

The fundamental frequency (f0) distribution of American speakers in a spontaneous speech corpus*

Byunggon Yang**

Department of English Education, Pusan National University, Busan, Korea

Abstract

The fundamental frequency (f0), representing an acoustic measure of vocal fold vibration, serves as an indicator of the speaker's emotional state and language-specific pattern in daily conversations. This study aimed to examine the f0 distribution in an English corpus of spontaneous speech, establishing normative data for American speakers. The corpus involved 40 participants engaging in free discussions on daily activities and personal viewpoints. Using Praat, f0 values were collected filtering outliers after removing nonspeech sounds and interviewer voices. Statistical analyses were performed with R. Results indicated a median f0 value of 145 Hz for all the speakers. The f0 values for all speakers exhibited a right-skewed, pointy distribution within a frequency range of 216 Hz from 75 Hz to 339 Hz. The female f0 range was wider than that of males, with a median of 113 Hz for males and 181 Hz for females. This spontaneous speech corpus provides valuable insights for linguists into f0 variation among individuals or groups in a language. Further research is encouraged to develop analytical and statistical measures for establishing reliable f0 standards for the general population.

Keywords: f0, distribution, statistics, variability, American English corpus, spontaneous speech

1. Introduction

According to Fant's (1973) source-filter theory, the fundamental frequency(f0) represents the sound source generated by the vibrations of the speaker's vocal folds. This source undergoes filtration through the speaker's vocal tract shapes, leading to the production of segmental sounds essential for communication. Throughout this process, the f0 fluctuates in response to the speaker's emotional state and the overall intonation patterns inherent in English sentence structure. The variability observed in everyday

English conversations may serve as a valuable criterion for assessing spoken English. Boothroyd (1986) provided a detailed summary of the physiological mechanism underlying vocal fold vibration, examining factors such as mass, length, and tension. Yang (1990, 2021) presented the explanations given by previous researchers regarding the associated muscle activities and the resulting f0 ranges observed for both males and females.

Numerous studies have been conducted to explore the distribution of f0 values in various language contexts (Hudson et al., 2007; Lennes et al., 2016; Lindh, 2006; Yang, 2021, 2023). Hudson et al.

* This work was supported by a Humanities·Social-Science Research Promotion of Pusan National University (2022).

** bgyang@pusan.ac.kr, Corresponding author

Received 30 January 2024; Revised 8 March 2024; Accepted 8 March 2024

© Copyright 2024 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://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.

(2007) examined 100 young male speakers of Southern British English during police dialogues, observing a generally normal f0 distribution. However, 60% of speakers had overlapping f0 values, posing limitations in speaker identification. Lennes et al. (2016) compared f0 distributions in two Finnish conversational speech corpora. The study included ten conversations among young, native Finnish-speaking adults and shorter dialogues involving eight adult men and women. The analysis focused on segments of the audio signals corresponding to utterances within the conversations. The reported overall median pitch for males was 117.5 Hz, and for females, it was 191.3 Hz. The f0 distributions for both male and female groups exhibited a more or less right-skewed pattern with a single peak of similar width. Lindh (2006) studied 109 Swedish male speakers' natural speech f0 values, reporting an average of 120.8 Hz with a median of 115.8 Hz. Skewness was 0.6, suggesting a preference for using median over mean due to extreme values. Yang (2021) analyzed f0 distribution in the Seoul Korean corpus of 40 Korean participants (Yun et al., 2015), finding a median f0 of 148 Hz with a wide range (65 Hz to 274 Hz) in daily speech. Gender differences were noted, with females exhibiting a median f0 almost twice that of males (200 Hz vs. 111 Hz). Regression analysis indicated a significant but low predictive accuracy of age on f0 values, leading to the suggestion for more diverse participant data analysis in corpora. In order to provide insights into normative f0 distribution for different Korean age and sex groups, Yang (2023) analyzed the f0 distribution in a dialogue speech corpus of 2,740 Korean speakers, using Praat and R. After filtering extreme values, he showed an average f0 of 185 Hz, a median of 187 Hz, and a positively skewed distribution. The pitch values in their daily conversations varied within a 238 Hz range. Gender-based analysis revealed distinct median f0 values (114 Hz for males, 199 Hz for females) with a significant difference. Skewness and kurtosis values indicated leptokurtic distribution for males. Regression analysis between median f0 and age showed a positive slope for males and a negative slope for females, suggesting a divergent relationship.

Some studies pursued cross-linguistic differences in fundamental frequency range (FFR) focusing on the impact of first language (L1) transfer and psychological factors (Mennen et al., 2014; Ordin & Mennen, 2017; Scharff-Rethfeldt et al., 2008; Zimmerer et al., 2014). Mennen et al. (2014) investigated German learners of English, discovering that their FFR values were intermediate between native German and English speakers. They noted the advantage of involving second language (L2) learners in controlling physiological differences but highlighted the potential for L1 transfer. Studies on Dutch learners of Greek and Czech learners of English suggested that compressed FFR might result from lack of confidence or anxiety associated with speaking a foreign language. Ordin & Mennen (2017) investigated cross-linguistic differences in FFR in 30 Welsh-English bilinguals, who read 35 semantically matched sentences in each language. They explored the behavioral pattern of FFR using measures of span (range of fundamental frequency covered by the speaker's voice) and level (overall height of fundamental frequency maxima, minima, and means of speaker's voice). They found that most female bilinguals showed distinct FFRs for each language but most male bilinguals maintained their FFR when switching languages. They discussed such possible sources for the differences as dissimilarities in intonational structure, cultural and social norms, ethological factors, and differences in the anatomy or physiology of speakers. In addition,

Scharff-Rethfeldt et al. (2008) found evidence of FFR transfer in German-English bilinguals, with higher FFR in their German than that in their English. Zimmerer et al. (2014) speculated that a trend for compressed pitch range in L2 speech production could be attributed to insecurity and lack of confidence. Those studies provide an insight into the language-inherent f0 values and further research on fundamental frequency values of bilingual or multilingual speakers.

To date, there is a limited number of studies providing an f0 distribution for American speakers engaged in daily conversations. This study seeks to address this gap by exploring the f0 values of American speakers in spontaneous speech, with the intention of facilitating a cross-linguistic comparison. The Buckeye Corpus had included such nonspeech sounds as the vocal noises and laughters. Moreover, the interviewers' voices in a very low amplitude were also present in the recording. There should be a method to remove those irrelevant sound data in order to examine any normative f0 values of the interviewees. The primary objectives of this research include: (1) examining the f0 data of American speakers; (2) analyzing both group and individual f0 variability; (3) investigating the relationship between f0 and age. The findings will offer valuable insights into the application of fundamental frequency in everyday conversations in American English.

2. Method

2.1. The Buckeye Corpus and Participants

The Buckeye Corpus of conversational speech comprises recordings from 40 American speakers, both male and female, who were born in or near Columbus, Ohio (Pitt et al., 2007). The study enlisted 20 men and 20 women, segregated into younger and older groups based on a threshold age of 40 years, each consisting of 20 participants. The participants engaged in hour-long friendly conversations with either a male postdoc or a female graduate assistant, covering topics such as politics and sports. To ensure a homogeneous target population, the corpus team screened individuals carefully through brief telephone interviews. Then, the team meticulously listened to the recordings, transcribing not only the interviewees' actual pronunciation using phonetic symbols but also capturing non-speech sounds like laughter, silence, noise, and the interviewer's voice.

2.2. Data Measurements and Analyses

Fundamental frequency values were gathered using Praat (v.6.3.14, Boersma & Weenink, 2023), and their distribution was subjected to statistical analyses using R (v.4.3.1, R Core Team, 2023).

Several Praat scripts were developed to refine the sound files and obtain accurate and reliable f0 values. Initially, a script was employed to silence non-speech sounds and the interviewer's voice by setting them to zero amplitude, referencing the labels of these segments in each transcript file. Another script measured f0 values for all files at 20 ms intervals within a range of 75 Hz to 600 Hz, storing the values on a notebook computer.

The initial f0 values of the forty speakers exhibited a broad distribution with numerous outliers. To address this, the author generated a boxplot in R and determined the interquartile range (IQR) for each dataset. Subsequently, the initial dataset underwent

filtration to eliminate values falling below and beyond 1.5 times the IQR. The following code illustrates this process, demonstrating the application to the original f_0 values of the first participant (S01) in the data frame (S01df) to create a new data frame (S01df_new) without outliers.

```
quartiles <- quantile(S01df$originalf0, probs = c(.25, .75),
  na.rm = TRUE)
IQR <- IQR(S01df$originalf0, na.rm = TRUE)
Lower <- quartiles[1] - 1.5 x IQR
Upper <- quartiles[2] + 1.5 x IQR
S01df_new <- subset(S01df$originalf0, S01df$originalf0 >
  Lower & S01df$originalf0 < Upper)
```

Figure 1 displays two boxplots representing the original data and the filtered data. Outliers are depicted as overlapped circles outside the second and third quartiles. While a few outliers are still noticeable in the filtered data, they appear negligible compared to those in the original dataset.

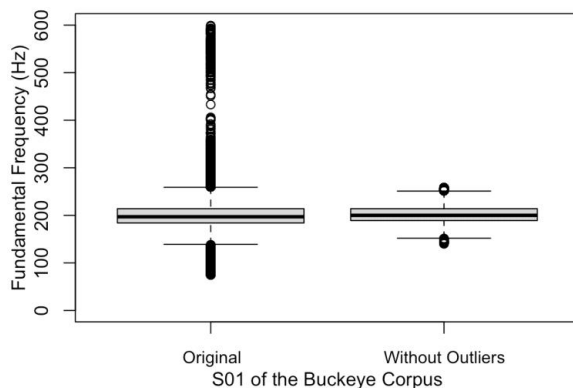


Figure 1. Boxplots of S01's original fundamental frequency (f_0) values (left) and outlier-filtered values (right).

Statistical analyses of the f_0 distribution were executed using R. An R code, proposed by Lennes et al. (2016), was adapted to generate probability density curves illustrating the f_0 distribution concerning age and sex groups. Key statistical parameters, including mode, median, mean, standard deviation (SD), and the minimum and maximum values were compiled to discuss both group and individual variability. Additional statistical measures, such as skewness and kurtosis, were acquired to assess group characteristics.

3. Results and Discussion

3.1. The f_0 Distribution of the American English Speakers

Table 1 provides a statistical summary of all the f_0 values from American speakers, comprising a final dataset of 1.774 million values.

Table 1. Statistics of all the f_0 values of forty American speakers in the Buckeye Corpus. n denotes the number of the f_0 values. Min indicates the minimum f_0 value while Max represents the maximum one (unit: Hz)

n	Mode	Median	Mean	SD	Min	Max
1,774,443	114	145	149	50	75	395

SD , standard deviation.

The mode in the Buckeye Corpus was observed at 114 Hz, with a median of 145 Hz and a mean of 149 Hz. Although the mean was the highest among the three major measurements, the difference between the median and mean values was small (4 Hz). In contrast, the difference between the mode and median was more substantial at 31 Hz, akin to the Korean corpus with a 46 Hz difference (Yang, 2021). These differences may be linked to the right-skewed f_0 values in the higher frequency region, where higher f_0 values could have influenced the mean and median. Yang (2021) reported a median f_0 value of 148 Hz from the Korean corpus, with a range from the maximum to the minimum f_0 values at 559 Hz. He reported 102 Hz for the Korean mode and 160 Hz for the Korean mean. Here the f_0 range of the American corpus records 320 Hz. The wide range difference between the American and Korean corpora might be related to the removal of the American outliers considering each individual IQR. On the other hand, Lennes et al. (2016:44) provided a statistical summary of the forty individual speakers. The average of their median values was 147 Hz spanning from 70 Hz to 199 Hz. The mode and mean of the Finnish speakers recorded 139 Hz, and 154 Hz. Considering extreme outliers derived from acoustic measurement errors, median values might best represent the inherent f_0 for a given language corpus. The skewness of the American corpus was 0.78, which indicated the frequent f_0 values at the lower end, and the tail pointed toward the higher end of the scale (Field, 2013). In addition, the kurtosis was 0.86 and there were three pointy peaks below 200 Hz. On the other hand, the skewness of the Korean corpus was 0.92, which represented a positive skew while that of the kurtosis was 0.94, which yielded a leptokurtic or pointy distribution.

Figure 2 presents a box plot of f_0 values collected from all the American speakers every 20 ms from nearly 40 hours of the American English corpus.

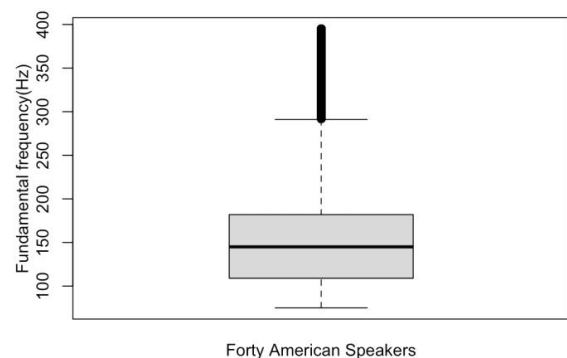


Figure 2. Boxplot of f_0 values (Hz) of forty speakers in the Buckeye Corpus.

In the boxplot, outliers are generally identified beyond these limits of a lower bound at 75 Hz and an upper bound at 291 Hz. The frequency band, spanning 216 Hz between these bounds, characterizes the typical f_0 range of American English speakers. Notably, there is a comparatively short tail of outliers around 100 Hz above the upper bound, which is half of the outlier range depicted in Figure 1. This limited range of outliers is likely a result of filtering individual data. The lower quartile registers at 109 Hz, while the upper quartile reaches 182 Hz, yielding an IQR of 73 Hz. To eliminate outliers further, an upper bound can be approximated by multiplying the IQR by 1.5 and adding the third quartile value, resulting in 291.5

Hz, which is only 0.5 Hz above the upper bound.

3.2. f0 Statistics of Sex Groups

Table 2 lists basic statistical measures of the American English data grouped by sex.

Table 2. Statistics of f0 values grouped by sex in the Buckeye Corpus. *n* denotes the total number of the measured f0 values. Min indicates the minimum f0 value while Max represents the maximum one (unit: Hz)

Group	n	Mode	Median	Mean	<i>SD</i>	Min	Max
Male	918,615	114	113	116	28	75	247
Female	855,828	175	181	185	43	75	395

SD, standard deviation.

The mode for male speakers registered at 114 Hz, aligning with the overall mode across all the speakers. Combining male and female data, the prevalence of higher peaks in the male data resulted in a shared mode for both groups. Specifically, female speakers exhibited a mode of 175 Hz, indicating a distinct 61 Hz contrast between the two groups.

In contrast, Korean speakers, as reported by Yang (2021), showcased a median of 111 Hz for the male group and 200 Hz for the female group. Also, the median of the twenty Finnish male speakers amounted to 110 Hz while that of the female speakers listed as 184 Hz (Lennes et al., 2016). These figures suggest slightly higher or lower values compared to their American counterparts. It is noteworthy that the Finnish participants predominantly consisted of young adults, and the potential clipping of f0 values below the threshold might have occurred due to the Praat software's lower limit of 75 Hz. This crosslinguistic distinction may stem fundamentally from anatomical differences between Americans and Koreans. Notably, the greater f0 in the female group could be linked to anatomical sex differences in the length and mass of vocal folds (Boothroyd, 1986).

Figure 3 illustrates two boxplots presenting the fundamental frequency values of male and female speakers collectively. A noticeable and substantial difference is apparent between the two groups. The boxplot for female speakers reveals a broader distribution range, featuring additional outliers identified by small circles positioned either above or below the quartile division.

Additionally, the f0 range for the female group appeared wider than that of the male group. For a meaningful crosslinguistic comparison, it is advisable to recruit participants within the same age range and control for speech style, be it reading or conversational. Failure to do so might introduce biases in any crosslinguistic comparison, influenced by these intervening factors.

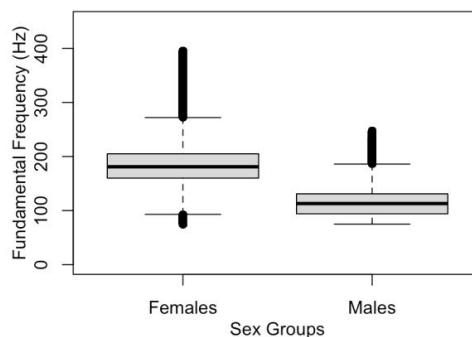


Figure 3. Boxplot of f0 values (Hz) of female (left) and male (right) groups in the Buckeye Corpus.

3.3. f0 Statistics of Individual Speakers

Table 3 provides a summary of the basic statistics for each American speaker.

Table 3. Statistical summary of f0 values of each individual speaker of the Buckeye Corpus (unit: Hz)

Subj	Mode	Median	Mean	<i>SD</i>	Subj	Mode	Median	Mean	<i>SD</i>
S01	191	200	203	19	S21	185	190	193	16
S02	156	172	178	38	S22	123	127	129	14
S03	120	128	131	23	S23	114	120	123	19
S04	188	194	196	21	S24	76	82	85	9
S05	153	157	161	19	S25	198	202	207	25
S06	108	110	110	7	S26	162	168	172	20
S07	177	199	205	63	S27	118	116	119	22
S08	178	193	198	29	S28	95	95	97	11
S09	152	154	156	13	S29	116	117	120	22
S10	126	134	138	23	S30	91	98	101	14
S11	99	102	104	12	S31	176	183	187	17
S12	200	229	240	54	S32	119	127	130	15
S13	81	83	84	5	S33	158	163	165	18
S14	166	170	173	16	S34	79	87	91	13
S15	112	118	121	16	S35	82	86	89	10
S16	167	177	182	23	S36	76	82	85	9
S17	148	158	161	36	S37	162	176	176	38
S18	162	172	176	24	S38	97	103	107	14
S19	124	139	148	34	S39	198	205	202	40
S20	157	158	160	24	S40	117	121	123	12

SD, standard deviation.

The mode range spans 124 Hz, ranging from 76 Hz (S24) to 200 Hz (S12). The lowest median, observed in S24, is 82 Hz, while the highest median, found in S12, reaches 229 Hz. Positioned between the mode and mean, the median offers a central tendency measure. In this context, the mean range is slightly broader than the preceding measures, extending from 84 Hz (S13) to 240 Hz (S12). The average *SD* is 21 Hz, ranging from 5 Hz (S13) to 63 Hz (S07). The relatively narrow *SD* may be attributed to the outlier filtering procedure using the IQR specific to each individual speaker. In comparison, Yang (2021) reported an *SD* range of 55 Hz, varying from 16 Hz to 71 Hz for the Korean corpus. Also, Lennes et al. (2016) provided an average *SD* of 31 Hz, covering from 17 Hz to 49 Hz.

In Figure 4, probability density curves delineate the f0 values of individual male and female speakers within the American English corpus. The author used distinct colors, with the f0 values of male speakers depicted in blue and those of female speakers in red. Generally the female data appear flat across a wide frequency range while those of the male speakers look more pointy peaks around 100

Hz, which we will check by the measure of kurtosis grouped by the sex and age factors in the following section.

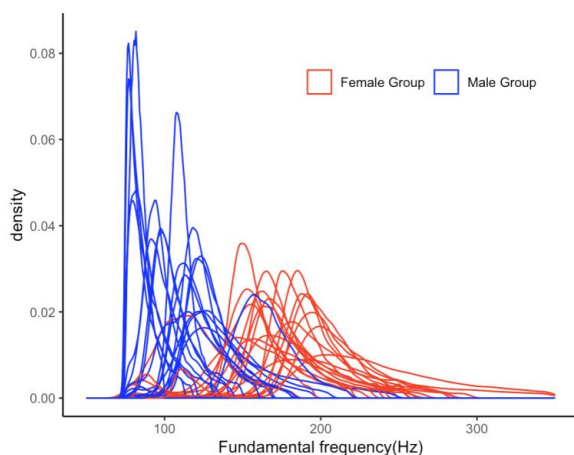


Figure 4. Density plot of f_0 values (Hz) of forty individual speakers colored by sex in the Buckeye Corpus.

In the plot, the skewed distribution observed can be partly attributed to the length of vocal folds, which sets a natural lower limit to glottal frequency. However, speakers possess the capability to extend their vocal folds, enabling the production of higher f_0 (Lennes et al., 2016). As a result, a prolonged tail is noticeable in each individual density curve. Yang (2021) advocated for the mode as a suitable measure to describe individual speaker characteristics but proposed the median as a better index, particularly considering the probability density graph.

Notably, two outliers are discernible within each sex group in Figure 4. Participant S33, a young male (YM), exhibits a median of 163 Hz, placing his data in the lower frequency range of the female group. Conversely, participant S27, an old female (OF), has a median f_0 value of 116 Hz, aligning more closely with the male data range. Consequently, defining a normative f_0 range for the two sex groups, inclusive of these outliers, might be misleading. It is recommended that further studies undertake a rigorous evaluation of the data before assigning a conclusive statistical summary to each group.

3.4. Relationship between Age and f_0

As the Buckeye Corpus did not specify participant ages numerically, we opted to explore the general association between age and f_0 values. To achieve this, we categorized speakers by both age and sex within the American English Corpus. Table 4 presents the statistical details of f_0 values for male and female speakers in these grouped categories.

Table 4. Median and other major statistics on f_0 values in Hz of male and female speakers by age (O: old; Y: young) and sex (M: male; F: female) in the Buckeye Corpus

Group	n	Median	Min	Max	Skew	Kurtosis
YM	455,397	106	75	219	1.10	1.20
OM	463,218	120	75	247	0.78	1.03
YF	363,058	192	85	395	1.16	3.31
OF	492,770	172	75	374	0.62	1.98

YM, young male; OM, old male; YF, young female; OF, old female.

The number of f_0 values across the three groups appears comparable, except for the young female (YF) group, which registers ten thousand counts less. Among these groups, the YM group exhibits the lowest median, while the young female group demonstrates the highest median. All skewness values are positive, signifying right-skewed f_0 values that taper toward the higher frequency region. The group difference range between young and old females is 0.54. Conversely, all kurtosis values are also positive. A kurtosis of zero typically indicates a bell-shaped or normal distribution, whereas the present kurtosis values exceeding one suggest a relatively peaked distribution. Figure 5 visually displays the density distributions for all four groups.

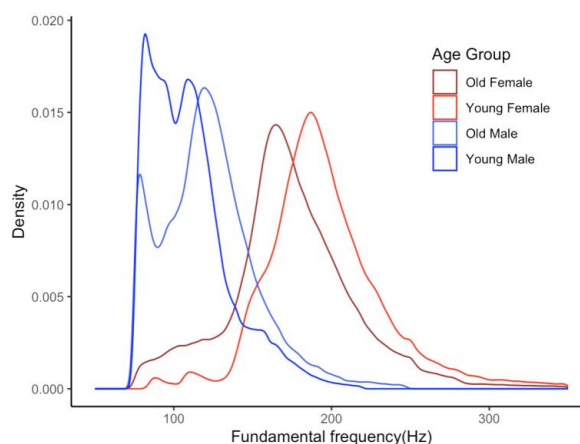


Figure 5. Probability density plot of f_0 values (Hz) of forty speakers grouped by age and sex in the Buckeye Corpus.

In Figure 5, it becomes apparent that the f_0 values of the YM group are lower than those of the old male group. Conversely, the f_0 values of the YF group are higher than those of the OF group. This age-group mismatch complicates any conclusive statement about older groups generally having higher f_0 values. Notably, two outliers within each sex group have already been identified. Yang's (2021, 2023) studies have also tentatively explored the relationship between age and f_0 values. These findings may be associated with the work of Hollien & Shipp (1972), who examined the correlation between age and f_0 in a sample of 175 individuals aged 20 to 89. Their results indicated a gradual decrease in f_0 values from 20 to 40 years, followed by an increase from 60 to 80 years. In contrast, male and female f_0 values regressed to the mean of the two groups, as female f_0 values decreased after middle age (Baken, 2005). Moreover, Reubold et al. (2010) investigated several aspects related to changes in fundamental frequency (f_0) and formant frequencies (F1) over time and their effects on age perception. Their findings revealed that f_0 and the first formant of schwas decreased over roughly 30 years in speakers, consistent with previous research indicating a decline in f_0 with increasing age. Additionally, f_0 and F1 changed at similar rates over a 50–60 year period in both male and female speakers, with evidence of a V-shaped trend in f_0 and a similar pattern in F1. As highlighted by Yang (2023), more comprehensive data collection, including biological data, might clarify the relationship between age and f_0 values.

One could deduce that the young groups are deviating from the group median value, while the old groups are converging towards the median. This trend might be more accurately described as a

regression toward the center of the group mean.

4. Summary and Conclusion

Variability in f0 during English conversations serves as a valuable criterion for spoken English assessment. Several studies have explored f0 distribution in various language contexts. However, there is a limited exploration of f0 distribution in American English daily conversations. This study aimed to fill this gap using the Buckeye Corpus. The method involved categorizing forty American speakers by age and sex, silencing non-speech sounds, and using Praat and R for f0 measurements and statistical analyses.

The results showed that the American speakers produced an f0 median of 145 Hz, which is quite comparable to 148 Hz of the Korean corpus, and 147 Hz of the Finnish corpus. The overall f0 values of the American corpus were right-skewed with a pointy distribution. Analysis by sex showed a notable difference between male and female f0 values, 113 Hz and 181 Hz, respectively with a wider range for females. Individual speaker statistics revealed two outliers in male and female groups. In addition, age-grouped f0 values revealed differences in median values between young and old groups. Young groups deviated from the median, while old groups converged towards it, representing a regression toward the center of the group mean.

In conclusion this research provides valuable insights into f0 distribution in American English daily conversations, emphasizing the need for more comprehensive data collection for cross-linguistic comparisons. The study contributes to the understanding of f0 variations, highlighting the impact of age and sex on speech patterns in American English.

References

- Baken, R. J. (2005). The aged voice: A new hypothesis. *Journal of Voice*, 19(3), 317-325.
- Boersma, P., & Weenink, D. (2023). Praat: Doing phonetics by computer (Version 6.3.14) [Computer Program]. Retrieved from <http://www.fon.hum.uva.nl/praat/>
- Boothroyd, A. (1986). *Speech acoustics and perception*. Austin, TX: Pro-ED.
- Fant, G. (1973). *Speech sounds and features*. Cambridge, MA: MIT Press.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London, UK: Sage Publications.
- Hollien, H., & Shipp, T. (1972). Speaking fundamental frequency and chronologic age in males. *Journal of Speech and Hearing Research*, 15(1), 155-159.
- Hudson, T., de Jong, G., McDougall, K., Harrison, P., & Nolan, F. (2007, August). F0 statistics for 100 young male speakers of Standard Southern British English. *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken, Germany.
- Lennes, M., Stevanovic, M., Aalto, D., & Palo, P. (2016). Comparing pitch distributions using Praat and R. *Phonetician*, 111(2), 35-53.
- Lindh, J. (2006). Preliminary descriptive f0-statistics for young male speakers. In G. Ambrazaitis, & S. Schötz (Eds.), *Working papers 52: Papers from Fonetik 2006* (pp. 89-92). Lund, Sweden: Department of Linguistics, Lund University.
- Mennen, I., Schaeffler, F., & Dickie, C. (2014). Second language acquisition of pitch range in German learners of English. *Studies in Second Language Acquisition*, 36(2), 303-329.
- Ordin, M., & Mennen, I. (2017). Cross-linguistic differences in bilinguals' fundamental frequency ranges. *Journal of Speech, Language, and Hearing Research*, 60(6), 1493-1506.
- Pitt, M., Dille, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye corpus of conversational speech* (2nd ed.). Columbus, OH: Ohio State University.
- R Core Team. (2023). R: A language and environment for statistical computing (version 4.3.1) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication*, 52(7-8), 638-651.
- Scharff-Rethfeldt, W., Miller, N., & Mennen, I. (2008). Unterschiede in der mittleren Sprechtonhöhe bei Deutsch/Englisch bilingualen Sprechern. *Sprache, Stimme, Gehör*, 32(3), 123-128.
- Yang, B. (1990). *Development of vowel normalization procedures: English and Korean* (Doctoral dissertation). The University of Texas, Austin, TX.
- Yang, B. (2021). The f0 distribution of Korean speakers in a spontaneous speech corpus. *Phonetics and Speech Sciences*, 13(3), 31-37.
- Yang, B. (2023). The fundamental frequency (f0) distribution of Korean speakers in a dialogue corpus using Praat and R. *Phonetics and Speech Sciences*, 15(3), 17-25.
- Yun, W., Yoon, K., Park, S., Lee, J., Cho, S., Kang, D., Byun, K., ... Kim, J. (2015). The Korean corpus of spontaneous speech. *Phonetics and Speech Sciences*, 7(2), 103-109.
- Zimmerer, F., Jügler, J., Andreeva, B., Möbius, B., & Trouvain, J. (2014, May). Too cautious to vary more? A comparison of pitch variation in native and non-native productions of French and German speakers. *Proceedings of Speech Prosody*. Dublin, Ireland.

• **Byunggon Yang**, Corresponding author
Professor, Dept. of English Education
Pusan National University
30 Changjundong, Keumjunggu, Busan 46261, Korea
Tel: 051-510-2619
Email: bgyang@pusan.ac.kr
Homepage: <http://fonetiks.info/bgyang>
Fields of interest: phonetics, phonology