Spectral analysis of english fricatives /s/ and /ʃ/ produced by people with profound hearing loss

Sungmin Lee

Department of Speech-Language Pathology and Audiology, Tongmyong University, Busan, Korea

Purpose: Hearing loss not only causes restricted speech understanding, but also alters speech production resulting in poor intelligibility. This study investigated whether English fricatives /s/ and /ʃ/ spoken by people with severe-to-profound hearing loss acoustically differ from those spoken by normal hearing (NH) individuals.

Methods: For hearing impaired (HI) group, we analyzed a part of deaf speech corpus collected at the Speech Perception Assessment Lab (SPAL) at the University of Memphis. From the big dataset, /aSa/ and /aSha/ recordings spoken by eight females were selected. For NH control group, Shannon's consonant set (1999) spoken by eight females was used. A set of spectral moment analysis (Mean, Variance, Skewness, Kurtosis) along with spectral peak (SP) and wiener entropy (WE) were used as measures of spectral properties of /s/ and /ʃ/.

Results: Variance and WE for HI were significantly higher than those for NH group ($p < 0.05$). Mean and SP for /s/ were significantly greater in NH group, while Mean and SP for /ʃ/ were significantly greater in HI group ($p < 0.05$). This trends suggest greater Mean and SP differences between /s/ and /ʃ/ for NH group than HI group.

Conclusions: Based on our finding, Variance and WE are spectral measures that are different between NH and HI individuals' spoken /s/ and /ʃ/. We assume that HI individuals likely pronounce /s/ and /ʃ/ without making clear distinction substituting by each other due probably to lack of auditory feedback or failure of motor control.

Keywords: Fricatives, Spectral moment analysis, Hearing loss, Deaf speech

Received: August 14, 2020

Revision: December 27, 2020

Accepted: December 31, 2020

Correspondence:

Sungmin Lee

Department of Speech-Language Pathology and Audiology, Tongmyong University, 428 Sinseon-ro, Nam-gu, Busan 45820, Korea
Tel: +82-51-629-2134
Fax: +82-51-629-2019
E-mail: slee18@tu.ac.kr

INTRODUCTION

Characteristics of deaf speech

Hearing loss not only causes restricted speech understanding, but also alters speech production resulting in poor intelligibility [1,2]. Due to the loss of auditory monitoring, the articulatory controls of deaf talkers are not regulated by auditory feedback and voice abnormalities in speech are sustained. Numerous studies have reported that the most prominent characteristic of deaf speech is excessive nasality [3,4]. In the beginning, the excessive nasality of deaf speech was thought to be associated with the slow rate of speech typical individuals with hearing loss represent [5]. However, contradictory findings have been reported by Fletcher and Daly [6], and the hyper-nasality has been found to be more related to the abnormal function of velopharyngeal valve caused by cul-de-sac resonance [7]. The cul-de-sac is caused by muffled airflows that block the resonance cavity. Furthermore, speech of the people with hearing loss is typi-

© 2020 The Korean Association of Speech-Language Pathologists

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

cally characterized by labored voice, breathy speech, slow rate, monotonic natures, devoiced stops, and abnormal rhythm [8].

Inadequate vowel articulation caused by nonflexible tongue movement which constricts the pharynx has been assumed to be a primary source of poor intelligibility as well [9]. Acoustic analyses have reported higher fundamental frequencies (F0) in deaf people in comparison with a normal control group [10]. It has been also found that formant frequencies for certain vowels differed between normal hearing (NH) group and deaf group; F2 of vowel /i/ was lower for deaf group [11]. Abnormal speech patterns of hearing impaired (HI) children do not differ greatly from those of HI adults. Liker et al. [12] described a study of Croatian children with cochlear implant (CI) and characterized their speech as having a smaller and frontal vowel space, indistinct frequency distributions between /s/ and /ʃ/, and longer affricates. Uchanski and Geers [13] examined a larger database of acoustic measures (voice onset time, second formant frequency, spectral moments, nasal manner metric, and durations) spoken by young CI users and they found acoustic properties of CI children that are not identical to those by NH children. In addition, they reported that a high percentage of CI children produced acoustic characteristics that are comparable to those produced by children with NH. In summary, the acoustic literature is helpful in understanding the bases of speech abnormalities concerning errors produced by HI talkers.

Spectral moment analysis

Acoustic analysis of deaf speech can be made by examining particular segment of phoneme using various measures. Of the several measures, spectral moment analysis (SMA) is considered as a quantification method that systematically analyze the acoustic characteristics of produced speech in a spectral domain. SMA consists of four primary moments: Mean, Variance, Skewness, and Kurtosis. The spectral Mean refers to the simple average of distributed energy in the spectrum. The second moment, Variance, is the deviation of frequencies represented in the spectrum (the amount of variance in terms of mean). The third moment, Skewness shows the degree to which the amount of energy accumulated at either end of the distribution. Positive (+) indicates higher energy at lower frequency, but negative (-) indicates higher energy at higher frequency. The last moment, Kurtosis, measures peakness of distribution. Higher kurtosis represents thin and peaked distribution, whereas a low kurtosis represents relatively fatty

and plat distribution.

Recently, Mendel et al. [14] analyzed four spectral moments (spectral mean, standard deviation, skewness, and kurtosis) for short passages spoken by three hearing impaired groups that were sorted based on perceptual intelligibility. In the results, spectral variance (standard deviation) increased with increases in intelligibility, while kurtosis decreased with increases in intelligibility, suggesting that speech that is more intelligible has a broader range of frequency distribution. Other two spectral moments (spectral mean and skewness) indicated that those with highly intelligible speech yields more energy in the high-frequency. This findings were in line with Raghavendra et al., [15] in which they systematically analyzed the /s/, /ʃ/, /f/, and /θ/ of the same dataset in Mendel et al., [14] by cutting them as a short duration of fricative segments. Another study [16] where Korean fricatives for children with hearing loss were analyzed using SMA reported that spectral mean, standard deviation, and kurtosis were found to be significant indicators showing clinical difference from NH children. They also noted that comparing the alveolar fricatives based on vowels (/a/, /i/, and /u/) do not affect the results.

For the acoustic analysis of speech segment, voice onset time (VOT), F0, F1, F2, duration are typical indexes to be examined in values. However, for the analysis of speech segment where SMA is effectively used is obstruent consonants [17]. Obstruent is yielded by release of noise with either turbulence (fricatives and affricates) or plosive burst (stop). SMA measures and analyzes such noise energy with a statistical procedure using a series of fast Fourier transforms (FFTs). Fricatives are a type of consonant produced by turbulent air flow at pass through narrow opening of oral cavity. Voiceless fricatives can be characterized by its spectral shape according manner and place of articulation. Alveolar sibilant /s/ which shapes a shorter anterior cavity produces higher frequency energies (4 to 5 kHz) compared to the palate-alveolar sibilant /ʃ/, non-sibilant /f/ (labiodental) and θ (dental) which have relatively flat frequency distribution without dominant peak in any particular frequency. Children tend to have difficulty producing fricatives and produce them at the last stage frequently replacing them with other consonants. CI children have more difficulty producing fricatives than other consonants [13]. This difficulty could probably be caused by articulatory complexity of producing fricatives as well as relatively weaker energy distribution of fricatives in higher frequency range where human auditory system does not handle sensitively.

Purpose

Although many studies on acoustic analysis of speech have examined HI children [16,18], few have examined HI adults particularly using SMA. With a great help of hearing assistive technology, patients with hearing difficulties greatly improved speech as the result of continued use of their devices. Nonetheless, it is frequently reported that these patients still speak with residual speech errors. In this proposed research, we analyzed comprehensive speech data of fricative speech /s/ and /ʃ/ produced by HI adults. We investigated if there is a significant difference in each spectral moment component between NH group and HI group. In addition to SMA, spectral peak and wiener entropy (WE) were also analyzed. Spectral peak detects a particular frequency that has the greatest energy. WE quantifies the spectral density of noise yielding whether acoustic energy is diffused across the frequency or energy is focused at one frequency. It is represented in a log scale, which ranges from 0 (power spectrum is flat) to minus infinity (power spectrum is infinitely narrow).

METHODS

Subjects

The present study was conducted by analyzing deaf corpus data that was obtained at the Speech Perception Assessment Lab (SPAL) at the University of Memphis. Among twenties hearing impaired subjects who have English as a first language in the deaf speech corpus, eight females aged from 22 to 68 (M : 50.62, SD : 13.61 years) were chosen for acoustic analysis of fricatives /s/ and /ʃ/. All of them were bilaterally diagnosed with the severe to profound hearing without his-

tory of neurological or cognitive deficits. Six of the participants were wearing either a hearing aid or a CI, but two of them did not use any hearing assistive device. Six of our participants were pre-lingually impaired, but two of them have hearing loss began after critical period of language learning. Table 1 shows demographic of eight HI participants. For the comparison of acoustic characteristics with NH control group, we extracted fricatives of eight normal female talkers from the widely used /aCa/ consonant sets by Shannon et al., [19]. The experimental and control group were those who had no noticeable regional accent.

Acoustic analysis

A set of fricatives /s/ and /ʃ/ spoken by eight females with severe-to-profound hearing loss and NH were taken from the non-sense syllable (aCa) dataset in the deaf speech corpus and in Shannon's consonant data set, respectively. We used Praat program [20] to analyze four major spectral moments (Mean, Variance, Skewness, and Kurtosis), spectral peak (SP) and wiener entropy (WE). First we extracted 40 msec of central position of each fricative from each speaker's /aSa/ and /aSHa/ set. Increasing acoustic energy in high frequency components with a pre-emphasis filter at 80 Hz was carried out. Then, the four SMAs and SP were computed based on the spectral properties in a given 40 ms window of each fricative. For computing the WE, a freely available custom Praat script developed by Gabriel J.L. and Beckers (2004; available online) was used. The procedure of WE analysis is the same, except that WE was estimated without pre-emphasizing high frequency components.

Table 1. Demographic of eight participants with severe-to-profound hearing loss or deaf

	Age (yr)	Age of onset (yr)	Starting age of using amplification (yr)		Type of amplification		Pure-tone threshold (Right)						Pure-tone threshold (Left)					
			Right	Left	Right	Left	250	500	1k	2k	4k	8k	250	500	1k	2k	4k	8k
1	59	At birth	42	52	CI	CI	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
2	48	5	5	5	HA	HA	90	85	95	90	95	105	90	90	120	NR	NR	NR
3	44	10	28	-	CI	NA	100	NR	105	110	NR	NR	105	105	110	NR	100	NR
4	22	0.7	0.6	2	HA	CI	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
5	53	4	34	34	HA	HA	85	85	85	75	60	90	90	95	95	90	75	95
6	67	35	38	38	CI	HA	NR	NR	NR	NR	NR	NR	90	80	95	75	80	90
7	55	At birth	-	-	NA	NA	90	95	110	120	NR	NR	100	110	NR	NR	NR	NR
8	53	At birth	-	-	NA	NA	85	105	110	115	NR	NR	NR	NR	120	NR	NR	NR

CI, cochlear implant; HA, hearing aid; NA, not available; NR, no response.

Statistical analysis

Two-way analysis of variance (ANOVA) was performed to determine the effect of groups and fricatives on obtained values. Two groups (NH and HI) and two fricatives (/s/ and /ʃ/) were independent variables assumed to be associated with the products obtained with the set of spectral analysis. Statistically significance level was set as $p < 0.05$.

RESULTS

Table 2 shows the group means and standard deviations (SDs) of spectrally analyzed numerical data for /s/ and /ʃ/ spoken by NH and HI groups. It is illustrated in Figure 1: The values of Mean, Variance, and SP were represented in Hz [(A) - /s/ and (C) - /ʃ/], and the other three spectral components, Skewness, Kurtosis, and WE, were derived in small scale units that fall between -2 and 6 [(B) - /s/ and (D) - /ʃ/].

A series of Two-way ANOVAs were conducted that examined the effect of groups and fricatives on each spectral analysis measures (see summary of ANOVA Table 3). Simple main effects of group were found only in Variance and WE outcomes. Variance and WE for HI were significantly higher than those for NH group ($p < 0.05$). There was no significant difference in Mean, SP, Skewness, and Kurtosis between two groups.

There were significant main effects of fricative in Mean and SP with interaction effects reporting that Mean and SP for /s/ are significantly higher in NH group, while Mean and SP for /ʃ/ are higher in HI group ($p < 0.05$). This trends suggest greater Mean and SP differences between /s/ and /ʃ/ for NH

group. On the other hands, there were comparatively smaller Mean and SP differences between /s/ and /ʃ/ for HI group as shown in Figure 2. Another notable result was that the SD of Kurtosis is larger than other spectral measures indicating a greater individual difference in wideness of peak.

DISCUSSION

This study investigated whether English fricatives /s/ and /ʃ/ spoken by people with severe-to-profound hearing loss acoustically differ from those spoken by NH individuals. Acoustic analysis, specifically SMA (Mean, Variance, Skewness, Kurtosis), SP, and WE were used as measures of spectral properties of /s/ and /ʃ/.

Overall, our spectral analysis results were consistent with other literatures that show /s/ has a greater mean, SP, kurtosis, but smaller skewness than /ʃ/ [13,21,22]. Again, the length of anterior cavity determines dominant frequency range of two fricatives /s/ and /ʃ/: Alveolar /s/ produced with a shorter anterior cavity has greater spectral energy at about 4 to 5 kHz, whereas palato-alveolar /ʃ/ has greater spectral energy at 3.5-5 kHz. Those spectral measures that segregate /s/ and /ʃ/ were the ones representing frequency regions which contain greater acoustic energy well enough to represent spectral characteristics of fricatives.

A significant Variance and WE difference between two groups was found. There was no other significant difference between two groups. Variance and WE are measures of spectral energy showing how much spectral energy is dispersed across the frequency. Our results suggest that spectral energy of /s/ and /ʃ/ spoken by HI group is more likely to be flat and scattered across a wide range of frequency. It is assumed that unclear and incomplete articulatory movement by HI individuals yielded such fricative energy spread widely rather than focusing at a certain point.

Notably, our finding showed that Mean and SP for /s/ were significantly higher for NH group than HI group, but those for /ʃ/ were significantly higher for HI group than NH group. In other words, the identical pattern was found between Mean and SP. It is reasonable that the first moment of SMA, Mean, is acoustically associated with SP in that they account for at which the greatest spectral energy is located. As shown in Figure 2, the /s/ and /ʃ/ difference in Mean and SP were quite clear for NH group, but not for HI group. Based on this finding, we assume that HI individuals likely pronounce /s/ and /ʃ/ without making clear distinction substituting by each

Table 2. Group mean and SD for six spectral analysis

		/s/		/ʃ/	
		NH	HI	NH	HI
Mean	Mean	9,492.297	8,507.527	5,599.042	7,318.515
	SD	1,813.579	1,573.958	780.697	1,637.968
Variance	Mean	1,614.241	2,682.331	2,325.671	2,659.015
	SD	473.403	1054.410	792.734	1,121.028
Skewness	Mean	0.860	1.278	1.386	1.091
	SD	0.903	1.049	0.592	1.00
Kurtosis	Mean	5.173	3.425	3.746	3.873
	SD	2.591	4.014	3.537	5.679
SP	Mean	9,250	7,781.25	4,031.250	7,000
	SD	1,899.247	2,777.259	1,267.291	2,394.189
WE	Mean	-1.8	-0.737	-1.417	-0.398
	SD	0.428	0.857	0.543	0.289

Figure 1. Group mean spectral measure differences between NH and HI groups for /s/ and /ʃ/. (A) represents Mean, Variance, and SP for /s/, (B) represents Skewness, Kurtosis, and WE for /s/, (C) represents Mean, Variance, and SP for /ʃ/, (D) represents Skewness, Kurtosis, and WE for /ʃ/. Error bars denote ± 1 SEM.

other due probably to the lack of auditory feedback or failure of motor control. The failure of motor control was specified as inability to place a tongue at right vocal position [13]. Although several studies have reported that lower spectra in fricatives (spectral Mean or SP) is negatively related to intelligibility of speech [16,18], our finding indicates that simply higher or lower pitch may not fully explain the different characteristics of deaf speech. This pattern is consistent with Yang et al., [23] where higher SP for /ʃ/ in CI children group over NH group was reported. Acoustic and phonetic analysis are

typically carried out in different manner across studies. Some variables, such as the manner of pre-processing, syllable structure, sex and age of speakers may cause the complexity of interpreting results. Thus, intensive care should be taken when comparatively describing the results across the studies.

Despite our effort, some limitations exist in current study. The present study recruited only eight females with the age-unmatched control group data set due to the limited amount of our original data. Thus, the homogeneity of our subjects in our study is unfortunately weak. Although it is not included

Table 3. ANOVA summary for six spectral analysis. Asterisks indicate significance with p -value of 0.05

	Source	Degree of Freedom	F-value	p -value		Source	Degree of Freedom	F-value	p -value
Mean	Group	1	0.477	0.496	Kurtosis	Group	1	0.311	0.581
	Fricative	1	22.810	0.000*		Fricative	1	0.113	0.739
	Group*Fricative	1	6.458	0.017*		Group*Fricative	1	0.417	0.524
	Error	28				Error	28		
Variance	Group	1	4.878	0.036*	SP	Group	1	0.965	0.334
	Fricative	1	1.176	0.287		Fricative	1	45.435	0.001*
	Group*Fricative	1	1.341	0.257		Group*Fricative	1	8.443	0.007*
	Error	28				Error	28		
Skewness	Group	1	0.037	0.848	WE	Group	1	26.689	0.000*
	Fricative	1	0.280	0.601		Fricative	1	3.205	0.084
	Group*Fricative	1	1.243	0.274		Group*Fricative	1	0.012	0.914
	Error	28				Error	28

Figure 2. Group mean spectral Mean and SP differences between NH and HI groups for /s/ and /j/.

here, our pilot study found significantly greater Mean and SP for females over males ($p < 0.05$). Another evidence of gender /s/ and /j/ difference has been also previously reported [24]. Thus English fricative /s/ and /j/ spoken by HI males could be different than those spoken by NH male. In addition, our spectral analysis is limited to sibilant /s/ and /j/. Although misarticulation of /s/ and /j/ is one of the common articulation errors that contribute to the speech quality and intelligibility of HI speaker [9], other errors and corresponding analysis also should be included for the better understanding of HI speech. Future studies would be more comprehensive and

sophisticated by controlling and considering other important variables with a larger sample size.

Acoustic analyses of the speech by HI talkers have yielded mixed results, with some studies indicating near normal spectral qualities, while other studies have suggested a more pervasive vocal degradation. The results of this study will provide acoustic/phonetic researchers with insight of spectrally distinguishable fricative properties in HI group. With a spectral analysis, quantifying speech intelligibility may be possible. Clinicians including speech therapists could benefit from our findings by using spectral measures as objective indicators for

monitoring and examining speech improvement in articulation for HI patients. Moreover, spectral variations in produced speech could be potential cues for prediction speech perception performance. Clarifying the relationship between acoustic properties of deaf speech and their speech perception will be needed in that regard in the future study.

ACKNOWLEDGMENTS

We thank Dr. Lisa Lucks Mendel for sharing the deaf speech corpus. This Research was supported by the Tongmyong University Research Grants 2020 (2020A014).

REFERENCES

- Porter KF, Bradley S. A comparison of three speech intelligibility measures for deaf students. *American Annals of the Deaf*. 1985; 130:514-525.
- Osberger, MJ, McGarr N. Speech production characteristics of the hearing impaired. In N. Lass (Ed.), *Speech and language: advances in basic research and practice*, New York: Academic Press. 1982.
- Hudgins CV, Numbers FC. An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology Monographs*. 1942;25: 289- 392.
- Sherman D. The merits of backward playing of connected speech in the scaling of voice quality disorders. *Journal of Speech and Hearing Disorders*. 1954;19:312-321.
- Colton RH, Cooker HS. Perceived nasality in the speech of the deaf. *Journal of Speech and Hearing Research*. 1968;11:553-559.
- Fletcher SG, Daly DA. Nasalance in utterances of hearing-impaired speakers. *Journal of Communication Disorders*. 1976;9:63-73.
- Boone DR. Modification of the voices of deaf children. *The Volta Review*. 1966;68:686-692.
- Nickerson RS. Characteristics of the speech of deaf persons. *The Volta Review*. 1975;77:342-362.
- Monsen RB. Voice quality and speech intelligibility among deaf children. *American Annals of the Deaf*. 1983;128:12-19.
- Gilbert H, Campbell M. Speaking fundamental frequency in three groups of hearing-impaired individuals. *Journal of Communication Disorders*. 1980;13:195-205.
- Monsen R. Normal and reduced phonological space: the production of English vowels by deaf adolescents. *Journal of Phonetics*. 1976;4:189-198.
- Liker M, Mildner V, Šindija B. Acoustic analysis of the speech of children with cochlear implants: A longitudinal study. *Clinical Linguistics & Phonetics*. 2007;21:1-11.
- Uchanski RM, Geers AE. Acoustic characteristics of the speech of young cochlear implant users: A comparison with normal hearing age-mates. *Ear and Hearing*. 2003;24:90S-105S.
- Mendel LL, Lee S, Pousson M, Patro C, McSorley S, Banerjee B, Najnin S. Corpus of deaf speech for acoustic and speech production research. *The Journal of the Acoustical Society of America*. 2017;142:EL102-EL107.
- Raghavendra S, Lee S, Hui J, Tan, C. Analysis of fricatives in speech produced by hearing impaired individuals aided with different assistive hearing devices. Presented at Conference on Implantable Auditory Prostheses. Lake Tahoe, USA; 2019 July;160
- Kim Y, Kim E, Jang S-J, Choi Y. Comparison of acoustic phonetic characteristics of Korean fricative sounds pronounced by hearing-impaired children and normal children. *Phonetics and Speech Sciences*. 2014;6:73-9.
- Forrest K, Weismer G, Milenkovic P, Dougall RN. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America*. 1988;84:115-123.
- Scarbel L, Vilain A, Loevenbruck H, Schmerber S. An acoustic study of speech production by French children wearing cochlear implants. 3rd Early Language Acquisition Conference, Dec 2012, Lyon, France.
- Shannon RV, Jensvold A, Padilla M, Robert ME, Wang X. Consonant recordings for speech testing. *The Journal of the Acoustical Society of America*. 1999;106:L71-L41.
- Boersma P, Weenink D. Praat: doing phonetics by computer [Computer program]. Version 6.1.16, retrieved 6 June 2020 from <http://www.praat.org/>
- Hernandez A, Lee H, Chung M. Acoustic analysis of fricatives in dysarthric speakers with cerebral palsy. *Phonetics and Speech Sciences*. 2019;11:23-29.
- Jongman A, Wayland R, Wong S. Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*. 2000;108:1252-1263.
- Jing Y, Jessica V, Zhigang Y, Chao-Yang L, Li X. Production of word-initial fricatives of Mandarin Chinese in prelingually deafened children with cochlear implants. *International Journal of Speech-Language Pathology*. 2017;19:153-164.
- Fox RA, Nissen SL. Sex-related acoustic changes in voiceless English fricatives. *Journal of Speech, Language, and Hearing Research*. 2005;48:753-765.