 that the breath group duration ranges from 1.7s to 5.9s (in *kilamen ser*). *Axavtin* and *stran* show short median breath group durations (1.5s and 2s respectively), while both kinds of *kilamen ser*—melodized and non-melodized—use longer durations: 3.3s and 3.5s, respectively. It appears that breath durations are very regular in *stran*, which is a probable consequence of the regular rhythmic pulse that is characteristic of the genre. However, the melodized "words about" also behave more regularly in this respect than *axavtin*. In both kinds of *kilamen ser*, breath groups are typically longer than in *stran* and *axavtin*.

Another aspect that we analysed was the relative loudness of all the utterances. The evolution of loudness over time was evaluated for each breath group,<sup>15</sup> and the median of each loudness contour was taken as a representative value for its respective breath group. Figure 5 presents the results. Each point represents a breath group and its associated loudness median. As for figure 4, boxes and lines have been added to highlight statistical distribution of the points: the

14. Statistics were computed for each of these categories. The category results are all the more robust if the number of breath groups is greater. In particular, the *stran* category contains only two breath groups, and it is therefore more difficult to be confident about the representativeness of the results computed for this category.

15. Loudness was computed by the root mean square method: each 10ms within a 100ms window with the results set out in decibels relative to the full scale (dB FS).

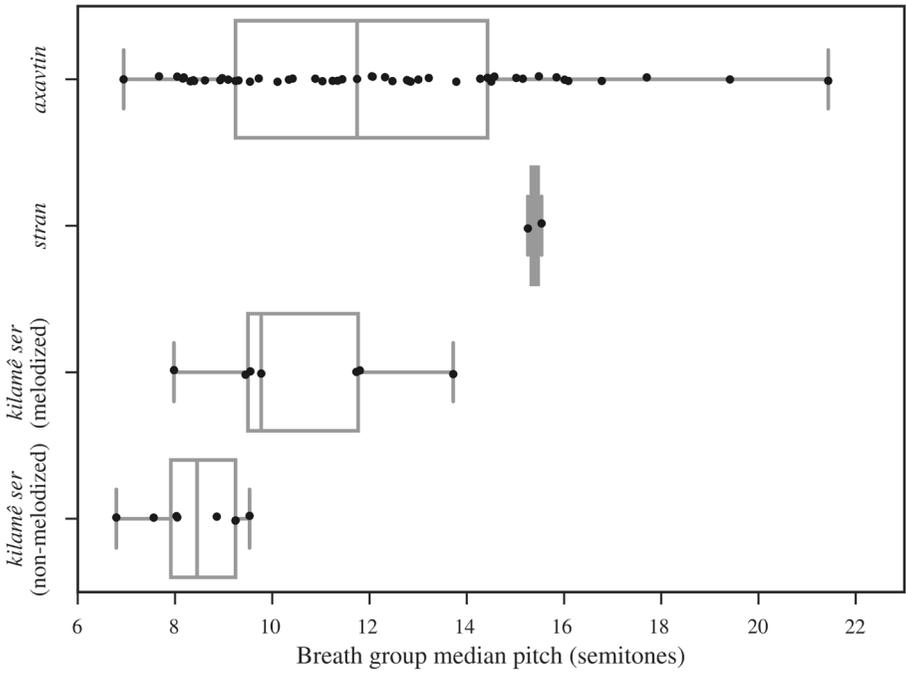


Figure 6. Distribution of breath group median pitches for four types of vocal production.

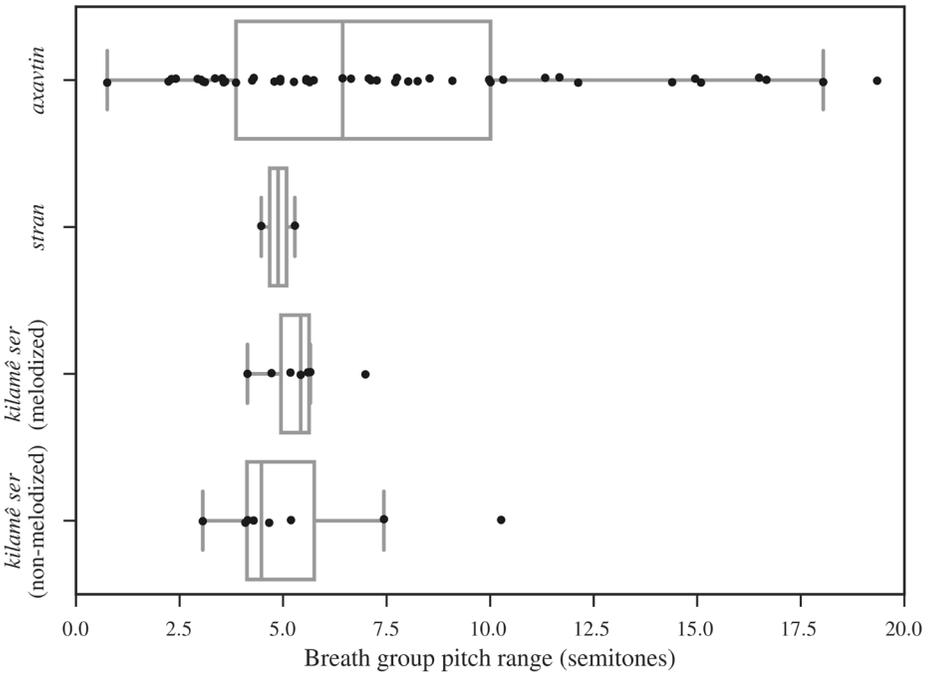


Figure 7. Distribution of breath group pitch range for four types of vocal production.

boxes represent half the data and the lines extending from the boxes represent the remaining data (except for the outliers); the vertical line inside the box corresponds to the median value of the breath group loudness medians. The *stran*'s median value (i.e., the vertical line) is greater than those for the other three categories, which means that the sound is, on average, louder. Here again, singing stands out as a controlled vocal genre with a very limited range of loudness. At the other extreme, *axavtin* has a large 20dB dispersion that exceeds the values of the three other categories. During *kilamen ser*, voice loudness is more uniform in this respect and typically softer than in the two other genres.

Third, we examined the melodic contour of Hbo's utterances through the application of two calculations. First, we obtained the median pitch for each breath group.<sup>16</sup> The results are displayed in figure 6 with the scale given in semitones with an arbitrary reference to the pitch, A2.<sup>17</sup> Each point represents a breath group and the associated median pitch value. Again, the boxes represent half the data and the lines extending from the boxes represent the remaining data (except for the outliers). The vertical line inside the box corresponds to the median value of the median pitch values. Once again, this demonstrates how *axavtin* is the most variable of all categories in the pitch domain, with *stran* proving the most constrained. *Kilamen ser* is in-between. Also of note, when considering only half the data centred on the median (the pitches depicted inside the boxes), sung breath groups feature the highest median pitches, three to five semitones higher than speech and "words about." *Axavtin* and *kilamen ser*, on the other hand, have similar pitch ranges (with the greatest variation for speech, as noted previously).

The second characteristic we computed from the melodic contour is the range in semitones (the difference between the highest and lowest pitch) for each breath group, with the results presented in figure 7. Each point represents a breath group and the associated pitch range. Again, the boxes represent half the data and the lines extending from the boxes represent the remaining data (except for the outliers). The vertical line inside the box corresponds to the median value of the breath group pitch ranges. For *axavtin*, the range varies greatly between breath groups, from nearly monotone to more than one octave. By contrast, the ranges of *stran* and *kilamen ser* nearly always register at around five semitones. The median pitch range of the breath groups is also higher in speech than in the other genres. This suggests that Hbo's speech voice has more modulations of pitch than her song voice.

---

16. Pitch was computed by using an autocorrelation-based algorithm (provided by Praat software). Each breath group's melodic profile was then manually checked and compared to the audio source file to detect and correct for any possible mistakes in the algorithm (for instance, the algorithm sometimes detects the pitch as an octave higher or lower).

17. This means, for instance, that 0 corresponds to A2, 5 corresponds to D3, and 12 corresponds to A3. This reference point was selected for convenience of comparison in order to provide an idea of the variation range; this choice does not suggest any "musical" interpretation.

The last acoustic characteristic we examined was the silence duration. This is calculated from the absence of harmonic partials relative to the total duration of a category.<sup>18</sup> We observed that the proportion of silence duration is similar during speech (35 per cent) and melodized “words about” (30 per cent), while it is twice less during non-melodized “words about” (17 per cent) and song (15 per cent). Moreover, *axavtin* presents frequent pauses and *kilamen ser* less frequent but longer pauses. The longer pauses in *kilamen ser* may be linked to a physiological necessity—remembering that the breath groups are also longer here. But they also highlight the conversational difference between *axavtin*, a mode of vocal production in which anyone can intervene at almost any time, and *kilamen ser*, when the speaker knows that she can take her time without being interrupted. Longer silences also allow more time for memories to emerge and for the many figures of speech to work their way through the listener’s emotional consciousness. Figure 8 summarizes the main results reported above. According to this analysis, there are acoustically quantifiable distinctions between Yezidi performances of *stranen*, on the one hand, and both *kilamen ser* and *axavtin*, on the other.

As a whole, Hbo’s voice is more constrained when she sings than when she speaks or delivers “words about.” This holds true for all of the features analysed, including time, pitch, and loudness domains. Furthermore, Hbo delivers her *kilamen ser* and her soft *axavtin* in a similar manner—that is, with the same loudness. Her median pitch and pitch range also are similar in her speech and “words about.” The distinction between *stran*, *axavtin*, and *kilamen ser* is also reflected in the relevance of the semantic content to listeners; linguistic meaning is of primary importance in both speech and “words about,” but only secondary in songs (Amy de la Bretèque 2012, 2013). Finally, the fact that one can produce non-melodized “words about”—non-melodized songs do not exist—further illustrates the connection of this genre to speech.

There are also a number of acoustic differences between *kilamen ser* and *axavtin*. For a start, there is much greater variability in speech than other types of vocal performance. This holds true for all acoustic parameters: the standard deviation (i.e., the statistical variability) of *axavtin* is two or three times greater than found in *kilamen ser*. For instance, the loudness of Hbo’s breath groups varies on a range of 20dB for speech, but only 9dB for “words about”—and only 4dB for the non-melodized version. Similarly, Hbo’s speech range is approximately two octaves, while her “words about” range is less than a seventh. The breath group durations are also much more regular for the *kilamen ser* (from 2.7s to 5s) than for *axavtin* (from 0.3s to 4.4s). Another difference is that in *axavtin* there are generally no sustained vowels, while, as can be seen in the musical

---

18. Strictly speaking, what we call a “silence duration” in our sample includes both silence and unvoiced phonemes.

	<b>AXAVTIN (Speech)</b>	<b>KILAMEN SER (“Words about”)</b> m = melodized; nm = non-melodized	<b>STRAN (Song)</b>
<b>Breath group duration</b>	Irregular Short	Rather regular (m) / irregular (nm) Long	Regular Short
<b>Breath group loudness</b>	Irregular Loud or soft	Regular Soft	Regular Loud
<b>Breath group median pitch</b>	Irregular Middle pitches	Rather regular Lower (m) / lowest (nm) pitches	Regular Higher pitches
<b>Breath group pitch range</b>	Irregular Large ranges	Rather regular Short ranges	Regular Short ranges
<b>Silence proportion</b>	35 per cent	30 per cent (m) / 17 per cent (nm)	15 per cent

Figure 8. Summary of results of acoustic analysis.

transcription (figure 3), Hbo’s “words about” contain—and end on—lengthened syllables. All of these acoustic characteristics indicate that “words about” are a more structured, coherent, and regular production than speech. They are akin, in this respect, to formalized speech, which in a sense “diminishes the propositional force of language” (Bloch 1974:67). Language has the ability to articulate almost any kind of content, but this rests on the speaker’s freedom to manipulate the syntax and prosody of his/her utterances. Narrowing down the combinatory possibilities logically results in narrower expressive ranges available to the speaker. What remains untouched, and is possibly increased, is the illocutionary force of the utterance. Ritual speech and other forms of “traditional authority” are examples of the inverse variation of propositional and performative capacities of language (ibid.). *Kilamen ser* are also linked to forms of “traditional authority” (like the authority of “burning hearts” who know how to suffer), which is why they can’t be interrupted or contradicted. *Kilamen ser* are powerfully emotional, but in an illocutionary, rather than semantic way. They don’t “tell” the sadness as much as they bring it into being.

### SHAPES OF SADNESS?

The question remains: Just why do Yezidis feel that *kilamen ser*, which deploy a narrower range of vocal possibilities than speech, better convey their feelings of sadness? One possible answer is that *kilamen ser* take on the shape typical of a sad utterance. The literature on expressing emotion in speech quite clearly establishes that listeners are able to infer affective states from various vocal cues: variations in pitch and vocal energy, the distribution of energy in the frequency spectrum, formant frequencies, and tempo and pausing (e.g., Banse and Scherer 1996). In the case of sadness, markers of affect are remarkably consistent across many languages and cultures in these studies. In many languages, sadness is manifested by a decrease in the mean pitch, in the pitch range, and in the mean energy coupled with a downward pitch contour. There also are related decreases in the

high frequency energy and in the rate of articulation. For singing productions, a similar tendency is observed for low arousal emotions such as sadness: low levels of waveform irregularity, lower loudness levels, and slower tempos are all observed (Scherer et al. 2015).

Hbo's *kilamen ser* also display these characteristics. Compared to her *axavtin*, the pitch mean and range decrease by three semitones, the mean intensity decreases by 3dB, and the intensity range is much smaller (4dB and 9dB compared to 20dB). The rate of articulation—measured as the ratio of the number of syllables over the breath group duration—is also smaller for *kilamen ser* than for *axavtin*. Finally, many acoustic parameters show much more regular behaviour for “words about” than for speech. Hbo's *kilamen ser*, therefore, share the same acoustic cues as the expressions of sadness, whether spoken or sung, studied in the literature on acoustic markers of affect. This brings us back to our initial question: Why are there special vocal categories, when one can just say (or sing) them in more or less the same way?

Answering this question requires paying attention to what these categories of vocal production encompass. *Kilamen ser* can be performed in the course of daily conversation, as in the case studied here. They are also typical of funeral contexts, where they are uttered around the deceased or by the grave. Epic narratives are also “words about” to Yezidis. The only slight difference is that they are not “words about the dead” (*kilamen ser mirya*), but, rather, “words about the hero” (*kilamen ser mêranîê*). Yezidi heroes (and Kurdish heroes more generally) are all dead, having perished in tragic conditions, which probably explains why epic songs share many semantic traits with lamenting utterances and, indeed, fall into the same vocal category (Amy de la Bretèque 2012, 2013).

Laments are a remarkable cross-cultural category (Feld 1990; Urban 1988). Beyond their acoustic shape, they share in common a kind of interactional indirection. Urban describes them as utterances without an actual addressee, a fact that “renders ... ritual wailing an overtly monologic or expressive device” (1988:386). According to Urban,

wailing is a process of making public the feelings of the person who is wailing. It is intended not to be heard in the ordinary linguistic sense, but to be overheard. Ritual wailing purports overtly not to engage an addressee, but to allow anyone within earshot access to something that would otherwise be private. (ibid.:392)

Contrary to a common assumption, laments do not necessarily imply catharsis either for the listener or for the audience.<sup>19</sup> Feld, for instance, points out that shedding individual sorrow is not relevant among the Kaluli of Papua New Guinea, where laments are better understood as a “‘pulling together’ of affects” (1990:257).

---

19. This assumption has been perpetuated, in part, in the ethnological literature. See, for example, Andreesco and Bacou (1986); Doubleday (1988); Khouri (1993); Tolbert (1990); and Xanthakou (1990).

Yezidis, and other Kurds in the Middle East, present us with an attitude quite opposite to what one would expect in catharsis (Amy de la Bretèque 2016). To start with, Hbo says that she does not experience any relief while (or after) performing *kilamê ser*. Performed on a daily basis, these utterances seem to keep old wounds open. As mentioned above, the status of *dilşewat* is intrinsically linked to this kind of voluntary martyrdom. Hbo says that she does not wish to get rid of her sufferings but, rather, to share them. In her willingness to utter *kilamê ser* daily (and not only in ritual contexts such as funerals), she shows no hesitation in inviting her neighbours, family, or visitors to join into the same mood. The question here is not how sadness is “expressed,” but, rather, how other people might be “pulled together” within its scope. To achieve this affect, an efficient vocal strategy might be to mould the words into a form that is not specifically personal. Replacing the prosody of normal speech with a shared acoustic shape helps to symbolically and pragmatically detach the utterance from the speaker. Because *kilamê ser* are more rigid than *axawtin*, they attenuate, to some extent, the vocal personality of the utterer—a prerequisite for their subjective reappropriation by others. Without overlooking the actual feelings of the speaker, their acoustic shape gives her (and occasionally him) the possibility of distancing herself from her own narrative. At the same time, the performer’s emotions are made available to listeners, who can invest their own memories in this shared acoustic space of sorrow.

*Kilamê ser* also retain more freedom, and therefore more potential for surprise, than *stranen*. Recall that “singing,” here, means a steady tone of voice, a steady beat, and rhyming lines. These qualities make it a highly predictable genre in comparison with *kilamê ser*. Yezidis have no sad *stranen*, possibly because sadness implies, for them, more unexpectedness than “singing” offers. There is, of course, more than one cliché and stereotyped expression in *kilamê ser*; these formulaic qualities, perhaps, require the vocal characteristics of the genre to remain flexible.

“Words about” sad things sound rather predictable compared to normal *axawtin*, but still less predictable than *stranen*. Listeners must pay close attention to grasp something of *kilamen ser*’s semantic content. While prosodic features normally facilitate the parsing of linguistic utterances, “incorrect” prosodic cues require greater cognitive effort from listeners (e.g., Wingfield and Klein 1971; Marslen-Wilson et al. 1992; see Cutler and Dahan 1997 for a comprehensive review). With their long breath groups, limited dynamics, and restrained variations of pitch, “words about” typically neglect many prosodic expectations of Kurmanji Kurdish. The listener is easily lost in their abundant use of reported speech, time shifts, and elaborate metaphors. The vocal shape of *kilamê ser* is an additional appeal to “let go” of any attempt to parse a coherent semantic meaning. Their only real coherence is affective, and this is granted by their homogeneous yet flexible vocal shape.

In the context of this curated section on “liminal utterances” and the colloquium out of which it grew (see Engelhardt and Amy de la Bretèque, this volume), it may be useful to conclude with a final word on liminality. Because speech and song are salient vocal categories in western European culture, there is a strong temptation to consider them as two ends of a continuous linear axis. Other vocal categories encountered in Europe and elsewhere (e.g., psalmody, chanting, *kilamê ser*) are described as falling somewhere “in-between.” In this sense, they are liminal utterances; neither speech nor song. A closer look reveals that while the acoustic features used to study Hbo’s vocal productions are continuous variables (frequencies, intensities, durations, ranges, means, etc.), no single parameter is useful in being able to discriminate between one category and another. On the contrary, many parameters from different dimensions are needed to adequately describe even basic categories such as song and speech. Understanding that acoustic space is multidimensional rather than linear, *kilamen ser* are not *between* speech and song; rather, they hold a third position among multiple variables on the map of vocal and emotional possibility.

## REFERENCES CITED

- Amy de la Bretèque, Estelle  
 2005 “Femmes mollah et cérémonies féminines de deuil en Azerbaïdjan.” *Cahiers de musiques traditionnelles* 18: 51–66.  
 2010 “Des affects entre guillemets: Mélodisation de la parole chez les Yézidis d’Arménie.” *Cahiers d’ethnomusicologie* 23: 131–45.  
 2012 “Voices of Sorrow: Melodized Speech, Laments and Heroic Narratives among the Yezidis of Armenia.” *Yearbook for Traditional Music* 44: 129–48.  
 2013 *Paroles mélodisées. Récits épiques et lamentations chez les Yézidis d’Arménie* [Melodized speech. Heroic narratives and laments among the Yezidis of Armenia]. Paris: Classiques-Garnier.  
 2016 “Self-Sacrifice, Womanhood, and Melodized Speech: Three Case Studies from the Caucasus and Anatolia.” *Asian Music* 47/1: 29–63.
- Andresco, Ioanna, and Mihaela Bacou  
 1986 *Mourir à l’ombre des Carpathes*. Paris: Payot.
- Banse, Rainer, and Klaus R. Scherer  
 1996 “Acoustic Profiles in Vocal Emotion Expression.” *Journal of Personality and Social Psychology* 70/3: 614–36.
- Bloch, Maurice  
 1974 “Symbols, Song, Dance and Features of Articulation: Is Religion an Extreme Form of Traditional Authority?” *European Journal of Sociology / Archives Européennes de Sociologie / Europäisches Archiv Für Soziologie* 15/1: 55–81.
- Charron, Claude  
 1978 “Toward Transcription and Analysis of Inuit Throat-Games: Micro-Structure.” *Ethnomusicology* 22/2: 245–59.

- Cutler, Anne, and Delphine Dahan.  
1997 "Prosody in the Comprehension of Spoken Language: A Literature Review." *Language and Speech* 40/2: 141–201.
- Delaporte, H el ene  
2010 "De l'enterrement   la f ete, parcours d'un texte fun ebre en  pire–Gr ce." In *La voix act ee: Pour une nouvelle ethnopo etique*, ed. Claude Calame, Florence Dupont, Bernard Lortat-Jacob, and Maria Manca, 195–210. Paris: Kim e.
- Doubleday, Veronica  
1988 *Three Women of Herat: A Memoir of Life, Love and Friendship in Afghanistan*. London: Jonathan Cape.
- Fales, Cornelia  
2002 "The Paradox of Timbre." *Ethnomusicology* 46/1: 56–95. doi:10.2307/852808.
- Feld, Steven  
1990 "Wept Thoughts: The Voicing of Kaluli Memories." *Oral Tradition* 5/2–3: 241–66.
- Goudbeek, Martijn, and Klaus R. Scherer  
2010 "Beyond Arousal: Valence and Potency/Control Cues in the Vocal Expression of Emotion." *Journal of the Acoustical Society of America* 128/3: 1322–36.
- Holst-Warhaft, Gail  
1992 *Dangerous Words: Women's Laments and Greek Literature*. London: Routledge.
- Khoury, Nagib  
1993 *Le feu et la cendre: Travail de deuil et rites fun eraires dans un village libanais*. Paris: L'Harmattan.
- Kreyenbroek, Philip G.  
1995 *Yezidism: Its Background, Observances and Textual Tradition*. Lewiston, NY: Edwin Mellen Press.
- Kreyenbroek, Philip G., and Khalil J. Rashow  
2005 *God and Sheikh Adi Are Perfect: Sacred Poems and Religious Narratives from the Yezidi Tradition*. Wiesbaden, DE: Harrassowitz Verlag.
- Marslen-Wilson, William D., Lorraine K. Tyler, Paul Warren, Paula Grenier, and Catherine S. Lee  
1992 "Prosodic Effects in Minimal Attachment." *The Quarterly Journal of Experimental Psychology* 45/1: 73–87.
- Scherer, Klaus R., Johan Sundberg, Lucas Tamarit, and Gl ucia L. Salom ao  
2015 "Comparing the Acoustic Expression of Emotion in the Speaking and the Singing Voice." *Computer Speech and Language* 29: 218–35.
- Tolbert, Elizabeth  
1990 "Magico-religious Power and Gender in the Karelian Lament." In *Music, Gender, and Culture*, ed. Susanne Ziegler and Marcia Herndon, 41–56. Wilhelmshaven, DE: Florian Noetzel Verlag.
- Urban, Greg  
1988 "Ritual Wailing in Amerindian Brazil." *American Anthropologist* 90/2: 385–400.
- Vaughn, Kathryn  
1990 "Exploring Emotion in Sub-Structural Aspects of Karelian Lament: Application of Time Series Analysis to Digitized Melody." *Yearbook for Traditional Music* 22: 106–22.
- Wilce, James M.  
1999 *Crying Shame: Metaculture, Modernity, and the Exaggerated Death of Lament*. Singapore: Wiley-Blackwell.

Wingfield, Arthur, and Jan F. Klein

1971 “Syntactic Structure and Acoustic Pattern in Speech Perception.” *Perception and Psychophysics* 9/1: 23–25.

Xanthakou, Margarita

1990 “Discours d’outre-tombe: le langage du mirolöi.” *Cahiers de littérature orale* 27: 137–62.

Zemp, Hugo

1996 Ed. *Les voix du monde. Une anthologie des expressions vocales / Voices of the World: An Anthology of Vocal Expression*. Recordings and notes by Gilles Léothaud, Bernard Lortat-Jacob, Trần Quang Hai, and Jean Schwarz. Collection CNRS-Musée de l’Homme. Le Chant du Monde CMX 374 1010.12. [http://archives.crem-cnrs.fr/archives/collections/CNRSMH\\_E\\_1996\\_013\\_001/](http://archives.crem-cnrs.fr/archives/collections/CNRSMH_E_1996_013_001/) (accessed 2 June 2017).

## LIST OF AUDIO FILES

Audio clips referenced in this paper are available online at <http://www.ictmusic.org/ytm2017> and <http://www.ebreteque.net/liminalutterancesandshapesofsadness>.

*sound\_entireFile.wav*

Description: Entire Hbo audio file

*sound\_part1\_speech.wav*

Description: speech (*axavtin*)

Relationship to entire Hbo audio file: 0:00.0 to 0:39.1

*sound\_part2\_song.wav*

Description: song (*stran*)

Relationship to entire Hbo audio file: 0:39.1 to 0:45.7

*sound\_part3\_speech.wav*

Description: speech (*axavtin*)

Relationship to entire Hbo audio file: 0:45.7 to 1:35.5

*sound\_part4\_lament.wav*

Description: “words about” (*kilamê ser*)

Relationship to entire Hbo audio file: 1:35.5 to 2:14.0

*sound\_part5\_speech.wav*

Description: speech (*axavtin*)

Relationship to entire Hbo audio file: 2:14.0 to 3:48.5

*sound\_part6\_nonMelodizedLament.wav*

Description: non-melodized “words about” (non-melodized *kilamê ser*)

Relationship to entire Hbo audio file: 3:48.5 to 4:26.5