Journal of Language Researches
Vol.12, No.2, Autumn & Winter 2021-2022

Formerly "the Journal of the Faculty of Literature and Humanities"... University of Tehran

Speech Rhythm Measures: Acoustic Cues for Speaker Identification

Homa Asadi[1] & Batool Alinezhad[2]
(29-49)

Received: 22, July, 2020 & Accepted: 2, January, 2021

10.22059/jolr.2020.307695.666628
Print ISSN: 2288-1026-Online ISSN: 3362-2676
https://jolr.ut.ac.ir

Abstract

Rhythmic characteristics of speech based on consonantal and vocalic intervals as well as syllabic intervals vary between speakers of the same language. Nonetheless, the rhythmicity of a speech signal is not solely dependent on the durational variability of phonetic intervals but it is also associated with the variability of the intensity patterns as well. Acoustic parameter of intensity is largely determined by the articulatory behaviors of the speech organs such as lip movement or mouth aperture. Therefore, it is plausible that speaker idiosyncrasy in movement of speech articulators and anatomical differences in individual's vocal tracts may influence the energy distribution across a speech signal which subsequently leads to the variability in the values of the intensity measures. Using experimental phonetics tools and from an explicitly speaker-specific perspective, the present research attempts to explore potential speaker-specific acoustic parameters of speech rhythm which are extracted from the intensity contours across Persian speakers. This research aims to discover whether intensity-based measures of speech rhythm are able to discriminate between speakers in Persian. Two types of acoustic rhythmic measures based on the mean syllable intensity (stdevM, varcoM, rPVIm, nPVIm) and peak syllable intensity (stdevP, varcoP, rPVIp, nPVIp)) were selected for this study. Speech data from 12 Persian male speakers were recorded non-contemporaneously in laboratory environment on two different occasions separated by one to two weeks. Speech tokens were acoustically measured with PRAAT version 5.2.34 and statistical analyses were carried out with SPSS version 21 and R version 3.3.3. Results of the study indicated that speech rhythm measures based on intensity fluctuations play an important role in between-speaker rhythmic variability. In addition, discriminatory power of intensity-based measures is not affected by the language-dependent characteristics of Persian. The results also showed that the peak syllable intensity measures carry more speaker-specific information compared to the mean syllable intensity measures.

Keywords: Experimental phonetics, intensity-based measures, speaker identification, speech rhythm, between-speaker variability.

1 .Email of the corresponding author: h.asadi@fgn.ui.ac.ir
Assistant Professor of Linguistics, University of Isfahan, Isfahan, Iran.
2 .Associate Professor of Linguistics, University of Isfahan, Isfahan, Iran.

1.Introduction

Speech is highly organized in time. In the present study, we study how suprasegmental temporal features can contribute to speaker-individuality in Persian. Temporal characteristic of speech is a newly developed method applied in forensic phonetics. The rationale behind this idea is that humans differ in terms of the anatomical dimensions of the articulators, which result in idiosyncratic temporal characteristic articulation, namely speaker-specific rhythm. This relationship, however, is very complex as numerous acquired and language-specific characteristics also influence the temporal structure of speech. Suprasegmental speech rhythm measures involve a tight interaction between speakers' anatomy and their learned behavior. It seems that a possible contribution of the individual kinematics of the articulators which might lead to individual temporal characteristics in the signal may influence aspects of the signal that stand in relation to its intensity. To date, there has only been research in languages like English and German whose speakers might vary considerably in the way they operate the articulators to produce complex phonotactics and vowel reductions. Therefore, it is conceivable that language-specific features of Persian influence on temporal characteristic of speech. This study therefore sets out to explore between-speaker rhythmic variability as well as within-speaker rhythmic variability across Persian speakers to see to what degree such temporal characteristics vary among Persian speakers. Certainly a comprehensive understanding of speaker-specific temporal characteristics can only be obtained via a contrastive analysis of such characteristics in numerous different languages with widely different phonological systems.

2.Literature Review

In this section, different numerous studies on speech rhythm measures were categorized based on the measures of speech interval duration and temporal characteristics of the amplitude envelope.

3.Methodology

To test between-and within-speaker variability, twelve male speakers of Standard Contemporary Persian were recorded on two different sessions, separated by a time-lapse of one to two weeks. Speakers were asked to read the 54 sentences one by one, with a pause, and in a natural way, without any marked intonation. The microphone was positioned approximately 20 cm away from the mouth of the speakers in a diagonal position. Speech tokens were analysed using Praat (version 5.2.34, Boersma and Weenink 2013). For this study, we selected suprasegmental intensity-based measures retrieved from mean and peak syllable units of speech signals. Statistical analysis of data was carried out using R (R core Team 2014) version 3.3.3 and the R package lme4 (Bates, Maechler, Bolker and Walker 2016) and SPSS (IBM Corp. 2012).

4.Results

In this section, we provide the results of different acoustical models i.e. two-way anova, multinomial logistic regression and principal component analysis that were employed on the collected speech data of Persian. In the present study, we explored potential speaker-specific acoustic parameters of speech rhythm in Persian from a forensic-acoustic perspective. Statistical analysis of speech data revealed that selected acoustic speech rhythm measures are able to discriminate between Persian speakers and among them the variation coefficient of syllable peak intensity levels (varcoP) could account better for variability across speakers.

5.Discussion

We discuss how temporal characteristics of speech signals based on intensity fluctuations vary between Persian speakers and which acoustic parameter of speech rhythm has more potential in showing between-speaker variability. Furthermore, we compared the results of the present study to the findings of previous studies to see in what way Persian has been similar or different from other investigated languages

6.Conclusions

Based on the results obtained we can conclude that speech rhythm measures can be potentially characterized as language-independent parameters which can be utilized as acoustic cues in cases where knowledge about speaker-specific rhythm is needed.