

Temporal and Frequency Characteristics of Turkish Vowels in Laryngectomized Speakers: Preliminary Study

Larenjektomili Konuşmacılarda Türkçe Ünlülerin Temporal ve Frekans Özellikleri: Ön Çalışma

Fatma ESEN AYDINLI , Maviş Emel KULAK KAYIKCI , Nilda SUSLU 

Ethics Committee Approval: This study approved by the Hacettepe University Clinical Studies Ethics Committee, 20 February 2018, GO 18/207-34.

Conflict of interest: The authors declare that they have no conflict of interest.

Funding: None.

Informed Consent: Informed consent was taken from all participants.

Cite as: Esen Aydinli F, Kulak Kayikci ME, Suslu N. Temporal and Frequency Characteristics of Turkish Vowels in Laryngectomized Speakers, Preliminary Study. Medeniyet Med J. 2019;34:149-59.

Received: 07.12.2018
Accepted: 17.04.2019
Online First: 10.06.2019

Corresponding Author:
F. Esen Aydinli
ORCID: 0000-0002-5624-267X
Hacettepe University Faculty of
Health Sciences,
Department of Speech and
Language Therapy,
Ankara, Turkey
✉ fesen04@gmail.com

M.E. Kulak Kayikci
ORCID: 0000-0002-1791-0647
Hacettepe University Faculty of
Health Sciences,
Department of Speech and
Language Therapy,
Ankara, Turkey

N. Suslu
ORCID: 0000-0001-9901-3044
Hacettepe University
Faculty of Medicine,
Department of Otorhinolaryngology,
Ankara, Turkey

ABSTRACT

Objective: In the alaryngeal speech acoustics, formant frequencies and durations of vowels are two most fundamental parameters. However, it is not clear if the vocal tract transmission properties and temporal characteristics of vowels vary related to alaryngeal speech type.

The main purpose of this study is to compare the formant frequency and duration characteristics of Turkish vowels between the esophageal, tracheoesophageal, and laryngeal speakers.

Methods: Formant frequency and duration values of 8 Turkish vowels have been defined and these values were compared between the esophageal, tracheoesophageal, laryngeal speakers. Detection of a formant was based on Linear Predictive Coding and Fast Fourier Transform.

Results: Except the vowels /y/ and /i/; F1 and F2 frequency values of alaryngeal speakers were higher than laryngeal speakers both in the initial and final position. Alaryngeal groups' vowel space areas have been diminished. The vowels in the final position had longer durations comparing to their initial position in all groups.

Conclusion: The results of this study generally supported the shorter vocal tract and fronted-higher tongue position in Turkish alaryngeal speakers. The minor differences between other studies are thought to arise from the differences in consonantal context and possible coarticulatory effects. Acoustic findings obtained in the present study, indicated that transesophageal speech is more similar to laryngeal speech as compared to esophageal speech. However, further studies including higher number of participants are needed to verify the findings of this study.

Keywords: Laryngeal speech, formant frequency, duration, vowels

ÖZ

Amaç: Alarengel konuşmanın akustiğinde en temel iki parametre; ünlülerin formant frekansları ve durasyon özellikleridir. Ancak, vokal yolun iletim özellikleri ve ünlülerin zamansal özellikleri alaryngeal konuşma türüne göre farklılık göstermektedir.

Bu çalışmanın temel amacı Türkçe ünlülerin formant frekans ve durasyon özelliklerinin özefageal, trakeoözefageal ve laryngeal konuşmacılar arasında karşılaştırılmasıdır.

Yöntem: Bu çalışmada, Türkçede yer alan 8 ünlünün formant frekans ve durasyon değerleri belirlenmiş ve bu değerler özefageal, trakeoözefageal ve laryngeal konuşmacılar arasında karşılaştırılmıştır. Formant Frekansın belirlenmesi, Linear Productive Coding ve Fast Fourier Transform yöntemiyle yapılmıştır.

Bulgular: Alaryngeal konuşmacıların /y/ ve /i/ ünlüleri hariç diğer ünlülerde 1. Formant Frekans (F1) ve 2. Formant Frekans (F2) değerleri, sözcük başı ve sözcük sonunda laryngeal konuşmacılardan daha yüksek bulunmuştur. Alaryngeal konuşmacı gruplarının ünlü alanları daralmıştır. Tüm konuşmacılarda sözcük sonundaki ünlülerin sözcük başındakine göre durasyonları daha uzun belirlenmiştir.

Sonuç: Bu çalışmanın sonuçları, genel olarak, alaryngeal konuşmacılarda vokal traktın daha kısa ve dilin yüksek-önde pozisyonlandığını desteklemektedir. Literatürdeki diğer çalışmalar arasındaki küçük farklılıkların, kullanılan ünsüzler ve koartikülasyon farklılıklarından kaynaklanabileceği düşünülmektedir.

Ayrıca elde edilen bulgular trakeoözefageal konuşmanın, özefageal konuşmaya kıyasla laryngeal konuşmaya daha benzer olduğunu işaret etmektedir. Ancak, bu çalışmanın bulgularını doğrulamak için daha fazla sayıda katılımcı içeren daha fazla çalışmaya gereksinim vardır.

Anahtar kelimeler: Alaryngeal konuşma, formant frekans, durasyon, ünlüler



INTRODUCTION

Tracheoesophageal speech (TES) and esophageal speech (ES) are two major methods of gaining speech after laryngectomy. Many researchers have compared these two methods with respect to speech acceptability, intelligibility, and acoustics¹⁻¹⁰. In the alaryngeal speech acoustics, vowels constitute one of the foremost issues^{2,11-12}. Formant frequencies (FF) and durations of vowels are two main parameters searched in this domain. The supralaryngeal airways always act as an acoustic filter, suppressing the transfer of sound energy at certain frequencies, letting maximum energy through at other frequencies. The frequencies at which local energy maxima may pass through the supralaryngeal vocal tract are called FF¹³. FFs have been studied for revealing vocal tract properties of alaryngeal speakers¹⁻⁶. The results of these studies were consistent with higher FFs in alaryngeal male speakers regarding laryngeal speakers¹⁻⁵. This finding is thought to arise from shortening vocal tract in alaryngeal speakers. In addition, a lower tongue position at rest and fronted and higher tongue positions during articulation were thought to be other possibilities responsible for the higher FFs^{2,3,5}. The literature on acoustical and perceptual speech research has given some insight into the differences between, TES, ES, and normal speech. When the two alaryngeal speech methods (TES and ES) are compared, several studies showed differences between the two groups concerning FF, but the results of the studies are not consistent in the literature^{2,3,8}.

Although FFs are thought to be related to the vocal tract, additional factors may contribute to discrepancies in alaryngeal speech. In one study, a correlation had been found with the individual differences such as postoperative complications and FF⁵. It is known that for laryngeal speakers, changing the acoustic impedance of the vocal tract affects glottal vibrational patterns and glottal flow pulses^{14,15}. Inversely, the source of vibration may affect the resonant characteristics of

the vocal tract. While the speakers use the same vibrating source of the pharyngoesophageal segment (PES) in alaryngeal speech, they drive different air sources during phonation. In addition to PES vibration variations among subjects, subpseudeoneoglottal pressure and mucosal waves can differ between the two speaker groups¹⁶. In conclusion, different myoelastic-aerodynamic behaviors may finalize with different vocal tract transmission characteristics in alaryngeal speakers. Although the difference of laryngectomized patients from laryngeal speakers is evident, there are confounding results across studies comparing ES and TES. It is not clear if the vocal tract transmission properties vary related to the alaryngeal speech type.

Vowel duration is a parameter in speech acoustics that is taught to be affected by two main parameters in laryngectomized subjects namely air reservoir and PES control¹⁷. Besides, ES is slower than TES. Several studies about the pauses, syllable rates per minute, maximum phonation durations, paragraph duration measurements, confirmed this consensus; which is acceptable because TES users generally use more air^{8,10,17-19}. Concerning vowel duration, Diedrich and Youngstrom²⁰, as well as Robbins and Kempster¹⁷, found the durations shorter in TES than ES in the consonant-vowel-consonant (CVC) context. Duration characteristics of vowels between the two alaryngeal speech groups vary across studies.

In the present study; it was aimed to focus on the research area of vocal tract transmission characteristics regarding Turkish alaryngectomized speakers. Because of the fact that the larynx and subglottal system have been thought to have minor effects on the FF^{21,22} fundamental frequency (F0) was not included in the present study. Therefore, the main purpose of the study was to compare the FF and duration characteristics of Turkish vowels between the esophageal, tracheoesophageal, and laryngeal speakers.

MATERIAL and METHODS

This study approved by the Hacettepe University Clinical Studies Ethic Committee, 20 February 2018, GO 18/207-34. Informed consent was taken from all participants.

Participants were divided into three groups as follows: alaryngeal speakers in the esophageal group (n=5), tracheoesophageal group (n=5) and laryngeal speakers (n=5). Alaryngeal speakers were males whose speech were judged as acceptable by two speech-language therapists (SLTs). Laryngeal speakers were non-smoking speakers with normal hearing status who had no history of any voice and speech disorders confirmed by two SLTs. The participants with different dialects were thought to have different vowel durations were excluded. The minimum duration after laryngectomy was one year and esophageal speakers had taken voice therapy minimum for three months. All the participants in TES group were using Provox™ voice prostheses for at least three months. All of them were using a digital occlusion of the stoma. The age characteristics of participants according to each group has been showed shown in Table 1.

Table 1. Participants' age characteristics according to the group.

Group			
Variable Speakers	Tracheoesophageal Speakers (n=5)	Esophageal Speakers (n=5)	Laryngeal (n=5)
Age (years)			
Mean ± SD	62.4 (6.26)	62.8 (5.35)	53 (2.34)
Median	60	62	52
Range	56-72	56-71	51-56

SD: Standard Deviation.

Speech material consisted of 111 words with each of the eight Turkish vowels included equally. All vowels were in two environments: (1) word-initial preceding a voiceless stop; (2) word-final following a voiceless stop. The database included 5-7 words each with the initial syllable word and

final syllables (appendix 1)²³. This database was chosen as speech sample because when the vowel is preceded or followed by a voiceless stop, it is easier to label reliably in the spectrogram.

The occurrences of the voiceless stops /p/, /t/, and /k/ were equally distributed over the database. All Turkish words studied in the database had word stress on their last syllable. Nine percent of the words were trisyllabic, 5% were monosyllabic, and the remaining were bisyllabic.

Measurement

The recordings were gathered by using the Computerized Speech Laboratory Model 4300B (Kay Elemetrics Corp, Lincoln Park, NJ) equipment. These recordings were taken in a sound-isolated room, the microphone (SHURE SM 48) was placed 15 cm away from the participants' lips and 22050 Hz sampling rate was used. Prior to recording, all the speech tasks were read by the SLTs and taught to participants.

Wideband spectrogram (323Hz) was used for defining vowels by labeling starting and finishing points of vowels from the temporal 20 msn window; linear predictive coding (LPC) and Fast Fourier Transform (FFT) analyses were performed^{24,25}. The signal was multiplied in a Hamming window, and LPC coefficients were computed by using an autocorrelation method^{2,3}. For FF analysis, an anti-aliasing filter with an 11025 Hz cutoff frequency was used before the analog to digital conversion. Filter order number was 20, and preemphasis factor was set to 0.9²⁴. Detection of a formant was based on the visual examination of LPC and FFT spectra. The first two peaks from the LPC spectrum were chosen as F1 and F2²⁶. Because of the fact that studies conducted with similar methodology as used in the present study have reported difficulties in measuring the third formant frequency (F3) and higher FF in alaryngectomized patients, the present study did not include measuring of F3 and higher FF^{2,3}.

The starting and ending points of the vowels were determined by visual inspection of the vertical striations on the spectrogram²⁴. A general knowledge of acoustic phonetics also helped when estimating vowel formants.

Statistical analysis

The IBM SPSS Statistics 20 program (IBM Corporation, Armonk, NY, USA) was used. Nonparametric tests were used in the analyses. Numeric variables were represented by median values and interquartile ranges. Quantitative differences between the three groups were analysed using the Kruskal-Wallis test. Corrected p values have been used to reveal pairwise comparison values. The significance value (corrected p-value) was accepted as 0.05.

RESULTS

Formant Frequency Measurement

Initial syllable

Median and interquartile range of the first and se-

cond formant values of vowels produced by each group are shown in Table 2. Table 3 shows the differences between groups for all acoustic parameters.

The FF patterns gathered in laryngeal speakers for each vowel is compatible with the literature (Table 2). According to Table 3, the FF of /a/, /i/, /o/, /u/ and /y/ vowels were found to be significantly different in alaryngeal groups than laryngeal speakers. Regarding the F1 values, not only the /a/, /i/, /o/, /u/ and /y/ vowels but also F1 values of all the eight vowels were found to be higher in the alaryngeal speech groups as compared to the laryngeal speakers. Similarly, F2 values were found to be higher in the alaryngeal speech groups except for the vowels /y/ and /i/. When comparing two alaryngeal speech groups, it has been seen that the F2 values of /y/, /ε/ in TES group were found to be significantly higher than ES. Also, F1 and F2 values of /α/ vowel and F2 of the /ω/ vowel were significantly higher in ES than TES. Figure 1 shows the FF of all vowels in three groups.

Table 2. Median and Interquartile Range of the First and Second Formant Values (in Hz) of Vowels Produced By Esophageal, Tracheoseophageal and Laryngeal Speakers (Initial Syllable).

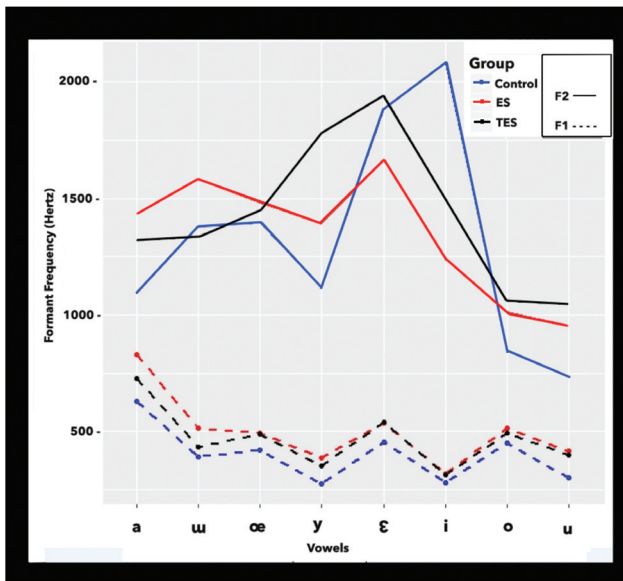
Vowel	Formant Frequency Type	Speaker Group		
		Esophageal Speakers Median±IQR	Tracheoseophageal Speakers Median±IQR	Laryngeal Speakers Median±IQR
/ε/	F1	539 (125)	489 (312)	460 (73)
	F2	1695 (293)	2023 (902)	1896 (268)
/a/	F1	834 (211)	730 (183)	650 (103)
	F2	1452 (198)	1287 (280)	1126 (107)
/o/	F1	509 (97)	491 (139)	458 (43)
	F2	990 (218)	1022 (149)	833 (125)
/æ/	F1	474 (67)	485 (129)	428 (79)
	F2	1444 (415)	1412 (250)	1394 (128)
/u/	F1	407 (94)	405 (125)	323 (140)
	F2	968 (274)	1087 (263)	766 (176)
/y/	F1	364 (108)	362 (99)	275 (83)
	F2	1318 (537)	1767 (187)	1439 (145)
/ω/	F1	529 (120)	429 (153)	410 (115)
	F2	1528 (257)	1306 (254)	1349 (144)
/i/	F1	304 (62)	300 (62)	275 (59)
	F2	1222 (516)	1202 (683)	2133 (190)

IQR= Interquartile Range

Table 3. The differences between groups for all acoustic parameters in the initial syllable.

Vowel	Parameter	Comparable Groups		
		Tracheosephageal-Esophageal Speakers p value	Esophageal Speakers-Laryngeal Speakers p value	Tracheosephageal-Laryngeal Speakers p value
/ε/	F1	0.438	<0.05*	0.13
	F2	<0.05*	<0.05*	1.000
	Duration	0.122	0.483	0.374
/a/	F1	<0.01**	<0.01**	<0.05*
	F2	<0.01**	<0.01**	<0.05*
	Duration	0.592	<0.05*	0.513
/o/	F1	0.306	<0.01**	0.140
	F2	0.431	<0.01**	<0.01**
	Duration	0.328	0.272	0.685
/æ/	F1	1.000	0.051	0.072
	F2	0.589	0.166	0.369
	Duration	0.160	<0.05*	0.494
/u/	F1	1.000	<0.01**	<0.01**
	F2	1.000	0.051	0.072
	Duration	0.412	0.078	0.979
/y/	F1	0.448	<0.01**	<0.01**
	F2	<0.05*	0.712	<0.01**
	Duration	0.445	0.152	0.339
/ω/	F1	0.072	<0.01**	0.522
	F2	<0.01**	<0.05*	1.000
	Duration	0.694	0.276	0.686
/i/	F1	1.000	0.056	0.074
	F2	0.284	<0.01**	<0.01**
	Duration	0.070	0.055	0.070

*refers to corrected p value <0.05, **refers to corrected p value <0.01

**Figure 1. Formant Frequencies of all vowels in three groups in the initial syllable position.**

Final syllable

The median and interquartile range of the first and second formant values of vowels produced by each group have been shown in Table 4. Table 5 shows the differences between groups for all acoustic parameters.

According to the data in Table 4 and Table 5, the FF of /a/, /o/, /y/, /ω/, /u/, /æ/ vowels were found to be significantly different in alaryngeal groups than laryngeal speakers. Additionally, except for the vowels /y/ and /i/, numerical values of F1 and F2 were higher in the alaryngeal speakers as compared to the laryngeal speakers. When comparing two alaryngeal speech groups, F2 values of /ε/ and /y/ in TES group were found to be significantly higher than ES group. In addition, F1 and

Table 4. Median and Interquartile Range of the First and Second Formant Values (in Hz) of Vowels Produced By Esophageal, Tracheosephageal and Laryngeal Speakers (Final Syllable).

Vowel	Formant Frequency Type	Esophageal Speakers Median±IQR	Tracheosephageal Speakers Median±IQR	Laryngeal Speakers Median±IQR
/ε/	F1	693 (121)	637 (146)	548 (58)
	F2	2121 (372)	1931 (274)	1671 (142)
/a/	F1	874 (209)	739 (154)	646 (71)
	F2	1567 (152)	1341 (165)	1235 (123)
/o/	F1	567 (171)	535 (137)	478 (52)
	F2	1042 (238)	1170 (249)	931 (123)
/æ/	F1	584 (122)	529 (127)	460 (67)
	F2	1565 (367)	1486 (442)	1396 (119)
/u/	F1	505 (84)	485 (75)	381 (73)
	F2	1176 (262)	1216 (250)	965 (115)
/y/	F1	494 (123)	505 (129)	395 (119)
	F2	1553 (440)	1590 (272)	155 (203)
/ω/	F1	657 (193)	556 (147)	470 (68)
	F2	1709 (329)	1391 (383)	1375 (126)
/i/	F1	490 (152)	491 (123)	343 (47)
	F2	1968 (923)	2055 (560)	2005 (1106)

IQR= Interquartile Range

Table 5. The differences between groups for all acoustic parameters in the final syllable.

Vowel	Parameter	Comparable Groups		
		Tracheosephageal- Esophageal Speakers p value	Esophageal Speakers- Laryngeal Speakers p value	Tracheosephageal- Laryngeal Speakers p value
/ε/	F1	<0.05*	<0.01**	0.066
	F2	<0.01**	<0.01**	0.479
	Duration	0.071	0.056	0.060
/a/	F1	<0.05*	<0.01**	<0.01**
	F2	<0.01**	<0.01**	<0.01**
	Duration	<0.01**	<0.01**	1.000
/o/	F1	<0.01**	<0.01**	0.081
	F2	<0.01**	<0.01**	<0.01**
	Duration	0.109	0.087	0.137
/æ/	F1	<0.01**	<0.01**	<0.05*
	F2	0.565	<0.01**	0.183
	Duration	0.106	<0.01**	1.000
/u/	F1	<0.01**	<0.01**	0.931
	F2	<0.01**	<0.01**	<0.05*
	Duration	<0.05*	<0.01**	0.885
/y/	F1	<0.01**	<0.01**	1.000
	F2	0.299	0.714	0.152
	Duration	0.061	<0.01**	0.079
/ω/	F1	<0.01**	<0.01**	<0.01**
	F2	0.493	<0.01**	<0.05*
	Duration	0.433	<0.01**	0.060
/i/	F1	<0.01**	<0.01**	1.000
	F2	0.227	0.348	0.155
	Duration	<0.01**	<0.01**	1.000

*refers to corrected p value <0.05, **refers to corrected p value <0.01

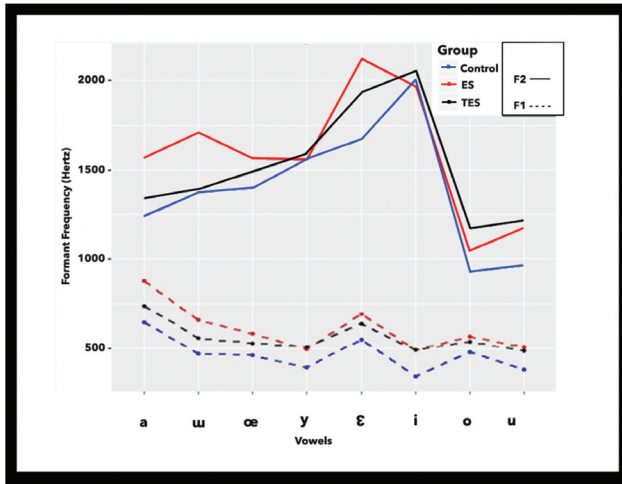


Figure 2. Formant Frequencies of all vowels in three groups in the final syllable position.

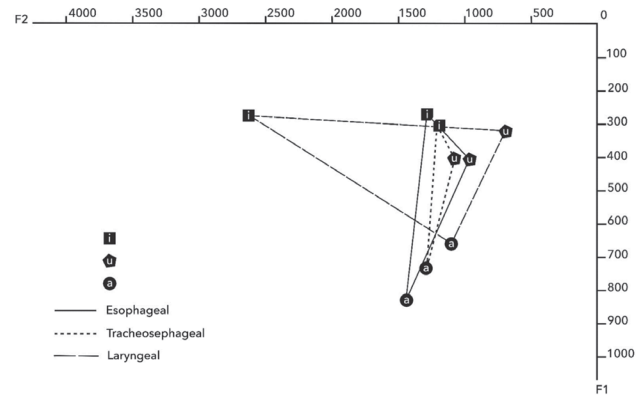


Figure 3. Vowel space including the three corner vowels /u, i, a/ for all the three group.

Table 6. Median and Interquartile range of vowel durations in milliseconds in all groups for both positions.

Vowel	Syllable	Group		
		Esophageal Speakers Median±IQR (msn)	Tracheoesophageal Speakers Median±IQR (msn)	Laryngeal Speakers Median±IQR (msn)
/ε/	Initial	84.7 (.05)	83.2 (.04)	87.0±30
	Final	121.8 (.06)	141.7 (.03)	158.8±90
/a/	Initial	75.2 (.02)	69.5 (.03)	84.7±20
	Final	102.9 (.04)	146.4 (0.4)	160.8±90
/o/	Initial	70.8 (.05)	83.1 (.04)	84.7±20
	Final	126.5 (.06)	132.3 (.09)	154.5±130
/æ/	Initial	63.2 (.03)	74.1 (.03)	84.3±20
	Final	127.0 (.08)	158.7(27)	185.1±130
/u/	Initial	72.5 (.05)	63.2 (16)	57.4 (20)
	Final	103.5 (4)	130.2 (.03)	155.0 (50)
/y/	Initial	72.0 (.03)	63.4 (.03)	62.1 (40)
	Final	105.6 (.05)	135.2 (.04)	147.1(50)
/ω/	Initial	71.6 (.04)	70 (03)	62.7 (30)
	Final	118.7 (.05)	134.9 (.03)	161.4 (70)
/i/	Initial	75.7 (.03)	52.9 (.03)	61.6 (30)
	Final	105.8 (.04)	136.9 (.06)	148.2 (70)

IQR= Interquartile Range

F2 values of /a/ vowel and F2 of the /ω/ vowel were significantly higher in ES than TES. F1 and F2 values of /ε/, /o/, /æ/ and /u/ were significantly higher in ES as compared to the laryngeal speakers. Similarly, the F1 value of the /y/ is significantly higher in ES. Comparing TES to the laryngeal speakers, F1 and F2 of the /ω/ in TES group and F2 values of /o/ and /u/ in TES were found to be higher. Figure 2 shows the FF of all vowels in three groups.

Vowel space including the three corner vowels /u/, /i/ and /a/ has been shown in Figure 3.

According to the data in Figure 3, the alaryngeal groups' vowel space areas have been diminished. The vowel space pattern of TES was more like laryngeal speakers as compared to ES. It was noticed in alaryngeal speakers that /a/ and /u/ vowels have been produced in a higher and fronted position.

Duration

As seen in Table 6 numeric values of durations in both syllable positions are longer for laryngeal speakers relative to alaryngeal speakers. As it can be seen in Table 3, in the initial position, the duration of /a/, /æ/ and /u/ in ES was found to be significantly shorter than in laryngeal speakers. In the final syllable position, the ES group had significantly shorter durations as compared to the other groups. That is, for /a/, /æ/, /u/, /y/, /ω/ and /i/, the durations were significantly shorter than laryngeal, and TES groups for /a/, /u/, /i/ vowels (Table 5).

The duration pattern for all groups is compatible with the pattern in Turkish laryngeal speakers which means that the duration of a vowel sound in the initial syllable is shorter than the final syllable in multisyllabic words²³.

DISCUSSION

In alaryngeal speech, one of the problems that affects speech intelligibility is related to vowels¹¹. In recognition of the vowels, the FF and durations are important acoustic parameters^{5,17}. A vowel's formant pattern can be used to identify a vowel and even to establish relationships between acoustic and perceptual parameters²⁷. Any anatomical structural difference in the vocal tract may cause changing of the FF characteristics²⁷. In general, F1 is thought to be related to the tongue height, and F2 is thought to be related to the tongue advancement and lip rounding²⁷. F1 decreases with increases in tongue height, and back-to-front movements also result in decreased F1. As the constriction moves from back to front, the frequency of F2 increases dramatically. As the lips become more rounded, all formant frequencies decrease^{25,26}. In the earlier studies of voice production in alaryngeal speech, it was thought that the changing of the vibration source is the only responsible mechanism³; however, later studies on vowel formants revealed that the vocal tract transmission properties might also change^{2,5-6}.

In the literature regarding the FF of vowels in alaryngectomized patients, Kyatta et al.²⁴ searched for F1 and F2 values of Finnish ES speakers in CVC context, and Liu et al. searched for the F1, F2, and F3 values of Mandarin ES speakers in isolated vowels⁴. Kazi et al.⁵ documented F1, F2, and F3 values in isolated vowels in English, and Dalatri et al.¹ presented F1 and F2 values in Italian TE speakers in bisyllabic word /papa/. Sisty et al.³ searched formant frequencies in English monosyllabic sequences, and Cervera et al.² worked with 24 Spanish words in consonant-vowel/consonant-vowel (CVCV) concept. In the study of Cervera et al.² while the three groups were different in F1, the ES group was not different from the laryngeal group in F2. Sisty et al.³ found higher formant values in the TES group as compared to laryngeal and ES groups. To conclude, these studies reveal some differences between the ES and the TES groups in resultant vocal tract resonance conflicting with the idea that removing the sound source shouldn't affect filtering characteristics of the vocal tract. In a review of Turkish speech acoustics in alaryngectomized patients, Şahin M. et al.²⁸ included Fundamental Frequency (F0), jitter as a frequency perturbation parameter, shimmer as an amplitude perturbation parameter and noise-to-harmonic ratio as a spectral parameter. They found mean F0 to equal 72.9 Hz in esophageal speakers. In the same study, maximum phonation time, counting and the number of syllables per minute were taken into account in esophageal speakers.

In another study concerning Turkish alaryngectomized speakers who use electrolarynx, the aim was to develop speech coding algorithms which would reduce noise of the device²⁹. The results of the present study show that the general formant frequency pattern in alaryngeal speakers is compatible with the literature except for the /y/ and /i/ vowels. Therefore, the F1 and F2 values of alaryngeal speakers were higher than laryngeal speakers. These findings support the idea that shorter vocal tracts are one of the factors that change

resonant cavity transmission characteristics^{2,3,30}. According to Figure 1, it is seen that /a/, /ε/, /u/ vowels have a great difference in patterns of alaryngeal speech as compared to laryngeal speech in F1 and F2. However, /i/ has not been found higher in one of the alaryngeal groups. Similarly, in a recent Polish study, F2 of the /i/ vowel was not found higher in TES than laryngeal speakers³¹. In the present study, when both the vowel space and individual vowel FF patterns were taken into account, results indicated that vowels have been in a more front and high position except for the vowels /i/ and /y/. This finding supports the idea that “alaryngeal participants have a fronted and higher tongue position during articulation” as reported in the literature^{2-3,5}. In Turkish, /i/ and /y/ are front-high vowels, and there is not any systematic reason thought for this difference. However, especially for the vowel /i/, it is found more difficult to detect formant frequency in the FFT as the peak of the F2 values were not as clear for the /i/ vowel as for some of the other vowels. In addition to this impression, high interquartile range (IQR) values for these two vowels supported this finding. One of the limitations of the present study arises at this point. Although we have many word samples for each participant, and we have investigated all vowels in Turkish both in the initial and final positions, we had a limited number of participants. Since the FF measurement is directly related to the quality of the speech being analysed, the FF analysis is harder in alaryngeal speakers²⁴. In Kazi et al.’s⁵ study, it was reported that the FF were variably assessable for the alaryngeal speakers. In addition, when the many possible variabilities, such as chemoradiotherapy and postoperative complications are taken into account, the results of the present study about the exceptions of vowel /i/ and /y/ should be carefully commented and need to be investigated in future studies conducted with a higher number of participants. The higher IQR values in ES group, the difficulties confronted in the FF analysis, and different vowel space patterns indicate that ES had more variables. Similarly, in the literature,

it has been found that good ES and moderate ES are found to have different formant characteristics as compared to the laryngeal speakers⁶. In future studies, more participants should be included and also stricter inclusion criteria according to the audio-perceptual evaluation should be applied especially for esophageal speakers. According to the audio-perceptual evaluations, it is recommended that the participants be separated into the bad, moderate, and good esophageal speakers so that we may provide more realistic and reliable information for ES.

Duration is also an important acoustic parameter in alaryngeal speech acoustics because significant correlations were found between sentence duration, syllable duration characteristics, speech acceptability and vowel duration.

In the literature, early studies reported shorter durations in TES than ES in the CVC context^{16,19}. They related this finding with the more controlled pulmonary air source used in TES. It is thought that limited control of the cricopharyngeal segment may be responsible for the difficulty in initiating and terminating voicing in the esophageal speakers and resulting in longer duration¹⁶. In a Hebrew study, moderate esophageal speakers were found to have longer vowel durations as compared to the TES and laryngeal speakers⁶. In that study, researchers measured ten vowels in initial syllables of short and long sentences. In contrast, in a Spanish study, TES had the longest duration as compared to ES and laryngeal speakers in the CVCV segments where total word duration was shorter in TES². The researchers concluded that because of the small air reservoir ES produced, a reduced second syllable of the vowel duration may be responsible for the shorter duration in ES².

In the present study, it was observed that the median durations of nonhigh vowels /a/, /ε/, /o/, /æ/ were longer than the median durations of high vowels /ω/, /i/, /u/, /y/ in Turkish in

all the three groups. As it is same for laryngeal speakers, both alaryngeal speakers showed the same duration pattern; that is nonhigh vowels had longer durations and vowels in the final position comparing to the initial position²³. However, ES had shorter durations in both initial and final positions than laryngeal speakers. Also, the durations of some vowels of ES were different from TES in the final position. These findings indicated that ES had deletions of the syllables more prominently in the final position. As we also noted, Cervere et al. reported that visual inspection of the spectrogram showed the deletion of the end syllable². Like Cervera et al.'s study², most of the speech samples were multisyllable in the present study. Turkish is known to have a simple stress assignment rule that places primary stress on the final syllable of a word irrespective of the length of the word and weight of the syllables²⁹. The differences in duration between TES and ES at the end of the word indicated the responsible mechanism for this is the aerodynamic differences between these two groups. It is thought that longer durations detected in ES in English can be related to the differences in the language characteristics^{17,20}. For example, English has a tense/lax distinction that associates with longer durations in tense vowels³². Additionally, it is thought that the variability of ES quality between the different studies may have caused different results as was the case with FF.

CONCLUSION

The results of this study generally supported the earlier findings of a shorter vocal tract and a more frontal and higher tongue position in Turkish alaryngeal speakers. The minor differences between other studies are thought to be due to the consonantal context and possible co-articulatory effects. Acoustic findings obtained in the present study indicated that TES was more like laryngeal speakers as compared to ES. However, further studies including higher number of participants are needed to verify the findings of this study.

REFERENCES

1. D'Alatri L, Bussu F, Scarano E, Paludetti G, Marchese MR. Objective and subjective assessment of tracheoesophageal prosthesis voice outcome. *J Voice*. 2012;26:607-13. [\[CrossRef\]](#)
2. Cervera T, Miralles JL, González-Álvarez J. Acoustical analysis of Spanish vowels produced by laryngectomized subjects. *J Speech Lang Hear Res*. 2001;44:988-96. [\[CrossRef\]](#)
3. Sisty NL, Weinberg B. Formant frequency characteristics of esophageal speech. *J Speech Hear Res*. 1972;15:439-48. [\[CrossRef\]](#)
4. Liu H, Ng ML. Formant characteristics of vowels produced by Mandarin esophageal speakers. *J Voice*. 2009;23:255-60. [\[CrossRef\]](#)
5. Kazi RA, Prasad VM, Kanagalingam J, Nutting CM, Clarke P, Rhys-Evans P, et al. Assessment of the Formant Frequencies in Normal and Laryngectomized Individuals Using Linear Predictive Coding. *J Voice*. 2007;21:661-8. [\[CrossRef\]](#)
6. Most T, Tobin Y, Mimran RC. Acoustic and Perceptual Characteristics of Esophageal and Tracheoesophageal Speech Production. *J Commun Disord*. 2000;33:165-81. [\[CrossRef\]](#)
7. Van As CJ, Hilgers FJ, Verdonck-de Leeuw IM, Koopmans-van Beinum FJ. Acoustical analysis and perceptual evaluation of tracheoesophageal prosthetic voice. *J Voice*. 1998;12:239-48. [\[CrossRef\]](#)
8. Sedory SE, Hamlet SL, Connor NP. Comparison of Perceptual and Acoustic Characteristics of Tracheoesophageal And Excellent Esophageal Speech. *J Speech Hear Disord*. 1989;54:209-14. [\[CrossRef\]](#)
9. Baggs TW, Pine SJ. Acoustic Characteristics: Tracheoesophageal Speech. *J Commun Disord*. 1983;16:299-307. [\[CrossRef\]](#)
10. Štajner-katušić S, Horga D, Mušura M, Globlek, D. Voice and speech after laryngectomy. *Clin Linguist Phoniatr*. 2006;20:195-203. [\[CrossRef\]](#)
11. Jongmans P, Wempe TG, van Tinteren H, Hilgers FJ, Pols LC, van As-Brooks CJ. Acoustic analysis of the voiced-voiceless distinction in dutch tracheoesophageal speech. *J Speech Lang Hear Res*. 2010;53:284-97. [\[CrossRef\]](#)
12. Jongmans P. The intelligibility of tracheoesophageal speech: An analytic and rehabilitation study/dissertation/. Amsterdam: University of Amsterdam; 2008.
13. Lieberman P, Blumstein SE. Speech physiology, speech perception, and acoustic phonetics. 1th ed. New York: Cambridge University Press; 1988. [\[CrossRef\]](#)
14. Story BH, Laukkanen AM, Titze IR. Acoustic Impedance of an Artificially Lengthened and Constricted Vocal Tract. *J Voice*. 2000;14:455-69. [\[CrossRef\]](#)
15. Titze IR, Laukkanen AM. Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study. *Logop Phoniater Vocol*. 2007;32:147-56. [\[CrossRef\]](#)
16. Moon JB, Weinberg B. Aerodynamic and myoelastic contributions to tracheoesophageal voice production. *J Speech Lang Hear Res*. 1987;30:387-95. [\[CrossRef\]](#)
17. Robbins J, Christensen J, Kempster G. Characteristics of speech production after tracheoesophageal puncture: voice onset time and vowel duration. *J Speech Lang Hear Res*. 1986;29:499-504. [\[CrossRef\]](#)

18. Debruyne F, Delaere P, Wouters J, Uwents P. Acoustic analysis of tracheo-oesophageal versus oesophageal speech. *J Laryngol Otol.* 1994;108:325-28. [\[CrossRef\]](#)
19. Globlek D, Stajner-Katusic S, Musura M, Horga D, Liker M. Comparison of alaryngeal voice and speech. *Logoped Phoniatr Vocol.* 2004;29:87-91. [\[CrossRef\]](#)
20. Diedrich WM, Youngstrom KA. Alaryngeal speech. 1th ed. Springfield: Charles C. Thomas Publisher; 1966.
21. Fant G. Acoustic Theory of Speech Production. 1th ed. The Hague, Paris; Mouton & Co.: 1970.
22. Flanagan JL. Speech analysis synthesis and perception. 2nd ed. New York: Springer-Verlag; 1972. [\[CrossRef\]](#)
23. Arısoy E, Arslan LM, Demiralp MN, et al. Duration of Turkish Vowels Revisited. In: 12th International Conference on Turkish Linguistics (ICTL 2004), Dokuz Eylül Üniversitesi, İzmir, Türkiye. 2004; 11-3.
24. Kent RD, Read C. The acoustic analysis of speech. 2nd ed. San Diego: Singular Publishing Group; 2002.
25. Hixon TJ, Weismer G, Hoit JD. Preclinical speech science. 1th ed. San Diego: Plural; 2008.
26. Fant G. Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations. 2nd ed. Paris: Walter de Gruyter; 1970. [\[CrossRef\]](#)
27. Ball MJ. Phonetics for Speech Pathology. 1th ed. London: Whurr Publishers Ltd; 1993.
28. Şahin M, Ögüt MF, Vardar R, Kirazlı T, Engin EZ, Bor S. Novel esophageal speech therapy method in total laryngectomized patients: biofeedback by intraesophageal impedance. *Dis Esophagus.* 2016;29:41-7. [\[CrossRef\]](#)
29. Akdeniz R, Oktay, MO. Comparison for speech coding algorithms for total laryngectomies. *Matter: Int J Sci Technol.* 2017;3(2). [\[CrossRef\]](#)
30. Kytta J. Finnish oesophageal speech after laryngectomy: sound spectro-graphic and cineradiographic studies. *Acta Otolaryngol.* 1964;195(Suppl 1):1-94. Available from: <https://europepmc.org/abstract/med/14267561>.
31. Mięsikowska M. Analysis of Polish Vowels of Tracheoesophageal Speakers. *J Voice.* 2017;31:263e5-11. [\[CrossRef\]](#)
32. Yavuz H, Balcı A. Turkish Vowels. In: Balpınar Z, ed. Turkish phonology and morphology. Eskişehir: TC Anadolu Üniversitesi Publisher; 2011: p. 32-45.