

A corpus study of durational rhythmic measures in the Kalhori variety of Kurdish

Nafiseh Taghva¹, *Amirsaeid Moloodi¹, Vahideh Abolhasanizadeh², Roya Tabei³

¹Department of Foreign Languages and Linguistics, Shiraz University, Shiraz, Iran

²Department of Foreign Languages, Shahid Bahonar University of Kerman, Kerman, Iran

³Department of English Language and Literature, Razi University, Kermanshah, Iran
taghvanafiseh@gmail.com ORCID: <https://orcid.org/0000-0001-9715-4505>

*Corresponding author: amirsaeid.moloodi@shirazu.ac.ir ORCID: <https://orcid.org/0000-0003-4682-3160> abolhasani@uk.ac.ir
ORCID: <https://orcid.org/0000-0002-5426-9369>
tabei.roya@razi.ac.ir ORCID: <https://orcid.org/0009-0009-4461-7807>

Enviado: 30/06/2023; Aceptado: 28/08/2023; Publicado en línea: 23/10/2023

Citation / Cómo citar este artículo: Taghva, N., Moloodi, A., Abolhasanizadeh, V., & Tabei, R. (2023). A corpus study of durational rhythmic measures in the Kalhori variety of Kurdish. *Loquens*, 10(1-2), e098, <https://doi.org/10.3989/loquens.2023.e098>.

ABSTRACT: In order to identify between-sentence and between-speaker variabilities, one of the methods used by phoneticians is studying durational rhythmic features. In the present research, to classify speech rhythm of Kalhori, a variety of Kurdish, and to find out about the most appropriate measures for between-sentence and between-speaker rhythmic variability in Kalhori, durational speech rhythmic measures were analyzed. To this end, two speaking styles (read and spontaneous) were explored. The analysis of the read corpus revealed that Kalhori Kurdish rhythm pattern is between stress-timed and syllable-timed. The results indicated that %V (proportion over which speech is vocalic) was the most significant measure for distinguishing between-sentence rhythmic variability in the read corpus, while %V and rateSyl (syllable rate) were the most efficient measures for identifying the between-speaker rhythmic variability in both the read and spontaneous corpus.

Keywords: durational variability, acoustic correlations, rhythmic measures, Kalhori, Kurdish

RESUMEN: *Un estudio del corpus de medidas de duración rítmica del dialecto kalhori del kurdo.* Uno de los métodos empleados en fonética para identificar la variabilidad entre oraciones y hablantes es el estudio de las características rítmicas. En este estudio, se han analizado algunas métricas temporales de ritmo en kalhori (una variedad del kurdo) para descubrir las que mejor explican la variabilidad rítmica entre oraciones y entre hablantes. Con este fin, se han utilizado dos estilos de habla: lectura y habla espontánea. El análisis del corpus de lectura demostró que el tipo de ritmo del kurdo kalhori se puede situar en el medio del continuo entre lenguas de ritmo acentual y lenguas de ritmo silábico. Los resultados indican que la métrica más adecuada para explicar la variabilidad rítmica entre oraciones en el corpus leído fue %V (proporción de vocales sobre el total de habla), mientras que %V y rateSyl (número de sílabas pronunciadas por minuto) fueron las métricas más eficientes para identificar la variabilidad rítmica entre hablantes, tanto en el corpus leído como en el espontáneo.

Palabras clave: variabilidad de duración, correlaciones acústicas, medidas rítmicas, kalhori, kurdo

1. INTRODUCTION

Speech rhythm has remained a controversial topic for a long time, especially in the field of acoustic phonetics. Examples of such controversial issues are (to name a few): a) how to define and measure speech rhythm and whether this concept is valid and/or useful (Arvaniti, 2012; Tilsen, 2016), b) if distinct rhythm classes or types such as stress-timed and syllable-timed languages exist and whether or not they reflect linguistic or perceptual categories (Dauer, 1983; Ramus et al., 1999; White & Mattys, 2007); and c) how speech rhythm relates and interacts with other prosodic or segmental features such as intonation, stress, vowel quality, and syllable structure (Grabe & Low, 2002; Arvaniti et al., 2008; Dellwo et al., 2015). The fact that such issues exist reflects the diverse and complex nature of studies done in this area, which, in turn, have served as the rationale behind the current study.

Moreover, one significant contribution of such studies to the body of literature has been the introduction of the above-mentioned issues in at least four different branches of applied phonetics, namely 1) Children's speech and second language learning (Lee et al. 2014; Polyanskaya & Ordin, 2015), 2) Speech technology (Barbosaand & Bailly, 1994, Gibbon, 2023), 3) Speech pathology (Leong & Goswami, 2014; Liss et al., 2009; White et al., 2010, Magne et al., 2016) and, 4) Forensic phonetics (Dellwo et al., 2012; Leeman et al., 2014; Asadi et al. 2018).

The distinction between speaker variability and language-dependent rhythmic characteristics is very relevant, as it helps determining the sources and effects of rhythmic variation in speech (Fuchs, 2016). Nonetheless, drawing a clear line between such concepts is not an easy task since they tend to interact and influence each other in complex ways (Mok & Dellwo, 2008). For example, some speaker-specific features may or may not be pronounced depending on the language or dialect spoken by the speaker (Leemann et al., 2014). while, some other language-specific features may be pronounced more strongly or weakly depending on the speaker's individual style or preference (Asadi et al., 2018).

Consequently, to distinguish between speaker variability and language-dependent rhythmic characteristics, both acoustic measurements and perceptual judgments of speech rhythm are needed. Acoustic measurements provide objective and quantitative data on temporal features of speech like duration, intensity, frequency, and variability (Gibbon, 2023).

That being said, to study the between speaker variability and language-dependent rhythmic characteristics, this study attempted to find out the appropriate measures for between-speaker and between-sentences rhythmic variability in two speaking styles in Kalhori. Kalhori was selected because it has unique features that may affect its speech rhythm, such as its complex syllable structure (C)(C)V(V)(C)(C), its stress pattern (penultimate or final), its vowel harmony system (front-back and round-unround), and its tonal accent system (Karimi-doostan, 2002; Krey-nbroek, 2005; Thackston, 2006). Therefore, segmental

intervals, consonant and vowel intervals, vocalic and consonantal intervals, voiced and unvoiced intervals, syllable intervals, and syllable peak intervals were examined in both spontaneous and read speech styles of Kalhori, a variety of Kurdish.

Kurdish is a covering term used to refer to a group of Northwestern Iranian languages spoken in parts of Turkey, Iran, Armenia, Iraq, Syria, and Azerbaijan (Windfuhr, 1989). Generally, there is no agreement on the classification of Kurdish dialects whether in Iran or other countries. McCarus (2009), for instance, believes that Kurdish cannot be located in a single group among Iranian languages because, according to Gharib (2011) and McCarus (1959), it shares syntactic and morphological similarities with Balouchi, Gilaki, Taleshi, and Farsi. Dabirmoghadam (2013) Daneshpazhouh (2010), Thackston (2006) and Krey-nbroek (2005) provide different classifications for Kurdish. It is, however, mainly divided into three main groups including Northern Kurdish or "Kurmanji", Central Kurdish or "Sorani" and Southern Kurdish. According to Fattah (2000), Southern Kurdish is spoken by three million people across an extensive region in Kermanshah, Ilam, Parts of Lorestan and Kurdistan Provinces in Iran and Khanaqin and Mandali in Iraq. As Figure 1 illustrates, the southern Kurdish consists of several varieties including Kermashani, Feyli, Laki and Kalhori. The data for the current study has been based upon Kalhori, one of the biggest tribes of Kermanshah and the second biggest tribe in Iran. Kalhori is the spoken variety in Iran's Kermanshah (in Eslamabad, Gilan-e-Gharb, southern part of Qasr-e-Shirin), Ilam (in Abdanan), Kurdistan (in Bijar and Qorveh) and Iraq (in Khaneqeyn, Kalar, Kofri and Di-yala). Figure 1 shows the Revised Map of the distribution of Southern Kurdish dialects (Belelli, 2019).

The current study aims to respond to the following research questions using two different styles: read and spontaneous speech.

Q1: What is the typology of Kalhori rhythm based on the read corpus?

Q2: How does sentence structure impact the rhythmic measures of Kalhori's read speech?

Q3: Which durational measures have a significant impact on the between-speaker rhythmic variability in read and spontaneous Kalhori speech?

Thence, the first question helped with documenting and describing the rhythmic typology of Kalhori, while the second question allowed for the analysis of between-sentence variation by revealing how the rhythmic measures varied across different sentences within same speakers and speech style. Lastly, the third question showed speakers' consistency and/or flexibility while producing and maintaining their speech rhythm, and the adaptation of their speech rhythm in different sentence structures, contents, and different styles.

2. LITERATURE REVIEW

Speech rhythm approaches can be roughly divided into three categories: durational, modulation, and prom-

Figure 1: Revised Map of the distribution of Southern Kurdish dialects from Belelli (2019, p. 3)



inence (He, 2022). Durational approaches measure the variability of different phonetic intervals in speech, especially vocalic and consonantal intervals (Ramus et al., 1999; Grabe & Low, 2002; White & Mattys, 2007; Dellwo 2009, 2010; Arvantini, 2012; Dellwo et al., 2015). Modulation approaches analyze the temporal envelope of speech and extract the recurring frequencies and phase relationships of different modulation rates such as syllable and stress (O’ Dell & Neiminen, 1999; Barbaso, 2002; Tilsen & Johnson, 2008; Leong et al. 2014; Malisz et al., 2017; Lancia et al., 2019; Gibbon, 2023). Prominence approaches examine the intensity or spectral variability of speech and use them to identify the rhythmic skeleton or pattern of speech (Todd, 1985; Cummins & Port, 1998; Lee & Todd, 2004).

Since this study aimed at exploring the durational aspects of the Kalhori variety rhythm, a brief overview of durational approach is provided. Durational approaches can be traced back to the theories of isochrony in stress-timed and syllable-timed languages. This theory was first proposed by Pike (1945), and James (1938, 1929) who claimed that “stress-timed” languages, such as English, German, and Dutch, had equal/periodic feet and “syllable-timed” languages such as French, Italian, and Spanish, had equal/periodic syllables. Nonetheless, such attempts proved that the isochrony or quasi-isochrony of durational intervals were not observable in several languages (Bertrán, 1999; Dauer, 1983; Pointon, 1980; Roach, 1982). Later on, other measures for speech rhythm were proposed by phoneticians. Standard deviation of vocalic and consonantal intervals (ΔC and ΔV) as well as the percentage of vocalic intervals (%V) were examined for each

sentence by Ramus et al. (1999) to determine the rhythmic typology of different languages. Data from Ramus et al. (1999) consisted of five 15- to 19-syllable sentences read by four native speakers in eight different languages. Their entire database contained 2,720 syllables, with each language consisting of 340 syllables. The results of this study indicated that English is a stress-timed language while French is a syllable-timed language based on ΔV and %V.

In order to measure durational variability between sequences of vocalic and consonantal intervals, Grabe and Low (2002) introduced the Pairwise variability index (nPVI-V and rPVI-C) in which they examined 16 languages, in each language a native speaker read the original text or translation of the story “North Wind and the Sun”. This story contained 141 syllables in the English version. Assuming that the average number of syllables in each version of each language is about 150 syllables, the total number of syllables examined in this study were 2256 syllables (16×150). Based on the results of this study, English rhythm shows patterns that are more closely aligned with stress-timed languages, while French leans closer to syllable-timed languages.

White and Mattys (2007) studied PVI, ΔC , ΔV , VarcoC, VarcoV and %V in English and Dutch as representative of stress-timed languages and Spanish and French as syllable-timed. Their database included 5 speakers from each language that read the text of a short story.

Varco coefficient and the natural logarithm that are other normalization methods on the speech rate were proposed by Dellwo (2009, 2010) by using Bonn Tempo corpus that consisted of 12 German speakers, 7 English

speakers and 7 French speakers at the time of the research.

Moreover, Arvantini (2012) investigated the repetition of acoustical information of syllables instead of segmental units by introducing the amplitude envelope measure of rhythm. Arvantini (2012) used 3 different styles in her study: story reading, spontaneous speech and sentence reading. Participants in her research were from six different languages of Greek, English, German, Spanish, Korean and Italian. Eight speakers of each language were present in this research. In the story reading section, the text of the story “North Wind and the Sun”, was recited for about one to two minutes in the form of spontaneous speech style, and in the sentence reading section, 5 sentences were read by each speaker.

Subsequent research has shown that vocalic and consonantal rhythm measures can vary significantly in a language based on the speaker’s performance (Wiget et al., 2010; Yoon, 2010, Loukina et al., 2011; Arvantini, 2012; Leeman et al., 2014). However, Wiget et al. (2010) indicated that %V and VarcoV are more variable than nPVI among different English speakers. While Dellwo and Fourcin (2013) proposed that speaker-specific information is also reported in the duration of voiced and unvoiced intervals in the German-Swiss language, Dellwo et al. (2015) suggested that between-speaker variability of speech rhythm measures is robust in different within-speaker situations by considering %V, ΔV (ln), ΔC (ln), $\Delta peak$ (ln) based on the speed of speech production organs movement and linguistic structures in which 12 German speakers read seven sentences in five different speech rates: very slow, slow, normal, fast and very fast.

Persian within-speaker and between-speaker differences with different speech rates have been studied by Asadi et al. (2018) where 10 Persian speakers read the story “The North Wind and the Sun” in 5 different speech rates. The results showed that %V is a robust parameter in distinguishing between-speaker factors. Taghva et al. (2021) studied a read text in Persian and indicated that VarcoC and %V are the robust measures in between-sentence differences in which ten Persian speakers read the story of “The North Wind and the Sun”.

Having studied the literature and to the best of the present researchers’ knowledge, no study has yet comprehensively investigated the quantitative rhythmic measures for Kurdish language and its varieties, which, as mentioned in “Section 1”, is being spoken in parts of Turkey, Iran, Armenia, Iraq, Syria, and Azerbaijan (Windfuhr, 1989). To fill this gap, this study examined the between-sentence and between-speaker rhythmic measures in two different speaking styles (read and spontaneous speech) in Kalhori, a variety of Kurdish.

3. METHOD

Ten native speakers of Kalhori variety who were originally from the same region (Kermanshah, which is the largest Kurdish-speaking city in Iran [Borjian, 2017]), including five males and five females, participated in

this study. Ages ranged between 21 and 40 with a mean of 31.72 years and SD of 8.81. To be of the same social group, all participants were recruited among Shiraz University students.

The experiment took place at Shiraz University’s acoustic room where the researchers were able to use Zoom h4 recorder. The recorder was positioned diagonally around 20cm away from the participants’ mouths using a base.

To move forward with the experiment, two sets of corpora were compiled. Gibbon (2022) had indicated that depending on the styles, the degree of rhythm may vary from being more rhythmical in the rhetoric of public speeches, poetry recitation and reading aloud to being more arhythmical in planning discussions. Therefore, in the first corpus, the participants read an identical story to determine the rhythmic typology of Kalhori variety and to express between-sentence and between-speaker rhythmic variability in Kalhori read speech. Following previous studies (Pellegrino, 2019; Gibbon, 2022; Asadi et al., 2018), in this study, the potential effects of age, style, and speech rate on the rhythmic metrics was eliminated by selecting participants of approximately the same age group (21-40) who read the same story at a normal speed. In the second corpus, participants were interviewed to observe between-speaker rhythmic variability in Kalhori spontaneous speech.

3.1. Experiment 1: Read corpus

To elicit precise instances of between-sentence and between-speaker diversities, in the first experiment, attempts were made to provide identical situations for all participants. As a result, we gave the Kalhori version of the “North Wind and the Sun” story (written with Persian orthography) to the participants before beginning the interview and asked them to read it at a normal speed. The reason for selecting this story was that it has been recognized as a standard for phonetic documentation of many languages by the International Phonetic Association, and it has been frequently utilized by speech scientists for analyzing both sound segments and prosodies (Baird et al., 2022). This story comprises seven complex Kalhori sentences, a total of 70 tokens (10 speakers \times 7 sentences). In the event that a mistake was made by the participants while reading the sentences during the interview, they were asked to read the sentences again.

3.2. Experiment 2: Spontaneous corpus

To devise the spontaneous corpus, we interviewed the participants by asking them six questions about the content of which they were unaware prior to the study. Then, 21 sentences were extracted from each participant’s speech. The selected sentences were grammatically meaningful; the speakers did not express them with hesitation and did not have any pronunciation problems. Eventually, the final set of data for this part of the experiment comprised 210 tokens (10 speakers \times 21 sentences).

Where *Invl* is vowel, consonant or peak intervals and *N* is the number of these intervals.

3.5. Data analysis

To calculate all the rhythm measures in Praat, the script written by Dellwo (<https://www.cl.uzh.ch/de/people/team/phonetics/vdellw.html>) was used. Then, correlational measures were determined after running Pearson correlation analysis. Pearson correlation is a statistical method that measures the linear relationship between two continuous variables. It is useful for feature selection, the process of choosing the most relevant variables for analyzing and reducing the dimensionality of the data (James et al. 2013). As shown in Table 1, we calculated 68 durational rhythmic measures, which was a very large a number for effective analysis. We, therefore, applied Pearson correlation as a feature selection method to reduce the number of measures and retain the most relevant ones for speech rhythm analysis. Pearson correlation allowed us to examine the linear relationship between each pair of measures (He & Dellwo, 2016) and eliminate those that were highly correlated ($r > 0.5$) with others so that redundant information about speech rhythm would be avoided. Those measures that had low correlation ($r < 0.5$) were kept since they provided independent information about speech rhythm. Moreover, sentences and/or speakers were considered as an independent variable and the rhythmic measures as dependent variable.

Afterwards, since in the read corpus data were balanced and orthogonal, to ascertain Kalthori's between-sentence rhythmic measures variability, a one-way ANOVA test was run. ANOVA was used to see how language and method affect measures. It was also utilized to sentence types and determine whether means differ significantly and helped the authors understand the data's variability and patterns (see Arvaniti, 2012).

Furthermore, to explore Kalthori's between-speaker rhythmic measures variability, a mixed-design ANOVA or a MANOVA was used. MANOVA is a statistical method that compares the means of multiple dependent variables across different groups and conditions, while accounting for both between-subjects and within-subjects factors (Stevens, 1996). To interpret the results of MANOVA, both the multivariate tests and the univariate tests were studied. Multivariate tests determine the significance of the overall effects of the factors on combination of dependent variables; and, univariate tests show effects of the factors on each dependent variable (Stevens, 1996). In this study, MANOVA allowed simultaneous comparison of multiple dependent variables (rhythmic measures) across two independent variables i.e., styles and speakers. It showed whether any rhythmic measures differed significantly between the two styles when considered together.

4. RESULTS

To identify the typology of Kalthori rhythmic features, and determine the between-speaker and between-sentence

rhythmic variabilities, we used the measures investigated in previous studies (Ramus et al., 1999; White & Mattis, 2007; Grabe & Low, 2002; Dellwo, 2009, 2010; Dellwo et al., 2012; Dellwo et al., 2015).

4.1. Read corpus analysis

The sum of interval durations considered in the read experiment are shown in Table 2.

Table 2: Sum of considered intervals in the read corpus

Intervals	Sum
segmental intervals	3880
syllable intervals	1795
consonantal intervals	1632
vocalic intervals	1599
consonantal-vocalic intervals	3439
consonant intervals	1215
vowel intervals	1581
peak intervals	1658
voiced intervals	1599
unvoiced intervals	1632

4.1.1. The rhythmic typology of Kalthori

To determine the typology of rhythm in Kalthori variety ΔC , %V and nPVI-V (Ramus et al., 1999; Grabe & Low, 2002; Dellwo, 2010) were explored. The descriptive statistics are as follows (Table 3):

Table 3: The descriptive statistics of ΔC , %V and nPVI-V

	Mean	Std.	Skewness	Kurtosis
ΔC	.056	.01	.59	-.02
%V	42.28	5.61	.42	-.01
nPVI_V	47.36	9.02	.08	-.541

The comparison of the results of table 3 with Ramus et al (1999) – and Grabe and Low (2002) – shows that the mean value of ΔC is 0.056, which is relatively low compared to some stress-timed languages like English (0.07). The mean value of %V is 42.28 which is relatively high compared to some stress-timed languages like English (38.5), and the mean value of nPVI_V is 47.36, which is also relatively low compared to some stress-timed languages like English (52.1). Table 4 presents the standard deviation of %V, ΔC , and nPVI-V of Kalthori Kurdish in comparison with English as a stress-timed language and French as a syllable-timed language derived from Ramus et al. (1999), and Grabe and Low (2002).

Table 4: Classifying the rhythm of Kalhori Kurdish (English and French measures are derived from Ramus et al. 1999 and Grabe and Low, 2002)

	%V (std)	ΔC (std)	nPVI-V (mean)
English	5.4	1.63	54
French	4.5	0.74	43.05
Kalhori Kurdish	5.61	0.01	47.36

4.1.2. Between-sentence measures in read corpus

To answer the second question of this study and understand the impact of sentence structure on the rhythmic measures of read Kalhori speech, at first, Pearson correlation analysis was run to keep the measures with low correlation ($r < 0.5$). The results showed that rateSyl, ΔSylLn, VarcoC, nPVI-V, and %V are the least correlated measures in the read corpus. As mentioned in part (3.4), RateSyl measures the overall speech rate, ΔSylLn shows how the syllable lengths vary within an utterance, VarcoC reveals how consonantal intervals vary with regards to their average length, nPVI-V indicates how similar or different the vowel durations are from each other, and %V tells us how much of the utterance is occupied by vowels. The results of Pearson correlation analysis of these five measures are represented on Table 5.

Table 5: Pearson correlation analysis for read speech

	rateSyl	ΔSylLn	VarcoC	nPVI_V	%V
rateSyl	1	-.01	-.02	.27*	.39**
ΔSylLn	-.01	1	.12	.04	-.07
VarcoC	-.02	.12	1	-.01	.06
nPVI_V	.27*	.04	-.01	1	.33**
%V	.39**	-.07	.06	.33**	1

** Correlation is significant at the 0.01 level (2-tailed).
 * Correlation is significant at the 0.05 level (2-tailed).

Table 5 indicates that the selected measures (rateSyl, ΔSylLn, VarcoC, nPVI_V, and %V) are less related to each other compared to the rest of the measures because of their low correlation coefficients ($r < 0.5$) which suggests that they capture different aspects of speech rhythm and do not provide redundant information.

Afterwards, a one-way ANOVA test was used for the measures selected using Pearson correlation analysis. We considered the sentences of read corpus as the independent variable and the measures as the dependent variables (Table 6).

The results of the ANOVA one-way test (Table 6) indicate that VarcoC and %V are meaningfully significant. Although VarcoC is also significant, the significance level of VarcoC is 0.03, which is very close to 0.05, the usu-

Table 6: One-way ANOVA for between-sentence identification based on the read corpus

Measures	Sum of squares	F	Sig.
rateSyl	2.56	1.29	.27
ΔSylLn	.04	1.93	.08
VarcoC	.13	2.40	.03
nPVI-V	978.92	2.21	.05
%V	739.24	5.41	.00

al threshold for rejecting the null hypothesis. This means that VarcoC is only marginally significant. Moreover, the F-value of VarcoC is 2.40, which is much lower than the F-value of %V, which is 5.41. F-value is the ratio of the variance between groups to the variance within groups for each measure. Therefore, the higher the F-value, the greater the between-sentence variabilities of this measure. This means that VarcoC has a smaller ratio of variance between groups to variance within groups than %V, and it explains less of the total variation in the data than %V. Therefore, VarcoC is not as effective as %V in discriminating between sentences based on their speech rhythm. So, comparing the significant actions, %V (F-value=5.41) is the most efficient measure to reflect the Kalhori between-sentence variability based on this study's data. Figure 3 indicates the %V and VarcoC changes for the sentences of the study.

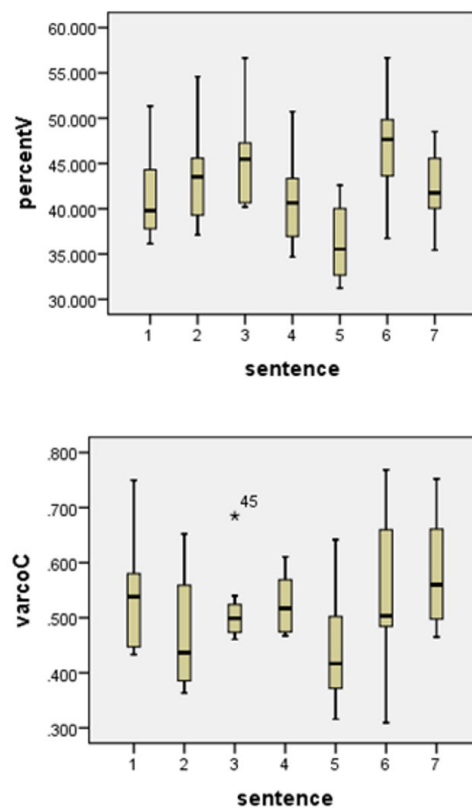


Figure 3: %V and VarcoC boxplots based on the sentences for the read corpus

Based on the boxplots in figure 3, comparing the sentences in terms of their %V values can be done. For example, we can see that sentence 5 has the lowest median %V value, suggesting that this sentence on average has fewer vowels than the other sentences. It also has the lowest variability in %V values, which means this sentence has less variation in vowel density compared to other sentences. Sentence 6 has the highest median %V value, which means that this sentence on average has more vowels than the other sentences. It also has the highest variability in %V values, which means that this sentence has more variation in vowel density than the other sentences.

Moreover, VarcoC comparison between the sentences indicates that sentence 2 has the lowest median VarcoC value, meaning this sentence on average and compared to others has less variability in consonant length. It also has the lowest variability in VarcoC values, which means that this sentence has more consistent consonant length than the other sentences. Sentence 7 on average and compared to others has the highest median VarcoC value, indicating more variability in consonant length. It also has the highest variability in VarcoC values, which means that this sentence has more variation in consonant length than the other sentences.

4.2. Spontaneous corpus analysis

We investigated 210 tokens of spontaneous Kalhori sentences (10 speakers × 21 sentences) in the second experiment. The sum of duration of intervals considered in this experiment is shown in table 7.

Table 7: Sum of considered intervals in the read corpus

Intervals	Sum
segmental intervals	4999
syllable intervals	2364
consonantal intervals	2109
vocalic intervals	2000
consonantal-vocalic intervals	4479
consonant intervals	1641
vowel intervals	1974
peak intervals	2208
voiced intervals	2000
unvoiced intervals	2109

According to the results for the Pearson correlation analysis, measures including rateSyl, ΔSylLn, VarcoC, nPVI-V and %V had low correlation. Table 8 shows the results of these five measures' Pearson correlation analysis.

Table 8 indicates that the selected measures (rateSyl, ΔSylLn, VarcoC, nPVI_V, and %V), which had low correlation coefficients ($r < 0.5$), are less related to each other compared to the rest of the measures. In other words, they capture different aspects of speech rhythm, and do not provide redundant information.

Table 8: Pearson correlation analysis for the spontaneous corpus

	rateSyl	ΔSylLn	VarcoC	nPVI_V	%V
rateSyl	1	.07	.05	-.10	.35**
ΔSylLn	.07	1	.20**	.14*	.00
VarcoC	.05	.20**	1	.02	.09
nPVI_V	-.10	.14*	.02	1	.09
%V	.35**	.00	.09	.09	1

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

4.3. Between-speaker measures in read and spontaneous corpus

To answer the third question, regarding which durational measures have a significant impact on the between-speaker rhythmic variability in read and spontaneous Kalhori speech, a MANOVA test was run on the data obtained from the results of the Pearson correlation analysis in section 4.1.2 and 4.2.2 for both corpora. In this study, the style and speaker were applied as independent variables and the rhythmic measures as the dependent variables. Table 9 presents the Multivariate Tests and Table 10 shows the tests of Between-Subjects Effects (univariate test).

Table 9: Multivariate Test showing the influence of style and speakers on the rhythmic measures

Effect		Value	F	Sig.
Intercept	Pillai's Trace	.98	3993.95b	.00
	Wilks' Lambda	.01	3993.95b	.00
Style	Pillai's Trace	.34	27.31b	.00
	Wilks' Lambda	.65	27.31b	.00
speakers	Pillai's Trace	.55	3.63	.00
	Wilks' Lambda	.53	3.86	.00
Style * speakers	Pillai's Trace	.19	1.14	.23
	Wilks' Lambda	.82	1.15	.22

The MANOVA results (Table 9) for the multivariate tests demonstrate that both "Styles" and "Speakers" have significant and individual impacts on the variations in rhythmic measures. While the interaction between styles and speakers may not be statistically significant, the main effects of styles and speakers are indeed significant and contribute to the observed variability in the dataset.

The results of s of Between-Subjects test Effects (Table 10) for the dependent variables (rhythmic measures) under the effect Intercept (rateSyl, ΔSylLn, VarcoC, nPVI_V, %V) all show significant p-values ($p < .001$). This suggests that these features are highly effective in distinguishing between the styles and speakers.

Table 10: Tests of Between-Subjects Effects (univariate test), showing the influence of styles and speakers on the rhythmic measures

Source	Dependent Variable	F	Sig.
Intercept	rateSyl	7532.33	.00
	ΔSylLn	2853.23	.00
	VarcoC	3837.55	.00
	nPVI_V	2638.98	.00
	%V	9540.12	.00
Styles	rateSyl	76.74	.00
	ΔSylLn	23.16	.00
	VarcoC	.06	.80
	nPVI_V	.73	.39
	%V	5.44	.02
Speakers	rateSyl	7.41	.00
	ΔSylLn	.78	.63
	VarcoC	.41	.92
	nPVI_V	2.43	.01
	%V	8.13	.00
Styles * speakers	rateSyl	1.92	.04
	ΔSylLn	.42	.92
	VarcoC	1.16	.32
	nPVI_V	.92	.50
	%V	.90	.52

For the “Styles” effect (Table 10), some features have significant p-values (rateSyl, ΔSylLn, %V), indicating their level of importance in distinguishing between the two speaking styles (read and spontaneous). However, VarcoC (p = .80) and nPVI-V (p = .39) do not indicate a significant effect, suggesting that they might not be as effective in differentiating styles.

Under the “Speakers” effect, the dependent variables rateSyl, nPVI_V, and %V show significant p-values (p < .001), suggesting significance in distinguishing between individual speakers. On the other hand, ΔSylLn and VarcoC have higher p-values (ΔSylLn: p = .11, VarcoC: p = .67), indicating that they might be less effective in differentiating individual speakers. This also means that these measures do not vary much across speakers in either read or spontaneous speech. However, the interaction between style and speakers are not statistically significant.

Based on these results, it can be concluded that %V and rateSyl are the rhythmic measures that can discriminate speakers the best followed by nPVI-V. These two rhythmic measures (%V and rateSyl) have significant effects of speaker at the 0.000 level, and have relatively large F-values compared to the other measures which means that they vary significantly across speakers in both read and spontaneous speech. Table 11 and Figure 4 show %V and rateSyl changes for the participants of the study.

Table 11: RateSyl and %V mean in both Read and Spontaneous (Spo) corpora

Speakers	rateSyl			%V		
	Read	Spo	Mean	Read	Spo	Mean
1	4.77	5.32	5.18	45.23	41.39	42.35
2	4.25	5.51	5.20	51.24	46.77	47.89
3	4.09	5.82	5.39	41.71	39.61	40.14
4	3.94	4.72	4.52	36.47	34.47	34.97
5	3.36	4.41	4.15	39.71	42.33	41.68
6	4.29	5.42	5.14	38.52	33.53	34.78
7	4.91	5.19	5.12	43.62	42.90	43.08
8	4.41	4.77	4.68	42.85	38.23	39.39
9	5.08	6.14	5.87	41.90	42.15	42.09
10	4.44	6.03	5.63	41.56	41.69	41.66

The comparison of rateSyl in both corpora, represented on Table 11, indicates that Speaker 9 exhibits the highest mean rateSyl value in both styles, implying the fastest speech rate on average compared to other speakers; while Speaker 5 displays the lowest mean rateSyl value in both styles, suggesting the slowest speech rate on average in comparison with other speakers. Speakers 2, 3, 6, and 8 have similar mean rateSyl values in both modes, indicating relatively consistent speech rates between their read and spontaneous speech. However, Speakers 1, 4, 7, and 10 have intermediate mean rateSyl values in both styles, which suggests moderate speech rates compared to the other speakers.

The mean %V value varies among the speakers in both read and spontaneous speech styles. Speaker 2 exhibits the highest mean %V value in both styles while speaker 4 displays the lowest mean %V. Speakers 1, 5, 7, 8, 9, and 10 have intermediate mean %V values in both styles. However, Speaker 6 shows a notable difference in mean %V value between read and spontaneous speech modes, with a lower value in spontaneous speech compared to read speech. The comparison of %V in both corpora is shown in Table 11 and Figure 4.

The boxplots for rateSyl (Figure 4) show how the 10 speakers differ in their speech rate in Kalhori speech. According to the plots, the range of rateSyl values for speaker 9 is from 4 to 7.2, meaning that this speaker sometimes speaks as slow as 4 syllables per second and sometimes as fast as 7.2 syllables per second. This is while other speakers’ ranges were from 2.8 to 6.8. Speaker 2 also produces the lowest variability in rateSyl values, as indicated by the width and shape of the box and whiskers. The range of rateSyl values for speaker 2 is from 3 to 4.6, which means that this speaker does not change their speech rate as much and speaks consistently around 3 to 4 syllables per second. This is a narrower range compared to others which range from 3 to 7.2. The other speakers have median rateSyl values ranging from 35 to 40, and variabilities ranging from low to high. Therefore, rateSyl varies signif-

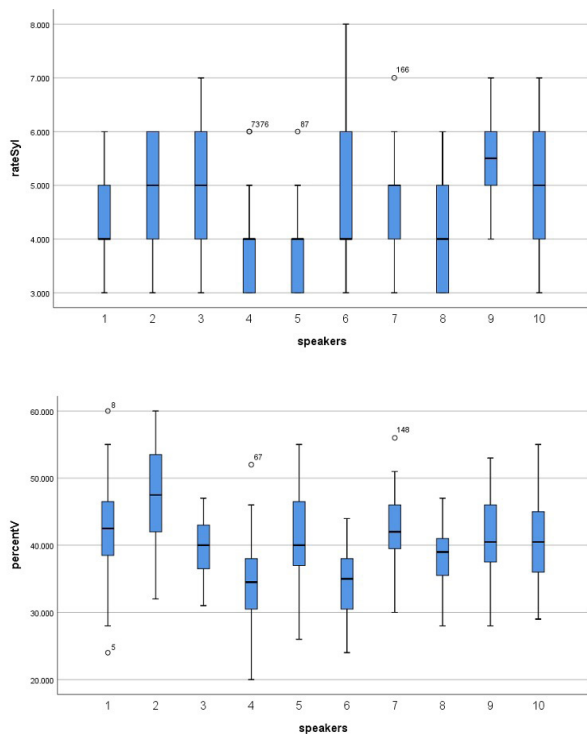


Figure 4: Boxplots of %V and rateSyl based on the speakers in both spontaneous and read speech

icantly between these 10 speakers, and it signifies different levels of variability, different medians, and different ranges across speakers.

The boxplots for %V (Figure 4) shows how the 10 speakers differ in their vocalic intervals in both spontaneous and read speech. Accordingly, speaker 1 has the highest median %V value, meaning that this speaker on average has more vowels in speech than the other speakers. It also has the highest variability in %V values, which means that this speaker, compared to others, produces more variation in vocalic intervals. Speaker 3 has the lowest median %V value. It also has the lowest variability in %V values. The other speakers' median %V values range from 35% to 40%, and their variabilities range from low to high. Some speakers also have outliers, extreme values that deviate from the rest of the data. These outliers indicate that some speakers in some cases produce very low or very high %V values. Therefore, %V varies significantly between these 10 speakers, as it shows different levels of variability, different medians, and different ranges across speakers.

5. DISCUSSION AND CONCLUSION

Documenting and describing languages, whether they are endangered or widely spoken, has many purposes, from conserving the inherited knowledge of the language community to exploring the range of structures and communication events the human mind can handle (Gibbon, 2022). One aspect of this range is how language relates

to other modes of communication, and one feature of this aspect is the specific rhythm patterns of speech that distinguish a language community, along with other regular events in daily life and culture (Gibbon, 2022).

To respond to the first research question (i.e., to study the rhythmic typology of Kalhori rhythm based on the read corpus), ΔC , %V and nPVI-V were analyzed. Ramus et al. (1999) by calculating ΔC , %V showed that English is a stress-timed language and French is a syllable-timed language while stress-timed languages demonstrated a high ΔC by reflecting high C-interval variability and low %V by reflecting high V-interval variability, and syllable-timed languages indicated a low ΔC and high %V. On the other hand, nPVI that were studied by Grabe and Low (2002) classified English as a stress-timed language and French as a syllable-timed language since the variability of consecutive vocalic intervals in stress-timed languages was higher than syllable-timed languages.

Findings of the descriptive analysis of read corpus (Table 3) demonstrate that the Kalhori nPVI-V is 47.36, std of %V is 5.61 and std of ΔC is 0.016. Table (4) compares ΔC , %V of French and English (derived from Ramus et. al, 1999), their nPVI-V (derived from Grabe & Low, 2002) to the finding of this study. These findings are comparable to the outcome of this study since both Ramus et al. (1999) and Grabe and Low (2002) studies used the story of “The North Wind and the Sun” to collect their data.

As lower value of %V shows more variability of vowel intervals, and a lower value of ΔC reflects less variability of consonant intervals (Dellwo, 2010), Table (4) presents that Kalhori Kurdish has less variability of vowel intervals and less variability of consonant intervals than English and French. Moreover, nPVI-V reflects the variability of successive vocalic intervals.

Drawing on Grabe and Low (2002), Kalhori read speech is placed among the stress-timed languages since Table (4) shows that the variability of vowel intervals in Kalhori Kurdish is higher than French, but lower than English. Consequently, the rhythm class of Kalhori Kurdish can be placed between stress-timed and syllable-timed based on the read corpus with the controlled situation in which participants of the same aged group read a story in a normal speed.

Furthermore, conducting the first experiment in read corpus allowed us to investigate the impact of sentence structure on the rhythmic measures of read Kalhori speech. Five measures of rateSyl, $\Delta Syll_n$, VarcoC, nPVI-V and %V were selected based on Pearson correlation analysis. The results indicate that only two of these measures (VarcoC and %V) are significantly different between sentences. While VarcoC is a measure of consonantal variability and reflects the degree of variation in the duration of consonantal intervals, %V is a measure of vocalic proportion and shows the percentage of vowel duration in the total duration of the utterance. These two measures are related to the syllable structure and the vowel-consonant ratio of the sentences (Dellwo, 2010). According to the results (Table 6), VarcoC is only marginally significant

VarcoC even while showing a low F-value, proposing a small part of the total variation in the data.

On the other hand, %V is highly significant, meaning that the difference between sentences is due to sentence structure rather than random variation. Moreover, %V has a high F-value (5.41), which is indicative of a large part of the total variation in the data. The results suggest that, based on data, %V is the best measure to determine the Kalhori between-sentence variability. In other words, sentences with different structures have different proportions of vowel duration in their total duration. This may be related to the phonological and morphological features of Kalhori, such as vowel harmony, vowel lengthening, and consonant clusters. Hence, the outcome of this study is aligned with the results of Taghva et al. (2021), who showed that VarcoC and %V are robust measures among Persian between-sentence differences.

To respond to the research question probing the most efficient durational rhythmic measures for between-speaker rhythmic variability in Kalhori speech, the read speech style as well as the spontaneous speech style were examined using five rhythmic measures selected by Pearson correlation analysis: ratesyl, Δ Sylln, VarcoC, nPVI-V, and %V. Therefore, a MANOVA (Table 9 and 10) was conducted to examine which rhythmic measure or measures best discriminated between-speakers. The results revealed that:

- RateSyl, %V and nPVI-V differed significantly between both speech styles and speakers. However, the F-value of nPVI-V (4.37) is less than RateSyl (11.036) and %V (11.121).
- Δ Sylln and VarcoC did not show significant differences between speakers.

Therefore, based on this analysis, the rhythmic measures that best discriminated between Kalhori speakers in both read and spontaneous speech styles were %V and rateSyl. These two measures identified individual speakers most effectively based on durational rhythmic analysis. Consequently, the rate of the syllable intervals together with the vocalic proportion of speech are the most useful features for identifying the speakers based on durational rhythmic measures. Findings of this study are in line with the findings of Asadi et al. (2018) and Dellwo et al. (2015) for Persian and German. Asadi et al. (2018) demonstrated the robustness of %V against both sources of within-speaker variability including time-lapsing and speech-rate variability.

In conclusion, the use of durational measures as a forensic cue may have important implications for the situations where speaker identification information is required (Arvaniti, 2012; Leeman et al., 2014; Dellwo et al., 2015; He & Dellwo, 2016; Asadi et al., 2018). Therefore, the findings of this study hold great potential for enhancing speaker identification in diverse forensic cases. Particularly, the identification of %V and rateSyl as the most distinguishing measures between speakers suggests their potential as valuable acoustic-prosodic features for forensic voice comparison tasks.

However, the comparison of the most discriminative measures for between-sentence variability (VarcoC and

%V) with those for between-speaker variability (rateSyl, %V) reveals that %V is influenced by both language-specific and speaker-specific factors, which may affect its variability between sentences and speakers. Hence, while rhythmic measures such as %V hold promise as effective discriminators between speakers, their performance can be influenced by factors other than the voice alone, including linguistic peculiarities. Therefore, forensic practitioners must exercise caution in adapting and validating speaker identification models to account for the specific linguistic and contextual characteristics of the language being investigated.

This study thus sheds light on the complex interplay between language-specific factors and speaker identification, highlighting the need for a nuanced and comprehensive approach to ensure the accuracy and reliability of forensic voice analysis techniques. Analyzing other varieties of Kurdish language could also serve as a fruitful area of study for future attempts.

6. REFERENCES

- Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of phonetics*, 26(1), 3-25. <https://doi.org/10.1006/jpho.1997.0063>
- Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, 40(3), 351-373. <https://doi.org/10.1016/j.wocn.2012.02.003>
- Asadi, H., Nourbakhsh, M., He, L., Pellegrino, E., & Dellwo, V. (2018). Between-speaker rhythmic variability is not dependent on language rhythm, as evidence from Persian reveals. *International Journal of Speech, Language and the Law*, 25(2), 151-174. <https://doi.org/10.1558/ijsl.37110>
- Baird, L., Evans, N., & Greenhill, S. J. (2022). Blowing in the wind: Using 'North Wind and the Sun' text-to-sample phoneme inventories. *Journal of the International Phonetic Association*, 52(3), 453-494. <https://doi.org/10.1017/S002510032000033X>
- Barbosa, P. A. (2002). Explaining cross-linguistic rhythmic variability via a coupled-oscillator model of rhythm production. In *Speech Prosody 2002*
- Barbosa, P., & Bailly, G. (1994). Characterisation of rhythmic patterns for text-to-speech synthesis. *Speech Communication*, 15(1-2), 127-137. [https://doi.org/10.1016/0167-6393\(94\)90047-7](https://doi.org/10.1016/0167-6393(94)90047-7)
- Belelli, S. (2019). Towards a dialectology of Southern Kurdish: Where to begin. *Current issues in Kurdish linguistics*, 73, 92.
- Bertrán, A. P. (1999). Prosodic typology: on the dichotomy between stress-timed and syllable-timed languages. *Language design: journal of theoretical and experimental linguistics*, 2, 103-130.
- Borjian, H. (2017). Kermanshah vii. languages and dialects. *İçinde Encyclopædia Iranica. Erişim tarihi: 14 Temmuz 2021*.
- Brown, K. (2005). *Encyclopedia of language and linguistics* (Vol. 1). Elsevier, 634-38.
- Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics*, 26(2), 145-171.
- Dabirmoghadam, M. (2013). *Typology of Iranian 2*. Tehran: SAMT.
- Daneshpazhouh, F. (2010). Ergativity in Kurmanji (Badini), Sorani and Hawrami Kurdish, Tehran: Dastan.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of phonetics*, 11(1), 51-62. [https://doi.org/10.1016/S0095-4470\(19\)30776-4](https://doi.org/10.1016/S0095-4470(19)30776-4)
- Dellwo, V., Huckvale, M., & Ashby, M. (2007). How is individuality expressed in voice? An introduction to speech production and description for speaker classification. *Speaker Classification I: Fundamentals, Features, and Methods*, 1-20. DOI: 10.1007/978-3-540-74200-5_1
- Dellwo, V., Schmid, S., Schwarzenbach, M., & Studer-Joho, D. (2009). Choosing the right rate normalization method for

- measurements of speech rhythm. <https://doi.org/10.5167/uzh-45236>
- Dellwo, V. (2010). *Influences of Speech Rate on the Acoustic Correlates of Speech Rhythm: An Experimental Phonetic Study Based on Acoustic and Perceptual Evidence*. Doctoral dissertation, Bonn University, Germany.
- Dellwo, V., Leemann, A., & Kolly, M. J. (2012, September). Speaker idiosyncratic rhythmic features in the speech signal. *Inter-speech Conference Proceedings*. <http://interspeech2012.org/accepted-abstract.html?id=1195>
- Dellwo, V., & Fourcin, A. (2013). Rhythmic characteristics of voice between and within languages. *Travaux neuchâtelois de linguistique*, (59), 87-107. <https://doi.org/10.26034/tranel.2013.2947>
- Dellwo, V., Leemann, A., & Kolly, M. J. (2015). Rhythmic variability between speakers: Articulatory, prosodic, and linguistic factors. *The Journal of the Acoustical Society of America*, 137(3), 1513-1528. <https://doi.org/10.1121/1.4906837>
- Fattah, I.K. (2000). Les dialectes kurdes méridionaux: *Étude linguistique et dialectologique* (Acta Iranica 37). Leuven: Peeters.
- Fuchs, R. (2016). The Concept and Measurement of Speech Rhythm. In: *Speech Rhythm in Varieties of English. Prosody, Phonology and Phonetics*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-47818-9_3
- Gharib, H. E. (2011). *Transitivity alternations in Sorani Kurdish* (Doctoral dissertation, University of Kansas). <http://hdl.handle.net/1808/9791>
- Gibbon, D. (2022). Speech rhythms: learning to discriminate speech styles. *Proc. Speech Prosody 2022*, 302-306.
- Gibbon, D. (2022). New Perspectives on Ibibio Speech Rhythm. In *Current Issues in Descriptive Linguistics and Digital Humanities: A Festschrift in Honor of Professor Eno-Abasi Essien Urua* (pp. 457-486). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-19-2932-8_34
- Gibbon, D. (2023). The rhythms of rhythm. *Journal of the International Phonetic Association*, 53(1), 233-265. doi:10.1017/S0025100321000086
- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology*, 7(1982), 515-546. <https://doi.org/10.1515/9783110197105>
- Harrington, J. (2010). *Phonetic analysis of speech corpora*. John Wiley & Sons.
- He, L. (2022). Characterizing first and second language rhythm in English using spectral coherence between temporal envelope and mouth opening-closing movements. *The Journal of the Acoustical Society of America*, 152(1), 567-579. <https://doi.org/10.1121/10.0012694>
- He, L., & Dellwo, V. (2016). The role of syllable intensity in between-speaker rhythmic variability. *International Journal of Speech, Language & the Law*, 23(2). doi: 10.1558/ijssl.v23i2.30345
- James, A. L. (1929). *Historical introduction to French Phonetics*, London: ULP.
- James, A. L. (1938). *Our spoken language* (No. 9). London: T. Nelson.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, p. 18). New York: springer.
- Karimi-Doostan, G. (2002). Syllable structure in Kurdish language. *Specialized Journal of Language and Literature of Mashhad Faculty of Literature and Human Sciences*, 35(2), 235-248.
- Kreyenbroek, P. G. (2005). On the Kurdish language. In *The Kurds* (pp. 62-73). Routledge.
- Lee, C. S., & Todd, N. P. M. (2004). Towards an auditory account of speech rhythm: application of a model of the auditory 'primal sketch' to two multi-language corpora. *Cognition*, 93(3), 225-254. <https://doi.org/10.1016/j.cognition.2003.10.012>
- Lee, C. S., Kitamura, C., Burnham, D., & McAngus Todd, N. P. (2014). On the rhythm of infant-versus adult-directed speech in Australian English. *The Journal of the Acoustical Society of America*, 136(1), 357-365. <https://doi.org/10.1121/1.4883479>
- Leemann, A., Kolly, M. J., & Dellwo, V. (2014). Speaker-individuality in suprasegmental temporal features: Implications for forensic voice comparison. *Forensic science international*, 238, 59-67. <https://doi.org/10.1016/j.forsciint.2014.02.019>
- Leong, V., & Goswami, U. (2014). Impaired extraction of speech rhythm from temporal modulation patterns in developmental dyslexia. *Frontiers in human neuroscience*, 8, 96. <https://doi.org/10.3389/fnhum.2014.00096>
- Leong, V., Stone, M. A., Turner, R. E., & Goswami, U. (2014). A role for amplitude modulation phase relationships in speech rhythm perception. *The Journal of the Acoustical Society of America*, 136(1), 366-381. <https://doi.org/10.1121/1.4883366>
- Liss, J. M., White, L., Mattys, S. L., Lansford, K., Lotto, A. J., Spitzer, S. M., & Cavinness, J. N. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of speech, language, and hearing research: JSLHR*, 52(5), 1334-1352. [https://doi.org/10.1044/1092-4388\(2009/08-0208\)](https://doi.org/10.1044/1092-4388(2009/08-0208))
- Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech. *The Journal of the Acoustical Society of America*, 129(5), 3258-3270. <https://doi.org/10.1121/1.3559709>
- Magne, C., Jordan, D. K., & Gordon, R. L. (2016). Speech rhythm sensitivity and musical aptitude: ERPs and individual differences. *Brain and Language*, 153-154, 13-19. <https://doi.org/10.1016/j.bandl.2016.01.001>
- MacKenzie, D. N. (1961). *Kurdish Dialect Studies, Volume I*. [2. Reprint, 1990, London: SOAS] London: Oxford University Press. Page references are to the reprint edition.
- McCarus, E. N. (1958). *A Kurdish Grammar: Descriptive Analysis of the Kurdish of Sulaimaniya, Iraq*. American Council of Learned Societies Program in Oriental Languages, Publications Series B-Aids-Number 10.
- McCarus, E. N. (2009). *Kurdish*. In G. Windfuhr Ed. London & New York: Routledge.
- McDougall, K. (2004). Speaker-specific formant dynamics: an experiment on Australian English/al. *International Journal of Speech, Language and the Law*, 11(1), 103-130. <https://doi.org/10.1558/sll.2004.11.1.103>
- McDougall, K. (2006). Dynamic features of speech and the characterization of speakers: Toward a new approach using formant frequencies. *International Journal of Speech, Language and the Law*, 13(1), 89-126. <https://doi.org/10.1558/ijssl.v13i1.89>
- Malisz, Z., O'Dell, M., Nieminen, T., & Wagner, P. (2017). Perspectives on speech timing: Coupled oscillator modeling of Polish and Finnish. *Phonetica*, 73(3-4), 229-255. <https://doi.org/10.1159/000450829>
- Mok, P., & Dellwo, V. (2008, May). Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. In: *Speech Prosody 2008, Campinas, Brazil*. ISCA, 423-426.
- Nolan, F. (2002). Intonation in speaker identification: an experiment on pitch alignment features. *International Journal of Speech, Language and the Law*, 9(1), 1-21. <https://doi.org/10.1558/sll.2002.9.1.1>
- O'Dell, M., & Nieminen, T. (1999, August). Coupled oscillator model of speech rhythm. In *Proceedings of the XIVth international congress of phonetic sciences* (Vol. 2, pp. 1075-1078). Berkeley: University of California.
- Pellegrino, E. (2019). The effect of healthy aging on within-speaker rhythmic variability: A case study on Noam Chomsky. *Loquens*, 6(1), e060. <https://doi.org/10.3989/loquens.2019.060>
- Pointon, G. E. (1980). Is Spanish really syllable-timed? *Journal of Phonetics*, 8(3), 293-304. [https://doi.org/10.1016/S0095-4470\(19\)31479-2](https://doi.org/10.1016/S0095-4470(19)31479-2)
- Polyanskaya, L., & Ordín, M. (2015). Acquisition of speech rhythm in first language. *The Journal of the Acoustical Society of America*, 138(3), EL199-EL204. <https://doi.org/10.1121/1.4929616>
- Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press.
- Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265-292. [https://doi.org/10.1016/S0010-0277\(99\)00058-X](https://doi.org/10.1016/S0010-0277(99)00058-X)
- Roach, P. (1982). On the distinction between 'stress-timed' and 'syllable-timed' languages. Linguistic controversies. In Crystal, Linguistic controversies: essays in linguistic theory and practice in honour of F.R. Palmer, 73-79.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Mahway, NJ: Lawrence Erlbaum.
- Thackston, W. M. (2006). *Kurmanji Kurdish: A Reference Grammar with Selected Readings*. Renas Media.
- Taghva, N., Abolhasani Zade, V., and Moloodi, A., (2021). Durational rhythmic variability of Persian between-sentence, Presented at SLE 2021, Societies Linguistica Europia, 223-224.

- Tilsen, S., & Johnson, K. (2008). Low-frequency Fourier analysis of speech rhythm. *The Journal of the Acoustical Society of America*, 124(2), EL34-EL39. <https://doi.org/10.1121/1.2947626>
- Tilsen, S. (2016). Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics*, 55, 53-77. <https://doi.org/10.1016/j.wocn.2015.11.005>
- Tilsen, S., & Arvaniti, A. (2013). Speech rhythm analysis with decomposition of the amplitude envelope: Characterizing rhythmic patterns within and across languages. *The Journal of the Acoustical Society of America*, 134(1), 628-639. <https://doi.org/10.1121/1.4807565>
- Todd, N. P. M. (1994). The auditory “primal sketch”: A multiscale model of rhythmic grouping. *Journal of new music Research*, 23(1), 25-70.
- White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501-522. <https://doi.org/10.1016/j.wocn.2007.02.003>
- White, L., Liss, J. M. & Dellwo, V. (2010). Assessment of rhythm. In A. Lowit and R.D. Kent ed. *Assessment of Motor Speech Disorders*, 312–352, San Diego: Plural Publishing.
- Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm? *The Journal of the Acoustical Society of America*, 127(3), 1559-1569. <https://doi.org/10.1121/1.3293004>
- Windfuhr, G. (1989). West Iranian Dialects. In R. Schmitt, ed. *Compendium Liguarum Iranicum*, 294-295, Wiesbaden: Reichert.
- Yoon, T. J. (2010). Capturing inter-speaker invariance using statistical measures of rhythm. *Speech Prosody 2010*, 100201, 1-4.