

Harmonic Amplitude Measures to Note Gender Difference

Manjunath Narra* Thejaswi. D, Anu. C. C Sneha Mareen Varghese T. Dattatreya
Department of Speech Language Pathology and Audiology, Nitte Institute of Speech and Hearing, Deralakatte,
Mangalore-575018, Karnataka, INDIA

Abstract

The present study was taken up with an objective to find whether harmonic amplitude measures can be used to note gender differences. A total of sixty subjects were divided into two groups. Group I consisted of 30 male subjects and group II consisted of 30 female subjects. They were instructed to produce three trials of sustained vowels (/a/, /i/ and /u/). The harmonic amplitude difference measures (H1-H2; H1-A1; H1-A2 and H1-A3) were extracted from Praat software. Results revealed significant difference between male and female subjects for all the harmonic amplitude difference measures. On measuring the difference between male and female subjects, the values ranged from 3.09 to 24.04. The results are discussed with respect to the underlying physiological differences and glottal chink during phonation across gender.

Keywords: Spectrum, FFT, Speech synthesis, Gender, Praat

1. Introduction

Human voice is a rich source of information for listeners as well as the medium of communication in oral language, in its non-linguistic role listeners draw inferences from the voice regarding age, gender, socio-economic origin, education and occupation (Ryan, Giles & Sebastian, 1982) and emotional state (Scherer, Banse, & Wallnott, 2001). The cues that are used to assess the gender of the speaker are related to both the fundamental and formant frequencies (Whiteside, 1998; Childers & Wu, 1991). However in most speech-analysis tools and speech applications, the synthesis of female speech tends to sound less natural than the synthesized male speech (Karlsson, 1992). Thus, it would seem that features of the acoustic sound-pressure waveform other than fundamental frequency and formant frequencies contribute to gender characteristics of speech (Hanson & Chuang, 1996).

In the past, researchers studied female voice source characteristics to differentiate it from male voice (Titze, 1989; Klatt & Klatt, 1990; Karlsson, 1992). Holmberg, Hillman, & Perkell (1988, 1989) identified gender differences in glottal-waveform characteristics like larger open quotients, more gradual rise and fall in glottal flow in females that affect perceived voice quality. Observation of the glottis during phonation has suggested that the presence of a posterior glottal opening that persists throughout a vibratory cycle is common for female speakers, but occurs much less frequently among male speakers (Sodersten & Lindestad, 1990). Also, a careful control over the glottal characteristics had shown to improve the naturalness in the synthesized speech sample of female speakers (Hanson, 1995b). Recently, Simpson (2009) reported breathiness to be a differentiating factor between male and female voice. Presence of posterior glottal opening during phonation has been observed in females resulting in a more breathy voice than in males. Thus, glottal configuration and its effects on voice-source characteristics play a significant role in the perception of gender.

The above mentioned characteristics of the glottal waveform are difficult to estimate quantitatively. Earlier technique like inverse filtering of the acoustic sound pressure and oral airflow were employed to study the glottal waveform characteristics. However, both are sensitive to experimental error during data recording and require strict conditions and special equipment during data acquisition (Hanson, & Chuang, 1999). Thus, voice-source characteristics may prove useful for improving applications such as, speech recognition, speaker recognition and, speech synthesis. The analysis techniques that are robust and easily automated are necessary in order to make their use practical. To overcome this problem, Hanson (1995a) reported an alternative to these techniques to make measurements directly on the acoustic sound-pressure waveform and spectrum called harmonic amplitude measures, which requires only simple microphone recordings. He employed harmonic amplitude difference measures (H1-H2, H1-A1, H1-A2 and H1-A3) on 22 female subjects. Preliminary evidence based on breathiness ratings and on fiberscope images suggested that the acoustic measures could be used to categorize the speakers by glottal configuration.

Harmonic amplitude measures are sensitive to varying degrees of vocal fold adduction in normal speakers. According to the theoretical models, the harmonic amplitude measures are related to existence and size of glottal chink. Differences in amplitudes of the first and second harmonics (H1, H2) and the harmonic amplitudes located at the first, second, and third formant frequencies (A1, A2, A3) of the voice spectrum are useful for quantification of degree of glottal adduction in different voices perceived. The amplitude of the first harmonic (H1) compared with the amplitude of the second harmonic (H2) is an indicator of the relative length of the opening phase of the glottal pulse (Hanson, 1997). The amplitude of the H1 compared with the strongest harmonic in the first formant (A1) reflects the first formant bandwidth (formant bandwidths increase with the

presence of a posterior glottal chink, common in breathy voices and spectral tilt). The amplitude of the H1 compared with the amplitude of the strongest harmonic in the second formant (A2) is an indicator of spectral tilts at the mid formant frequencies. The amplitude of the H1 compared with the amplitude of the strongest harmonic in the third formant (A3) is an indicator of spectral tilt at the higher formant frequencies.

Hanson (1995) and Hanson & Chuang (1999) extracted harmonic amplitude difference measures for normal phonation samples with different degrees of breathiness and to note gender differences respectively. Klatt & Klatt (1990) obtained a H1-H2 difference of 5.7 dB between males and females. Similarly, Henton & Bladon (1985) also found that H1-H2 difference was about 6 dB greater for female speakers. Hanson & Chuang (1999) reported significant differences between genders in native American English speakers for H1-H2, H1-A1 and H1-A3. Recently, harmonic amplitude differences have been shown to be clinically useful in differentiating breathiness in individuals with unilateral vocal fold paralysis (Hartl, Hans, Vaissiere, Riquet, & Brasnu, 2001), vocal nodules (Hanson, Stevens, Hong-Kwang, Chen, & Slifka, 2001) and adductor spasmodic dysphonia (Michael, Eugene, & Lesya, 2004).

Sodersten & Lindestad (1990) reported that in females the presence of a posterior glottal opening throughout a glottal cycle introduces additional modification to the spectrum. Firstly, the first formant bandwidths were increased due to additional energy loss at the glottis (Stevens, 1998). The second consequence of the glottal chink is the introduction of additional spectral tilt. This additional tilt is due to the fact that the airflow through the glottal opening cannot undergo a discontinuous change because of the acoustic mass of the moving air in the system. The change in airflow at glottal closure has a time constant which can be used to calculate the cutoff frequency of the effective low-pass filter. Above this frequency, the spectral tilt increases by 6 dB per octave. The third acoustic consequence of a glottal opening is the generation of turbulence noise in the vicinity of the glottis. When the glottal opening is large, the harmonic amplitude of the noise becomes comparable to the harmonic amplitude of the periodic source at high frequencies. Therefore the harmonic amplitude measures are good indicators of glottal status.

Most of the studies conducted to note gender differences are based on the basic acoustic parameters like intensity, fundamental frequencies and formant frequencies. There has been a relatively small amount of research examining the gender differences on harmonic amplitude difference measures. But there is dearth of studies highlighting this on Indian context. Hence the present study was taken up with an objective to find whether harmonic amplitude difference measures can be used to note gender differences and to establish a normative database for Indian population.

2. Method

2.1 Subjects

A total of sixty subjects in the age range of 18 to 25 years (mean age 21.3 years) participated in the study. They were divided into two groups. Group I consisted of 30 normal male subjects (mean age 21.4 years) and group II consisted of 30 normal, age matched female subjects (mean age 21.2 years). All the subjects were native speakers of Malayalam language. The subjects were non-smokers and had normal hearing and oro-facial structures. Those who had a history of vocal abuse or misuse, exposure to toxic chemicals, or any upper respiratory tract infection were excluded from the study.

2.2 Test Material

The test material for the study involved sustained phonation of vowels - /a/, /i/ and /u/.

2.3 Procedure

The subjects were seated comfortably in a quiet room. All subjects were instructed to produce three trials of sustained vowels (/a/, /i/ and /u/) at their most comfortable pitch and loudness level. All phonation samples were recorded using Sony digital recorder ICD-U60 which was placed 10 cm away from the speaker's mouth in a relatively low noise environment. Total of 120 tokens (3 vowels* 40 subjects) were acoustically analyzed to extract the harmonic amplitude difference measures (H1-H2; H1-A1; H1-A2 and H1-A3).

2.4 Acoustic analysis

The recorded phonation samples were fed into the Praat software (Boersama & Weenink, 2010; Version 5.1.43) digitally and sampled at 44 kHz. The samples were subjected to acoustic analysis by Praat script developed by Remijsen (2004) to extract the harmonic amplitude difference measures. This script calculates F1 and F2 at the midpoint of a specific segment as marked by the researcher. Only mid 3- second portion of the phonation sample was selected to extract the amplitude measures. F1 and F2 were calculated using 'To formant (burg)' and the tracker. Both of these algorithms set parameters as a function of speaker sex. This parameter is controlled by the researcher. The script displays the spectrum (F1 & F2), the Fast Fourier transform (H1 & H2) and the LPC spectrum (A1, A2, & A3). The LPC spectrum uses the 'autocorrelation' algorithm, by which the values are

computed. The script puts in markers at H1, H2, A1, A2, A3, and writes the values for harmonic amplitude difference measures (H1-H2; H1-A1; H1-A2 and H1-A3) to the information window.

2.5 Statistical analysis

Statistical analysis was carried out using Statistical Package for the Social Sciences (SPSS 17.00; SPSS Inc., Chicago, IL). The mean and standard deviation values were calculated and tabulated for each harmonic amplitude difference measure. Comparison of harmonic amplitude difference measures between male and female subjects were done using independent t-test for the phonation samples /a/, /i/ and /u/.

3. Results and discussion

Table 1 and Figure 1, 2 & 3 shows the mean and standard deviation (SD) values obtained for male and female subjects for all the amplitude difference measures across phonation samples (/a/, /i/ and /u/). Results revealed that females had higher mean values compared to that of male counterparts irrespective of harmonic amplitude differences measure extracted (H1-H2, H1-A1, H1-A2 and H1-A3). The lowest value (H1-H2=5.41) was obtained for males and highest value (H1-A3=36.31) was obtained in female subjects noted for /u/ phonation sample. On measuring the difference between male and female subjects, the values ranged from 3.09 to 24.04. The lowest difference value was obtained for amplitude measure H1-A2 (3.09) for /a/ phonation sample and highest difference value was obtained for amplitude measure H1-A1 (24.04) for /u/ phonation sample. Overall, the male and female difference values were lower for /a/ phonation sample followed by /i/ and; /u/ had higher values.

Table 2 depicts the independent t test *p* values obtained on comparing phonation samples across gender for harmonic amplitude differences. Independent sample test was administered to find the difference across the gender. The statistical analysis results showed significant difference between male and female subjects for all the harmonic amplitude difference measures across phonation samples at the level 0.05.

Results obtained showed considerable differences between the mean values for female and male subjects for all amplitude measures extracted. The results are in support with the findings of Hanson & Chuang (1999), who reported higher amplitude difference measures in normal female speakers compared to normal male speakers (Simpson, 2009). On an average for H1-H2 amplitude difference, male and female speakers differed by 5.53 dB (mean) in the current study (for /a/=4.31; /i/=6.62 and /u/=5.53) and the results are in consonance with the findings of Klatt & Klatt (1990), who found a difference of 5.7 dB between the genders, and; Henton & Bladon (1985) also found that H1-H2 was about 6 dB greater for female speakers. Thus, the present study demonstrated that spectral amplitude difference measures could be reliably used to measure gender differences.

The amplitude difference H1-H2 is an indicator of the open quotient. In the present study, H1-H2 amplitude differences were higher for female subjects. Larger relative amplitudes of the first harmonics in females may be due to the presence of posterior chink resulting in larger first formant. This is in consonance with the studies that reported the evidence of higher open quotient in females (Soderson & Lindestad, 1990, Holmberg et al, 1988; Hanson, 1997). Higher H1-A1 amplitude differences values in females indicate the larger first formant bandwidth. The first formant bandwidths increase with the presence of a posterior glottal chink (Hanson, 1997) which is common in breathy voices. Females are usually presented with glottal chink; this may be attributed to an increased H1-A1 amplitude difference in this group. This result is in support with the findings of Soderson & Lindestad, (1990) and Hanson & Chuang (1999), who reported higher H1-A1 values in normal female speakers.

The amplitude of the H1 compared with the strongest harmonic in the second formant (A2) and the strongest harmonic in the third formant (A3) is an indicator of spectral tilts at the mid and high formant frequencies respectively. The reason attributed for increased H1- A2 in females could be because of the introduction of additional spectral tilt at the mid formant frequencies (Hanson et al, 1999). On inspection, it is clear that females had higher H1-A3 amplitude reflecting weaker high frequency content in their speech signal. This result is in agreement with the findings of Perkell, Hillman, & Holmberg (2004) who noted higher maximum flow declination rate (MFDR) in males than female speakers. Results emphasizes that voice is a multidimensional phenomenon and could be best described by harmonic amplitude difference measures (H1-H2, H1-A1, H1-A2 and H1-A3). These measures are sensitive to varying degrees of vocal fold adduction and abduction in normal speakers and would give a better clinical picture of voice status.

4. Conclusions

The present study was designed to study if harmonic amplitude difference measures can be used to note gender differences among Indian population. Results revealed significant differences for all the measures extracted across gender and phonation samples recorded. These differences are explained with respect to the physiological variations and size of glottal chink. Hence it can be concluded that harmonic amplitude is a reliable tool to note gender differences. Moreover, as the study illustrate normal variation of the source spectrum characteristics

among individual speakers, and between males and females, they may be useful for applications such as speech synthesis, speech recognition, and speaker recognition. However, as sample size is comparatively small, further studies can be carried out to validate the results. Furthermore, the future research can be directed across different age groups and voice disorders to obtain harmonic amplitude difference measures.

5. Acknowledgements

The authors wish to express their deepest gratitude to the Nitte Education Trust & Nitte University for permitting to conduct the study and also authors express sincere thanks to the Prof T. Dattatreya, Principal, Nitte Institute of Speech and Hearing, Mangalore for supporting us to conduct the study.

6. References

- Boersama, P., & Weanik, D. (2010). Praat (Version 5.1.43). The Netherlands: University of Amsterdam. Retrieved from <http://www.praat.org/>
- Childers, D. G., & Wu, K. (1991). Gender recognition from speech. Part II: Fine analysis. *Journal of Acoustic Society of America*, 90, 1841–1856. doi:10.1121/1.401664
- Henton, C. G., & Bladon R. A. W. (1985). Breathiness in normal female speech: Inefficiency versus desirability. *Language and Communication*, 5, 221–227. doi: 10.1016/0271-5309(85)90012-6
- Hanson, H. M., & Chuang, E. S. (1996). Glottal characteristics of male speakers: acoustic correlates and comparison with female data. *Journal of Acoustic Society of America*, 106(2), 1064–1077. doi: 10.1121/1.427116
- Hartl, D. H., Hans, S., Vaissiere, J., Riquet, M., & Brasnu, D. F. (2001). Objective voice quality analysis before and after onset of unilateral vocal fold paralysis. *Journal of Voice*, 15 (3):351–361.
- Hanson, H. M., Stevens, K. N., Hong-Kwang, J. K., Chen, M. Y., & Slifka, J. (2001). Towards models of phonation. *Journal of Phoniatics*, 29, 451–480.
- Hanson, H. M. (1995a). Glottal characteristics of female speakers. Ph.D.thesis, Harvard University, Cambridge, MA
- Hanson, H. M. (1995b). Synthesis of female speech using the Klatt synthesizer. *Speech Communication Group Working Papers 10*, Research Laboratory of Electronics, M.I.T., 84–103.
- Hanson, H. M. (1997). Glottal characteristics of female speakers: Acoustic Correlates. *Journal of Acoustic Society of America*, 101, 466–481. doi:10.1121/1.417991
- Hanson, H. M., & Chuang, E. S. (1999) Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Journal of American Speech and Hearing Association*, 106 (2), 1064–1077. doi:10.1121/1.427116
- Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal and loud voice. *Journal of Acoustic Society of America*, 84, 511–529. doi:10.1121/1.396829.
- Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1989). Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch. *Journal of Voice*, 4, 294–305. doi:10.1016/S0892-1997(89)80051-7.
- Karlsson, I. (1992). Analysis and synthesis of different voices with emphasis on female speech. Ph.D. thesis, Royal Institute of Technology, Stockholm.
- Karlsson, I. (1992). Evaluations of acoustic differences between male and female voices: A pilot study. *Speech Trans. Lab. Q. Prog. Stat. Report 1*, Royal Institute of Technology, Stockholm, pp. 19–31.
- Klatt, D., & Klatt, L. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of Acoustic Society of America*, 87, 820-57. doi:10.1121/1.398894
- Michael, P. C., Eugene, H. B., & Lesya, B. C. (2004). Spectral Amplitude Measures of Adductor Spasmodic Dysphonic Speech. *Journal of Voice*, 19 (3), 391–410.
- Nygaard, L. C., Sommers, M. S., & Posoni, D. B. (1995). Effects of stimulus variability on perception and representation of speaking words in memory. *Perception and Psychophysics*, 57, 989-1001. doi: 10.3758/BF03205458
- Perkell, J. S., Hillman, R. E., & Holmberg, E. B. (1994). Group differences in measures of voice production and revised values of maximum airflow declination rate. *Journal of Acoustic Society of America*, 96, 695–698. doi:10.1121/1.410307
- Scherer, K., Banse, R., & Wallbott, H. (2001). Emotional inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, 32 (1), 76-92. doi:10.1177/0022022101032001009
- Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, 3, 2, 621–640. doi: 10.1111/j.1749-818X.2009.00125.x

- Sodersten, M., & Lindestad, P. A. (1990). Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech Language Hearing Research*, 33, 601–611. Retrieved from <http://jshlr.asha.org/cgi/content/abstract/33/3/601>
- Stevens, K. N. (1998). *Acoustic Phonetics*, MIT Press, Cambridge, MA.
- Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *Journal of Acoustic Society of America*, 85, 1699–1707. doi: 10.1121/1.397959
- Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: a synthesis of acoustic and anatomic data. *Journal of Speech Language Hearing Research*, 50 (60), 1510-1545. doi:10.1044/1092-4388(2007/104)
- Whiteside, S. P. (1998). The identification of a speaker's sex from synthesized vowels. *Perception and Motor Skill*, 87, 595-600. doi: 10.2466/pms.1998.86.2.579.

TABLES

Table 1. Mean and standard deviation (SD) values for all the amplitude difference measures for phonation sample across males and females.

Groups		Males		Females		Difference (F-M)
Phonation Sample	Measures	Mean	SD	Mean	SD	
/a/	H1-H2	7.18	3.70	11.49	2.73	4.31
	H1-A1	6.70	2.53	11.17	4.54	4.47
	H1-A2	9.64	4.76	12.73	3.00	3.09
	H1-A3	24.53	6.06	28.79	5.41	4.26
/i/	H1-H2	4.47	2.88	11.09	3.52	6.62
	H1-A1	4.55	4.95	26.35	8.36	21.8
	H1-A2	21.15	9.42	28.00	9.09	6.85
	H1-A3	17.92	5.11	29.56	4.07	11.64
/u/	H1-H2	5.41	4.54	10.94	2.37	5.53
	H1-A1	5.05	3.61	29.09	7.69	24.04
	H1-A2	13.89	5.73	26.42	6.07	12.53
	H1-A3	27.57	6.95	36.31	7.32	8.74

Table 2. Independent t test values of the harmonic amplitude difference measures across gender for all phonation samples.

Groups		t-test for equality of means	
Phonation Sample	Measures	t	Sig. (2-tailed)
/a/	H1-H2	-3.844	.001
	H1-A1	-2.935	.007
	H1-A2	-2.394	.022
	H1-A3	-2.170	.038
/i/	H1-H2	-13.94	.000
	H1-A1	-12.44	.000
	H1-A2	-2.239	.032
	H1-A3	-5.608	.000
/u/	H1-H2	-10.32	.000
	H1-A1	-15.63	.000
	H1-A2	-3.554	.001
	H1-A3	-3.825	.000

FIGURES

Figure 1. Mean values for all the amplitude difference measures for phonation sample /a/ across males and females.

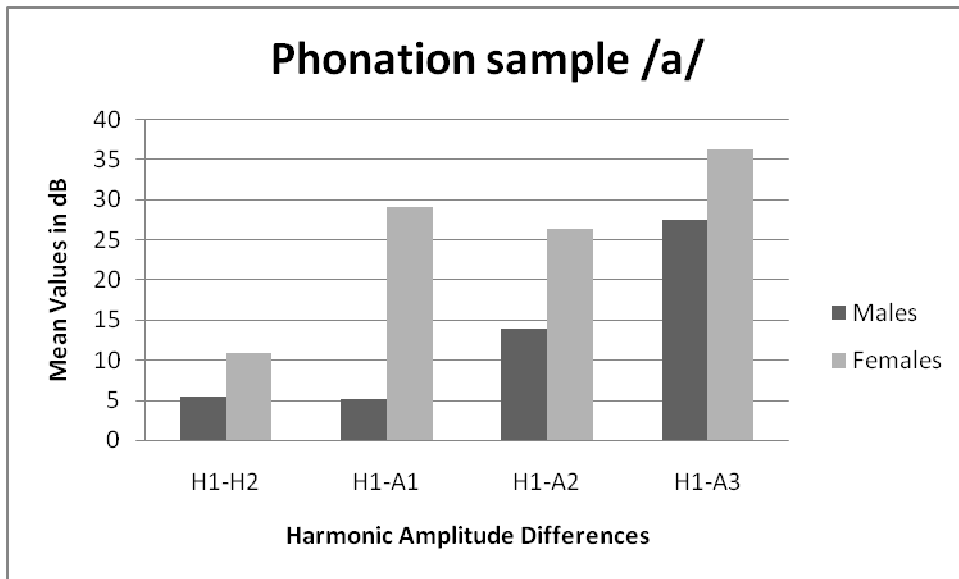


Figure 2. Mean values for all the amplitude difference measures for phonation sample /i/ across males and females.

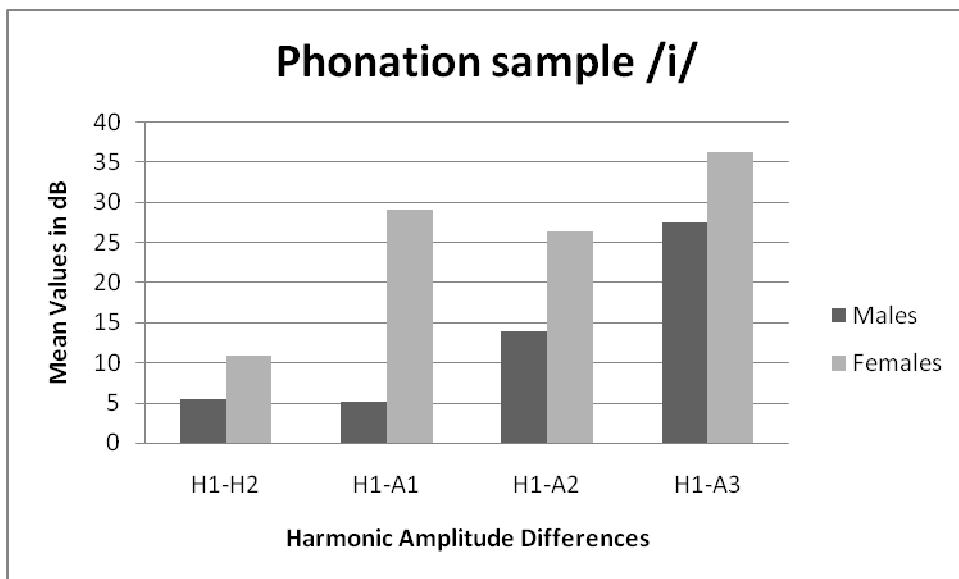
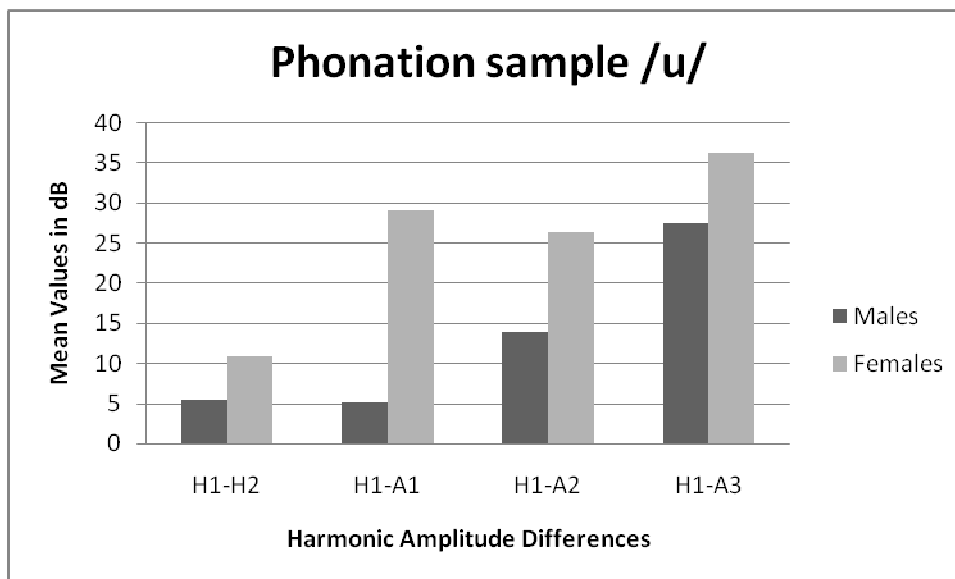


Figure 3. Mean values for all the amplitude difference measures for phonation sample /u/ across males and females.



The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:

<http://www.iiste.org>

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: <http://www.iiste.org/journals/> All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: <http://www.iiste.org/book/>

Academic conference: <http://www.iiste.org/conference/upcoming-conferences-call-for-paper/>

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digital Library, NewJour, Google Scholar

