Linguistic Portfolios

Volume 8 | Article 5

2019

A Longitudinal Acoustic Phonetic Study of English Vowels by a Panamanian Speaker

Ettien Koffi
St. Cloud State University, enkoffi@stcloudstate.edu

Fernando G. Lesniak
St Cloud State University, gofe1001@stcloudstate.edu

A LONGITUDINAL ACOUSTIC PHONETIC STUDY OF ENGLISH VOWELS PRODUCED BY A PANAMANIAN SPEAKER

ETTIEN KOFFI AND FERNANDO GONZALEZ LESNIAK [1]

ABSTRACT

Longitudinal acquisitions of English vowels have been previously studied by Monroe (2008), Lai (2010), and others. Most of these studies rely primarily on an impressionistic methodology, i.e., native speaking judges listen to oral inputs by non-native speakers and rate the intelligibility of their vowels on a Likert scale. This is not so for this study. The assessment relies primarily on the speech signals emitted by Author 2 when reading vowels in citation form and in running speech. The F1 and F2 correlates of his vowels are measured at three different intervals: in 2011, in 2017, and in 2018. The measured correlates are compared and contrasted with each other, and with the prototypical formant values in Peterson and Barney (1952). Masking thresholds, Just Noticeable Difference (JND) limens, and relative functional load (RFL) calculations are used to determine which of Author 2's vowels have become native-like and which ones have remained intractable over the course of seven years. The acoustic measurements from 2017 and 2018 also help investigate whether or not F1 and F2 correlates change or remain the same in citation form and running speech.

1.0 Introduction

This study examines how Author 2's Panamanian Spanish-accented English vowels have evolved over a period of seven years, from 2011 to 2018. During these years, he has tracked his vowels by performing acoustic phonetic analyses on them. The seven years are divided into three data collection points. The first set was collected in 2011 when Author 2 enrolled in Author 1's acoustic phonetics course. The second data set came five years later, when Author 2 collected the same corpus and performed the same acoustic phonetic analyses in a graduate sociolinguistics course taught by Author 1. Subsequently, he recorded himself reading an elicitation paragraph containing the same vowels. The measurements obtained from all three types of data are discussed in the paper which is organized around these data points. At each juncture, acoustic phonetic measurements comparing the correlates of Author 2's vowels and those of General American English (GAE) are presented and discussed. The discussions in each section contain comparative acoustic vowel spaces where his vowels are plotted together with GAE vowels. Additionally, internal and external masking calculations are performed to gauge the intelligibility of Author 2's vowels over the course of seven years.

2.0 Historical Background

Author 2 was born and raised in the Republic of Panama. Spanish is his native language. He began learning English in kindergarten and continued all the way through high school and college and majored in English. In fall 2010, he received a scholarship through the U.S. Department of State's Bureau of Educational and Cultural Affairs and came to the USA where he attended Saint Cloud State University (SCSU) in Minnesota. While at SCSU, he attended

[1] **Authorship Responsibilities:** Author 2 provided the acoustic phonetic data, the measurements, and all but the last two acoustic vowel spaces. He wrote an initial report. Author 1 has rewritten and expanded the analyses on the basis of the data provided by Author 2 who has vouched for the accuracy of all the measurements. To the extent that the acoustic measurements are accurate, Author 1 assumes full responsibility for any erroneous interpretation of the data.

numerous courses, including linguistic courses taught by Author 1. In fact, it is in Author 1's acoustic phonetics course, that he recorded his English vowels for a course project. Between 2011 and 2012, Author 2 returned to Panama to continue his program, but he decided to move back to the USA in 2012 to finish his undergraduate degree at Saint Cloud State University. After graduation, he started working as a Spanish teacher, and later took a position as an English as a Second Language (ESL) teacher. In 2017, Author 2 enrolled in the Master's Degree in Teaching English as a Second Language at Saint Cloud State University. All in all, he has been learning English for 27 years. For the past six years English has become his dominant language of everyday communication. He uses it professionally and with people in his social network (see Appendix 3).

3.0 Replication Study

In Author 1's course, students are asked to replicate Peterson and Barney's 1952 seminal study of General American English (GAE) vowels. The original study involved 76 participants (33 men, 28 women, and 15 children). They produced 10 words containing English monophthong vowels in /hVd/ frames. The words are: <heed>, <hid>, <head>, <had>, <hawed>, <hod>, <hood>, <who'd>, <hud>, and <heard>. Their study was replicated in 1995 by Hillenbrand et al. They had 105 participants (45 men, 48 women, and 12 children). Whereas Peterson and Barney's study had participants from various areas of the USA, Hillenbrand et al.'s participants were from the Midwest, and especially from Michigan's lower peninsula. In fact, 87% of all the participants came from this area of Michigan. Peterson and Barney's excluded /e/ and /o/ but Hillenbrand et al. included them in their study. Author 2 recorded himself reading the following words: <heed>, <hid>, <hayed>, <head>, <had>, <hawed>, <hoed>, <hod>, <hood>, <who'd>, and <hud> in their citation form. This list contains the 11 phonemic vowels of English, namely, /i/, /ɪ/, /e/, /ɛ/, /æ/, /ɑ/, /o/, /ɔ/, /ʊ/, /u/ and /ʌ/. The vowel /ɚ/ is not included in this study because it is not a phoneme, but rather an allophone of various unstressed vowel phonemes in GAE. In 2018, he recorded himself reading the elicitation paragraph in section 8.0. For nearly a decade, Author 1 and his students have replicated the same methodology to investigate the intelligibility of L2-accented English vowels. In all such studies, Peterson and Barney's measurements are taken as the standard against which L2 vowels are compared and contrasted.

4.0 Methodology

The methodology outlined in the previous section was followed scrupulously in the 2011 and 2017 recordings. Author 2 generated spectrographs of each word and measured F1 and F2 of vowels in the first project. For the second project (2017), he measured six acoustic correlates: F0, F1, F2, F3, intensity, and duration. The measurements were taken from the steady state portion of vowels, as illustrated by the annotations in Figure 1.

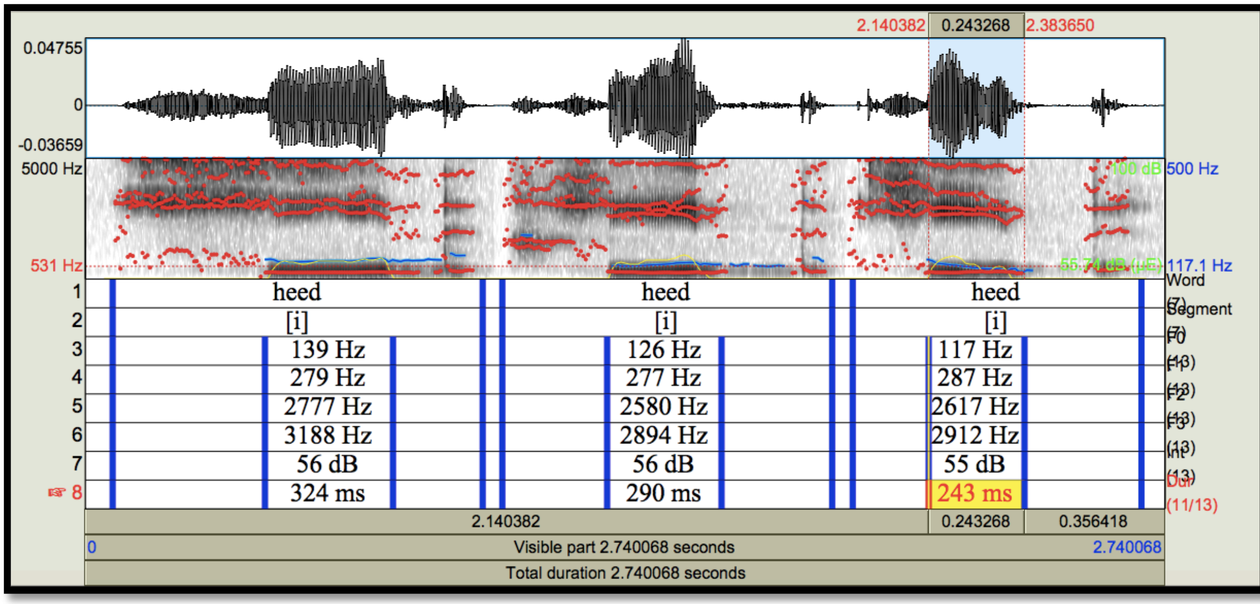

Figure 1: Sample Spectrograph

In 2011, Author 2 measured 66 vowel tokens (11 vowels x 3 repetitions x 2 correlates). In 2017 and 2008, he measured 198 vowel tokens (11 vowels x 3 repetitions x 6 correlates). The total number of tokens investigated in the seven year period amount to 462. All the data were collected and analyzed using Praat, an open source software created by Boersna and Weenink.[2] The acoustic vowel space diagrams used in this paper are generated in NORM, another open-source software created by Thomas and Kendall (2007).[3]

5.0 An Overview of the Instrumental Assessment of Intelligibility

The main focus of Author 1's research is the instrumental assessment of speech intelligibility. He trains his students to do acoustic phonetic fieldwork by teaching them various instrumental techniques. In this framework, the intelligibility of vowels is gauged by taking into account Just Noticeable Differences (JND), internal and external masking levels, and relative functional loads (RFL). The consensus JND for the intelligibility of vowels is 60 Hz. Several seminal studies have confirmed this JNDs (Mermelstein 1978: 578, Hawks 1994:1079, Labov et al. 2013:43) among others. This is also the threshold used in *Atlas of North American English*. If the acoustic distance between two phonetically contrastive front, back, central, or low vowels mask is <60 Hz, masking is likely. Conversely, if the acoustic distance between two phonemically contrastive vowels is >60 Hz, no masking is expected. In addition to F1 frequency, the RFL calculations by Catford (1987:89-90) play an important role in the overall assessment of intelligibility. Koffi (2019:68, 93) has proposed the severity scale in Table 1 to account for the various factors that impact intelligibility:

---

[2] Source: http://www.fon.hum.uva.nl/praat/
[3] Source: http://lingtools.uoregon.edu/norm/norm1.php

N0	F1 Distance	Masking Levels	RFL	Intelligibility Rating
1.	> 60 Hz	No masking	0-24%	Good intelligibility
2.	41 Hz – 60 Hz	Slight masking	25-49%	Fair intelligibility
3.	21 Hz – 40 Hz	Moderate masking	50–74%	Average intelligibility
4.	0 Hz – 20 Hz	Complete masking	75–100%	Poor intelligibility

Table 1: Correlations between Masking, Acoustic Distance, RFL, and Intelligibility

A quick explanation of Table 1 is in order. Various acoustic phonetic experiments have reported that humans cannot perceive any difference between two speech signals if the acoustic distance between them is ≤ 20 Hz (Flanagan 1955: 616 and Thomas 2011:56), to mention only these two authors. In other words, masking is complete if the acoustic distance between two phonemically contrastive vowels is ≤ 20 Hz. However, all vowels do not have the same RFL. Consequently, their impact on intelligibility must be weighed using Catford's percentages. Furthermore, in estimating intelligibility, we focus exclusively on F1 because this formant alone contains 80% of the acoustic energy found in vowels (Ladefoged and Johnson 2015:207). The information summarized in Table 1 is the interpretive tools that we will use in gauging the intelligibility of Author 2's English vowels. We proceed in three stages: we first focus on Author 2's 2011 vowels, then move to 2017 vowels, and conclude with 2018 vowels.

6.0 Acoustic Phonetic Analysis of 2011 Vowels

The acoustic measurements obtained from the 2011 recordings are displayed in Table 2:

Words		**heed**	**hid**	**Hayed**	**Head**	**Had**	**hod**	**hawed**	**hoed**	**Hood**	**who'd**	**Hud**
Vowels		[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
GAE	F1	270	390	476	530	660	730	570	497	440	300	640
LESN1	F1	290	528	375	649	775	631	621	405	322	335	424
DIFF	F1	**20**	**138**	**101**	**119**	**115**	**99**	**51**	**92**	**118**	**35**	**16**
GAE	F2	2290	1990	2089	1840	1720	1090	840	910	1020	870	1190
LESN1	F2	1373	2060	2082	2000	1204	1100	983	884	991	1111	1565
DIFF	F2	**917**	**70**	**7**	**160**	**516**	**10**	**143**	**26**	**29**	**241**	**375**

Table 2: F1 and F2 of 2011 Data Set

When a person produces the **same** vowel several times, variations are expected. However, variations in F1 are not expected to exceed 60 Hz on the F1 frequency band (Ladefoged 2003:123, Kent and Read 2002:110). Similarly, a person produces the same vowel; variations are expected, but they do not ordinarily exceed 200 Hz on the F2 frequency band. The F1 threshold measures intelligibility, while the F2 threshold measures accentedness (Kent and Read 2002:111). We can use these two JNDs to access both the intelligibility and the accentedness of Author 2's speech. By the F1 criterion, when we compare Author 2's vowels with their equivalents in GAE, if the acoustic distance between them is ≥ 60 Hz, we can say that the vowels in question are likely to compromise intelligibility. According to the F2 criterion, if his vowels and their equivalents in GAE are ≥ 200 Hz, we can say that his vowels are accented. However, as Ladefoged and Disner (2012:43) have famously stated, "Everybody has an accent. It's impossible to talk without one. Your accent is usually the same as that of the people you grew up with. As long as you are among them you don't stand out, and nobody thinks that you have an accent. But as soon as you go away and live with other people, it's your way of speaking that sounds different, and people will say that you have an accent." For this reason, we **do** expect to see variations between the F2 of Author 2's vowels and those of GAE speakers.

A casual look at the information in Table 2 shows that Author 2 produces the F1 of his vowels [ɪ], [e], [ɛ], [æ], [ɑ], [o], and [ʊ] rather differently from GAE speakers. The F2 measurements also show that his [i], [æ], and [ʌ] are substantially different from those in GAE. The fact that 10 of his 11 vowels differ significantly from those of GAE talkers is an indication that his English was strongly accented in 2011 and less intelligible, as depicted in Figure 2:

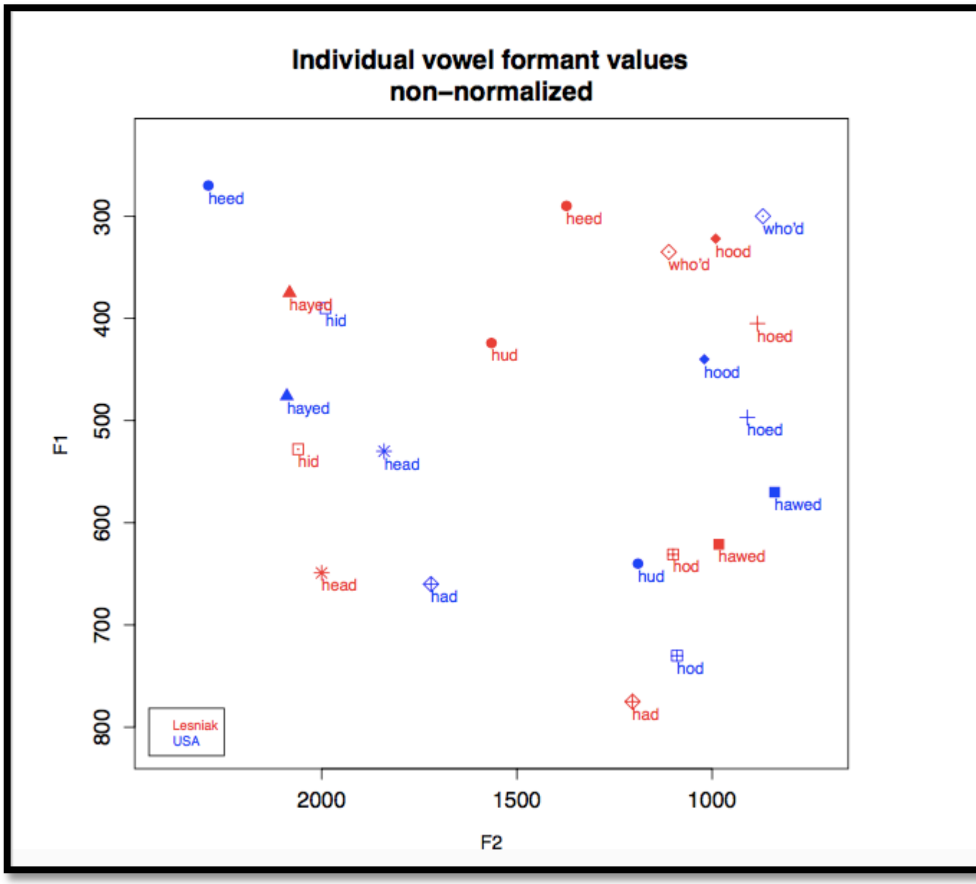

Figure 2: Comparative Acoustic Space of Vowels (2011) [color coordinated]

6.1 Segmental Intelligibility Analysis of the 2011 Data

The JND of 60 Hz is an important threshold in assessing intelligibility. Yet, contextual cues can help mitigate its effects. The true bar for absolute unintelligibility is ≤20 Hz. If the acoustic distance between two phonemically contrastive front, back, or low vowels is below this JND, then masking is absolute. In other words, the human ear cannot auditorily discriminate between such segments. This has been the consensus all along until Kweley-Port and Watson (1994) found otherwise. Their speech synthesis experiments led them to conclude that people can discriminate between segments with frequencies as low as 12 Hz. Their laboratory experiments and findings elicited the following response from Ladefoged and Disner (2012:176) "Of course in the rough and tumble of every day speech, differences as small as this are unlikely to be used for distinguishing words. Even if we only set about to distinguish the subtleties of accent within a language, we probably need differences that are twice or three times as long as the set of just noticeable differences." In this paper, we stick to the JND of ≤ 20 Hz in determining lack of intelligibility because when people interact in real life situations, the conversations take place in free field environments, not in speech laboratories. In word others, unintelligibility is absolute at ≤ 20 Hz, and intelligibility is compromised between 20 and 40 Hz.

6.2 Masking and Intelligibility Analysis of the 2011 Data

When Author 2 pronounced [ɑ] (631 Hz) and [ɔ] (621 Hz), his two vowels masked each other absolutely because the acoustic distance between them is just 10 Hz. However, the RFL between them is only 26%. Consequently, this masking has only a marginal effect on intelligibility. Furthermore, this masking would go unnoticed by many GAE hearers because these two vowels have merged in a number of American English dialects. His [u] (335 Hz) and [ʊ] (322 Hz) masked each because the acoustic distance between them is only 12 Hz. However, intelligibility is not compromised at all because these two vowels no longer contrast phonemically in English. They are in free variations in many dialects of American English. As for external masking, we see that when Author 2 pronounces [ɑ] (631 Hz), GAE hearers may think that he is producing [æ] (660 Hz) because his [æ] masked [ɑ] in GAE by 29 Hz. The RFL of these two vowels is 76%, which is high. Consequently, this confusion is likely to lead to poor intelligibility. Also, his [ɪ] (528 Hz) externally masks [ɛ] (530 Hz) in GAE. The distance between them is only 2 Hz. This confusion leads to average intelligibility because the RFL between them is 54%. Last and also least, his [ʊ] (322 Hz) masks [u] (300 Hz) in GAE by 22 Hz. However, as noted earlier, this confusion does not cause any intelligibility problem.

6.3 Phonological Processes in the 2011 Data

The acoustic vowel space in Figure 2 allows us to visualize how Author 2's vowels relate to those of GAE speakers. A few phonological processes are at play here. First, his vowel [ɪ] (528 Hz) is lowered by 138 Hz compared to [ɪ] (390 Hz) in GAE. Secondly, his [ʊ] (322 Hz) have the same height. In other words, we see [ʊ] raising in his pronunciation. His [ʌ] is also raised (424 Hz) in comparison with the one in GAE (640 Hz). Another important feature of Author 2's pronunciation is the strong centralization of [i]. The F2 of his [i] is 1373 Hz compared with 2290 Hz in GAE. Another important difference between Author 2's vowels and those of GAE is seen in his pronunciation of [æ] (1204 Hz). He produced it as a back vowel whereas its counterpart in GAE is at 1720 Hz. The huge F2 distances between his vowels and those of GAE speakers underscore his strongly foreign-accented English in 2011.

7.0 Acoustic Phonetic Analysis of the 2017 Data

Author 2 recorded himself again in 2017 for Author 1's graduate sociolinguistics course. By this time, Author 2 had been living USA continuously for 5 years. He read the same words containing the same 11 phonemic English vowels as he did in the 2011. The words in Table 3 were annotated and measured under the same conditions as discussed previously. The only difference this time is that many more acoustic correlates were in investigated (see Appendices 1 and 2).

Words		heed	hid	Hayed	Head	Had	hod	hawed	hoed	Hood	who'd	Hud
Vowels		[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
GAE	F1	270	390	476	530	660	730	570	497	440	300	640
LESN2	F1	281	437	358	461	743	600	585	352	311	332	583
DIFF	**F1**	**11**	**47**	**118**	**69**	**83**	**130**	**15**	**145**	**97**	**32**	**57**
GAE	F2	2290	1990	2089	1840	1720	1090	840	910	1020	870	1190
LESN2	F2	2658	2106	2389	2069	1571	1066	974	800	916	1019	1214
DIFF	**F2**	**368**	**116**	**300**	**229**	**149**	**24**	**134**	**110**	**244**	**149**	**24**

Table 3: F1 and F2 of 2017

In the following sections, we discuss the changes that have taken place in Author 2's pronunciation of English vowels.

7.1 The Phonological Process of Vowel Fronting

The most spectacular changes in Author 2's pronunciation from 2011 to 2017 have to do with the vowels [i] and [æ]. These changes are highlighted in Figure 3:

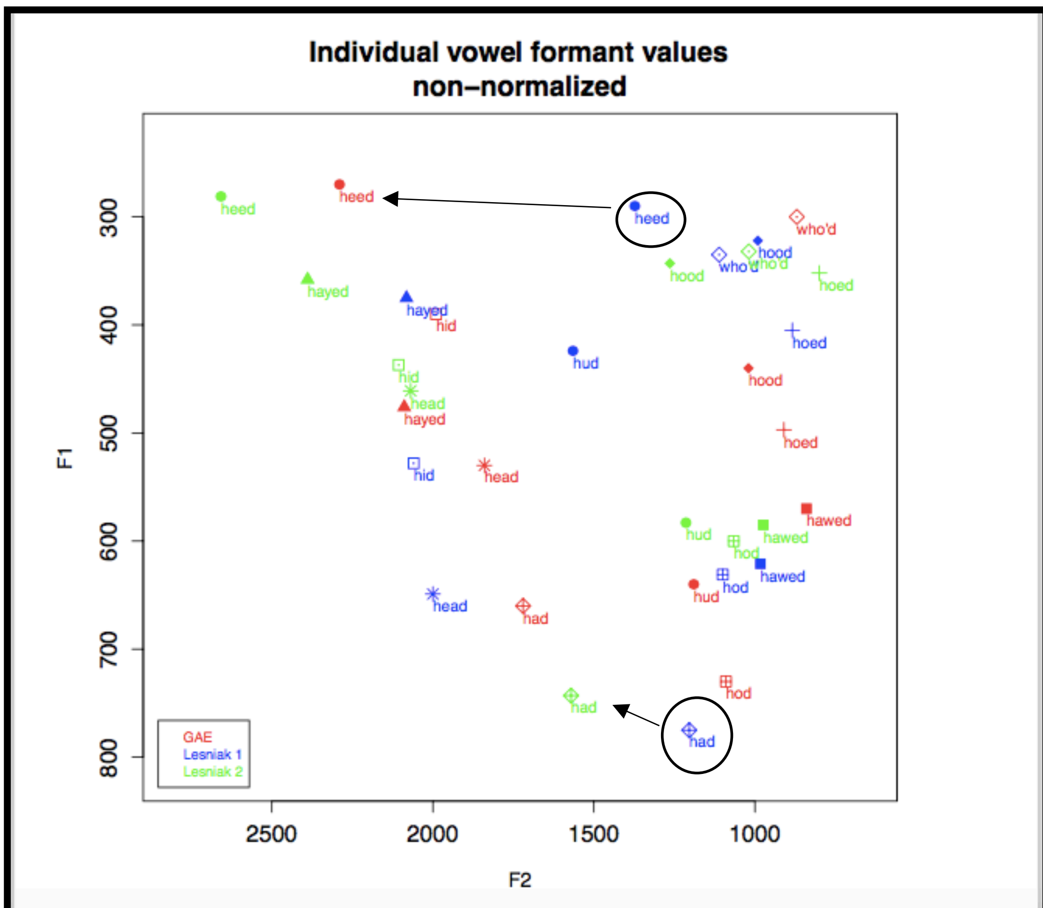

Figure 3: The Fronting of [i] and [æ] (color coordinated)

Whereas his [i] was strongly centralized with an F2 of 1373 Hz in 2011, this was no longer the case in 2017. By then, it has become fronted (2,658 Hz). This translates into a fronting gain of

1285 Hz. His new way of pronouncing [i] is well beyond the average of 1979 Hz reported in Giacomino (2012:102) for Hispanic speakers of English. In fronting his [i] in this manner, Author 2 has reduced the amount of accentedness displayed in 2011. His [æ] has also become more fronted but less spectacularly. It was 1204 Hz in 2011 but changed to 1571 Hz in 2017. The difference of 367 Hz is perceptible to the naked ear. Concurrently, on the F1 frequency band, Author 2 raised his [æ] from 775 Hz in 2011 to 743 Hz in 2017.[4] The gain of 32 Hz is barely perceptible to the naked ear.

7.1 The Phonological Process of Vowel Raising

Other noteworthy changes are also seen in Figure 4 :

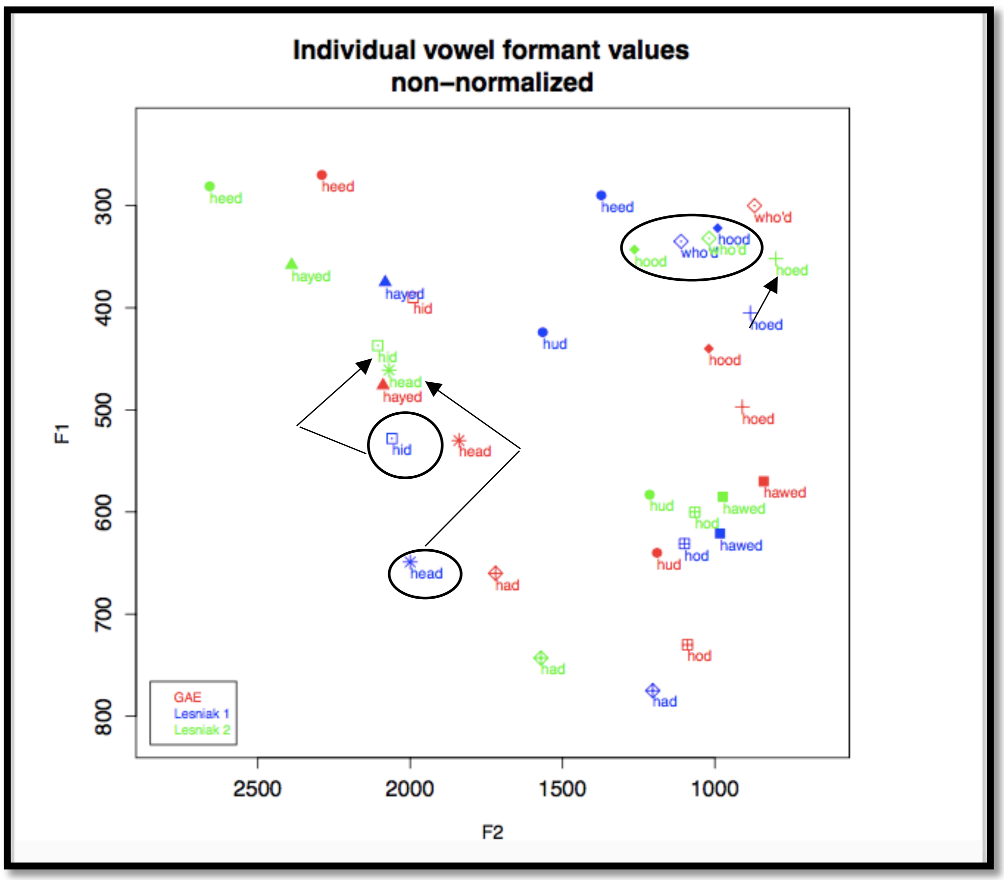

Figure 4: The Raising of [ɪ] and [ɛ] (color coordinated)

In 2011, the F1 of Author 2's [ɪ] was 528 Hz. In 2017, it is raised to 437 Hz. The raise of 91 Hz is perceptually salient. We see the same upward trend in his pronunciation of [ɛ]. Its F1 was 649 Hz in 2011, but it rose to 461 Hz in 2017. The gain of 188 Hz is also perceptually salient. The changes noted so far have had to do with front vowels. Overall, Author 2 did not make significant improvements in his pronunciation of back vowels. The F1 of his back vowel [o] was 405 Hz in 2011. By 2017, it has risen to 352 Hz. However, the gain is less than 60 Hz (53 Hz to be exact).

[4] In interpreting F1 values in regard to mouth aperture, it is important to keep in mind that values are inversely proportional. This means that bigger F1 values are indicative of vowel lowering, whereas smaller values correlate with vowel raising.

Consequently, it is not perceptually significant. His high back vowel [ʊ] did not budge either. It was 322 Hz in 2011. It rose to 311 Hz in 2017, but the change of 11 Hz is auditorily imperceptible. Consequently, his [ʊ] continued to mask [u] (300 Hz) in GAE.

7.2 The Phonological Process of Vowel Lowering

Another important change that Author 2 implemented between 2011 to 2017 has to do with his pronunciation of [ʌ]. This vowel underwent a drastic change. Its F1 in 2011 was 424 Hz. By 2017, it has lowered to 583 Hz. Author 2 has shaved off 159 Hz from his previous pronunciation. This change in F1 is perceptually significant. The Hispanic males in Giacomino's data produced [ʌ] with 531 Hz. This suggests that in Spanish-accented English this vowel is much higher than its counterpart in GAE (640 Hz). Author 2 made great progress with this vowel. By 2017, his pronunciation has become native-like as far as F2 is concerned. It was 1214 Hz in 2017 compared with 1194 Hz in GAE. The fact that the acoustic difference between his pronunciation and that of GAE speakers is only a 20 Hz attests to the fact that he has mastered the pronunciation of this vowel.

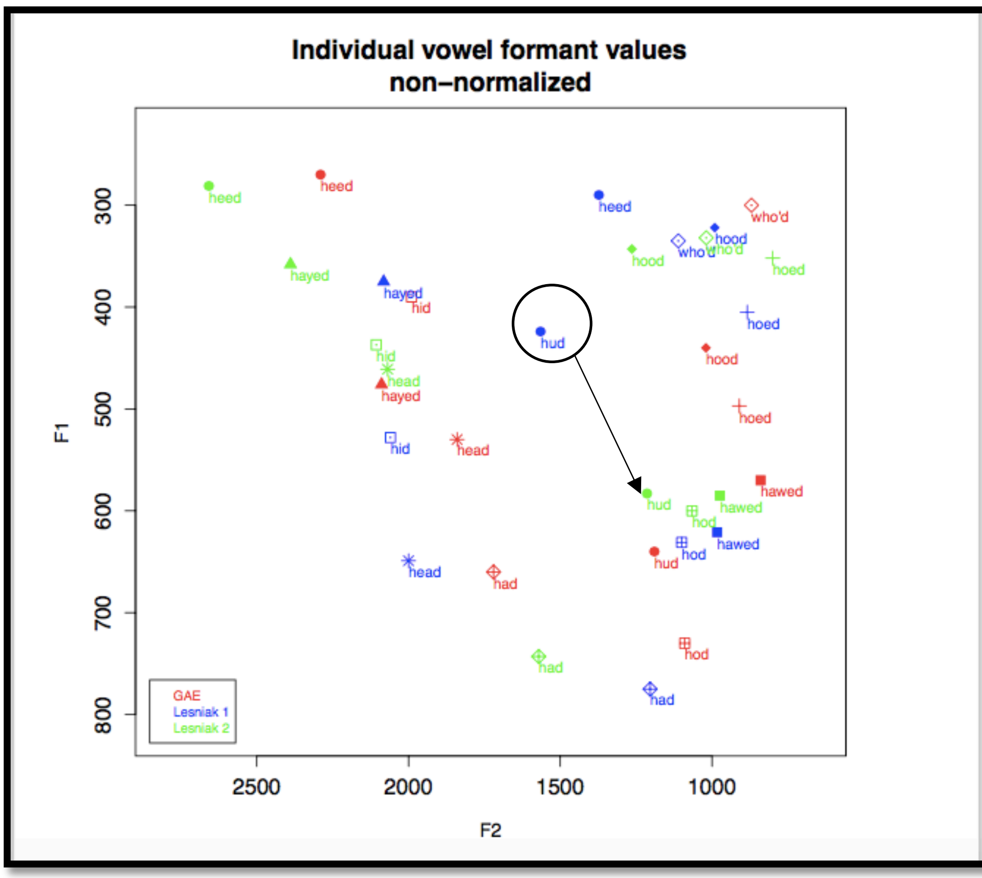

Figure 5: The Lowering of [ʌ] (color coordinated)

8.0 Acoustic Phonetic Analysis of the 2018 Vowels

The preceding analyses highlighted the changes in Author 2's pronunciation of English vowels between 2011 and 2017. These recordings represent the correlates of his vowels produced in citation form. We now turn our attention to the acoustic measurements of the vowels that he produced in 2018 while reading the text below:

> *Please call Stella. Ask her to bring these things with her from the store: Six good spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a footlong sandwich as a snack for her brother Bob. We also need a small plastic snake, a yellow book, a rubber duck, a paper I-pad, the dog video game, a big toy frog for the kids, but not the faked gun. She can scoop these things into three red bags, and two old backpacks, and we will go meet her, Jake, and Jenny Wednesday at the very last train station at the edge of the zoo near York's Treasure Bank.*

This text is a slightly expanded version of the elicitation paragraph found at the Speech Accent Archive (http://accent.gmu.edu/howto.php#cite). An expansion became necessary for three reasons. First, because the original text omits the vowel [ʊ]. Secondly, we want to make sure that each vowel phoneme would occur at least three times. Thirdly, in the original text, some consonants did not appear at all or appeared only in very restricted environments.[5]

8.1 Departing from Tradition

The classical methodology pioneered by Peterson and Barney (1952) consisted in studying vowels in citation form. Nearly all replication studies since then have followed their lead. However, there is a need to investigate whether or not L2 speakers' vowels are intelligible in running speech. This line of inquiry is in keeping with Lisker and Abramson's (1964:407) observation that "The ultimate usefulness of measuring voice onset time depends on how effectively it enables us to identify stops in running speech." Clearly, this study is not about VOT, but nothing prevents us from expanding Lisker and Abramson's observation to the study of vowels (or any speech segment) in running speech. In so doing, we seek to provide answers to the following questions:

1. Do the acoustic correlates of segments (in this case vowels) change from citation form to running speech?
2. If they change, do the changes affect intelligibility?

Do vowels produced in citation form versus in running speech affect intelligibility with regard to duration? Casual observations of everyday speech tell us speech velocity has a direct impact on intelligibility. But how does this manifest itself acoustically? Here are some answers. Researchers have documented that segments produced in citation form are longer that those produced in running speech (Thomas 2011:138-9, 174-5, 292-3). We see this in Author 2's pronunciation of vowels as well. In his 2017 data, the duration of his 11 vowels was 2915 ms. On average, each vowel lasted 265 ms. In running speech, his 11 vowels lasted 2117 ms, that is, 193 ms for each vowel. His speech rate in running speech comes as close as possible to ideal speech rate described in Pena-Brooks and Hedge (2000:107). Everest and Pohlmann (2015:492) note that, generally speaking, consonants have an average duration of 65 ms, vowels about 100 ms, syllables between 300 to 400 ms, and 600 to 900 for whole words, depending on delivery. Author 2's vowels in running speech lasted more than 100 ms. Consequently, the intelligibility of his vowels would not be affected. Furthermore, it is well attested in acoustic phonetic circles that a minimum of 35 ms

[5]As of 31 January 2019, the original text has been read by 2783 people (http://accent.gmu.edu/index.php).

is all the duration that humans require to fully perceive a speech segment (Everest and Pohlmann 2015:60).

Do vowels produced in citation form versus in running speech affect intelligibility with regard to formant frequency? Ladefoged et al. (1976:230) examined this question by measuring F1 and F2 frequencies in various speech rates and styles. They summed their findings as follows: "It may be seen that across all styles of speech, we can estimate that two-thirds of all first formant frequencies are within about 40 Hz of the mean, and two-thirds of all second formant frequencies are within about 90 Hz of the mean." What does this mean for the intelligibility? As noted earlier, formant frequencies of less than 60 Hz do not affect intelligibility on the F1 frequency band. Also, formant frequencies of less than 200 Hz on the F2 bandwidth are not perceptible to the naked ears. In other words, Ladefoged et al.'s (1976) findings indicate that producing vowels in isolation and producing them in running speech have no impact on intelligibility.

The duration and frequency thresholds examined above tell us that it is acoustically permissible to compare and contrast vowels in citation form and those in running speech. Doing so is not like comparing "apples and oranges." We will therefore proceed with comparing Author 2's pronunciation of vowels in 2018 with the formants vowels that GAE speakers produced in citation form in order to assess the intelligibility of his vowels during his seven years in the USA.

8.2 Vowels Frequencies in Citation Form vs. Running Speech

Table 4 contrasts the measurements of 2017 with those of 2018. "LESN2" stands for vowels in citation form, while "LESN3" represents vowels in running speech. Furthermore, "LESN3" is contrasted with GAE vowels produced in citation form.

Words		**heed**	**hid**	**hayed**	**head**	**had**	**hod**	**hawed**	**hoed**	**hood**	**who'd**	**hud**
Vowels		[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
LESN2	F1	281	437	358	561	743	600	585	352	311	332	583
LESN3	F1	287	444	420	602	839	724	505	472	374	323	708
Difference		**6**	**7**	**62**	**41**	**96**	**124**	**80**	**120**	**63**	**9**	**125**
GAE	F1	270	390	476	530	660	730	570	497	440	300	640
LESN3	F1	287	444	420	602	839	724	505	472	374	323	708
Difference		**17**	**54**	**56**	**73**	**179**	**16**	**65**	**25**	**66**	**23**	**68**
LESN2	F2	2658	2106	2389	2069	1571	1066	974	800	916	1019	1214
LESN3	F2	2508	2117	2181	1574	1548	1097	1028	837	1225	1169	1208
Difference		**150**	**11**	**208**	**495**	**23**	**31**	**54**	**37**	**309**	**150**	**6**
GAE	F2	2290	1990	2089	1840	1720	1090	840	910	1020	870	1190
LESN3	F2	2508	2117	2181	1574	1548	1097	1028	837	1225	1169	1208
Difference		**218**	**127**	**92**	**266**	**172**	**7**	**188**	**73**	**205**	**299**	**18**

Table 4: LESN2 vs. LESN3

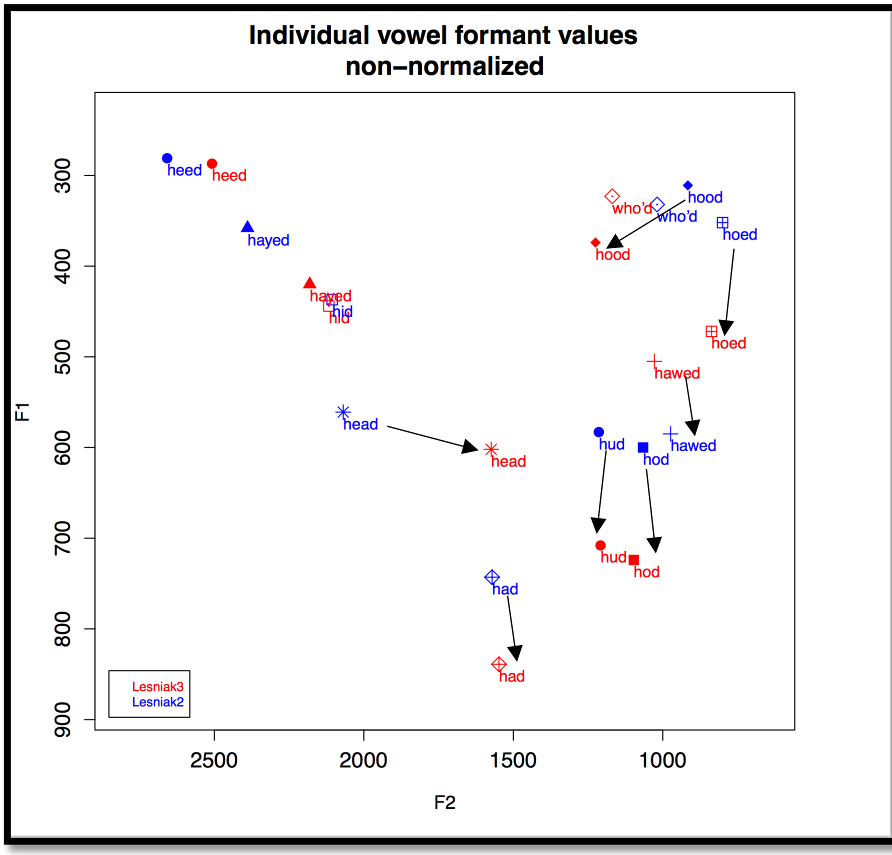

Figure 7: LESN2 vs. LESN3 (color coordinated)

The most important observation regarding LESN2 and LESN3 is that Author 2's acoustic vowel space is more spread out. In fact, his pronunciation in running speech (LESN3) is more aligned with GAE vowels than his pronunciation of vowels in their citation form, as seen clearly in Figure 8:

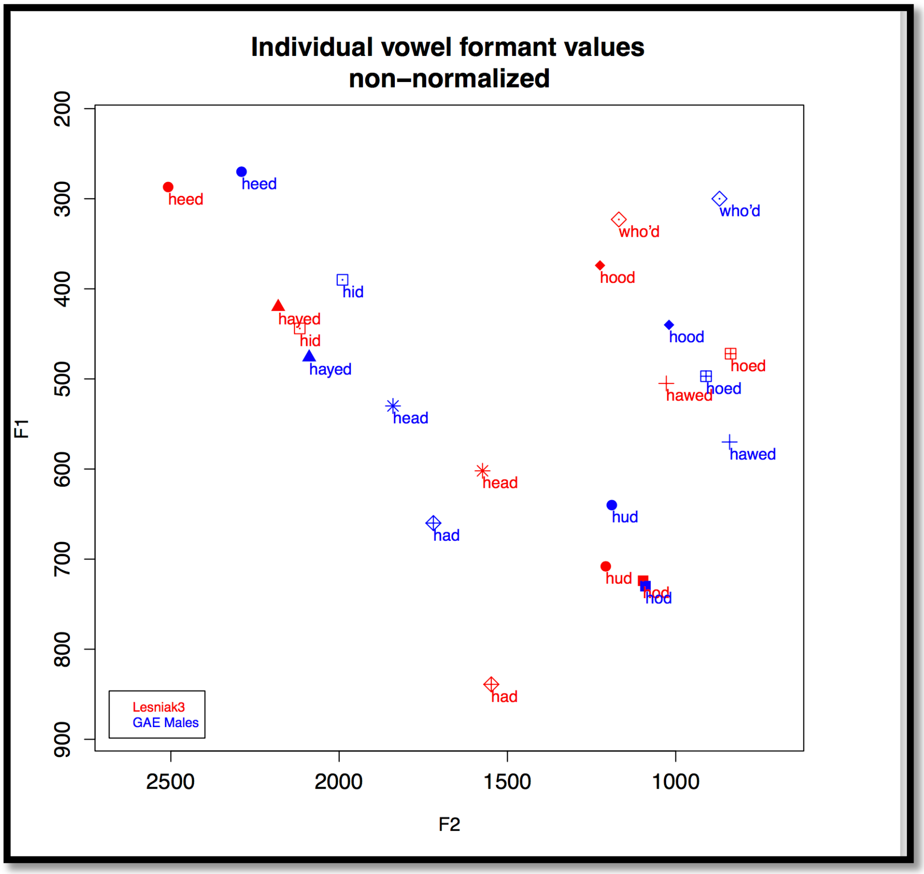

Figure 6: Author 2 vs. GAE (color coordinated)

In other words, Author's 2 vowels are more intelligible to GAE hearers in running speech than they are in citation form. Only four vowels appear to be mildly problematic. GAE hearers may on occasion confuse his [ɪ] (444 Hz) with his [e] (420 Hz) because the acoustic distance between them is 24 Hz. With an RFL of 80%, this confusion can lead to unintelligibility at times. His [o] (472 Hz), [ɔ] (505 Hz), and [ʊ] (440 Hz) can also lead to poor intelligibility. We also see that his [ɔ] (505 Hz) masks [o] (497 Hz) in GAE by 8 Hz. Since their RFL is 88%, intelligibility can be obscured. His [o] (472 Hz) masks [ʊ] (440 Hz) in GAE by 32 Hz. But this confusion is inconsequential for intelligibility because the RFL of the two vowels is only 12%. Last but not least, his [ʌ] (708 Hz) masks [ɑ] (730 Hz) in GAE by 22 Hz. With an RFL of 65%, his pronunciation can compromise intelligibility.

9.0 Discussions

Within the span of seven years, Author 2 has implemented (consciously and subconsciously) a number of phonological changes that have had a huge beneficial impact on the intelligibility of his vowels. During these seven years, he has learned to front his [i], to raise his [ɛ], and to control the excessive masking of his low vowels [æ], [ɔ], and [ɑ]. He has also managed to lower his [ʌ]. In spite of his best effort, his [ʊ] has been resistant to change. It is still acoustically indistinguishable from [u]. His achievements and his remaining challenge are highlighted in Table 5:

Words		heed	hid	Hayed	Head	Had	hod	hawed	hoed	Hood	who'd	Hud
Vowels		[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
LESN1	F1	290	528	375	649	775	631	621	405	322	335	424
LESN2	F1	281	437	358	561	743	600	585	352	311	332	583
LESN3	F1	287	444	420	602	839	724	505	472	374	323	708
GAE	**F1**	**270**	**390**	**476**	**530**	**660**	**730**	**570**	**497**	**440**	**300**	**640**
LESN1	F2	1373	2060	2082	2000	1204	1100	983	884	991	1111	1565
LESN2	F2	2658	2106	2389	2069	1571	1066	974	800	916	1019	1214
LESN3	F2	2508	2117	2181	1574	1548	1097	1028	837	1225	1169	1208
GAE	**F2**	**2290**	**1990**	**2089**	**1840**	**1720**	**1090**	**840**	**910**	**1020**	**870**	**1190**

Table 5: Intractable Vowels

What has contributed to Author 2's success? Two factors readily come to mind. The first is his social network. Since moving back to the USA in 2012, the language used by the people in his social network is English. He got married to a native speaker of American English who provides him with both quantity and quality of input. His professional language is also English. An analysis of his social network in Appendix 3 shows that he spends 41% of his weekly interactions with native speakers of GAE. Most of the remaining 54% is spent with a very fluent non-native English speaker. Only a small fraction of his weekly interactional hours are spent using Spanish. The second important factor is the acoustic phonetic project that Author 2 did while he was in Author 1's acoustic phonetic course in 2011. This course gave him the opportunity to visualize his pronunciation of English vowels and to see where his vowel pronunciation problems lay. Ladefoged and Johnson (2015:234) recommend making acoustic vowel spaces as a way of teaching English vowels to non-native speakers:

> … When teaching English as a second language, one might use the vowels of the first language of the students as reference points for comparison with the dialect of English that one is trying to teach. If a chart of the vowels of this language is not available, then the instructor's first step should be to make one.

Being able to see for himself which ones of his vowels mask each other internally and which ones mask GAE vowels in acoustic vowel spaces such as the ones described throughout this paper has empowered Author 2 to make the changes needed to enhance the intelligibility of his vowels.

10. Summary

Several insights can be gathered from this longitudinal idiolectal acoustic phonetic study of vowels. One insight is that with some English vowels, success comes easily. With others, progress is slow, while with others progress is almost impossible. Munroe's (2008) longitudinal acquisition of English vowels by Mandarin and Slavic speakers confirms Author 2's experience. He reported that the vowels [ɪ, æ, ɛ, ʌ, ɑ, ʊ] were particularly difficult for the participants in his research (pp. 489-493). He noted that improvements were made in the production of [ɪ], but the pronunciation of [ʊ] was impervious to change. Similar observations can be made about Author 2's acquisition of GAE vowels. In spite of remarkable progress, the vowels [ʌ] and [ʊ] remain intractable. His [ʌ] still masks [ɑ] in GAE. He has not yet found the secret of how to prevent his [ʊ] and [u] from masking each other. The other important insight that one can glean from this study has to do with the acoustic phonetics of vowels in general. This study has shown, albeit tentatively, that L2-accented vowels may be more intelligible to native speakers in running speech than in citation form. More research is needed. The third significant insight is that the speaker's social network makes a huge difference. Last but not least, the improved intelligibility of Author 2's vowels can also be attributed to visualization as suggested by Ladefoged and Johnson.

ABOUT THE AUTHORS

Ettien Koffi, Ph.D. in linguistics from Indiana University, teaches linguistics at Saint Cloud State University, MN. Author of many peer-reviewed articles on various topics in linguistics and of four books: *Language and Society in Biblical Times* (1996), *Paradigm Shift in Language Planning and Policy: Game Theoretic Solutions* (2012), *Applied English Syntax* (2010, 2015), and the *New Testament in Anyi Morofu* (2017), a task which took over 25 years. Specializing in acoustic phonetics, dialect variation, and emergent orthographies, his current research centers on speech acoustics of L2 English (within the Speech Intelligibility Framework), Central Minnesota English, and Anyi. He can be reached at enkoffi@stcloudstate.edu.

Fernando Lesniak received his Bachelor of Arts in Linguistics and Spanish from Saint Cloud State University in December of 2015. After graduation, he began teaching Spanish and later English as a Second Language (ESL) in several school districts in Central Minnesota. Currently, he is pursuing his Master's degree in Applied Linguistics at Saint Cloud State University and writing his thesis under the supervision of Dr. Ettien Koffi. He has been admitted into the Ph.D. program at the University of Minnesota, Second Language Education Program in the Curriculum and Instruction Department. He can be reached at FernandoLesniak@gmail.com.

Reference

Boersma, Paul & Weenink, David (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.17, retrieved 21 April 2016 from http://www.praat.org/

Everest, Alton F. and Ken C. Pohlmann. 2015. *Master Handbook of Acoustics.* 6th edition. New York: McGraw Hill Education.

Flanagan, James L. 1955. A Difference Limen for Vowel Formant Frequency. *The Journal of the Acoustical Society of America* 27 (3):613-617.

Fromkin, V. R. Rodman, and N. Hyams. 2014. An Introduction to Language. 10th edition. Boston, MA: Cengage Learning.

Giacomino, Lindsay. 2012. Comparative Analysis of Vowel Space of L1 Spanish Speakers of General American English. *Linguistic Portfolios* 1:98-112.

Hawks, John W. 1994. Difference Limens for Formant Patterns of Vowels Sounds. *Journal of the Acoustical Society of America* 95 (2): 1074-1084.

Hillenbrand, James, Laura A. Getty, Michael J. Clark, and Kimberlee Wheeler. 1995. Acoustic Characteristics of American English Vowels. *Journal of the American Acoustical Society*, 97 (5) 3099-3111.

Kent, Ray D. and Charles Read. 2002. *Acoustic Analysis of Speech*. Second Edition. Clifton Park, NY: Delmar Cengage Learning.

Kewley-Port, Diane and Charles S. Watson. 1994. Formant-Frequency Discrimination for Isolated English Vowels. *Journal of the Acoustical Society of America* 95 (1):485-496.

Labov, William; Ingrid Rosenfelder; and Josef Fruehwadl. 2013. One Hundred Years of Sound Change in Philadelphia. *Language* 48. 30-65.

Ladefoged, Peter and Keith Johnson. 2015. *A Course in Phonetics.* Seventh Edition. Malden, MA: Cengage Learning.

Ladefoged, Peter and Sandra Ferrari Disner. 2012. *Vowels and Consonants.* Third Edition. Malden, MA: Wiley-Blackwell.

Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques.* Malden, MA: Blackwell Publishing.

Ladefoged, Peter, Iris Kemeny, and William Brackenridge. 1976. Acoustic Effects of Style of Speech. *Journal of the Acoustical Society of America* 59 (1):228-231.

Lai, Yi-hsiu. 2010. English Vowel Discrimination and Assimilation by Chinese-Speaking Learners of English. *Concentric: Studies in Linguistics* 36 (2):157-182.

Lisker, Leigh and Arthur Abramson. 1964. A Cross-Language Study of Voicing in Initials Stops: Acoustical Measurements. *Word 20*: 384-422.

Mermelstein, Paul. 1978. Difference Limens for Formant Frequency of Steady-state and Consonant-bound Vowels. *Journal of the Acoustical Society of America* 63 (2):527-680.

Munroe, Murray J. 2008. Segmental Acquisition in Adult ESL Learners: A Longitudinal Study of Vowel Production. *Language Learning* 58 (3):479-502.

Pena-Brooks, Adriana and M. N. Hedge. 2000. *Assessment and Treatment Articulation and Phonological Disorders in Children*. Austin, Texas: Pro.Ed, An International Publisher.

Peterson, Gordon F and Harold L. Barney. 1952. Control Methods Used in a Study of the Vowels. *Journal of the American Acoustical Society*, 24 (2): 175-184.

Thomas, Erik R. and Tyler Kendall. 2007. NORM: The vowel normalization and plotting suite. [Online Resource: http://ncslaap.lib.ncsu.edu/tools/norm/]

Weinberger, Steven. (2019). *Speech Accent Archive*. George Mason University. Retrieved from http://accent.gmu.edu

Words	**heed**	**hid**	**Hayed**	**Head**	**Had**	**hod**	**hawed**	**hoed**	**Hood**	**who'd**	**Hud**
	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Duration	285	167	371	240	282	260	288	336	283	167	236
Intensity	55	55	55	54	53	52	56	57	64	60	54

Appendix 1: Duration and Intensity Measurements (2017)

Words	**heed**	**hid**	**hayed**	**head**	**had**	**hod**	**hawed**	**hoed**	**hood**	**who'd**	**hud**
Vowels	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Duration	250	157	228	121	121	225	129	213	248	324	101
Intensity	72	70	73	71	72	75	68	68	74	74	76

Appendix 2: Duration and Intensity Measurements (2018)

Participants	Justin	Cassandra	Kim	Nataliya
Countries	USA	USA	USA	Ukraine
Native Language	English	English	English	Ukrainian
Frequency of Weekly Interactions	1	3	3	5
Hours per Weekly Interactions	3	15	15	40
Interactional Contexts	Social Family Celebrations	Social Holidays Leisure	Social Holidays Leisure	Work Gym
Language(s) Used	English	English	English	English

Appendix 3: Author 2's Social Network Analysis