İSTANBUL UNIVERSITY PRESS

The Turkish Journal of Ear Nose and Throat,
Tr-ENT 2025, 35 (2): 71–79

https://doi.org/10.26650/Tr-ENT.2025.1630333

Submitted: 06.02.2025
Revision Requested: 28.02.2025
Last Revision Received: 24.03.2025
Accepted: 16.04.2025
Published Online 14.05.2025

The Turkish Journal of Ear Nose and Throat

Research Article

Open Access

The Effect of Sex on the Turkish Voice Spectrum

Cemil Yılmaz [1], Erhan, Demirhan [2], Şerife Nur Biçen [3]

[1] Kütahya Health Sciences University, Faculty of Health Sciences, Department of Language and Speech Therapy, Kütahya, Turkiye
[2] Private Practice, Otorhinolaryngology, Head and Neck Surgery, İzmir, Turkiye
[3] Anadolu University, Graduate Schools and Institutes, Department of Speech and Language Therapy, Eskişehir, Turkiye

Abstract

Objective: The Long-Term Average Spectrum (LTAS) of voice is a functional method used to discriminate various voice qualities. In this study, the voice quality characteristics of Turkish female and male speakers were examined and compared through voice spectra created with LTAS.

Material and Methods: 30 male and 30 female speakers with typical voices, aged 18 to 50, were required to read a text at a habitual pitch and loudness. The recorded speech samples were used to create voice spectra containing 75 bands with a bandwidth of 160 Hz and covering the range from 0 to 12 kHz. The spectral tilt (ST) of the voice spectra was also compared between the sexes.

Results: Multiple comparisons revealed that there were significant intensity differences between sexes at the band regions of 0–160 Hz, 160–320 Hz, 640–960 Hz, 2880–3040, 4000–5120 Hz where females had higher intensities than males except for the band of 0–160 Hz. STs of females were found to be significantly lower than those of males.

Conclusion: All significant band differences between the sexes were below 6 kHz, and both sexes had a rather flat-spectrum above 6 kHz. Females had significantly higher intensities around the region of 2880–3040 Hz, because of aspiration noise caused by the possible posterior glottal gap during phonation. Females had shallower STs than males, indicating that females had greater hyperfunctional voice quality. The voice spectrum characteristics of Turkish speakers were compared with those of other studies and found more similarities with Spanish than Persian.

Keywords Voice · Spectrum Analysis · Voice Quality · Sex

Citation: Yılmaz C, Demirhan E, Biçen ŞN. The Effect of Sex on the Turkish Voice Spectrum The Turkish Journal of Ear Nose and Throat 2025;35(2):71-79. https://doi.org/10.26650/Tr-ENT.2025.1630333

Corresponding author: Cemil Yılmaz cemil.yilmaz@ksbu.edu.tr

INTRODUCTION

Acoustic analyses of voice have been used to understand voice quality variations, including normal and abnormal voices. The voice spectrum or measures of spectral shape is one of the acoustic measures of the vocal function (1). The voice spectrum was shown to reflect the characteristics of the voice that are specific to the sex of the speaker (2, 3). The anatomical and physiological differences between sexes create these characteristic features, such as the longer vocal fold length of males, which is the primary factor for the lower mean fundamental frequencies and the higher source strength than females (4, 5). However, the smaller larynx size, shorter vocal fold length and thinner vocal fold thickness result in the female voice radiating at higher frequencies (4). The voice spectra of females are generally found to contain characteristic features of more breathy phonation than those of males, although it should also be noted that there may be individual differences (2). Breathy voice spectra include a strong fundamental component, an aspiration noise component that replaces the energy of the harmonic components at high frequencies (around third formant region) and wider bandwidths for the first formant region (6). It is assumed that one of the anatomical or physiological causes of breathiness is the higher incidence of the posterior glottal chink - meaning a posterior gap during phonation, and it is observed at a higher rate in females than in males (2, 6).

LTAS, which is an acronym for Long Term Average Spectrum, is an acoustic analysis method and is frequently used to analyse voice source spectra in several languages (7-11). LTAS is based on constructing an average spectrum by averaging short-term spectra taken repeatedly from a long-time speech sample. When the duration of the speech sample is long enough (about 30–40 seconds), the direct and instant effect of the vocal tract on the voice source spectrum is eliminated (7, 12).

Löfqvist and Mandersson claimed that unvoiced speech segments could distort the voice spectrum (12). Based on this proposal, various studies have filtered out the unvoiced speech segments in their LTAS analysis for voice spectrum research (11, 13-18), while other studies have kept unvoiced speech segments in their LTAS analyses, but this time for speech spectrum research (19-23). If the purpose of a study is voice spectrum analysis, then the unvoiced speech segments such as noises of fricatives and affricatives and bursts of plosives should be removed from the recordings. Löfqvist and Mandersson stated that the duration of the only voiced portion of the speech signal is roughly half of the actual speech signal including pauses and unvoiced segments (12). Löfqvist and Mandersson presented only voiced portions of the recordings that were around 20 seconds in duration and halved the analysis time and found virtually no effect. However, when the duration of the analysis was one-third of the original signal, the variations occurred between the measurements.

Pittam calls on researchers to consider the LTAS on a “macroacoustic” level, as spectral movements are represented on a rather gross scale (24). The spectrum obtained because of LTAS shows a range of sound energy alterations from low frequencies to high frequencies. The spectrum shape is usually defined by spectral moments such as the spectral mean (centre of gravity) showing the amplitude of the centre of the spectrum, the spectral standard deviation showing the degree of deviation of the frequencies from the spectral mean, the spectral skewness showing the degree of the sound energy below the spectral mean, and the spectral kurtosis showing the asymmetry of the sound energy distribution around the spectral mean. It has been observed that the spectra created because of the LTAS procedure contain descriptive information about the voice source qualities (7). Tanner et al. researched various spectral moments of LTAS and found that the spectral mean and spectral standard deviation of the average spectra were significantly affected by the voice quality changes due to voice therapy (25). Lowell et al. showed moderate to strong correlations between the spectral mean, skewness and kurtosis of LTAS and overall voice severities (13). These studies, along with other studies, show that the LTAS has spectral features that are sensitive to the changes in voice qualities (8, 26). In addition to the spectral properties, various parameters were derived from the LTAS to analyse different voice qualities (12, 27). The alpha ratio, measured from the ratio of the energies of the LTAS between 0–1 kHz to 1–5 kHz, has been found to be associated with hyper-functional or hypo-functional voice source characteristics (12). The term alpha ratio can also be defined as the level difference between the bands of 0–1 kHz to 1–5 kHz, which refers to the spectral tilt (14, 15). A high spectral tilt or low alpha ratio indicates that the energy at the low-frequency band of 0–1 kHz is relatively greater than that at high frequencies, and this is correlated with the spectral features of a hypofunctional voice source.

Mendoza et al. analysed the voice quality differences between males and females using the method of LTAS for Spanish (9). The intensities of the female spectra were found to be significantly higher at the frequency bands of 640–960 Hz, 2720–3040 Hz and 4000–4960 Hz than the male spectra. These significant differences also resulted in a lower spectral tilt (shallower spectrum) for the spectra of females, indicating a more hyperfunctional voice source (9). It was stated in the study of Mendoza et al. that the significant difference between sexes at the frequency band of 2720–3040 Hz was the result of aspiration noise, which was shown in the study of Klatt and Klatt that aspiration noise could be around the third formant region (2, 9). Moradi et al. performed a similar research for Persian and found that the intensities of the female voice spectra were significantly higher than those of the males at the bands of 800–960 Hz, 3200–3840 Hz and 5760–5920 Hz, while these significant differences were interpreted as being a more breathy voice source for females (28). Smith and Goberman analysed the differences in voice quality between individuals with Parkinson's disease and controls with LTAS and found that females had lower ST values than males within both groups (16). Leino et al. compared the voice spectrum of Finnish university students with the LTAS (29). The intensities of the males' spectra were found to be significantly higher than those of the females at 3–4 kHz, while the intensities of the females' spectra were significantly higher than those of the males at 4–5 kHz, where the greatest differences were seen. Leino et al. divided the voices of students into non-pathological breathy and pressed phonation qualities. Although there were significant differences in spectral comparisons below 5 kHz, there was no significant difference at 3–4 kHz between the females having breathy and pressed phonation qualities (29). While there was a significant difference between the male voices labelled as breathy and pressed at 3–4 kHz, the absence of this significant difference between female voices labelled as breathy and pressed was not explained in the study of Leino et al. (29). However, it is suggested that this frequency band of 3–4 kHz is related to the aspiration noise caused by a breathy phonation, and it can be seen that the effect of aspiration noise around 3–4 kHz does not depend on whether the phonation is more breathy or more pressed for females (2).

Because of the sex effect on the voice spectrum, various other factors such as age, loudness and language have had to be controlled in relation to the factor of sex (10, 17, 30). Linville showed that the spectra of elderly females had significantly higher intensities around 6–7 kHz than those of young females (10). Nordenberg and Sundberg showed that an increase in the loudness of the voice resulted in greater gains for females than for males (30). Ng et al. researched the voice spectrum of Cantonese-English bilingual speakers and showed that females have been found to have lower mean spectral tilts than males, and this difference was greater when using Cantonese (17). LTAS analysis of voice quality has shown that the differences between females and males were affected by the language spoken (18, 31).

In this study, we aimed to analyse the effect of sex on the voice spectrum of Turkish speakers using the LTAS method. The findings of this study were compared with the findings of other studies regarding the female and male voice spectra of various languages. The results of this study could make a valuable contribution to the knowledge on the Turkish voice spectrum.

MATERIAL AND METHODS

Subjects

Monolingual Turkish speakers between the ages of 18 to 50 without any speech and language disorders were included in this study. Subjects underwent an examination of anterior rhinoscopy, oropharynx and otoscopy by an otorhinolaryngologist. The inclusion criteria included being a non-smoker, being healthy on the day of the recording, not having a neurological disorder or pulmonary disease, and not taking medication for gastroesophageal reflux disease. In addition, the criteria of not being a professional voice user, not having received training on voice, and having no dialectal differences were also applied to the subjects. A total of 66 subjects were recruited from the population of Izmir, Turkey, and 6 subjects were excluded due to high age, bilingualism, or speech impairment. The speech recordings of the remaining 60 subjects (30 male with a mean age of 36.1; 30 female with a mean age of 37.2) were analysed in the study. One subject included in the study had an upper respiratory tract infection within 2 weeks but was examined to be healthy on the day of the recording. One subject underwent septal surgery for nasal septal deviation and another subject underwent concha reduction surgery, and these two subjects were included in the study. The age of the subjects was analysed for an effect on the voice spectrum at the limitations part of the study and no significant effect was noted. All participants were required to read and sign the research participant consent form, and this study was approved by the ethical comity of Health Sciences University Izmir Tepecik Training and Research Hospital (Date: 07.03.2017, No: 13).

Voice recordings

Voice recordings were made with a Samson Go Portable USB condenser microphone and a MacBook Pro laptop in the silent audiometric booth. The cardioid microphone had a flat frequency response between 200 Hz and 12 kHz within ±2 dB. The mouth-to-microphone distance was set to 20 centimetres (between the 4 to 18 inches advised range for the microphone) and the azimuth angle was set to 45 degrees. Voice recordings without any popping sound were sampled at 44.100 Hz and a 16-bit rate. Since there was still no phonetically balanced text in Turkish at the time of the research, another text in which all speech sounds were distributed evenly was used.

The subjects were asked to read the text containing 149 words at a habitual pitch and loudness. The reading duration of the text was approximately 70 seconds.

Acoustic analysis

Speech recordings were high-pass filtered at 50 Hz (with a 10 Hz smoothing factor beginning from 40 Hz) to eliminate any possible low-frequency noise components (32). A 40-second long speech segment was cut out from the middle of the recordings. Breathing-in noises and any silent regions were cut out from the recordings because these parts did not provide direct information about phonation. After this edit, the mean duration was approximately 32 seconds. This duration was sufficient for reliable analysis because measurements of speech spectra became stable over speech of approximately 30 seconds or longer (33).

This study focused on the voice spectrum analysis so that the unvoiced speech segments such as noises of fricatives and affricatives and bursts of plosives were removed from the recordings. Voiced fricatives and affricatives were also removed from the recordings because these speech sounds contain the aperiodic sounds produced at the articulators as well as the periodic sound of phonation, and this noise can distort the acoustic information on phonation. This procedure was performed manually by two acoustic analysis experts by searching for the aperiodic waveforms in the recordings, cross-checking from spectrograms, listening to each speech segment before filtering out and cutting the segments at zero crossings without causing any artefacts. The LTAS produced from the recordings containing only the voiced speech segments and not containing articulator-produced noise were labelled as "LTAS". After the final editing of the recordings, the mean duration of only the voiced segments was 22 seconds. This duration was consistent with the duration values given in the procedure of Löfqvist and Mandersson, who stated that the duration of the only voiced portion of the speech signal was roughly half of the actual speech signal, including pauses and unvoiced segments (12).

LTAS between 0 and 12 kHz with bands having a 160 Hz bandwidth was generated from the spectra obtained through Fast Fourier Transformation (FFT) of the recordings. The bandwidth of 160 Hz has been used in various research so this bandwidth was used in this study as it would be more effective in comparing the data with other studies (9-11, 28). The highest frequency was accepted as 12 kHz because it has been shown that high frequencies have a functional role in the perception of the quality of voice (34).

The effect of intensity on LTAS was researched and it was found that this effect has variations by sexes (30). In this study, the recordings were performed by subjects using the habitual loudness and pitch of their daily voice, and thus voice spectra representing their habitual voices were obtained. However, the subjects used different vocal intensities while recording their speech. In this case, the effect of sound intensity must be normalised in order to reliably compare the voice spectra. The LTAS were normalised by placing the interval with the highest energy at 0 dB and the other intervals at negative values accordingly (35). After normalisation, mean intensities of the 75 bands measured for males and females were represented graphically and all LTAS bands were statistically compared between sexes. This normalisation procedure has been used in various research (8, 12, 28), and the study of Leino stated that speakers can be compared to each other with this normalisation regardless of the absolute Leq (equivalent sound level) values of the samples (8). All the acoustic analyses were done with Praat software (36).

Measurements

LTAS consisted of 75 bands with a bandwidth of 160 Hz. In order to measure ST with an exact cutoff frequency of 1 kHz, which indicates the energy difference between 0–1 and 1–5 kHz, high-pass filtered spectra from FFTs were also converted to 1-to-1 LTAS without bands. This parameter was measured from all the 1-to-1 LTAS created.

Reliability for the voices of the subjects

Auditory perceptual evaluation of the subjects' voices was performed by a speech-language pathologist and an otorhinolaryngologist using GRBAS, an acronym for Grade, Roughness, Breathiness, Asthenia, and Strain, and none of the voices were rated 1 or higher (37).

The voices of the subjects were also evaluated using an objective method based on the acoustics analysis of continuous speech. The smoothed cepstral peak prominence (CPPS) is an acoustic marker for the periodicity of an acoustic signal and is used to measure voice quality (38). In a meta-analysis for the acoustic measurement of voice quality, CPPS was shown to be the only acoustic metric that provided sufficient validity in both sustained vowel and continuous speech (39). The recorded first sentences of the text of all subjects were analysed with Praat software according to the protocol of Maryn and Weenink (38). The mean CPPS in this study was measured to be 10.62 dB (standard deviation of 1.31 dB) for females and 10.07 dB (standard deviation of 1.42 dB) for males. These values were found to be in accordance with the normative CPPS values in the study of Buckley et al., which used the same protocol for CPPS measurement (40). The CPPS values of females in this study were also found to be similar to the CPPS values of healthy female teachers in the study of Phadke et al., which used the same measurement protocol for CPPS (41). Auditory perceptual and objective acoustic analyses both confirmed that the subjects in this study had typical voices.

Reliability for the editing of recordings

Filtering out the unvoiced speech segments from the recordings for the analysis of LTAS was a critical editing step. Manual editing of 20% of the 60 voice samples was re-done by the same author, and the durations of the edited voice samples were used for reliability analysis as in the study of Goberman and Robb (42). Intra-rater reliability was checked with Pearson correlation and was found to be 0.957 (p<0.01), while the ratio of mean duration difference between the edited and re-edited samples was found to be 2.9%. The same 12 voice samples were edited by the third author and the inter-rater reliability was found to be 0.983 (p<0.01), while the ratio of the mean duration difference between the edited and re-edited samples was found to be %3.7. Both correlations were significant at the 2-tailed 0.01 level.

Since the durations of the edited voice samples might not be sufficient for reliability, the intensities of the LTAS bands were also analysed for intra- and inter-rater reliability. The intensities of the LTAS bands were checked with Pearson correlation, and the product-moment correlation was found to be between 1.000 and 0.999 for intra-rating and inter-rating (p<0.01). The mean band intensity difference was found to be 0.181 dB for the intra-rating and 0.308 dB for the inter-rating. These analyses showed that the manual editing of recordings was effective and had high reliability.

Statistical analysis

The effect of sex on the LTAS band intensities was analysed with repeated measures ANOVA. All statistical analyses were performed with IBM SPSS Statistics for Windows, Version 24.0 (IBM Corp., Armonk, NY, USA).

The intensities of the 75 bands in the LTAS of 60 participants were compared between sexes. A strict Bonferroni correction (p<0.00067 was accepted for these repeated comparisons while p was noted as appropriate elsewhere) was used to eliminate false positives because of multiple tests. These comparisons were made with the independent samples t-test, but when even one of the group variables was not normally distributed, Mann–Whitney U test was used.

RESULTS

Because of the full factorial repeated measures ANOVA, a significant interaction effect was found between the LTAS bands and the sex of the speaker (F(74,4292)=8.115, p=0.000). This indicates that the LTAS band intensities of different frequencies differed in females and males, and these bands were explored with multiple comparisons. However, the overall effect of sex on the LTAS bands from 0 to 12 kHz was found to be non-significant (F(1,58)=2.288, p=0.136).

Table 1. The descriptive statistics of ST for LTAS and statistical comparison result by sex

Parameter	Sex	Mean (dB)	Median (dB)	SD (dB)	Test	p value
ST	Female	−17.3	−18.1	3.0	U=312	**0.041***
	Male	−19.9	−19.6	4.4		

ST: Spectral Tilt, LTAS: Long-Term Average Spectrum, dB: Decibel, SD: Standard Deviation, U: Mann–Whitney U Test

The multiple band comparison analysis showed that there were significant intensity differences between female and male voices at the LTAS band regions of 0–160 Hz, 160–320 Hz, 640–960 Hz, 2880–3040, and 4000–5120 Hz (Figure 1). Female voice intensities were higher than males' at all those significant bands except for the 0–160 Hz where male voice intensity was found to be higher. It was found that the STs of female LTAS were significantly lower than those of males, meaning that the STs of females were shallower (Table 1).

The mean speaking fundamental frequency was measured from the recordings for both sexes and was found to be 129.3 Hz for males and 218.7 Hz for females.

DISCUSSION

Analysing voice source characteristics of Turkish female and male speakers at a "macro acoustic" level was possible with LTAS, and at this level the frequency region of 6–12 kHz was clearly separated from the region of 0–6 kHz because 6–12 kHz had a rather flat spectrum (24). Sergeant and Welch suggested that the maturation effect may have played a role in the formation of this difference, and according to this maturation effect, the intensities of the frequencies below 5.75 kHz increased relative to the upper frequencies through the maturation of children from 4 to 11 years old (43, 44). Although this study found that the effect of sex factor on voice spectrums was not significant, all significant band differences between sexes were found to be below 6 kHz. The finding that the band intensities of females and males were more similar above around 6–12 kHz, may have caused the sex factor to be non-significant.

It was only at the band of 0–160 Hz that the LTAS band intensity of male voices was found to be significantly higher than that of females. This band region was dominated by the fundamental frequency of males (mean 129.3 Hz), and this situation was

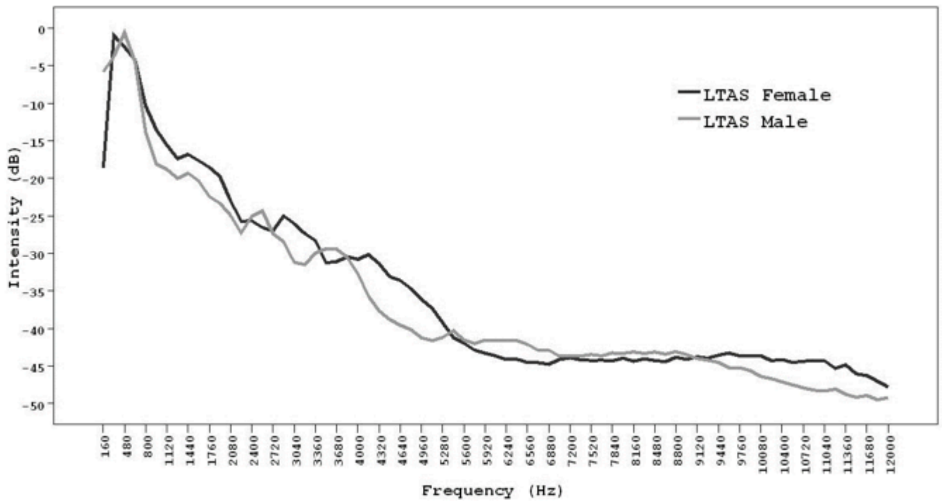

Figure 1. LTAS graphs showing the average intensity values of the female and male groups
LTAS: Long-Term Average Spectrum, dB: Decibel, Hz: Hertz

reversed at the band of 160–320 Hz because the fundamental frequency of females (mean 218.7 Hz) became dominant in this band. The mean intensity values of female LTAS were significantly higher than that of males at the band regions of 640–960 Hz, 2880–3040 and 4000–5120 Hz. As a result of these significant differences, the STs of the females were significantly lower than those of the males, which meant that the voice spectrum of the females had higher intensities in the frequency region of 1–5 kHz relative to 0–1 kHz. A shallower ST was associated with a more hyperfunctional voice source characteristic, and it was physiologically explained with a faster closing time of vocal folds in relation to the open time (12, 14, 15). As a result, the females in this study were found to have a more hyperfunctional voice source characteristic than the males. A similar result was found by Mendoza et al. for Spanish, by Smith and Goberman for American English, and by Ng et al. for Cantonese that females had significantly lower STs than males (9, 16, 17).

In this study, it was observed that females had significantly higher intensity than males around the region of 3 kHz (third formant region), as in the study of Mendoza et al. (9). Mendoza et al. reported that the higher intensity for females around the third formant region may be related to the aspiration noise found in breathy voice quality, as shown in the study of Klatt and Klatt, and this relationship was related to the presence of a posterior glottal gap during phonation (2, 9). Thus, Turkish females having an aspiration noise around the third formant region were associated with a more breathy voice, possibly resulting from a posterior glottal gap. Therefore, it could be argued that the effects of breathy voice quality differed between females and males in certain frequency regions. However, caution should always be taken when analysing LTAS measurements, as LTAS does not differentiate between the harmonic and inharmonic components in the voice source (12).

When the LTAS of males were separately analysed, a common local peak was found at around 3.5 kHz, unlike females. It has been observed that this peak seen in the males was in good agreement with one of the typical voice features of males so called “the actor’s formant” or so called “speaker’s formant” (8, 46-48) found in the frequency region of 3–4 kHz. In this study, the peak at 3.5 kHz could be expressed as one of the distinctive voice spectrum characteristics of male speakers, different from female speakers.

There were differences between females and males in terms of LTAS between the frequency region of 6–12 kHz, although not significantly. As stated by Shoji et al., the harmonic component in the high-frequency region was lowered and replaced by the aspiration noise produced at the glottis, and thus the energy decrease level at high frequencies was not similar to the dampening of the intensities in the low formant-frequency regions (49). Consistent with Shoji et al., in this study both LTAS of females and males presented a rather flat-spectrum around 6 kHz and above (49). While it was found that sex-specific characteristics in the voice were seen even above 10 kHz (49), Monson et al. found that male speech had higher intensities between around 5–8 kHz, but then voice intensities of females surpassed males’ after 8 kHz and this difference became significant above 11 kHz (32). Similar to the findings of Monson et al., in this study the glottal aspiration noise component in the voice spectrum above around 6 kHz increased for both sexes, and the mean glottal aspiration noise components of Turkish female voices were generally higher than those of males above 9 kHz (32).

It has been observed that the differences between the voice spectra of females and males vary according to the language. While it was found in Spanish that the intensities of the female LTAS bands were significantly higher than those of the males between 4 and 5 kHz, a similar outcome of this study was also valid for Turkish (9). However, significant sex differences for Persian were not found between 4 and 5 kHz but were found between 3 and 4 kHz where the intensity of female LTAS bands were significantly higher than that of males' (28). In addition, frequency regions below 3 kHz where females had significantly higher mean intensity values than males were found for both Spanish and Turkish, and the regions 640–960 Hz and 2880–3040 Hz were common to both languages. However, the opposite of this situation was observed for Persian and it was seen that the mean LTAS band intensity values of males were higher than those of females below 3 kHz. Although the spectral tilt was not measured for Persian, it was found that the males were significantly higher than the females in the spectral tilt values measured for the languages of Turkish and Spanish (9, 28). In this respect, the sex effect on the voice spectrum of Turkish was similar to its effect on Spanish rather than its effect on Persian. Here, the language effect might have also manifested itself on variations in the voice spectra of speakers for different languages, because in various studies on bilinguals, voice spectra were analysed and a significant language effect on voice spectra was found (17, 18, 31). Another factor that might affect the voice spectra similarities between languages is the variation in the intonation. The effect of intonation on the voice source has been shown to have a complex mechanism in relation to the subglottal pressure and open quotient of vocal fold vibration, but more studies are needed to understand the effect of intonation on the voice source (50).

Lastly, this study analysed the Turkish voice spectrum while Yüksel and Gündüz analysed the Turkish speech spectrum and included unvoiced speech segments in the LTAS analyses (22). Yüksel and Gündüz found significant differences between sexes at the frequency bands of 62.5–125 Hz, 625–1125 Hz, 2875–3125 Hz, 4125–4625 Hz and 7375–7500 Hz (22). The significant differences between LTAS bands seen in the study of Yüksel and Gündüz and this study were mostly similar below 4.5 kHz as unvoiced speech segments were found to mostly affect the spectrum above 4–5 kHz (12, 22, 30). Because of this effect, no significant differences were found between the spectral tilt values of females and males in the study of Yüksel and Gündüz (22). However, the spectral tilt of the voice spectrum of females and males in this study were found to be significantly different from each other, and the female voice spectrum was shown to have the features of a more hyperfunctional voice source characteristics than males.

Limitations of the study

The age range of the participants in this study was 18-50 years and the shown effect of age on LTAS band intensities (10, 11) was considered as a limitation for this study. The age effect might have compromised the analyses of sex. Three age groups of 18-34 years (11 male, 9 female), 35-39 years (9 male, 9 female) and 40-50 years (10 male, 12 female) were formed and analysed by a full factorial repeated measures ANOVA as age and sex were the factors. The effect of age on the LTAS band intensities was found to be non-significant ($F(2,54)=0.648$, $p=0.527$). The interaction between the effects of age and sex ($F(2,54)=1.132$, $p=0.330$) was also found to be non-significant. Lastly, the interaction effect between the LTAS bands and age was also non-significant ($F(148,3996)=0.662$, $p=0.520$) which indicated the intensities of different LTAS bands were similar between females and males. Through these analyses, the effect of age on the intensities of the LTAS bands was presumed to be at a minimum level.

CONCLUSION

This study analysed voice spectrum of Turkish female and male speakers using the LTAS method. All significant band differences between sexes were found to be below 6 kHz, and both sexes had a rather flat-spectrum above 6 kHz. Females had significantly higher intensities than males around the regions of 640–960 Hz, 2880–3040, and 4000–5120 Hz. These significant differences caused females to have shallower STs than males, indicating greater hyperfunctional voice quality in females. The significant difference in the 2880–3040 Hz region is considered due to aspiration noise occurring during phonation, and this noise is likely due to the posterior glottal gap in females. Finally, the voice spectrum characteristics of Turkish speakers were compared with those of other studies and more similarities were found with Spanish than Persian.

Acknowledgements: This article is the revised and developed version of the unpublished conference presentation entitled "Anadili Türkçe Olan Bireylerin Seslerinin LTAS ile Analizi", orally delivered at the 9. Ulusal Larengoloji Kongresi.

Ethics Committee Approval: This study was approved by the Health Sciences University Izmir Tepecik Training and Research Hospital (Date: 07.03.2017, Approval No: 13).

Informed Consent: Written informed consent was obtained from all participants.

Peer Review: Externally peer-reviewed.

Author Contributions: Conception/Design of Study- C.Y.; Data Acquisition- E.D.; Data Analysis/Interpretation- C.Y., Ş.N.B.;

Drafting Manuscript- C.Y.; Critical Revision of Manuscript- C.Y., E.D., Ş.N.B.; Final Approval and Accountability - C.Y., E.D., Ş.N.B.; Technical Support- C.Y., E.D.; Supervision- E.D., C.Y.

Conflict of Interest The authors have no conflict of interest to declare.

Financial Disclosure The authors declared that this study has received no financial support.

Author Details

Cemil Yılmaz

[1] Kütahya Health Sciences University, Faculty of Health Sciences, Department of Language and Speech Therapy, Kütahya, Turkiye

0000-0003-1821-7171 cemil.yilmaz@ksbu.edu.tr

Erhan, Demirhan

[2] Private Practice, Otorhinolaryngology, Head and Neck Surgery, İzmir, Turkiye

0000-0001-8871-0821

Şerife Nur Biçen

[3] Anadolu University, Graduate Schools and Institutes, Department of Speech and Language Therapy, Eskişehir, Turkiye

0009-0005-0976-7886

REFERENCES

1 Hillenbrand JM. Acoustic analysis of voice: a tutorial. Perspect Speech Sci Orofac Disord 2011;21(2):31-43.

2 Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. JASA 1990;87(2):820-57.

3 Hanson HM, Chuang ES. Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. JASA 1999;106(2):1064-77.

4 Titze IR. Physiologic and acoustic differences between male and female voices. JASA 1989;85(4):1699-707.

5 Zhang Z. Contribution of laryngeal size to differences between male and female voice production. JASA 2021;150(6):4511-21.

6 Hanson HM. Glottal characteristics of female speakers: Acoustic correlates. JASA 1997;101(1):466-81.

7 Pittam J. The long-term spectral measurement of voice quality as a social and personality marker: A review. Lang Speech 1987;30(1):1-12.

8 Leino T. Long-Term Average Spectrum in Screening of Voice Quality in Speech: Untrained Male University Students. J Voice 2009;23(6):671-6.

9 Mendoza E, Valencia N, Munoz J, Trujillo H. Differences in Voice Quality Between Men and Women: Use of the Long-Term Average Spectrum (LTAS). J Voice 1996;10(1):59-66.

10 Linville SE. Source Characteristics of Aged Voice Assessed from Long-Term Average Spectra. J Voice 2002;16(4):472-9.

11 da Silva PT, Master S, Andreoni S, Pontes P, Ramos LR. Acoustic and Long-Term Average Spectrum Measures to Detect Vocal Aging in Women. J Voice 2011;25(4):411-9.

12 Löfqvist A, Mandersson B. Long-Time Average Speech Spectrum of Speech and Voice Analysis. Folia Phoniatr Logop 1987;39(5):221-9.

13 Lowell SY, Colton RH, Kelley RT, Hahn YC. Spectral- and Cepstral-Based Measures During Continuous Speech: Capacity to Distinguish Dysphonia and Consistency Within a Speaker. J Voice 2011;25(5):e223-32.

14 Guzman M, Angulo M, Munoz D, Mayerhoff R. Effect on long-term average spectrum of pop singers' vocal warm-up with vocal function exercises. Int J Speech-Lang Pathol 2013;15(2):127-35.

15 Guzman M, Higueras D, Fincheira C, Munoz D, Guajardo C, Dowdall J. Immediate acoustic effects of straw phonation exercises in subjects with dysphonic voices. Logoped Phoniatr Vocol 2013;38(1):35-45.

16 Smith LK, Goberman AM. Long-time average spectrum in individuals with Parkinson disease. NeuroRehabilitation. 2014;35(1):77-88.

17 Ng ML, Chen Y, Chan EYK. Differences in Vocal Characteristics Between Cantonese and English Produced by Proficient Cantonese-English Bilingual Speakers-A Long-Term Average Spectral Analysis. J Voice 2012;26(4):e171-6.

18 Zhu S, Chong S, Chen Y, Wang T, Ng ML. Effect of Language on Voice Quality: An Acoustic Study of Bilingual Speakers of Mandarin Chinese and English. Folia Phoniatr Logop 2022;74(6):421-30.

19 Boullosa RR, Perez Ruiz SJ. Long-Term Average Spectrum of Spanish. Appl Acoust 1984;17(3):201-8.

20 Byrne D, Dillon H, Tran K, Arlinger S, Wilbraham K, Cox R, et al. An international comparison of long-term average speech spectra. JASA 1994;96(4):2108-20.

21 Dong L, Kong J, Sundberg J. Long-term-average spectrum characteristics of Kunqu Opera singers' speaking, singing and stage speech. Logoped Phoniatr Vocol 2014;39(2):72-80.

22 Yüksel M, Gündüz B. Long term average speech spectra of Turkish. Logoped Phoniatr Vocol 2017;43(3):101-5.

23 Narne VK, Sreejith VS, Tiwari N. Long-Term Average Speech Spectra and Dynamic Ranges of 17 Indian Languages. Am J Audiol. 2021;30(4):1096-107.

24 Pittam J. The long term measurement of voice quality: A comparison of acoustic measures. In: Proceedings of the First Australian Conference on Speech Science and Technology; 1986; Australian National University, Canberra, Australia. pp. 234-9.

25 Tanner K, Roy N, Ash A, Buder E. Spectral Moments of the Long-term Average Spectrum: Sensitive Indices of Voice Change After Therapy? J Voice 2005;19(2):211-22.

26 Houtz DR, Roy N, Merrill RM, Smith ME. Differential Diagnosis of Muscle Tension Dysphonia and Adductor Spasmodic Dysphonia Using Spectral Moments of the Long-Term Average Spectrum. Laryngoscope 2010;120(4):749-57.

27 Kitzing P. LTAS criteria pertinent to the measurement of voice quality. J Phon 1986;14(3-4):477-82.

28 Moradi N, Maroufi N, Bijankhan M, Nik TH, Salavati M, Jalayer T, et al. Long-term average spectra of adult Iranian speakers' voice. J Voice 2014;28(3):305-10.

29 Leino, T, Laukkanen AM, Ilomaki I, Maki E. Assessment of Vocal Capacity of Finnish University Students. Folia Phoniatr Logop 2008;60(4):199-209.

30 Nordenberg M, Sundberg J. Effect on LTAS of vocal loudness variation. Logoped Phoniatr Vocol 2004;29(4):183-91.

31 Bahmanbiglu SA, Mojiri F, Abnavi F. The Impact of Language on Voice: An LTAS Study. J Voice 2017;31(2):249.e9-249.e12.

32 Monson BB, Lotto AJ, Story BH. Analysis of high-frequency energy in long-term average spectra of singing, speech, and voiceless fricatives. JASA 2012;132(3):1754-64.

33 Li KP, Hughes GW, House AS. Correlation Characteristics and Dimensionality of Speech Spectra. JASA 1969;46(4B):1019-25.

34 Monson BB, Hunter EJ, Lotto AJ, Story BH. The perceptual significance of high-frequency energy in the human voice. Front Psychol 2014;5:587.

35 Master S. The long-term average spectrum in research and in the clinical practice of speech therapists. Pró-Fono Revista de Atualização Científica 2006;18(1):111-20.

36 Boersma P, Weenink D. Praat Manual: doing phonetics by computer (V. 6.1.41). (Computer program). Available at: http://www.praat.org/ (accessed 12 April 2021).

37 Hirano M. Clinical Examination of Voice. New York (NY): Springer-Verlag Press; 1981.

38 Maryn Y, Weenink D. Objective dysphonia measures in the program Praat: Smoothed cepstral peak prominence and acoustic voice quality index. J Voice 2015;29(1):35-43.

39 Maryn Y, Roy N, De Bodt M, Cauwenberge PV, Corthals P. Acoustic measurement of overall voice quality: A meta-analysis. JASA 2009;126(5):2619-34.

40 Buckley DP, Abur D, Stepp CE. Normative values of cepstral peak prominence measures in typical speakers by sex, speech stimuli, and software type across the life span. AJSLP 2023;32(4):1565-77.

41 Phadke KV, Laukkanen AM, Ilomaki I, Kankare E, Geneid A, Svec JG, et al. Cepstral and perceptual investigations in female teachers with functionally healthy voice. J Voice 2020;34(3):485.e33-43.

42 Goberman AM, Robb MP. Acoustic examination of preterm and full-term infant cries: The long-time average spectrum. JSLHR 1999;42(4):850-61.

43 Sergeant D, Welch GF. Age-related changes in long-term average spectra of children's voices. J Voice 2008;22(6):658-70.

44 Sergeant D, Welch GF. Gender Differences in Long-Term Average Spectra of Children's Singing Voices. J Voice 2009;23(3):319-36.

45 Whitfield JA, Goberman AM. Physiological Aspects and Methodological Considerations of Long-Term Average Spectrum Measures. In: Proceedings of the 10th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research; Cincinnati, OH, USA. pp. 101-2.

46 Master S, De Biase N, Chiari BM, Laukkanen AM. Acoustic and perceptual analyses of Brazilian male actors' and nonactors' voices: long-term average spectrum and the "actor's formant". J Voice 2008;22(2):146-54.

47 Nawka T, Anders LC, Cebulla M, Zurakowski D. The speaker's formant in male voices. J Voice 1997;11(4):422-8.

48 Bele IV. The speaker's formant. J Voice 2006;20(4):555-78.

49 Shoji K, Regenbogen E, Yu JD, Blaugrund SM. High-frequency components of normal voice. J Voice 1991;5(1):29-35.

50 Pierrehumbers J. A preiliminary study of the consequences of intonation for the voice source. STL-QPSR 1989;30(4):23-36.