Journal of Convergence for Information Technology
Vol. 9. No. 9, pp. 158-165, 2019
e-ISSN 2586-4440
DOI : https://doi.org/10.22156/CS4SMB.2019.9.9.158

PRAAT Software: A Spech Interaction Tool to Analyze Teacher Voices

Ella Jane Kidd
Student, Department of British & American Language and Culture, Kyunghee University

PRAAT 소프트웨어: 교사 목소리 분석을 위한 맞춤법 상호작용 도구

엘라 키드
경희대학교 영미어학과 박사과정

Abstract Through the use of speech software technology, this paper examines the effects of voice interactions within the inner circle of English. The fundamental frequency (F0) was obtained by analyzing native speakers (aged 30-55) speech effects based on nationality, age, and gender. The findings within this study reveal that the Caucasian British female (age 33) and the Caucasian American male (age 55) produced the most interactive speech. The contributing factor is the students' experience with various language styles throughout their language acquisition studies. The results of this study are compatible with Traunmüller & Eriksson (1995) and previous studies which agree that continuous speech above average is paramount towards student engagement and interactions.

Key Words : Interaction, Frequency, Language learners, Native speakers, Language education

요 약 본 논문은 음성 소프트웨어 기술의 사용을 통해 영어의 내부 영역 내에서 음성 상호 작용의 영향을 검토한다. 기본주파수(F0)는 국적, 연령, 성별을 기준으로 원어민(연령 30-55세) 음성 효과를 분석해 얻었다. 이 연구에서 밝혀진 바에 따르면, 백인 영국 여성(33세)과 백인 미국인 남성(55세)이 가장 많은 대화형 연설을 했다는 것이다. 기여 요인은 학생들이 언어 습득 연구를 통해 다양한 언어 스타일을 경험한 것이다. 이 연구의 결과는 평균 이상의 연속적인 말이 학생의 참여와 상호작용에 가장 중요하다는 데 동의한 Traunmüeller & Eriksson(1995) 및 이전 연구와 양립할 수 있다.

주제어 : 상호작용, 빈도, 언어 학습자, 원어민, 언어 교육

1. Introduction

The theoretical concepts within this paper examine the fundamental frequency of speech (F0) during connected speech between American and British speakers in the English as a Foreign Language (hereafter, EFL) classroom. There are two main objectives of this paper. The first is to examine British and American speaker's speech effects in Hz, according to the data. The second is to investigate mean speech in sustained vowels according to nationality, age, and gender to determine the most liveliest expressions spoken.

Similar and recent data have been published examining several language varieties between various types of discourse. Most of the data are usually expressed in Hz, and the average measure for F0 is usually included [1]. According to [1], the

*This paper was funded by the academic research fund of Kyung Hee University in 2019.
*Corresponding Author : Ella Jane Kidd(ekidd2000@hotmail.com)
Received May 13, 2019 Revised August 14, 2019
Accepted September 20, 2019 Published September 28, 2019

typical F0 for men is 120 Hz, and for women, it is 210 Hz — the mean average changes slightly with age. To determine the effects of teacher interactions on learner involvement, the influence of real-time communications on the participation in the EFL classroom, the language must be understood. When teacher interaction is taken into consideration, the importance of socio-cultural theory becomes evident as the repetition of peer-to-peer and peer-to- teacher communications. The quality of classroom interaction has been the topic of extensive research in education [2]. The work of Lev Vygotsky's social-cultural theory in regards to the second language (hereafter, L2) learning environment is related to standard approaches to second language acquisition (henceforth, SLA)[2].

In Korea, effective teaching strategies and methodologies for the acquisition of English language communications skills are non-restrictive topics which have resulted in frequent modifications by the Ministry of Education. The current 7th National English Curriculum has increased attention towards teaching the English language in Korea, where educators and politicians have become interested in improving the English proficiency of students in public and private institutions. At a private university in Daejeon where the data for this study was compiled, students are now more exposed to different varieties of English. But in the classroom, the motivation for English acquisition still tends to lack ambition. However, low motivation could be a result of voice naturalness from the teacher and teacher voice quality, which can influence the credibility of speech, pitch, perceptions, identifiable loudness, and phonetic category. It also includes all of the perceptual dimensions of the spectral envelope, which changes over time [3].

This paper examines voice interaction and the vocal effects of male and female American and British native speakers. The PRAAT speech software was used to determine the mean values of speakers aged 30-55 years old [4].

2. Literature Review

2.1 The importance of motivation in the EFL classroom

Motivating students in the English Foreign Language (EFL) classroom is often mentioned as a complex and somewhat challenging task. Dornyei characterized motivation as dynamic human behavior and provided direction [5]. On the other hand, motivation can include the involvement of multiple psycho-sociological and linguistic factors [5,6].

The prominent psychologist [7], asserted that student language success is indicative of student attitudes towards a specific language and the incorporation of various aspects of that language into meaning. In other words, students' extrapolate their experiences into the real-world; this could arouse several socio-affective factors and characterize student experiences.

Although there are other coined definitions for motivation from previous researchers [8], but in this paper, motivation refers to the overall effect of liveliness or teacher interactions in the EFL classroom. Psychologists' [9,10], [11] are among the pioneers of studies who view L2 motivation, as the influence of learners' attitudes on the social perceptions of an L2 learner, its speakers, their inter-ethnic contact, and overall degree of linguistic self-confidence [5]. Why is motivation important in the EFL classroom?

Additionally, a large number of English teachers can verify that interactions significantly impact their pedagogical contribution to learner's EFL environment.

To conclude, the interaction component of learning is a vital element towards increasing student proficiency elevation [2,7].

2.2 Interactive teaching methods

According to linguist [12], he compared conventional and non-interactive methods with minimal student interactions and interventions; his research states that the pedagogical methods of interaction foster the comprehension and communicative competence of university students. The findings concluded with teachers who incorporate interaction within their lessons are aware of student learning goals. On the other hand, teachers who do not incorporate communication within their classroom pedagogical approach felt pressure. In the area of comprehension and interaction, [13], reports a specific method of how EFL learners' extract meaning from the sounds they hear through the interactive process at the onset of the smallest acoustic message, such as phonemes, comprehension begins. Henceforth, this process creates a combination and formation of audible signals that form words, phrase, clauses, and sentences. In like manner to the previous researches mentioned, [14], states that interactive classes are preferred by university students who prefer the active naturalistic use of the language. The data for [12] study was collected from students of four colleges, the college of business, primary education, technological and health sciences.

As a result, the more interaction implemented in the classroom the more participation received from the students. On the other hand, teachers who do not incorporate interaction within their lesson approach tend to feel more pressure in obtaining student engagement and interaction because the presence of student participation is minuscule. This study intertwines native speakers' voice interactions based on the motivation variable and its influence that voice frequency has on student interactions in the context of teaching EFL Korean University Students.

2.3 The significance of F0, SD variations in discourse type

Phonetics is the study of sounds made during human speech. In this study, the linguistic phonetic value functions as a demonstrative, and original value, which can produce various acoustic properties of speech sounds when reading aloud. According to [15], the type of text read aloud there is a significant effect on the SD of F0. However, the various types of discourse, such as 'conversation' compared with 'acting' reveal more significant results in a different SD [16].

The type of discourse in the university classroom is affected by the degree of liveliness. For this study, the researcher used speech materials, and the vowel segments were trimmed based on a single morphology, and the gender of the speakers. The cut vowel segments by speakers can be shown vis-a-vis in Table 1.

Table 1. Trimmed vowel segments by speaker's

Nationality	Trim vowel Segments
Caucasian American Male	One /wən/
African-American Male	One /wən/
Caucasian British Male	One /wən/
Caucasian British Female	You /o͞o/
African-American Female	You /o͞o/
Caucasian American Female	You /o͞o/

2.4 The implementation of PRAAT in the EFL classroom

Programmers have made multiple attempts to develop computer software to teach language. In the past years, computer software has augmented in the EFL field, specifically in audioliguistics. Through a speech tool called Praat software, it can be used to teach stress, intonation, record student speech, and provide feedback. However, it can also

display the stress and intonation which determines the result of voice interaction [17].

According to [17], a recent research was conducted to explore the prosodic feature through PRAAT software.

The results concluded that students were successful in terms of learning prosodic features. Previous research shows that students who feel a sense of accomplishment are more motivated. However, this study does not examine PRAAT software as a teaching tool, but as a voice examination tool for EFL teachers in terms of prosody (stress) and frequency. The outcome of this paper will show high spoken frequency contribution to student involvement. The frequency results are shown in the test results section of this paper.

2.5 Native and non-native speakers' interaction

The work of Lev Vygotsky partly sought out to compare and contrast native and non-native teacher interactive patterns within the linguistic paradigm. The Vygotskian Theory posits that students are in the hands of the "expert," making it possible to absorb a significant amount of language exposure. Therefore, the acquisition of L2 learning is a shared experience and not individual-based. On the other hand, when it comes to being placed in the hands of an expert, native, and non-native teachers tend to respond to the needs of the students differently [14]. According to one linguist [18], native speakers (NS) and non-native speakers (NNS) discourse tend to differ in relationship between interlocutors. A linguist [19], defined the linguistic differences of NS's and NNS's as "two different species" with various levels of proficiency and language behavior (p. 76). For example, the NS's have more experience in turn-taking and specific subject matters. On the other hand, NNS's do not respond to the speaker quickly when it comes to self-expression and ideas. Overall, NN's and NNS's are experts in the L2 classroom; the purpose of this study is to compare, contrast, and examine the interactive aspects of L2 learning.

3. Method

3.1 Experimental Technique

There were 80 participants' in this study, all who are university students aged 22-24 studying various majors. The volunteer speakers nationalities consisted of six native English speaker volunteers were Caucasian-British, Caucasian American, and African-American for the three females and Caucasian British, Caucasian American, and African-American for the three male speakers. Each speaker read from the students' required textbook. The pre-recorded files were saved in WAV format and analyzed in Praat [4]. The data collection were uploaded in Praat. Afterward, spectrograms were created to identify the frequency output for each speaker. A spectrogram also referred to as a spectrum waterfall or voice print, is a visual representation of the spectrum of sound frequencies. Can be used to identify phonetic spoken words. In this study, the spoken script samples were clipped to test the prolonged vowel sounds /o/, /a/, and /e/, as well as the frequencies of each speaker's entire recording. Table 1 shows the trimmed vowels according to gender. The word "One /wən/" was used for examining the male speakers, and "you /o͞o/" for the female speakers. Tables 2 & 3 present the scripts that were read by the female and male speakers.

Table 2. Native speaker dialogues females

You're in great shape. What do you do to keep fit?
I go swimming, or I do yoga.
How often do you do that?
Every morning, I love it.
Wow! How often do you go to the gym?
Ugh! I never go to the gym. I hate it. It's too crowded.
I see.
And what about you? What do you do to keep fit?
I usually go to the gym after work, and I play tennis about twice a week

Table 3. Native speaker dialogues males

Hey! That's Angelia Jolie.
Who?
Angelia Jolie, the actress. She's on this show.
Which one is she? What does she look like?
She's tall, and she has dark hair.
Oh, I see her. Do you know her?
Yes, I do. I was in high school with her.
Really? What's she like?
She's really nice. And she's very cool
I see. Was she your girlfriend in high school?
I wish!

3.2 Description of the procedure

The English prose for the listening files includes six volunteer native speakers from the United States and the United Kingdom. The pre-recorded files were saved in WAV format and uploaded into the Pratt speech analysis device [4]. Analyses of the volunteers' dialogues are presented in spectrograms (Figs. 1-6). A spectrogram shows a visual representation of sound frequencies, which can be used to identify spoken words phonetically.

The spectrograms in Figures 1-6 represent each volunteer speaker's sound recording. The arrow mark categorizes a segment of the audio and is divided by a blue and green line. The beginning vowel sound is the dark shaded path located in the frequency contour (F0)window or the bottom half of the reading. The arrows point to the beginning vowel segment.

The spectrograms' show that the Caucasian American Male (CAM) stressed the /wən/ vowel sound. The energy is visible to the last vowel ending. The Caucasian British Female (CBF) placed stress on /o/, /e/, and /i/ vowels throughout the segment. The arrow represents the beginning second vowel of the segment. The African-American Female (AAF) placed the greatest stress on the /o͞o / vowel segment. The African-American Male (AAM) placed very little vowel stress on the letter /ä/ sound. The Caucasian British Male (CBM) placed less stress on /wən/ than the Caucasian American-Female (AAF).

3.3 Sound Recordings

At the beginning of a standard voice assessment, every subject was asked not to sustain the vowels (/o/, /e/ and /i/). The recorded voice samples included the use of digital technology with an attached microphone (i.e., mobile device or laptop) and saved in WAV format. The vowel samples used in this study were edited to include only the middle three seconds. The read text and continuous speech samples were not modified. However, through Praat [4], the voice samples were concatenated in the following order: text segment and three-second sustained vowel segment. The provided examples of the concatenated waveforms can be seen in Figs. 1-6. The following were used in or created for this study: 28*sentences, 28*mid-vowel segments, and 28*concatenated files.

Fig. 1. displays the mean frequency values of the spoken text produced by the Caucasian-American male. The arrow points to the whispered vowel /wən/.

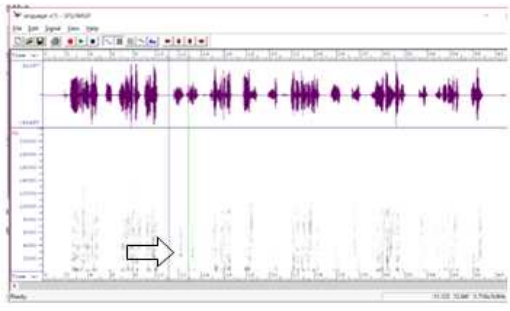

Fig. 2. displays the mean frequency values of the spoken text produced by the African-American Male. The arrow points to the whispered vowel /wən/.

Fig. 3. displays the mean values of the frequencies of the spoken text produced by the Caucasian-British male. The arrow points to the whispered vowel /wɐn/.

Fig. 4. displays the mean frequency value of the spoken text produced by the Caucasian-British female. The arrow points to the whispered vowel /o͞o/.

Fig. 5. displays the mean frequency value of the spoken text produced by the African-American female. The arrow points to the whispered vowel /o͞o/.

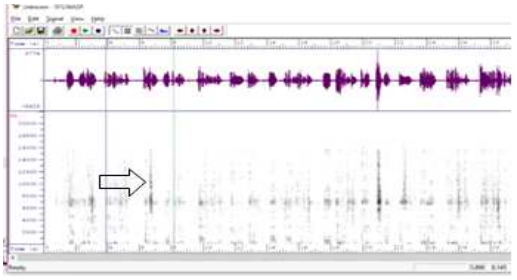

Fig. 6. displays the mean frequency values of the spoken text produced by the Caucasian-American female. The arrow points to the whispered vowel /o͞o/.

According to the pre-recorded questionnaire, the participants judged five speakers' speech rate as moderate speed. However, 15 respondents decided the CBF as "don't understand"; fourteen respondents rated the CBM as "don't understand"; the CAF ranked third amongst the other native speakers; the AAM ranked as the second-highest speaker with moderate speech, and the CAF and AAF was ranked as average.

4. Test Results and Discussion

A review of previous studies found similarities in the interactional patterns of males, but not for females. According to [1], speech increases slightly with age based on the measurement of the F0-excursions. In contrast to the findings of [1], the liveliest interactive effect in this study was produced by the youngest female. The results of this study are similar to the previous results where the older CAF produced the liveliest interactive effect.

In this study, the CBF speaker ranked above average on liveliness. Simultaneously, the CAF speaker ranked above average on liveliness, which exemplifies the significance of interaction as a vital supplement to the foreign language classroom. Whereby, minimizing the monotony in discourse [17, an increasing student engagement [19].

The following are the mean averages (F0) from lowest to highest. The CBM (age 33) and the CAF (age 55) produce the most interactive speech. The mean F0 produce by the AAF (age 47) was 205 Hz; for the CAF (age 37), it was 236 Hz, and for the CBF (age 33), it was 271 Hz. The mean speech (F0) produced by the AAM (age 34) was 113 Hz, for the CBM (age 51), it was 171 Hz, and for the CAF (age 55), it was 184 Hz. These results corroborate those of [1], namely that "typical" female speech might be perceived as being livelier or less lively than "typical" male speech.

Following the data that calculates the mean average (F0) from lowest to highest, the Caucasian

British Female (age 33) and the Caucasian American Male (age 55) produce the most interactive speech. The African-American Female (age 47) produces a mean speech (F0) of 205 Hz; the Caucasian American Female (age 37) produces a mean speech (F0) of 236 Hz, and the Caucasian British Female (age 33) produces a mean speech (F0) of 271 Hz. The African-American Male (age 34), produces a mean speech (F0) of 113 Hz; the Caucasian British Male (age 51) produces a mean speech (F0) of 171 Hz; the Caucasian American Male (age 55), produces a mean speech (F0) of 184 Hz. This result is compatible with [1], the 'typical' female speech might still be perceived as more lively or less lively than 'typical' male speech. The data results in this study reveal the Caucasian British Female speech is above the average measure of liveliness. However, the Caucasian American Male speaker is also above the average standard of liveliness. As a result, the participants in this study, are exposed to all kinds of voice qualities which could de facto base their experiences on various language styles. Their experiences are subsequent of their influence towards a specific vowel or continuous speech type. But, most importantly, the pedagogical approach, which includes constant speech with a frequency above average is vital in the role of interaction in the language classroom.

Lastly, the outcome of this study reveals that the voice interactions of native speakers spoken in the classroom at an (F0) of 170 Hz or greater for Male and Female tends to produce a lively effect, thereby breaking the simplicity in classroom discourse. This study implies that the sampling of teacher voice frequency through technology can predetermine the overall impact of student engagement in the Foreign language classroom. The idea of voice sampling could also assist in the potential search of prime teacher candidates for employment. This use of technology would show a concerted effort of efficiency to lawmakers, students, and parents.

To conclude, high voice frequency will increase student engagement, student motivation, and more extensive student conversation towards an increase in second language fluency. Therefore, further research in this field is necessary in this field.

REFERENCES

[1] H. Traunmüller & E. Anders. (1995). *The frequency range of the voice fundamental in the speech of male and female adults.* Unpublished Manuscript.

[2] B. Inan. (2012). A comparison of classroom interaction patterns of native and non-native EFL teachers. *Procedia Social and Behavioral Sciences, 46,* 2419-2423.

[3] Y. Maryn & N. Roy. (2012). Sustained vowels and continuous speech in the auditory-perceptual evaluation of dysphonia severity. *Journal de Sociedade Brasileira de Fonoaudiologia, 24(2),* 107-112.

[4] P. Boersma & D. Weenink. (2013). *Praat: doing phonetics by computer [windows 10]. Version 5.3.51,* retrieved 3 June 2017 from http://www.praat.org/.

[5] Z. Dornyei. (2010a). *The relationship between language aptitude and language learning motivation: Individual differences from a dynamic systems perspective.* In E. Macaro (Ed.) Continuum companion to second language acquisition (pp. 247-267). London: Continuum.

[6] Z. Dornyei & I. Otto. (1998). Motivation in action: A process model of L2 motivation. *Working Papers In Applied Linguistics. 4,* 43-69.

[7] R. C. Gardner. (1985). *Social psychology and second language learning: The role of attitudes and motivation.* London: Arnold.

[7] R. Bahous, N. N. Bacha & M Nabhani. (2011). Motivating Students in the EFL Classroom: A Case Study of Perspectives. *English Language Teaching, 4(3),* 33-43, 2011.

[8] A. Liuoliene & R. Metiuniene. (2006). Second language learning motivation. Santalka. *Filologija Edulologija, 14(2),* 93-98.

[9] H. Gardner. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

[10] R. C. Gardner. (1985). *Social psychology and second*

language learning: The role of attitudes and motivation. London: Arnold.

[11] R. Clement. (1980). *Ethnicity, contact, and communicative competence in a second language.* In E. Giles, D. Robinson, & P.M. Smith (Eds.) Language: Social psychology perspectives (pp. 146-154). Oxford: Pergamon.

[12] M. O. Akita. (2014). *The Effectiveness of Interactive Teaching Methods in EFL Classrooms: A Comparison with Bottom-Up and Top-Down Methods.* Applied Linguistics, 5, 2014.

[13] T. Morell. (2004). Interactive lecture discourse for university EFL students. *English for Specific Purposes, 23,* 325 - 338.
DOI : 10.1016/S0889-4906(03)00029-2.

[14] B. Inan. (2012). A comparison of classroom interaction patterns of native and non-native EFL teachers. *Social and Behavioral Sciences, 46,* 2419-2423.

[15] D. Graddol. (1986). *Discourse specific pitch behaviour," in Intonation in Discourse.* edited by Catherine Johns-Lewis (Croom Helm, London and Sidney), pp. 221-237.

[16] B. Gorjian, H. Abdolmajid & P. Pourkhon. (2013). Using Praat software in teaching prosodic features to EFL learners. *Social and Behavioral Sciences. 84,* 34-40.

[17] P. Medgyes. (1994). *The non-native teacher.* London: Macmillan.

[18] C. Johns-Lewis. (1986). *Prosodic differentiation of discourse modes in Intonation in Discourse.* edited by Catherine Johns-Lewis (Croom Helm, London and Sidney), pp. 199-219.

[19] Wiberg, E. (2003). The interactional context of L2 dialogue. Journal of Pragmatics. 35(3), 389-407.

엘라 키드(Ella Jane Kidd) **[정회원]**

· 2006년 12월 : Grambling State University, Social Science (BS)
· 2011년 5월 : Grambling State University, Education (MS)
· 2012년 9월 ~ 현재 : 우송대학교
· 관심분야 : 언어교육, 제2외국어
· E-Mail : Ekidd2000@hotmail.com