Articulation: Journal of Language, Linguistics, and Literature
Vol. 1 No. 1, June 2024, Hal. 36-46
DOI: https://doi.org/10.29313/jal.v1i1.4191
https://journals.unisba.ac.id/index.php/articulation/index

COMPARISON OF INTENSITY PATTERNS; CASE STUDY OF PHONETICS IN ENGLISH LEARNERS

Fauzan Novaldy Pratama [a, b]
[a]Program Studi D3 Bahasa Inggris, Fakultas Sastra,
Universitas Nasional PASIM
[b]Program Studi Linguistik, Sekolah Pascasarjana,
Universitas Pendidikan Indensia
fauzan.novaldy92@upi.edu

ABSTRACT

This descriptive-comparative research focuses on acoustic aspects in the field of phonological studies on sound production by L2 English learners in pronouncing words that contain diphthongs and have two syllables. The PRAAT application is used as an approach to describe and compare intensity patterns from digital voice recordings and L2 English learners. This research shows that of the eight diphthong words that were uttered, only two sound recordings from L2 whose intensity patterns were close to digital sound recordings. This research shows that PRAAT can be used as an approach to measure fluency, where further research can be directed to other aspects such as formant and pitch, as well as the use of L1 English sounds as a benchmark.

Keywords: intensity, diphthong, PRAAT

INTRODUCTION

In general, L2 learners, especially English, will face one challenge, namely language acquisition in terms of phonology, or how to pronounce a word correctly. This difficulty has the potential to arise as an impact of mastering L1 which is used first and more often (Schwartz, 2019), or in other words as a form of interference from L1 (Heryono, 2019), such as errors in the use of stress on a particular word (Kallio et al., 2020). According to Arias, Yoma, & Vivanco (2010), L2 learners themselves will find it difficult to match existing standards, such as fluency and naturalness of speech. Moreover, regional accents will have an impact on variations in sound production (Najafian & Russell, 2020). These phenomena explain cases where there are people who have difficulty producing sounds in English correctly. In fact, in the language system, especially communication, speech is an essential basis in the process of exchanging information. So, in terms of English L2 learners, there needs to be an awareness of having adequate pronunciation skills in an 'organic' manner.

Basically, the process of sound formation starts from the chest, throat, then head (Roach, 2001). When in the head, the variation of sounds produced is the result of the movement of the tongue and lips (Ladefoged & Johnson, 2011). another version, generation (vibration of the vocal cord and vocal fold) and filtering (passing through

the oral cavity as an articulator) are stages in sound formation (Suyudi & Saptano, 2015). In general, the sound produced by humans will be in the form of speech that has applied a certain grammatical system. The sound product (output), if captured and visualized, will be in the form of waves with a certain pattern.

In the case of L2 English learners, with an L1 background (for example) of Indonesian, pronunciation will feel different. This is because L1 Indonesian tends to sound the word the same as what is written. This is different from English, where what is written is not necessarily pronounced, especially vowel sounds. For example, the word 'break' is read [breɪk], where there is a reading like the letter 'i', while in the source word it is written as the letter 'a'. This example is a type of diphthong vowel, where the vowel sound involves a sound change from one vowel sound to another (Ladefoged & Johnson, 2011).

Diphthongs themselves are considered a unique type of vowel, because they are different from single vowels. In some parent literature such as by Ladefoged & Johnson (2011), Rachael & Knight (2012), Roach (1998), and Carr (2013), This type of vowel has its own sub-chapter of discussion. Diphthongs have different production locations at the beginning and end of the vowel. Sometimes, the intonation will even sound different. The pronunciation is like swinging from the beginning vowel to the end vowel. Roach (2001) notes that there are eight types of diphthongs in English, namely /eI/, /aI/, /ɔɪ/, /ɪə/, /eə/, /aʊ/, /oʊ/.

The usefulness of the discipline of phonology (and phonetics) is not only in the area of learning as a science that describes sound. The study of the realm of sound has been utilized in the field of computing technology as automation of operational work. SIRI on Apple, Bixby on Samsung, Alexa on Amazon, Sophia as a humanoid, and all kinds of sound-based technology will utilize the principles of phonetics and phonology, especially in the discussion of acoustic waves.

The sound produced by humans through their mouths (output) is basically a vibration that propagates in the air, which if visualized, will form a patterned wave. In terms of pronunciation patterns, an L2 English learner will produce sounds that are different from someone with an L1 English background. In the study of phonetics and phonology, these differences can be measured in terms of pitch (frequency/high-low), intensity (decibels/loud-soft), and formant (density) (Heryono, 2019). These wave patterns can be captured with certain software, such as open-source software called PRAAT.

The PRAAT system is able to elaborate and visually measure the pitch, intensity, and formant elements of a single sound recording in a single spectrogram (Boersma & Heuven, 2001) or waveform (Roach, 2001). The spectrogram that appears will provide a detailed picture of how the sound is produced. The elements of the spectrogram can show the diversity of one speaker to another in saying something, Roach (2001) calls it spectral analysis.

Articulation: Journal of Language, Linguistics, and Literature
Vol. 1 No. 1, June 2024, Hal. 36-46
DOI: https://doi.org/10.29313/jal.v1i1.4191
https://journals.unisba.ac.id/index.php/articulation/index

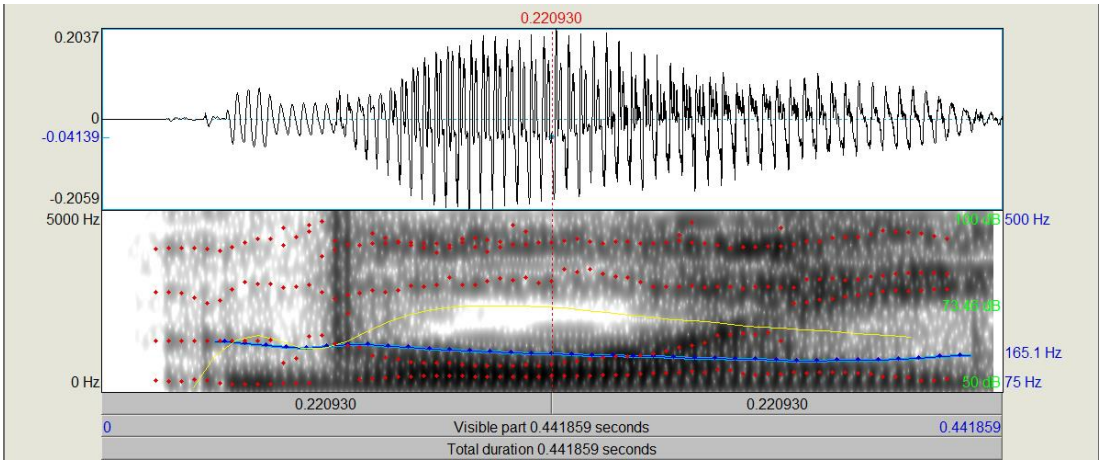


Figure 1. Example of a spectrogram for the pronunciation of 'Boy';
Bottom image: red dot = formant, yellow line = intensity, blue line = pitch, black line = spectrum

In the process of learning L2 English, PRAAT can be used as an approach to measure whether the L2 speaker has approached the 'degree' of L1 English speakers or even far different, this concept is usually called fluency (Derwing, Thomson, & Munro, 2006). The problem is, L1 English speech has various variations, while there needs to be a standard as a comparison. By utilizing technology such as voice on google translate or on an electronic dictionary with this feature, a 'standard' voice can be obtained, or a website that provides a converter from text to voice such as www.text2voice.com. So, by paying attention to the differences between the speech of L2 learners and the standard voice, acquisition can be maximized by seeing which elements need to be improved.

PRAAT AS A RESEARCH APPROACH

This study focuses on the comparison of intensity patterns between the recorded voices of L2 English learners and standard speech sounds taken from digital conversion from text to speech on the page www.text2voice.com. PRAAT is used as a dissector of voice recordings, both from students and the voice feature on Google Translate. Then, this study compares the wave patterns of the three spectrogram elements, whether the spectrum of the student's voice recording is in line with the standard recording or not. The object of speech in this study is a diphthong with one syllable, to make it easier for the speaker to pronounce the chosen word.

Wilson (2005) said that pitch and intensity can be produced from the production of diphthong and vowel sounds. PRAAT is also used by Wempe & De Jong (2009) to measure the speech rate in a particular syllable. Heryono (2019) uses the PRAAT approach to see which pitch and intensity are the highest among certain diphthongs. Variations in sound frequency and intensity were studied by Vogel et al. (2009) based on the gender of the speaker. In the medical world, this software can be used to detect speech disorders, such as research by Jong & Wempe (2009), Maryn (2017), Maryn,

Morsomme, & Bodt (2017), and Sampaio et al. (2020). In addition, in the forensic scope, PRAAT can be used to measure the authenticity of a sound recording as court evidence (Aligarh & Hidayanto, 2016). In the world of teaching, PRAAT is also used in a method called CALL (computer assisted language learning) to help acquire pronunciation (Gorjian & Hayati, 2012). In addition, there are also studies that utilize this tool for language acquisition purposes, such as Behr (2022), Osatananda & Thinchan (2021), Rahmatunisa & Syarifudin (2021), and Saito (2017).

This study and the explanation of the previous studies above show the potential of PRAAT as an effective tool to be an approach in the analysis of acoustic sound waves, as the creators themselves justify that many possibilities can be done using the system (Boersma & Heuven, 2001). The difference between this study and previous studies is in two things, namely the focus on word units containing diphthongs and data sources originating from the domestic L1.

PRAAT, DIFTONG, AND INTENSITY

PRAAT is a system developed by Boersma and Weenink at the Department of Phonetics of the University of Amsterdam (Boersma & Heuven, 2001). This software can be used to analyze, synthesize, and manipulate a sound recording. Until 2001, this system had been used by more than 5000 users in almost a hundred countries in the world. By utilizing this software, sound waves can be captured and visualized clearly which is called a Spectrogram. The default PRAAT display provides features in the form of two left and right dialog boxes; the left is for processing sound while the right is for displaying visuals (Lieshout, 2020).

Roach (1998, 2001), Ladefoged & Keith (2011), and Carr (2013) generally say that diphthongs are basically vowel sounds that contain changes such as 'swinging' because they combine two different vowel sounds. In phonetic studies, this diphthong is said to be a single vowel type (Heryono, 2019). In English, there are at least eight types of diphthongs; /eI/, /aI/, /ɔɪ/, /ɪə/, /eə/, /aʊ/, /oʊ/. Das (2014) states that the pronunciation of diphthongs by an L2 learner will be influenced by his L1. This type of vowel is divided into two types (Jones, 1972), namely rising and falling; it is said to rise if the initial sound is lower than the final sound and is said to fall if vice versa.

Intensity is the strength or weakness of sound measured in dB (decibels) (Hayward, 2000). The output of PRAAT on intensity parsing can be two things; overall dB numbers and intensity graphs. For example, Figure 2 shows how intensity patterns can be displayed visually by PRAAT. It can be seen in the intensity pattern in Figure 2 that there is a dynamic (up and down) that appears, where the horizontal axis is time and the vertical axis is a measure of the strength of the sound in dB. This shows that there is indeed a certain intensity pattern that appears from the pronunciation of sounds in certain words by someone.

Articulation: Journal of Language, Linguistics, and Literature
Vol. 1 No. 1, June 2024, Hal. 36-46
DOI: https://doi.org/10.29313/jal.v1i1.4191
https://journals.unisba.ac.id/index.php/articulation/index

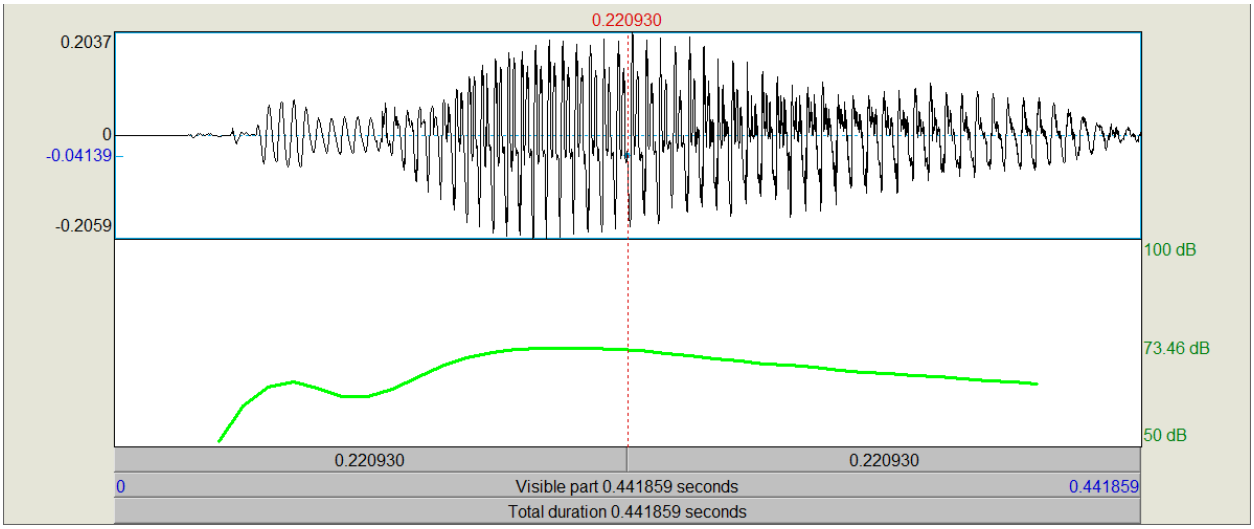


Figure 2. Example of intensity (green line) in the pronunciation of 'Boy'

METHOD

This study is a case study of whether a female student learning L2 English can match the intensity pattern carried out by the output of the digital process. This is done to see the learning potential in terms of pronunciation. The focus of this study is on the pronunciation of English words with one syllable and having diphthongs. For this reason, this study adopts a quantitative method by utilizing an instrumental approach to obtain numbers from the measurement process (Narhan, Sholihatun, & Syarfina, 2023). Quantitative itself deals with calculating numbers to test certain hypotheses, where the study focuses on the numerical extraction of the results of data collection from random samples. The instrumental approach plays a role as a supporter of the data extraction process into certain units (Pranoto, 2024). In this case, the instrument used is PRAAT, a software developed by Boersama and Weenink which is able to elaborate acoustic sound recordings into calculations in certain units, one of which is intensity.

This study aims to show the potential that PRAAT can be used as an approach for L2 learning, in this case English from its intensity pattern. When making a comparison, the x-axis (time) is not taken into account. The type of recording (black acoustic graph at the top of the image), both mono and stereo are considered the same, because basically the stereo type only separates the output source between left and right, does not change the quality of the recording. This study focuses on a brief comparison of the similarity of intensity graphics between digital audio recordings and English L2 learners.

Articulation: Journal of Language, Linguistics, and Literature
Vol. 1 No. 1, June 2024, Hal. 36-46
DOI: https://doi.org/10.29313/jal.v1i1.4191
https://journals.unisba.ac.id/index.php/articulation/index

FINDINGS AND DISCUSSION

This research uses sound sources, both from digital conversion and students, with female gender. In addition, a study says that women's intensity is usually higher than men (Flipsen, 1999). The words chosen to represent each diphthong analyzed are say (/eI/), go (/əʊ/), five (/aI/), now (/aʊ/), boy (/ɔɪ/), near (/ɪə/), hair (/eə/), and pure (/ʊə/).

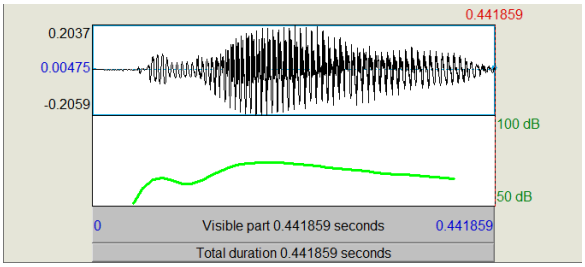
The results of the study can be seen that there are graphs of the intensity of digital and L2 voice recordings that have similarities and those that do not. Of the eight recordings, only two graphs look similar. The following data findings are presented in the table.

Table 1 shows the intensity similarity themes

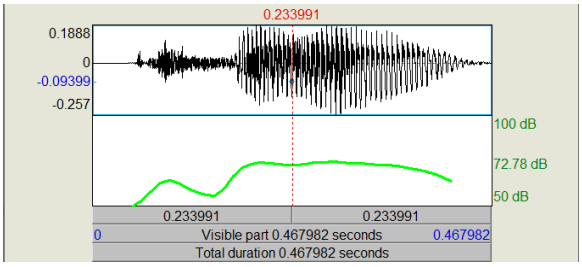
Status/Word	say (/eI/)	go (/əʊ/)	five (/aI/)	now (/aʊ/)	boy (/ɔɪ/)	near (/ɪə/)	hair (/eə/)	pure (/ʊə/)
Similar		√	√					
Not Simolar	√			√	√	√	√	√

A. Intensity Pattern of Digital Sound

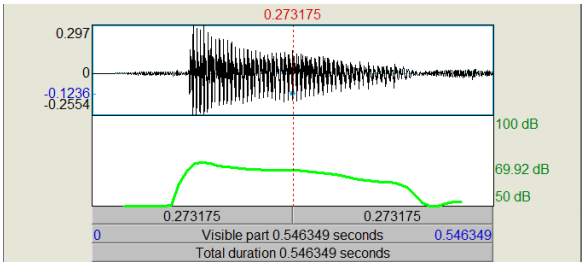
The following is the intensity pattern of digital sound that appears according to the selected words, complete with the mean dB of each sound (shown in the green line).



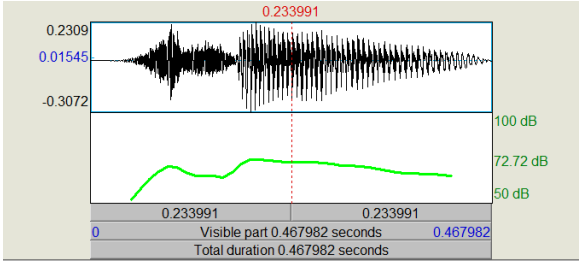
Boy /ɔɪ/ (mean dB = 69.96)



Pure /ʊə/ (mean dB = 69.96)

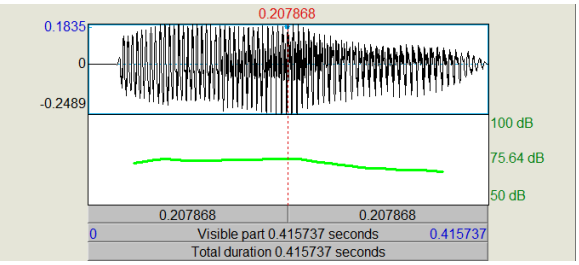


Five /aI/ (mean dB = 67.83)

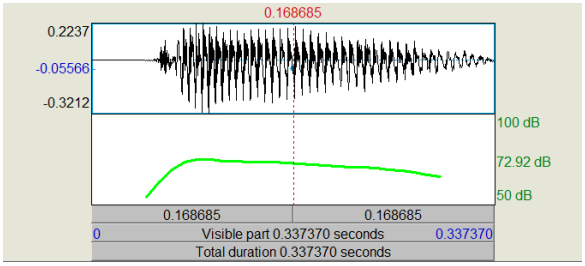


Say /eI/ (mean dB = 69.96)

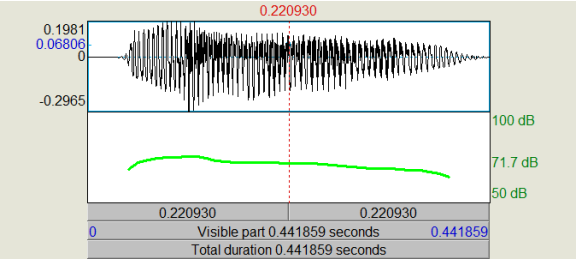
Articulation: Journal of Language, Linguistics, and Literature
Vol. 1 No. 1, June 2024, Hal. 36-46
DOI: https://doi.org/10.29313/jal.v1i1.4191
https://journals.unisba.ac.id/index.php/articulation/index



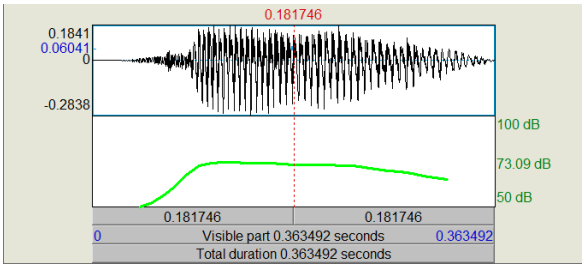
Near /ɪə/ (mean dB = 69.96)



Go /əʊ/ (mean dB = 71.82)



Now /aʊ/ (mean dB = 69.96)

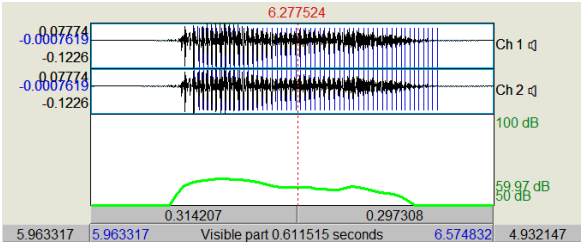


Hair /eə/ (mean dB = 69.96)

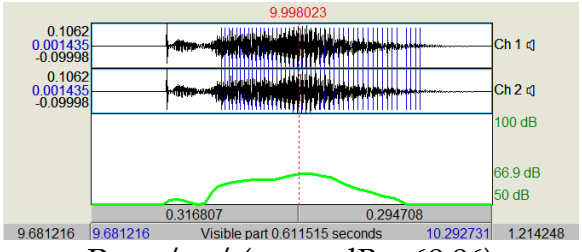
Figure 3. Intensity patterns based on selected diphthongs/ words from digital audio recordings

B. Pola *Intensity* dari Suara Mahasiswi

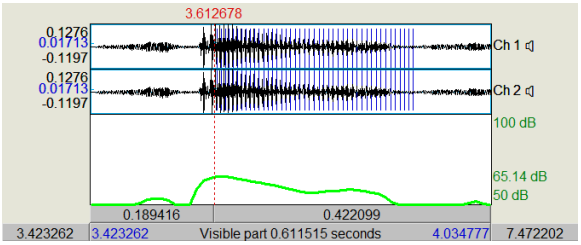
Berikut adalah pola *intensity* oleh mahasiswi yang muncul sesuai dengan kata-kata pilihan dilengkapi dengan *mean* dB dari masing-masing bunyi (diperlihatkan pada garis hijau).



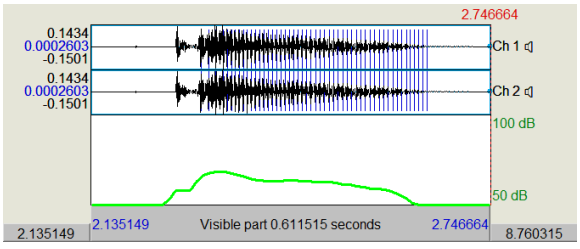
Boy /ɔɪ/ (mean dB = 59.98)



Pure /ʊə/ (mean dB = 69.96)

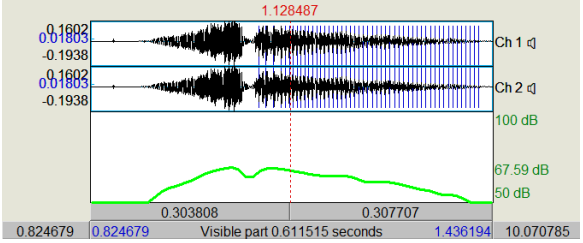


Five /aI/ (mean dB = 57.37)

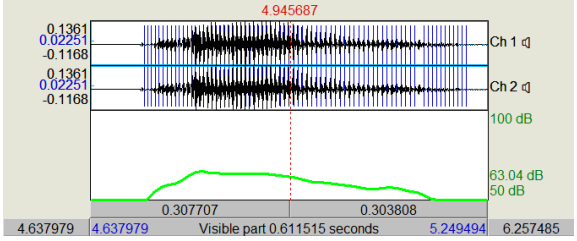


Go /əʊ/ (mean dB = 63.18)

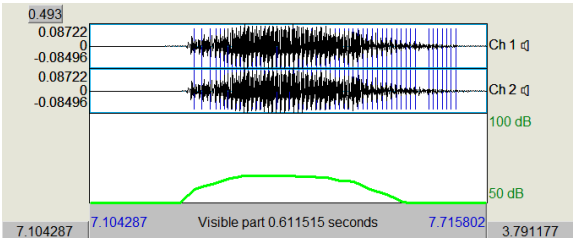
Articulation: Journal of Language, Linguistics, and Literature
Vol. 1 No. 1, June 2024, Hal. 36-46
DOI: https://doi.org/10.29313/jal.v1i1.4191
https://journals.unisba.ac.id/index.php/articulation/index



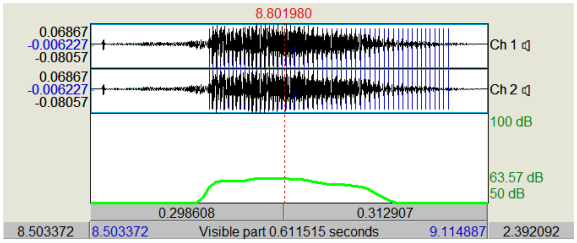
Say /eI/ (mean dB = 64.12)



Now /aʊ/ (mean dB = 60.79)



Near /ɪə/ (mean dB = 69.96)



Hair /eə/ (mean dB = 69.96)

Figure 4. Intensity patterns based on selected diphthongs/ words from L2 voice recordings

C. Comparison of Digital Audio Recordings and L2 English

After getting the graphs of the digital sound recording and L2 above, a comparison can be made by aligning the intensity graphs of the digital sound recording with L2 as follows (top digital sound, bottom L2 sound). The upper and lower graphs for the word 'go' appear to have a low decibel start, then peak and slope. It can be said that L2 has an indication that resembles the digital sound. While the upper and lower graphs for the word 'near' have differences, where the upper graph shows a nearly constant transverse line, while the lower graph appears to start with a low decibel, rise, then fall.

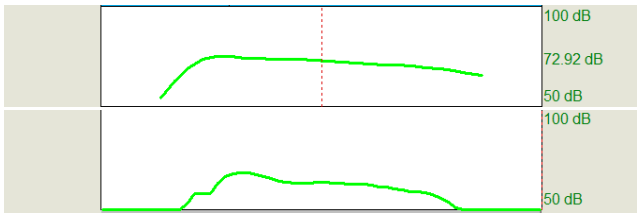


Figure 5, Go /aʊ/

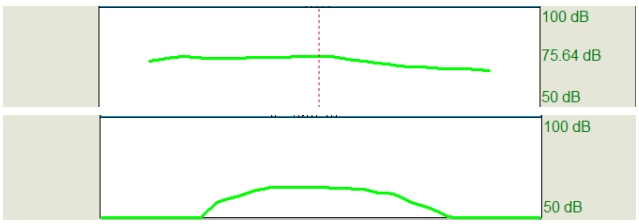


Figure 6, Near /ɪə/

CONCLUSION

From the findings above, it can be seen that L2 learners have the potential to acquire English with the help of PRAAT software to detect which factors need to be improved. In this case, a comparison of intensity patterns, namely the loudness and

softness of the voice recording, can show different and similar patterns in the pronunciation of diphthongs. If the pronunciation of a particular diphthong is already quite similar, then other diphthongs that have significant differences can be trained to match the digital voice.

This study can be said to be an initial step for further research, where the potential that can be developed is the use of L1 English voice as a standard voice, measurements that are not only limited to intensity, but also on the aspects of pitch and formant, and ultimately the compilation of functions to measure the level of fluency.

REFERENCES

Arias, J. P., Yoma, N. B., & Vivanco, H. (2010). Automatic Intonation Assessment for Computer Aided Language Learning. *Speech Communication*, *52*, 254–267.

Behr, N. S. (2022). English Diphthong Characteristics Produced by Thai EFL Learners: Individual Practice Using PRAAT. *Call-Ej*, *23*(1), 401–424.

Boersma, P., & Heuven, V. . (2001). Speak and unSpeak with PRAAT. *Glot International*, *5*(9/10), 341–347.

Carr, P. (2013). *English Phonetics and Phonology*. Chichester: Blackwell Publishing Ltd.

Das, K. (2014). Production of Central Vowels and Centering Diphthongs by Assamese Speakers of English. *Northeast Journal of Contemporary Research*.

Derwing, T. M., Thomson, R. I., & Munro, M. J. (2006). English Pronunciation and Fluency Development in Mnadarin and Slavic Speakers. *System*, *34*, 183–193.

Flipsen, P. J. (1999). Acoustic characteristics of /s/ in adolescents. *Journal of Speech, Language, and Hearing Research*, *42*(3), 663–677.

Gorjian, B., & Hayati, A. (2012). Using PRAAT Software In Teaching Prosodic Features To EFL Learners. *Procedia - Social and Behavioral Sciences*, 34–40.

Hayward, K. (2000). *Experimental Phonetics*. London: Longman.

Heryono, H. (2019). *Pengukuran Pitch dan Intensity Diftong Tertinggi Menggunakan Program PRAAT*. Jurnal Linguistik Komputasional.

Jones, D. (1972). *The Pronunciation of English*. London: Cambridge University Press.

Jong, N. H., & Wempe, T. (2009). *PRAAT Script to Detect Syllable Nuclei and Measure Speech Rate Automatically*. Behavior.

Kallio, H., Suni, A., Simko, J., & Vainio, M. (2020). Analyzing Second Language Proficiency Using Wavelet-Based Prominence Estimates. *Journal of Phonetics*, *80*, 100966–100977.

Ladefoged, P., & Johnson, K. (2011). *A Course in Phonetics* (Sixth). Boston: Wadshorth.

Lieshout, P. . (2020). *PRAAT Short Tutorial*. Retrieved from https://web.stanford.edu/dept/linguistics/corpora/material/PRAAT_workshop_manual_v421.pdf

Maryn, Y. (2017). Practical Acoustics in Clinical Voice Assessment: A Praat Primer. *Perspectives of the ASHA Special Interest Groups*, *2*(3), 14–32.

https://doi.org/10.1044/persp2.sig3.14

Maryn, Y., Morsomme, D., & Bodt, M. De. (2017). Measuring the Dysphonia Severity Index (DSI) in the. *Journal of Voice*, 1–12. https://doi.org/10.1016/j.jvoice.2017.01.002

Najafian, M., & Russell, M. (2020). Automatic Accent Identification as an Analytical Tool for Accent Robust Automatic Speech Recognition. *Speech Communication*, *122*, 44–55.

Narhan, R., Sholihatun, P., & Syarfina, T. (2023). Analisis Frekuensi, Intensitas, dan Durasi pada Bahasa Turki oleh Native Speaker dan Non-Native Speaker menggunakan Praat. *Lingua*, *20*(2), 3–51. https://doi.org/10.30957/lingua.v20i2.840.Analisis

Osatananda, V., & Thinchan, W. (2021). Using Praat for English Pronunciation Self-practice outside the Classroom: Strengths, Weaknesses, and its Application Varisa. *Journal: Language Education and Acquisition Research Network*, *14*(2), 372–396. Retrieved from https://so04.tci-thaijo.org/index.php/LEARN/index

Pranoto, M. sholihin. (2024). Analisis frekuensi, durasi, dan intensitas suara laki-laki dan perempuan bahasa Banjar menggunakan perangkat lunak Praat. *AKSARA: Jurnal Bahasa Dan Sastra*, *25*(1), 190–199. https://doi.org/10.23960/aksara/v25i1.pp185-200

Rachael, & Knight, A. (2012). *Phonetics: A Coursebook*. New York: Cambridge University Press.

Rahmatunisa, W., & Syarifudin, S. (2021). The use of PRAAT in learning English debate in Indonesian EFL classroom. *Indonesian EFL Journal*, *7*(1), 43–50. Retrieved from https://journal.uniku.ac.id/index.php/IEFLJ/index

Roach, P. (1998). *English Phonetics and Phonology, A Practical Course Second Edition*. Cambridge, Melbourn, New York: Cambridge University Press.

Roach, P. (2001). *Phonetics*. Oxford, New York: Oxford University Press.

Saito, K. (2017). Effects of Sound, Vocabulary, and Grammar Learning Aptitude on Adult Second Language Speech Attainment in Foreign Language Classrooms. *Language Learning*, *67*(3), 665–693. https://doi.org/10.1111/lang.12244

Sampaio, M., Vaz Masson, M. L., de Paula Soares, M. F., Bohlender, J. E., & Brockmann-Bauser, M. (2020). Effects of Fundamental Frequency, Vocal Intensity, Sample Duration, and Vowel Context in Cepstral and Spectral Measures of Dysphonic Voices. *Journal of Speech, Language, and Hearing Research*, *63*(5), 1326–1339. https://doi.org/10.1044/2020_JSLHR-19-00049

Schwartz, G. (2019). *Refining representations for L2 phonology*. Second Language Research.

Suyudi, I., & Saptano, D. (2015). Analisis Vocal Tract pada Kajian Akustik Vokal Bahasa Indonesia. In *Prosiding PESAT (Psikologi, Ekonomi* (pp. 20–21). Sastra: Arsitektur, & Teknik Sipil) Universitas Gunadarma.

Vogel, A. P., Maruff, P., Snyder, P. J., & Mundt, J. C. (2009). Standardization of pitch-range settings in voice acoustic analysis. *Behavior Research Method*, 318–324.

Wilson, I. (2005). Using Praat and Moodle for Teaching Segmental and Suprasegmental Prunonciation. *TESOL Quarterly*, *3*, 33–43.